Cognitive Science and Technology

Amit Kumar
Gheorghita Ghinea
Suresh Merugu *Editors*

Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1

ICCIC 2023, 8–9 December, Hyderabad, India

Springer

Cognitive Science and Technology

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 Third International Conference on Cognitive and Intelligent Computing, Volume 1

ICCIC 2023, 8–9 December, Hyderabad, India

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
University of Southampton Malaysia
Johor, Malaysia

ISSN 2195-3988 ISSN 2195-3996 (electronic)
Cognitive Science and Technology
ISBN 978-981-97-9261-0 ISBN 978-981-97-9262-7 (eBook)
https://doi.org/10.1007/978-981-97-9262-7

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

Bus Route Detection Using YOLOv5 and PaddleOCR

B. Komal Adithya Reddy, Mohammad Sohail, and Morarjee Kolla

Abstract Recognizing buses and other public transport vehicles can be a difficult task, especially in crowded and noisy environments. The primary objective of this research is to develop a robust and efficient system that can accurately detect bus routes and assist people in navigating public transportation. Our proposed system integrates a convolutional neural network-based deep learning model with optical character recognition technologies to detect bus routes efficiently in real-world scenarios. We employed YOLOv5 for object detection, which provides enhanced efficiency and accuracy in identifying buses and their routes. Additionally, we used PaddleOCR for precise character recognition to extract textual information from images, such as route numbers and destination names. By combining YOLOv5 and PaddleOCR, we have created a system that significantly improves the detection and recognition of bus routes compared to other competitive models. Our results demonstrate that the integrated approach not only enhances detection accuracy but also proves the efficacy of our model in practical applications. This research contributes to the development of intelligent transportation systems, making public transit more accessible and user-friendly for everyone.

Keywords Convolutional neural network · Deep learning · Optical character recognition · YOLOv5 · PaddleOCR intelligent transportation systems

1 Introduction

Public bus transportation is a fundamental aspect of daily commuting, offering an affordable and widespread means of travel. However, commuters often face challenges in accurately identifying specific bus routes, which impacts their overall travel experience. Our paper delves into the innovative combination of You Only Look Once

B. Komal Adithya Reddy · M. Sohail · M. Kolla (✉)
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: morarjeek_cse@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_1

(YOLOv5) [3] and Optical Character Recognition (PaddleOCR) technologies to ease the process of bus route detection.

In 2020, around 2.2 million buses were owned collectively by the public and private sectors across India. According to government data, approximately 0.23 million buses are run by state transport undertakings. Union Transport Minister Nitin Gadkari highlighted that while China has about six buses for every 1000 people, India, has only four buses for every 10,000 people [1]. To address this disparity, the government plans to increase the count by 50,000 electric buses with US support [2]. Therefore, it is even more crucial to have a system that can accurately detect bus routes to enhance the travel experience for such large crowds.

In the complex and dynamic landscape of urban environments, correctly identifying the appropriate bus route poses a significant challenge for commuters. Existing methods often fall short in providing an efficient and user-friendly solution, particularly for individuals who may not be proficient with smartphone applications.

In India, a significant portion of the population faces challenges in navigating public bus transportation due to a lack of awareness about bus routes. Many individuals may not be proficient in reading or understanding bus route information displayed on panels, contributing to difficulties during their travel. This underscores the need for an accessible and user-friendly solution that can empower a diverse range of commuters to navigate the public bus system effectively, addressing the prevalent issues of route awareness and comprehension. The integration of advanced technologies, as explored in this implementation, seeks to bridge this knowledge gap, and enhance the overall travel experience for a broad spectrum of individuals across the country.

The incorporation of YOLOv5, an advanced object detection algorithm, introduces real-time precision to the detection of buses and their respective routes. By complementing this with PaddleOCR, a sophisticated optical character recognition technology, the system gains the ability to extract textual information from bus panels, including route numbers and other pertinent details.

The significance of this implementation lies in its potential to reshape the user experience with public bus transportation. The seamless integration of YOLOv5 and PaddleOCR aims to establish a comprehensive solution that not only ensures the accurate identification of bus routes but also prioritizes accessibility for a diverse user base.

This paper comprehensively details the methodology and results derived from the application of YOLOv5 and PaddleOCR. Through this system, we aim to address the specific challenges faced by commuters in our diverse transportation landscape. By navigating through the intricacies of implementation, our ultimate objective is to contribute to a more inclusive, user-friendly, and efficient public transportation experience for individuals from various walks of life.

2 Related Work

Various technological solutions have been developed to tackle challenges in transportation and accessibility. These systems span from RFID-based bus identification to mobile applications providing real-time information, each contributing to the evolution of efficient and user-friendly public transportation networks. Key contributions are described below:

Harini et al. [4] introduced an RFID-based bus detection system to aid visually impaired individuals in identifying and boarding buses at stops. This system features RFID readers at bus stations and a dedicated bus detection device. Their methodology describes the process in detail, demonstrating successful RFID token reading and triggering of audio messages, emphasizing RFID's potential to enhance accessibility. Future improvements include sensor technology advancements and broader accessibility considerations.

Suriyakala et al. [5] developed a NodeMCU-powered bus identification system designed to assist visually impaired individuals. The system features subsystems for bus tracking and communication with a remote server, enabling voice announcements upon bus arrival. The architecture includes components such as a tag, interrogator, controller, and antenna, showcasing the system's effectiveness through practical demonstrations.

Kim et al. [6] proposed a novel model for optimizing visual-auditory sensory substitution using a Generative Adversarial Network (GAN). This model converts visual data into audio signals and generates synthesized visual data using a GAN. Experiments showed the efficiency of a wide frequency range in visual-auditory conversion and a new Mel-scale frequency mapping scheme that outperformed traditional functions.

Baseer et al. [7] created a GPS-based bus tracking system to enhance bus management systems. Using technologies like MongoDB, ExpressJS, AngularJS, and NodeJS, their system includes User and Admin modules with features such as route selection, tracking, and login. The implementation resulted in improved system effectiveness and performance.

López et al. [8] designed an architecture comprising a core application server, a public transport network server, and a mobile application. This system handles data updates, external device access, vehicle location calculations, and notifications using GTFS data. It implements a RESTful API and push notifications for efficient communication, ensuring scalability through replication and distribution mechanisms.

Manikandan et al. [9] proposed a system with two detection subsystems: the Personal Assistance Segment (PAS) for visually impaired individuals and the Bus Driver Segment (BDS). The system includes a controller, RFID, RF module, and Voice Playback module. It creates a wireless connection between the visually impaired person and the driver, allowing for independent boarding.

More et al. [10] developed a bus detection system to empower visually impaired individuals in navigating public transportation independently. The system provides comprehensive information about bus presence, arrivals, and routes, aligned with the Digital India initiative. It ensures safe boarding through real-time announcements, user alerts, and obstacle detection.

Stephanie et al. [11] conducted interviews to gather main requirements and validate functionalities for a user-friendly app for bus navigation. The app, developed with Flutter in Dart, utilizes REST APIs and interfaces with map services and SMS, prioritizing security, ease of use, and responsiveness to users' needs.

Tan et al. [12] addressed bus boarding challenges for visually impaired individuals with a support system utilizing an ego-camera system called MY VISION. This system detects buses and identifies route numbers through a multi-stage process, demonstrating effectiveness in experiments.

Agarwal et al. [13] proposed a Bluetooth communication-based system for bus identification. This system allows visually impaired individuals to access public transport through a microcontroller and ultrasonic sensor setup, with suggestions for future improvements like a mobile app with image processing for bus recognition.

Martínez-cruz et al. [14] introduced the SUBE system, a mobile application integrated with BLE beacons in transport vehicles and bus stops, providing real-time information through verbal instructions. Tested in real-city environments, SUBE effectively guides users in boarding specific buses and independently descending at stops.

Murugeshwari et al. [15] described a system with a Transmitter and Receiver Section. A PIC16F887 microcontroller in the bus transmits bus name information to a Blind walking stick via an RF transmitter. The system provides auditory cues for bus identification and incorporates a vibration sensor for enhanced safety.

Wongta et al. [16] utilized TensorFlow's Faster RCNN with a pre-trained ResNet101 model for bus panel detection. Their approach includes perspective correction and contrast optimization, integrating Google Cloud Vision API's OCR capabilities for precise bus number identification.

Kaushik et al. [17] introduced a framework with Bus, Server, and User modules. The Bus Module, installed on the conductor's smartphone, initiates journeys and transmits real-time GPS coordinates to the Server Module. The User Module offers interfaces for both visually impaired users and normal users, enhancing public transportation accessibility and efficiency.

These studies collectively highlight diverse innovations aimed at improving public transportation accessibility for the visually impaired. Approaches include sensory substitution through GANs, user-friendly architectures, wireless connection systems, mobile applications, Bluetooth communication, BLE beacons, and advanced image processing. These methodologies aim to enhance independence, safety, and the overall travel experience for visually impaired individuals using public transportation.

3 Methodology

Our proposed system is based on the integration of the YOLOv5 algorithm and PaddleOCR, combining powerful object detection and optical character recognition capabilities. The YOLOv5 algorithm is widely recognized for its real-time object detection prowess, efficiently identifying, and detecting objects in images. On the other hand, PaddleOCR enhances the system's ability to extract textual information from detected objects. In our approach, we leverage the strengths of both the YOLOv5 algorithm and PaddleOCR, which together form the foundation of our framework, providing a versatile and effective solution for object recognition and data extraction tasks. The architecture of our proposed system is illustrated in Fig. 1.

The process commences with capturing images of bus panels using a mounted camera module either onboard buses or at bus stops. These images serve as the primary input for subsequent processing stages, allowing for the analysis and extraction of pertinent details, such as bus route numbers.

Following image capture, an essential preprocessing step involves converting the images from the Blue-Green-Red (BGR) color space to Red-Green-Blue (RGB) format. This conversion ensures uniformity in color representation and compatibility with subsequent image processing algorithms.

To streamline computational efficiency and resource utilization, the captured images undergo resizing to a standardized resolution. This resizing process minimizes the computational burden of subsequent processing stages while retaining essential details necessary for accurate bus route detection.

Utilizing computer vision techniques, such as object detection and image segmentation, the system identifies and localizes bus panels and number panels within the resized images. This step is critical for isolating the areas of interest containing bus route information.

The localized number panels are then processed through the PaddleOCR system, leveraging deep learning models to accurately recognize text in various languages

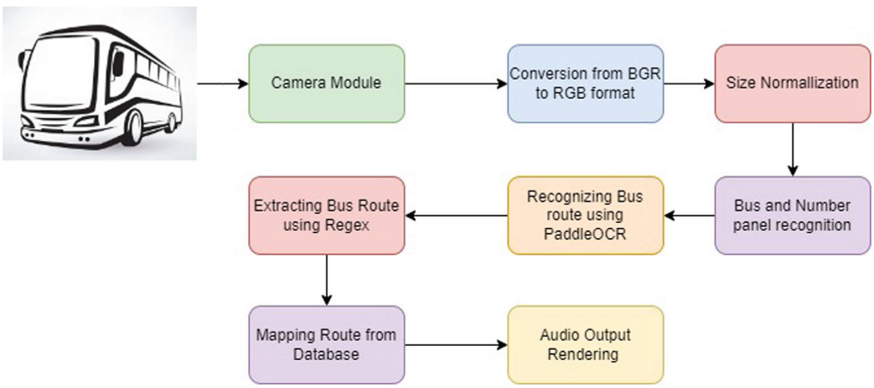


Fig. 1 Block diagram of bus route detection using YOLOv5 and PaddleOCR

and fonts. PaddleOCR plays a pivotal role in extracting textual information embedded within the images, with a specific focus on bus route numbers.

Following text recognition, the extracted bus route information undergoes further processing using regular expressions (regex). These patterns enable the system to isolate and extract relevant route numbers or identifiers from the recognized text, facilitating precise route identification.

To improve the efficiency of the OCR model, the Levenshtein algorithm has been employed to correct city names, thereby enhancing the accuracy of results.

The extracted bus route information is subsequently mapped to corresponding routes stored in a database. This database contains comprehensive information, including route numbers, destinations, and schedules, enabling accurate association of the detected routes with their corresponding details.

Finally, the identified bus route is converted into audio format, which is rendered through speakers or audio output devices. This audio announcement provides passengers with clear and concise information about the upcoming bus route, thereby enhancing accessibility and convenience for all commuters.

4 Implementation Details

For our experiments, we conducted the computational tasks on a system equipped with an Intel i5 processor, an Nvidia K80 GPU, and utilized Python version 3.10 as the programming language. PyTorch, a popular deep learning library, was used by us to implement and execute the proposed system. This setup allowed for a practical and accessible exploration of our framework's performance in a suitable environment.

The dataset utilized in this project is an in-house creation, meticulously compiled and augmented by our team. It comprises images of local RTC buses and college buses, initially totaling over 300 instances. Through systematic application of data augmentation techniques, we have significantly expanded the dataset to include more than 1100 images. Our dataset comprises images capturing a diverse array of local Road Transport Corporation (RTC) buses operating within the targeted area as illustrated in Fig. 2. Additionally, we have included images of buses affiliated with our college, constituting a significant portion of the dataset. To enhance the diversity and robustness of our dataset, we applied various augmentation techniques. These techniques include horizontally mirroring original images to simulate different viewing perspectives and angles, applying rotations to simulate diverse orientations and viewpoints, introducing variations in brightness levels to account for differing lighting conditions, incorporating zooming in and out of images to simulate varying observation distances, and applying shearing transformations to introduce distortions, replicating real-world scenarios.

In our comparison methods, we tried out different versions of YOLO–like YOLO 3, 4, 5, 6, 7, and 8. However, due to discontinued support for YOLO versions 3 and 4, we encountered challenges in running our models on these versions. So, we could

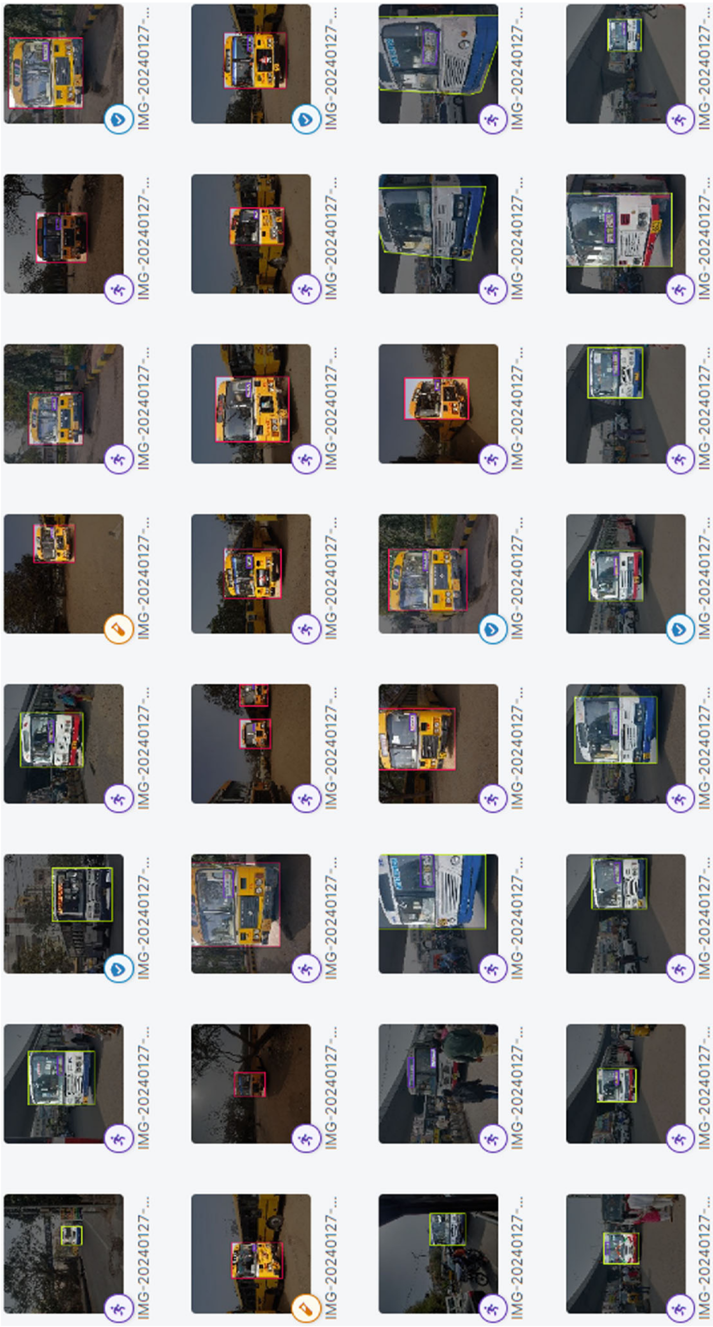


Fig. 2 Dataset of local RTC and college buses

not run our models on these versions. We still got results from the other versions–YOLO 5, 6, 7, and 8. We're looking at how well our method works compared to these different YOLO versions to figure out which one is doing the best job and delivers the best results.

The model considers a learning rate of 0.01 based on Stochastic gradient descent which is an optimization algorithm often used in machine learning applications to find the model parameters that correspond to the best fit between predicted and actual outputs producing minibatch size of 16. We have implemented the system until 150 epochs with early stopping.

In our evaluation, we are basically checking how accurate and precise our models are, using metrics such as mAP50, Class Loss and Runtime. In Figs. 3, 4 and 5 we can the detection outputs of the system. We will establish the result in form of plots. We are doing this across different time points (i.e., epochs). It will help us track the performance of our models as they learn from the data over time. We are looking at which model does the best job as we let them learn more and more from the dataset. This helps us figure out how good our models become with more training.

Fig. 3 Bus and number panel detection by YOLOv5

Fig. 4 Confidence on detection of bus and bus panel

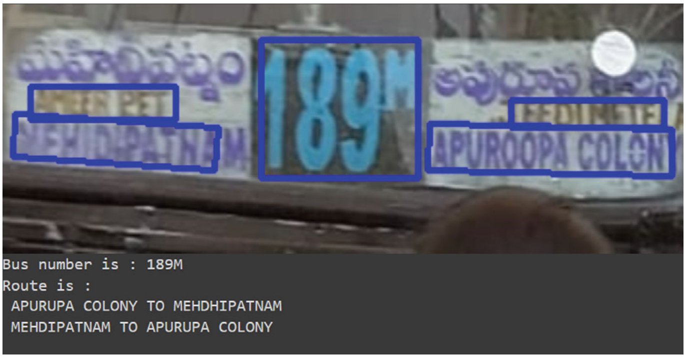


Fig. 5 Information extracted from bus panel by PaddleOCR

5 Result Analysis

In our study, we have compared how well different versions of YOLO model performs against each other. Our results are summarized in Table 1. We specifically looked at the mAP50 (measures the accuracy of a model by calculating the average precision across different levels of confidence with a 50\% confidence threshold), Class Loss (quantifies the difference between predicted and actual class labels), and Run Time (s). The results of Fig. 6 tell us that YOLOv8 has high accuracy with its consistent plot across different time points while other models display an upward trend, indicating that their performance increases with time and training. In Fig. 7, we can see the class loss measure of the models with YOLOv5 and YOLOv7 giving the best results. In

Table 1 Performance comparison of different YOLO models

Model	mAP50 (%)	Class loss (%)	Run time (s)
YOLOv5	90.1	0.081	3218.4
YOLOv6	91.7	56.29	3891.6
YOLOv7	94.72	0.1316	19306.8
YOLOv8	91.08	89.74	10750.3

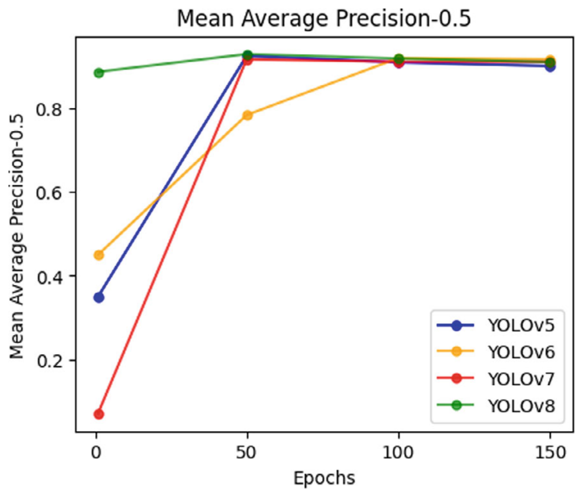


Fig. 6 Comparison of MAP scores of different YOLO models

Fig. 8, we see that YOLOv5 takes the least time to produce results, while YOLOv7 takes the highest time.

We ran tests with different training epochs—50, 100, 150. YOLOv8 is consistent in terms of accuracy but with training YOLOv5 has produced the best results. YOLOv5 has least class loss and run time as well, making it the choice for users as summarized in Table 1.

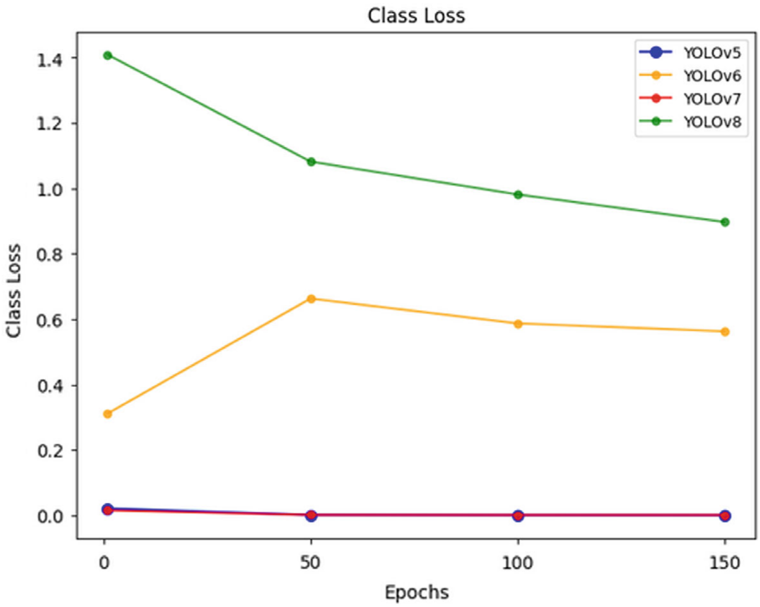


Fig. 7 Comparison of class loss of different YOLO models

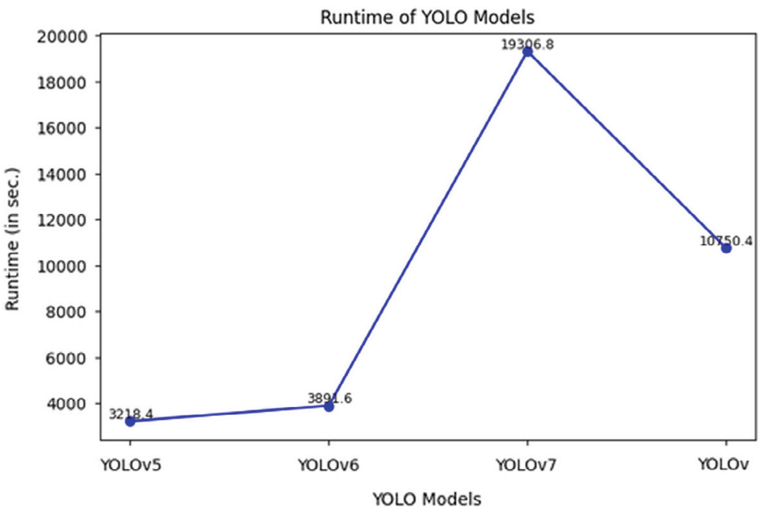


Fig. 8 Comparison of runtimes of different YOLO models

6 Conclusion

In conclusion, our proposed system represents a significant advancement in enhancing daily transportation experiences by seamlessly navigating public transportation. Leveraging cutting-edge technologies such as computer vision and the YOLO technique for efficient bus identification, our system streamlines the process of accessing bus route information. By translating route numbers into accessible formats, such as captions or audio cues, we aim to improve accessibility for diverse commuters and reduce reliance on smartphones.

Despite the promising capabilities demonstrated by the system, it has its inherent limitations. One significant constraint is the system's reliance on daylight conditions for optimal performance. During nighttime or low-light scenarios, when visibility is limited, the accuracy of bus detection may be compromised, affecting the overall effectiveness of the system. Additionally, the current scope of our system is confined to buses and bus stops within Hyderabad, Telangana. This geographical limitation restricts the applicability of the system to other cities or regions, potentially hindering its widespread adoption and impact.

Looking ahead, there are several avenues for expanding and improving our system. One significant area of future development involves enhancing the system's performance under low-light conditions to ensure consistent usability throughout the day and night. Additionally, scaling the system to cater to the needs of different cities across the country is paramount for broader adoption and impact. This expansion would require careful consideration of regional variations and infrastructure nuances to maintain performance and reliability. Furthermore, there is potential for extending the system's functionality to provide specialized interfaces tailored to visually impaired individuals, ensuring inclusivity and accessibility for all commuters. These future endeavors will contribute to realizing our vision of creating a more accessible, user-friendly, and inclusive public transportation system that meets the diverse needs of commuters across India.

References

1. Dash DK (2018) India needs 30 lakh buses for transport, has only 3 lakhs. Times of India. https://timesofindia.indiatimes.com/india/india-needs-30-lakh-buses-for-transport-has-only-3-lakh/articleshow/65726517.cms
2. Del Bello L (2023) India's plan for 50000 e-buses on its roads gets US support. The Economic Times. https://economictimes.indiatimes.com/industry/renewables/indias-plan-for-50000-e-buses-on-its-roads-gets-us-support/articleshow/105874584.cms?from=mdr
3. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR)
4. Harini N, Chandrasahoda S, Praneetha G, Nandini K (2023) Bus detection device for blind. IJCRT 11(11)
5. Suriyakala V, Stella R, Malligeshwari N (2023) Smart bus station boarding system for visually impaired passengers. IJERT 11(03)

6. Kim M, Park Y, Moon K, Jeong CY (2023) Deep learning-based optimization of visual–auditory sensory substitution. IEEE Access 11:14169–14180
7. Baseer KK, Jahir Pasha M, Jayasunitha M, Swetha K, Maria Joseph B (2023) Bus management system through MEAN technologies. In: 2023 International conference on intelligent systems for communication, IoT and security (ICISCoIS)
8. López AF, Ohmura R (2022) Public transport guiding system for visually impaired users easy to deploy, maintain and extend. In: 2022 IEEE international smart cities conference
9. Manikandan J, Bharathi J, Gayathri M, Lavanya G (2022) Bus identification system for visually impaired using Zigbee technology. In: 2022 International conference on communication, computing and internet of things
10. More CS, Bhadouria H, Aisha C, Nirne DD (2022) Bus identification & detection system for visually impaired people. IJSREM 06(07)
11. Mendonça S, Kunze V, Tejera V, Mayr C (2022) Assisted mobility: an application to help the blind to use public transport. In: 2022 IEEE global humanitarian technology conference (GHTC)
12. Tan JK, Hamasaki Y, Zhou Y, Kazuma I, Benitez-rangel JP, Camarillo-gómez KA (2021) A method of identifying a public bus route number employing MY VISION. Atl Press J Robot Netw Artif Life 8
13. Agarwal A, Agarwal K, Agarwal R, Patra AK, Mishra AK, Nahak N (2021) Wireless bus identification system for visually impaired person. In: 2021 1st Odisha international conference on electrical power engineering, communication and computing technology (ODICON)
14. Martínez-cruz S, Morales-hernández LA, Gerardo P-S, Benitez-rangel JP, Camarillo-gómez KA (2021) An outdoor navigation assistance system for visually impaired people in public transportation. IEEE Access J 9:130767–130777
15. Murugeshwari B, Monika Dharshni D, Jaya SP, Kavi Kumar V, Srinivasa Perumal M, Prema S (2020) An instrument to identify bus names for visually impaired person. In: 2020 6th international conference on advanced computing and communication systems (ICACCS)
16. Wongta P, Chalidabhongse TH (2018) Vision-based bus route number reader for visually impaired travelers. In: 2018 2nd international conference on imaging, signal processing and communication
17. Kaushik C, Kumar S, Gandhi S, Gandhi N, Rajput NK (2018) Automated public bus identification system for visually impaired. In: 2018 International conference on advances in computing, communications, and informatics (ICACCI)

Sentiment-Based Recommendation System for YouTube Comments

Vamshi Krishna Bathula, Rajavardhan Pagidipala, and Morarjee Kolla

Abstract In the ever-expanding digital landscape, content creators rely heavily on user feedback to understand audience sentiments and preferences. This research proposes to enhance sentiment analysis by integrating multi-class classification, emoji detection, and a recommendation system. This involves the development of a comprehensive workflow for processing user comments. Initially, a diverse dataset is collected, encompassing various sentiments, topics, and emoji expressions. The data undergoes thorough preprocessing, including text cleaning, tokenization, and emoji transformation. For sentiment analysis, models employing Support Vector Machines (SVM) and Random Forest classifiers are trained on the preprocessed text data. To further enrich user experience, a recommendation system is implemented, offering personalized content suggestions based on user engagement and feedback. The recommendation system employs content-based filtering algorithms, ensuring content creators receive tailored insights into their audience's preferences. This paper not only advances sentiment analysis capabilities but also empowers content creators with actionable insights for refining their content strategy. The combination of multi-class classification, emoji detection, and recommendation systems offers a holistic approach to understanding and engaging with user sentiments in the ever-changing field of producing digital content.

Keywords Random forest · Recommendation system · Sentiment analysis · Support vector machine · Tokenization

V. K. Bathula · R. Pagidipala · M. Kolla (✉)
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: morarjeek_cse@cbit.ac.in

V. K. Bathula
e-mail: ugs205147_cse.vamshi@cbit.org.in

R. Pagidipala
e-mail: ugs205166_cse.rajavardhan@cbit.org.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_2

1 Introduction

In the realm of digital content creation, understanding and interpreting user sentiments are paramount for content creators striving to connect with their audience. Our work addresses this critical need by presenting an innovative and comprehensive solution. The paper recognizes the evolving nature of user expression, encompassing not only textual sentiments but also the emotive richness conveyed through emojis. By amalgamating sentiment analysis with multi-class classification and emoji detection, we aim to provide content creators with a nuanced understanding of user feedback [1]. Our approach involves leveraging techniques including Support Vector Machines (SVM) Classifiers, to analyze and classify user comments [2]. These models go beyond traditional sentiment analysis by discerning multiple classes and incorporating the unique dynamics of emoji usage within textual content. Furthermore, the paper introduces a recommendation system tailored for content creators. This system not only gauges user sentiments but also offers personalized content suggestions based on audience engagement and feedback. Through content-based filtering algorithms, content creators gain valuable insights into the preferences of their audience, enabling them to refine and optimize their content strategy. The paper culminates in a unified system that seamlessly integrates sentiment analysis, multi-class classification, emoji detection, and recommendation features. A user-friendly dashboard visualizes the insights generated, empowering content creators with a holistic view of sentiment distribution, popular topics, emoji usage patterns, and personalized recommendations.

2 Related Work

Fang et al. [3] investigate sentiment analysis on YouTube social media, focusing on Nokia product comments. With the prevalence of opinions on social media, the study utilizes Decision Tree and Random Forest algorithms to classify sentiments as positivity, negativity, or neutrality. The research emphasizes the impact of social media, particularly YouTube, on product promotion. Nokia's journey, from being a leading telecommunications company to its decline in 2013 and subsequent revival, provides the backdrop for sentiment analysis. The study collects comments, translates them, and employs Natural Language Processing techniques for data preparation. The sentiment classification is based on the Decision Tree and Random Forest algorithms, with Decision Tree showing higher accuracy (89.4) compared to Random Forest (88.2).

Bouazizi et al. [4] focus on sentiment analysis in YouTube videos highlights three main levels of analysis: simple, complex, and advanced sentiment analysis. Studies employ various techniques, such as Naive Bayes, SVM, KNN, C4.5, word embedding, CNN, and SVM, across multiple languages like English, Indonesian, Portuguese, Arabic, and Bangla. While the majority of research focuses on simple sentiment analysis, there is a growing interest in advanced sentiment analysis, exploring correlations between sentiment, video duration, and user engagement. Notably, opportunities for future research include expanding multilingual sentiment analysis on YouTube and delving into the impact of sentiment on video popularity and hate crime detection [5].

Braig et al. [6] focus on opinion mining and sentiment analysis in the context of Twitter. It discusses the prevalence of research in binary and ternary sentiment classification (positive, negative, and neutral), emphasizing the challenges posed by the informal language, non-textual content, and slang expressions in microblogging data. The motivation for the study is to move beyond binary and ternary classifications and delve into multi-class sentiment analysis, aiming to distinguish between various sentiment classes. The authors introduce their novel approach, emphasizing the importance of classifying tweets into more than just positive, negative, and neutral categories. They propose a pattern-based method that divides tweets into seven sentiment classes based on writing styles and unique unigrams.

Qin et al. [7] propose a new method for generating a YouTube Recommender Network (YRN) that connects videos based on user comments. They build a recommender system on this network using intricate network analysis. The authors argue that their approach is more effective than existing systems that rely on textual features like video tags or user profiles. They also suggest that leveraging social networks power can improve recommender systems in various fields beyond video websites.

Liu et al. [8] proposed an emoji-embedding architecture named CEmo-LSTM to enhance the accuracy of sentiment identification and classification in sentiment analysis tasks. They aimed to investigate the impact of introducing emojis on sentiment analysis accuracy in both traditional rule-based and supervised learning algorithms. The study evaluated the effectiveness of embedding emojis in sentiment analysis algorithms and compared the performance of the CEmo-LSTM model with other mainstream sentiment analysis models in various experimental settings.

Kim et al. [9] propose alternative metrics and text mining methods to capture customer sentiment from online reviews. They suggest that relying solely on traditional metrics like online review ratings and volume may not provide an accurate representation of customer opinions due to limitations such as extremity bias and lack of depth in understanding customer sentiment. Specifically, the authors propose Sentiment Indicators from Text Mining, Word Count as a Metric and Utilization of Lexicon-Based Sentiment Analysis.

Mu et al. [10] discusses about a model called MSSSA-LSTM for predicting stock prices. The model combines multi-source data, deep learning, swarm intelligence algorithms, and sentiment analysis. The sentiment index is calculated using data from the East Money forum posts, and the Sparrow Search Algorithm optimizes the LSTM hyperparameters. The model integrates the fundamental trading data and mood index to predict stock prices. Experimental results show that the MS-SSA-LSTM model performs better than other models and has a broad range of applications.

Osman et al. [11] propose to integrate sentiment analysis into collaborative filtering (CF) recommendation methods to address the issues of data sparsity and cold-start users in recommender systems. By leveraging textual reviews and sentiment analysis, the study aims to enhance the accuracy and performance of recommender systems, particularly in domains where conventional recommender techniques face challenges in providing precise recommendations.

Ashok et al. [12] propose a personalized recommender system that uses machine learning and sentiment analysis to extract user reviews and comments on social media platforms, such as restaurants and points of interest, to provide personalized and relevant recommendations based on user preferences. The system aims to enhance the user experience by keeping information fresh and offering much-needed personalization. The authors also suggest incorporating preference- and context-based searching to improve accuracy and help users take advantage of the knowledge found in vast amounts of social data.

3 Proposed System

The proposed system of the paper consists of four major steps. In first step, the dataset is collected from Kaggle which contains 4 columns video id, comment, ratings and sentiment. The second step is preprocessing data, which involves cleaning the dataset by removing comments and eliminating non-essential characters such as newline, tabs, and spaces. Collecting data associated with each comment such as videoed, sentiment, and rating. The third step is training the model. CNNs are chosen for sentiment analysis due to their effectiveness in capturing spatial patterns and features from sequential data like text. CNNs can automatically learn hierarchical representations of text date. A sequential model is defined using Keras, Convolutional layers, Max Pooling layers, and Dense layers are added to the model architecture. Training process includes specifying batch size, number of epochs, and validation data for monitoring model performance. The final step is evaluating the model performance using Root Mean Square Error (RMSE) and predicting the sentiment of new comments.

3.1 Convolutional Neural Networks (CNNs)

CNN is a method for image and video processing. It is good at finding patterns and features in data with graphs similar to images, such as pictures. This network is made up of individual processes, each of which performs a specific task, such as feature recognition or image segmentation. Convolution programs are used to connect layers and use the input data to extract features. Additionally, this network has a layer used to reduce data size and improve the network's ability to identify hotspots. Prediction or classification of input data is the end of CNN.

3.2 Support Vector Machine (SVM)

A flexible and reliable supervised learning approach for both regression and classification applications is the Support Vector Machine (SVM). Finding the ideal line or decision boundary that divides an n-dimensional space into different classes is the main goal of support vector machines (SVMs). This makes it possible to classify new data points accurately in the future. The approach uses the most important points in the dataset, or support vectors, to calculate the ideal decision boundary, or hyperplane. The support vector machine technique gets its name from these support vectors, which play a crucial role in establishing the hyperplane.

3.3 K-Nearest Neighbors (KNNs)

A tracking technique for retrieval and classification is K-Nearest Neighbors (KNNs). When KNN is used to predict new data points, the method uses a defined distance measure (usually Euclidean distance) to find K training samples that are close together. In the classification function, most of the K nearest neighbors of the newly added data can predict its group. The expected result for the regression function is usually the average of the output values of the neighboring K. The number of neighbors, K, is one of the key variables influencing the model's success. KNN is a lazy learning method because it is flexible and unstructured during learning. However, it will depend on the distance scale and K selection.

The diverse approaches put out by different people to address the issue of sentiment analysis in YouTube comments. On the other hand, DL models provide integrated feature extraction and make the task simpler. Feature extraction is a necessary step in many machine learning techniques so that the models can be trained to categorize the data. The sentiment analysis framework has been easily linked with the hybrid recommendation system. It makes content recommendations based on the user's emotional preferences and shared interests.

4 Implementation Details

The dataset will enable users to build a predictive model that recommends which videos to the user based on their mood which is predicted through sentiment of their comments. This dataset contains 4 columns videoid, comment, ratings, and sentiment. The model analyzes the relation between the comments and sentiment and predict the videos. The training process undertaken to train the model to find out the sentiment for a comment. Clean the dataset by removing comments and eliminating non-essential characters such as newline, tabs, and spaces. CNNs are chosen for sentiment analysis due to their effectiveness in capturing spatial pattern and features from sequential data like text. CNNs can automatically learn hierarchical representations of text data, making them suitable for sentiment analysis tasks.

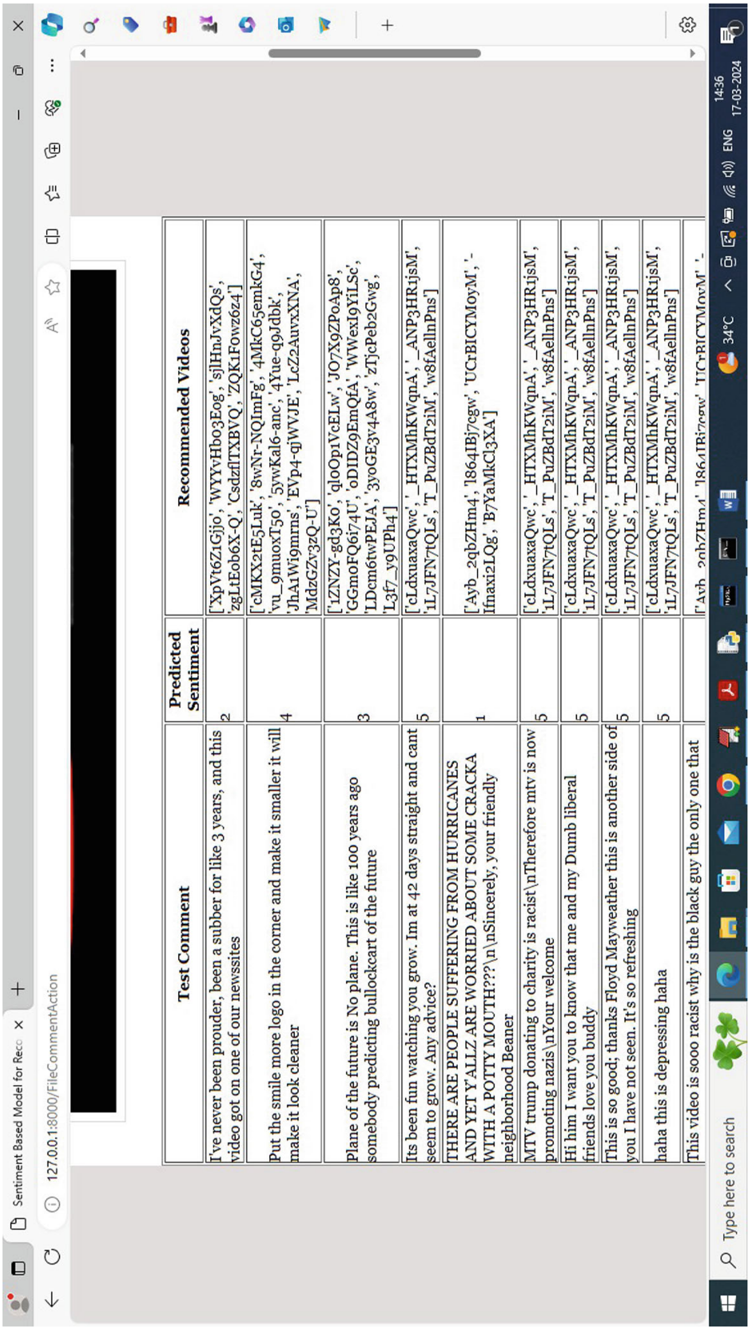
The dataset containing YouTube comments and their corresponding sentiment labels is preprocessed, missing values handled, and data is scaled using MinMax Scaler to ensure all features are within the same range. The preprocessed data is split into training and testing sets using train_test_split from scikit-learn. A sequential model is defined using Keras, Convolutional layers, MaxPooling layers, and Dense layers which are added to the model architecture. Training process includes specifying batch size, number of epochs, and validation data for monitoring model performance. Mean squared error (RMSE) is calculated to measure the accuracy of the model in predicting sentiment. The trained CNN model is used to predict the sentiment of new comments.

5 Result Analysis

The Proposed system presents a user-friendly interface for analyzing YouTube comments and recommending videos based on sentiment analysis. Leveraging machine learning algorithms such as CNN, the system accurately predicts sentiment scores for each comment and generates personalized video recommendations tailored to user's preferences and sentiments as shown in Figs. 1 and 2. Users can interact with the system by inputting comments and receiving real-time recommendations, enhancing their YouTube browsing experience and engagement with relevant content.

6 Conclusion

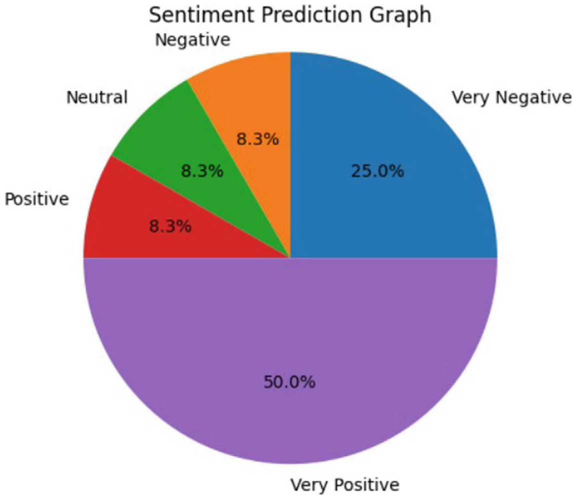
In conclusion, our approach of sentiment analysis system for YouTube represents a robust and multifaceted approach toward understanding user sentiments within the vast ocean of online content. The integration of multi-class sentiment classification, emoji detection, and a recommendation system will yield noteworthy insights into the dynamics of user engagement and content perception on the platform. The multi-class sentiment classification component will demonstrate commendable performance metrics. The recommendation system, seamlessly integrated into our sentiment analysis framework, represents a significant stride toward personalizing user experiences. By leveraging sentiment insights, the recommendation system suggests content that aligns with users' emotional preferences, thereby fostering a more tailored and engaging environment. As we reflect on the outcomes of our paper, it is crucial to acknowledge the potential applications of our sentiment analysis system. Beyond its immediate utility for content creators and advertisers on YouTube, our approach holds promise for shaping the landscape of sentiment analysis in other social media platforms. The adaptability and scalability of our model



Test Comment	Predicted Sentiment	Recommended Videos
I've never been prouder, been a subber for like 3 years, and this video got on one of our newssites	2	['XpVt6Z1Gjjo', 'WYYvHbo3Eog', 'sjlHnJvXdQs', 'zgLtEob6X-Q', 'CsdzflTXBVQ', 'ZQK1F0wz6z4']
Put the smile more logo in the corner and make it smaller it will make it look cleaner	4	['cMKX2tE5Luk', '8wNr-NQImFg', '4MkC65emkG4', 'vu_9muoxT5o', '5ywKal6-anc', '4Yue-q9Jdbk', 'JhA1Wigmrns', 'EVp4-qjWVJE', 'LcZ2AuvxXNA', 'MdzGZv3zQ-U']
Plane of the future is No plane. This is like 100 years ago somebody predicting bullockart of the future	3	['1ZNZY-gd3Ko', 'ql0Op1VcELw', 'JO7X9ZPoAp8', 'GGmoFQ6i74U', 'oDIDZ9EmQfA', 'WWexI9YiLSc', 'LDem6twPEJA', '3yoGE3v4A8w', 'zTjcPeb2Gwg', 'L3f7_y9UPh4']
Its been fun watching you grow. Im at 42 days straight and cant seem to grow. Any advice?	5	['cLdxuaxaQwc', '_HTXMhKWqnA', '_ANP3HR1jsM', '1L7JFN7tQLs', 'T_PuZBdT2iM', 'w8fAellnPns']
THERE ARE PEOPLE SUFFERING FROM HURRICANES AND YET Y'ALLZ ARE WORRIED ABOUT SOME CRACKA WITH A POTTY MOUTH???\n\nSincerely, your friendly neighborhood Beaner	1	['Ayb_2qbZHm4', 'l864IBj7cgw', 'UCrBICYMoyM', '-Ifnaxi2LQg', 'B7YaMkCl3XA']
MTV trump donating to charity is racist\nTherefore mtv is now promoting nazis\nYour welcome	5	['cLdxuaxaQwc', '_HTXMhKWqnA', '_ANP3HR1jsM', '1L7JFN7tQLs', 'T_PuZBdT2iM', 'w8fAellnPns']
Hi him I want you to know that me and my Dumb liberal friends love you buddy	5	['cLdxuaxaQwc', '_HTXMhKWqnA', '_ANP3HR1jsM', '1L7JFN7tQLs', 'T_PuZBdT2iM', 'w8fAellnPns']
This is so good; thanks Floyd Mayweather this is another side of you I have not seen. It's so refreshing	5	['cLdxuaxaQwc', '_HTXMhKWqnA', '_ANP3HR1jsM', '1L7JFN7tQLs', 'T_PuZBdT2iM', 'w8fAellnPns']
haha this is depressing haha	5	['cLdxuaxaQwc', '_HTXMhKWqnA', '_ANP3HR1jsM', '1L7JFN7tQLs', 'T_PuZBdT2iM', 'w8fAellnPns']
This video is sooo racist why is the black guy the only one that		['Ayb_2qbZHm4', 'l864IBj7cgw', 'UCrBICYMoyM', '-

Fig. 1 Output for the comments we have given

Fig. 2 Pie chart for sentiments of the comments



position it as a valuable tool in the hands of researchers and marketers seeking deeper insights into user sentiments in the digital realm.

References

1. Aufar M, Andreswari R, Pramesti D (2020) Sentiment analysis on youtube social media using decision tree and random forest algorithm: a case study. In: 2020 International conference on data science and its applications (ICoDSA), Bandung, Indonesia, 2020
2. Muhammad AN, Bukhori S, Pandunata P (2019) Sentiment analysis of positive and negative of YouTube comments using Naïve Bayes—support vector machine (NBSVM) classifier. In: 2019 International conference on computer science, information technology, and electrical engineering (ICOMITEE), Jember, Indonesia, 2019
3. Fang Y, Tan H, Zhang J (2018) Multi-strategy sentiment analysis of consumer reviews based on semantic fuzziness. Appl Sci
4. Bouazizi M, Ohtsuki T (2020) A pattern-based approach for multiclass sentiment analysis in twitter. Graduate school of science and technology. Keio University, 2020
5. Bouazizi M, Ohtsuki T (2018) Multi-class sentiment analysis in twitter: what if classification is not the answer. J Phys
6. Braig N, Benz A, Voth S, Breitenbach J, Buettner R (2023) Machine learning techniques for sentiment analysis of COVID-19-related twitter data. Agrono
7. Qin S, Menezes R, Silaghi M (2023) A recommender system for YouTube based on its network of reviewers. In: IEEE conference publication. IEEE, 2023
8. Liu C, Fang F, Lin X et al (2021) Improving sentiment analysis accuracy with emoji embedding. J Saf Sci Resil 246–252. ScienceDirect
9. Kim RY (2021) Using online reviews for customer sentiment analysis. IEEE, 2021
10. Mu G, Gao N, Wang Y, Dai L (2023) A stock price prediction model based on investor sentiment and optimized deep learning. IEEE, 2023
11. Osman NA, Noah SAM (2018) Sentiment-based model for recommender systems. In: 2018 4th international conference on information retrieval and knowledge management (CAMP). IEEE, 2018

12. Ashok M, Rajanna S, Joshi PV, Sowmya Kamath S (2016) A personalized recommender system using machine learning based sentiment analysis over social data. In: 2016 IEEE students' conference on electrical, electronics and computer science (SCEECS)

Analysing Twitter Data for Election Result Prediction

V. S. K. Harshavardhan Reddy, Vikas Swagath Rapelly, and Morarjee Kolla

Abstract Elections serve as a way for the people to express their opinions about the performance of a government and can alter their future to a great extent. In the last few years, there has been an exponential increase in the number of people openly expressing their political views through platforms like Twitter. This makes Twitter one of the best sources for analysing the public pulse. The major gaps observed in existing approaches include, inability to handle tweets in multiple languages, failing to handle fake Twitter handles, and failing to discard spam tweets. In addition, very few studies were performed with respect to Indian context. To overcome these gaps, we have come up with a solution that would specifically work in the Indian context and is capable of handling tweets in Hindi in addition to English. The data related to all major political parties is collected from Kaggle and various pre-processing steps are performed. The model uses Langdetect module to segregate Hindi tweets from the corpus. Subsequently, Googletrans module for translation. Our proposed system uses the RoBERTa model for performing opinion mining on the data sets. The results obtained will show the performance of our approach.

Keywords Elections · Kaggle · Indian context · Pre-processing · Twitter · Translation

1 Introduction

Elections play an important role in the day to day life of people. Elections are a way for people to express their opinion about the performance of a government. The elections impact the lives of people in many ways which may be good or bad. On the other hand, elections are also very important for the political parties. The ruling party tries to recapture the power, while the opposition tries to overthrow the existing

V. S. K. Harshavardhan Reddy · V. S. Rapelly · M. Kolla (✉)
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: morarjeek_cse@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_3

party. All the parties try to mine the opinion of the people before the polling to change their strategies accordingly. We observe that they are currently facing challenges in gauging public sentiment accurately. Conventionally, media reports and on-ground surveys used to help. But, after the unfortunate increase in the influence of money in politics, currently these reports are largely compromised. Almost every media institution is inclined towards some or the other political party. Therefore, leveraging machine learning-based techniques for sentiment analysis presents an opportunity to develop a more trustworthy and largely impartial methodology [1]. By analysing large amounts of data and using algorithms, machine learning can provide a more objective assessment understand and respond to the needs and desires of the people.

Twitter (the name of the company that owns it) was founded in 2006 which is an online social networking tool and has since expanded into social media. Today, it claims to have over 500 million monthly active users. On Twitter, users can share tweets that are similar to posts on Facebook or Instagram. These tweets contain words, images and videos In addition, registered users can post community notes that will provide more background information on posts by other registered users. The terminologies used were tweeting, retweeting and quote retweeting under the name of Twitter [2]. Twitter has become a platform where people express their opinions on various issues that are political, social, technological, educational, etc. Tweets of a person can be replied to retweeted or liked which will give more reach to the tweet on the platform and will show up in the accounts of the most active people or people who tweet about that particular topic. The news about any particular incident is immediately posted by the users on the platform. These users can be the common public, media or government servants. People generally put forward their opinion on a party or an individual of the party during elections and we can also see campaigning of a party by its party members or supporters on Twitter. Twitter differentiates from other social media primarily because of its wide acceptance as a formal platform for sharing announcements, political views and useful resources with others. Unlike other social media platforms that focus more on personal connections and casual interactions, Twitter's emphasis on concise and real-time updates allows it to serve as a powerful tool for analysing public sentiment. With its massive user base and the ability to track trending topics and hashtags, researchers and analysts can gain valuable insights into public opinions, attitudes and reactions. This makes Twitter an invaluable resource for understanding the pulse of society and gauging public sentiment on various issues.

Many studies have been done on the impact of Twitter and tweets on the elections and how the sentiments of the tweets can help in determining the outcome of the election. Studies have been done on the elections held in various countries with most of the studies focusing on the Presidential Elections of the United States, while other studies included countries like India, Pakistan, Nigeria and a few of the countries belonging to South America which include Argentina, Brazil and Mexico. Many algorithms have included Convolutional Neural Network (CNN) or an ensemble model that contains the CNN algorithm in their studies to show the prediction of the results and how well they can predict the results of the elections [3]. The studies also included various other machine learning and deep learning algorithms namely Naive

Bayes, Support Vector Machine, Regression, Long Short-Term Memory (LSTM), Bidirectional Encoder Representations from Transformers (BERT), Neural Networks to predict the sentiment of tweets that are being considered as datasets [4]. The studies have also proposed frameworks that integrate two or more algorithms to predict the election result more accurately such as CNN and clustering algorithms have been merged together to predict the sentiments of the tweets and thereby predict the results [5]. Datasets are built by scraping from Twitter using Twitter API or data sets that are available already are used. Pre-processing is done using Bag of N-words, Tokenization, TF-IDF, Vectorization, etc.

2 Related Work

Mudasir et al. [6] proposed a system poliWeet to predict the election results using sentiment analysis on Twitter data. The data collected is first sent into the data pre-processing stage in which the techniques namely tokenization, stop words removal and converting slang into meaningful words were done. The data set was built by them which consisted of 25,000 tweets of various sentiments regarding the elections. Feature extraction was done using the Word2Vec. The model used was pre-trained Bidirectional Encoder Representations from Transformers (BERT). The system was used to predict the 2019 Lok Sabha elections and showed an accuracy of 78\%. A further extension to this can be done by implementing a system that considers real-time tweets while predicting the election results.

Michelle et al. [7] analysed the various characteristics that help in classifying the accounts either as bots, trolls or humans by considering the data from Twitter data sets, namely IO, IRA and also from various other studies. The process included feature binding and dimensionality reduction using the Uniform Manifold Approximation and Projection) and Hierarchical DBSCAN was used for clustering as it identifies clusters that have variable density degrees. The results were analysed by using the purity score, silhouette coefficient and Jaccard similarity index. The main features that differentiate the trolls and humans are the presence of the URLs in the tweets and the number of retweeted accounts. The reply rate also plays a major role in differentiating trolls and humans as humans have a higher reply rate as compared to trolls. Between trolls and bots, the trolls have more retweets and their tweets and retweets have more URLs as compared to bots. Bots and humans are differentiated using two features—characters (mean tweets characters), the source of the tweets. Bots generate tweets from various sources whereas humans usually rely on a few interfaces only.

Olusola et al. [8] came up with an approach to predict the Nigeria 2023 presidential elections. The data set was taken using Twitter API over four months starting from June to September 2022 and contained a vast number of tweets. The data then underwent pre-processing which had techniques like word tokenization, lemmatization, stemming, and LDA implemented on the data. The models that were used are Bidirectional Encoder Representations from Transformers (BERT), Linear Support

Vector Classifier (LSVC) and Long Short-Term Memory (LSTM) for sentiment analysis. The tweets that were considered were of 3 types—tweets of personal accounts of candidates, the opinion of the public on the candidates and the opinion of the public on elections. Among the three models used, BERT gave the best results since it had the highest accuracy, precision, F-Measure, recall and AUC.

Shevtsov et al. [9] implemented a novel approach for detecting bot accounts on Twitter, utilizing the power of eXtreme Gradient Boosting (XGBoost) for classification. To ensure optimal performance, hyperparameter tuning is conducted through K-fold cross-validation. The classification outcome is further explained using the SHAP method, providing insights into the factors influencing the prediction. The implemented system has achieved high performance in accurately identifying bot accounts for different data sets, showcasing its high generalization ability. One major downside is that the model was tested exclusively against the US presidential elections.

Perera et al. [10] proposed a new method for feature extraction by utilizing fuzzy rough sets. The proposed approach is compared with multinomial NB and SVM techniques. The results show that the novel approach outperforms conventional methods, demonstrating better performance. However, it is worth mentioning that the method failed to detect spam and fake tweets effectively. Additionally, the study focused solely on Twitter data in English and considered only the Sri Lankan presidential election data. This limited scope suggests the need for further research to validate the approach's applicability in different contexts and datasets.

Brito et al. [11] have taken into consideration not only the Twitter data but also the data from other social media platforms, namely Instagram and Facebook. The aim is to predict the election results of various Latin American countries by using the data from various social media platforms and hence they have proposed a framework for this. The data sets were collected using the APIs of the social media platforms. They have performed principal component analysis (PCA) to reduce the dimensionality of the SM features. The methods used were general regression neural network (GRNN) and multilayer perceptron with back propagation (MLP-BP) were the methods used in modelling. It has been applied to some of the main recent presidential elections in Latin America: Argentina, Brazil, Colombia and Mexico.

Ali et al. [12] introduced a novel approach by proposing a deep learning-based model to accurately predict the results of the Pakistan general elections in 2018. The model utilized the Twitter API for data extraction, allowing for comprehensive and real-time information retrieval. To compare the performance of different models, Rapidminer was employed. The proposed method outperformed the Naive Bayes (NB) and Support Vector Machine (SVM) models, demonstrating its effectiveness. A completely new data set was created for this study, requiring manual labelling of the data for comparison with SVM and NB models. Despite using a relatively smaller data set, the results obtained were highly promising and showcased the potential of the proposed deep learning model. The work can be extended by training the proposed model with a larger dataset.

Bibi et al. [13] talk about the concept-based methods that consider the semantic meanings of the text whether they are positive or negative or neutral tweets about a group that is under study. The paper proposes a framework that considers unsupervised methods combined with concept-based and hierarchical clustering. This method has given better accuracy than the traditional methods that involved statistics and the combination of supervised and concept-based methods. Among the various ensembles, the Complete Linkage-Concept-based ensemble gave the highest accuracy.

Myilvahanan et al. [14] highlight the crucial role that data quality plays in determining the overall performance of a model. In this research, the proposition of employing the K-Nearest Neighbours (KNNs) algorithm for predicting election results is put forward. The paper outlines the advantages of using KNN, such as its simplicity, interpretability and ability to handle both numerical and categorical data. It provides valuable insights into the limitations identified in previous research papers. The major limitation is the lack of implementation. Future research can explore the implementation aspect, evaluating the theoretical findings.

Cano-Marin et al. [15] present a systematic literature review (SLR) that aims to identify gaps and opportunities for developing predictive systems using Twitter data. The review addresses several research questions. Firstly, it investigates the extent of research conducted on using Twitter as a predictive system. Secondly, it explores the main topics covered by empirical studies in this area. Thirdly, it examines the proposed applications of Twitter as a predictive system in the literature. Lastly, it identifies the most impactful journals in this field. The review is based on a search query constructed with specific keywords like predict or forecast. A major shortcoming is that new research articles may have been published after the period chosen for this systematic review.

3 Proposed System

The proposed system consists of the high-level overview and the complete architecture of the model.

3.1 High-Level Overview of Proposed System

The first step is to acquire data. Data was identified several data sets on Kaggle containing tweets related to US presidential elections as well as Indian elections. The second step is pre-processing that involves removing the hashtags, user accounts, mentions, hyperlinks, stop words which are not necessary for the sentiment analysis and including them may result in the increase of cost and impact the performance of the model. The third step is to train the models with the training data and analyse the metrics of the models which could give us the potential best model among the models in the study. The fourth step is to test the models by providing the testing

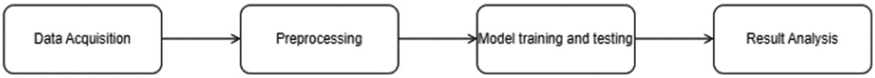


Fig. 1 High-level overview

data set and analyse the metrics by which we can conclude the best model among the given models as shown in Fig. 1.

3.2 Architecture

The data set is acquired after considering various sources including research papers and Kaggle. The data set chosen contains segregated tweets between two parties including the tweets in both Hindi and English languages. The tweets classified into positive and negative sentiment tweets on the basis of whether they are supporting or opposing the party. The pre-processing steps as shown in Fig. 2 involve removing the elements in the text that are not necessary for the sentiment analysis. The first step in the pre-processing involves the translation of the Hindi tweets into English as doing this later may involve in the loss of data that is required to classify the tweet. Langdetect was used to detect the tweets in Hindi and Googletrans to convert them into English. The further steps in the pre-processing involve the removing of URLs, user names or mentions, hashtags using the regular expressions. Stop words are removed by considering the stop words that are mentioned in the NLTK module. The text is then converted into lower case and the accented characters generally used for the pronunciation of the word are removed. The text is then tokenized by using the tokenizer that is most compatible with each of the model. The tokenizer involves the Byte Pair Encoding algorithm which is used for the model pre-training. After the pre-processing is completed, the data is divided into training and testing dataset. The models are trained with the training dataset and the training metrics are noted. The trained models are then given the testing dataset and the testing metrics are analysed that helps in identifying the best model. The best model is then saved and is used in predicting the result. The models in this study were first trained with specific hyperparameters on the train data. The best version of the model among the 200 epochs that were run was saved in the hugging face website and the test data was passed to it. The results on the test data were analysed and the best model was concluded.

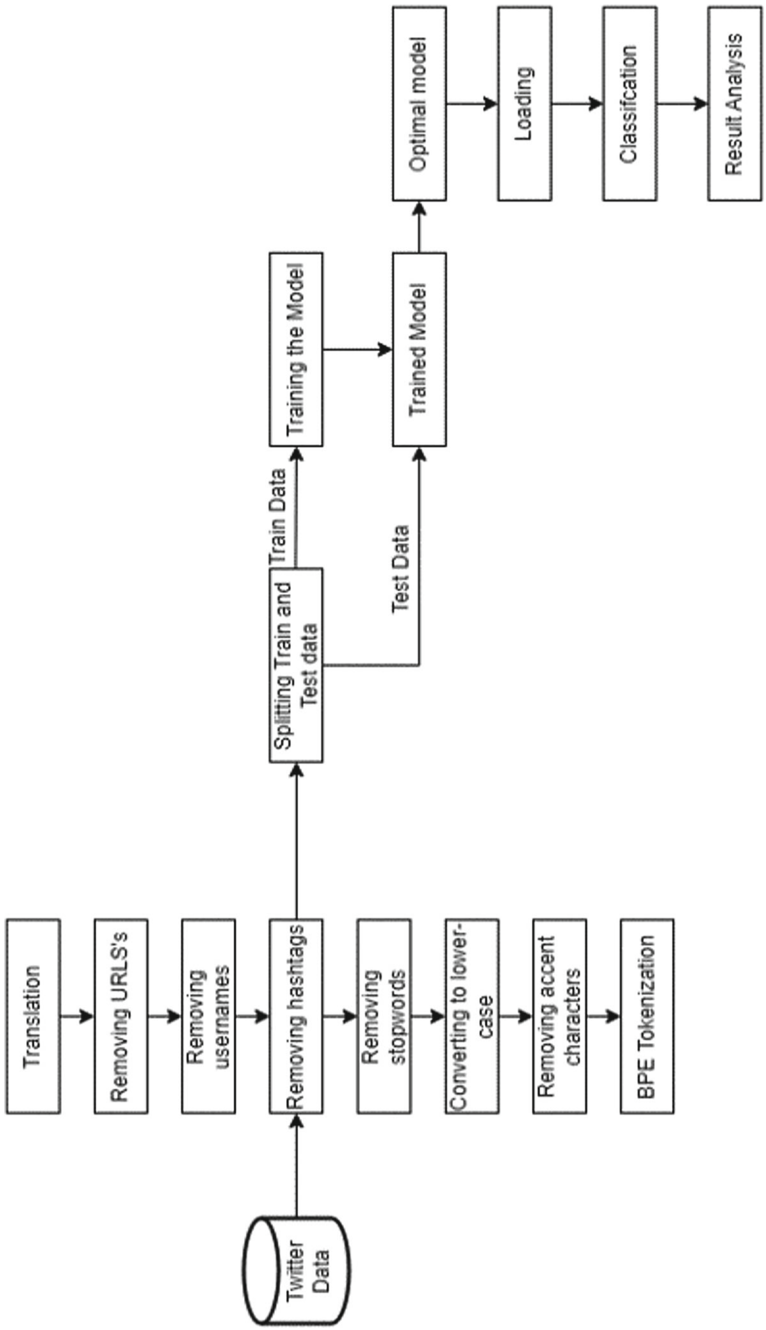


Fig. 2 Architecture of proposed system

4 Implementation and Result Analysis

The optimised version of RoBERTa was outperforming other models, achieving higher accuracy and lower loss. Therefore, RoBERTa model was used for classifying the tweets and predicting the election result. Comparison of the performance of various models against testing and training data is given in Tables 1 and 2, respectively.

Figs. 3 and 4 visualize the training accuracy and loss and Figs. 5 and 6 visualize the testing accuracy and loss, respectively.

The charts show the individual positive and negative tweets and the percentage of positive and negative tweets among the overall positive and negative tweets for both the parties. The performance of both parties is demonstrated in the following charts.

The charts shown in Fig. 7 demonstrate that among the tweets considered for Party-1, 63.2% of them were positive and for Party-2 68.1% were positive. Party-1 has less percentage of positive tweets compared to Party-2 and more number of negative tweets compared to Party-2. Therefore, we can conclude that Party-2 has an edge over Party-1.

Table 1 Comparison of election prediction system testing results

Model	Accuracy	Loss	Precision	Recall	F1-Score
RoBERTa-base	91.67	0.78	91.47	91.13	0.92
DistilBERT	90.33	0.81	89.88	90.17	0.90
BART	77.25	1.03	82.35	79.63	0.80
RoBERTa-latest	**93.67**	**0.59**	**93.2**	**93.88**	**0.93**

RoBERTa-latest is our model with best results compare to competitive models

Table 2 Comparison of election prediction system training results

Model	Accuracy	Loss	Precision	Recall	F1-Score
RoBERTa-base	100	0.006	100	100	1.0
DistilBERT	100	0.008	100	100	1.0
BART	98.93	0.03	92.53	89.82	0.91
RoBERTa-latest	**100**	**0.03**	**100**	**100**	**1.0**

RoBERTa-latest is our model with best results compare to competitive models

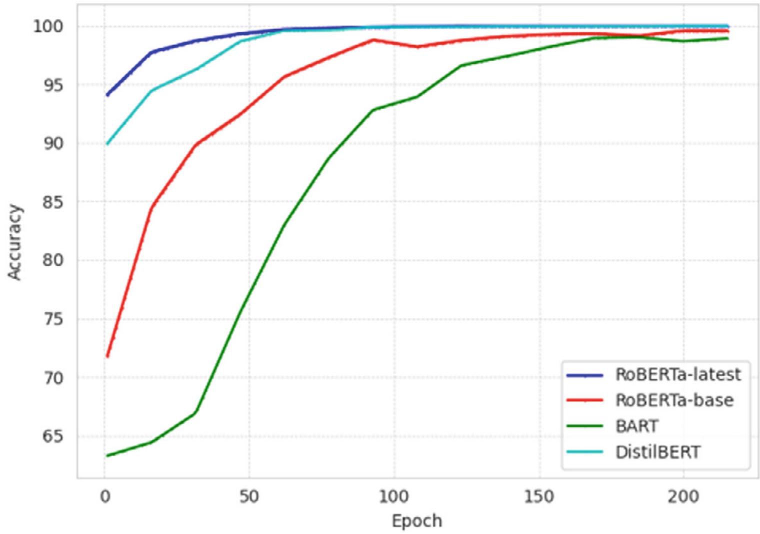


Fig. 3 Training accuracy versus epoch

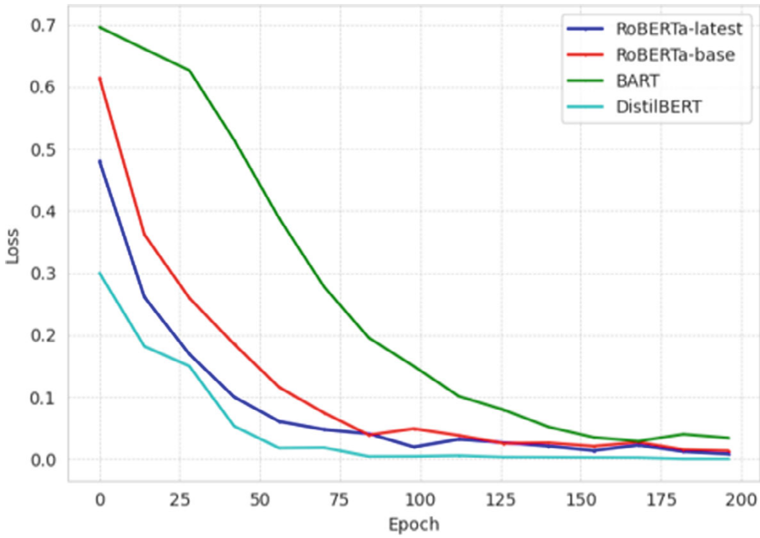


Fig. 4 Training loss versus epoch

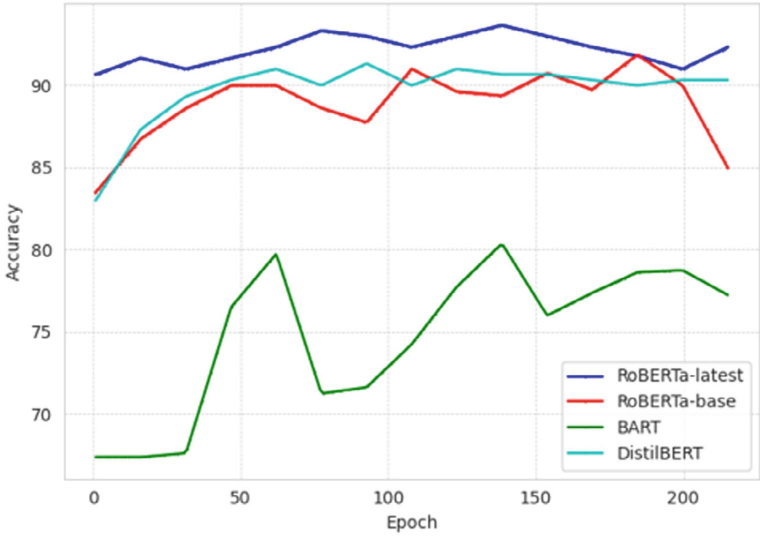


Fig. 5 Testing accuracy versus epoch

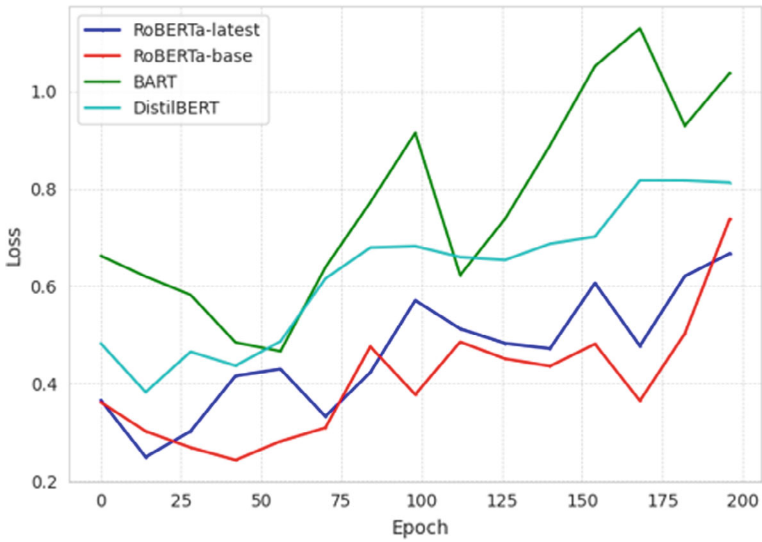


Fig. 6 Testing loss versus epoch

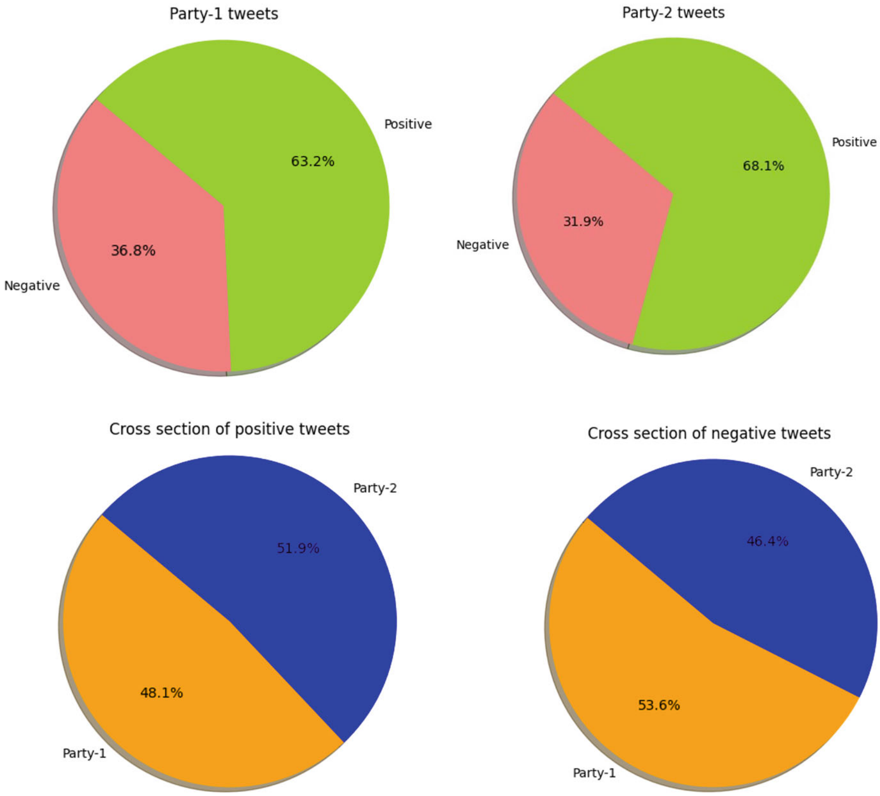


Fig. 7 Pie charts for result analysis

5 Conclusion

Election prediction is a difficult task. It is challenging to provide an accurate estimate with limited samples of data. To improve the quality of input data, we have proposed a novel approach to include tweets of local languages like Hindi. A rigorous comparative analysis was performed using state of the art deep learning models to identify the most suitable model for Indian context. The RoBERTa model was outperforming other models with better accuracy and less loss. A total of 10,000 tweets were passed to the RoBERTa model to classify the sentiment between positive and negative. The results were visualized using pie charts to predict the party that has better chances of winning the election. In future enhancement, this work can be extended with other sources of data sets like Facebook and Instagram. In addition to this, we can also include spam tweet detection to improve the quality of data set. Researchers with better computational resources can perform the hyper parameter optimization for the proposed model to further optimize the results.

References

1. Ramteke J, Shah S, Godhia D, Shaikh A (2016) Election result prediction using twitter sentiment analysis. In: 2016 International conference on inventive computation technologies (ICICT), Coimbatore, India, 2016, pp 1–5. https://doi.org/10.1109/INVENTIVE.2016.7823280
2. Khurana Batra P, Saxena A, Shruti, Goel C (2020) Election result prediction using twitter sentiments analysis. In: 2020 Sixth international conference on parallel, distributed and grid computing (PDGC), Waknaghat, India, 2020, pp 182–185. https://doi.org/10.1109/PDGC50313.2020.9315789
3. Tsai M-H, Wang Y, Kwak M, Rigole N (2019) A machine learning based strategy for election result prediction. In: 2019 International conference on computational science and computational intelligence (CSCI), Las Vegas, NV, USA, 2019, pp 1408–1410. https://doi.org/10.1109/CSCI49370.2019.00263
4. Limboi S, Dioşan L (2022) An unsupervised approach for twitter sentiment analysis of USA 2020 presidential election. In: 2022 International conference on innovations in intelligent systems and applications (INISTA), Biarritz, France, 2022, pp 1–6. https://doi.org/10.1109/INISTA55318.2022.9894264
5. Kaimakamidis A, Pitas I (2024) Political tweet sentiment analysis for public opinion polling. In: ICASSP 2024–2024 IEEE international conference on acoustics, speech and signal processing (ICASSP), Seoul, Korea, Republic of 2024, pp 121–125. https://doi.org/10.1109/ICASSP48485.2024.10448327
6. Mohd M, Javeed S, Wani M, Khanday H, Wani A, Mir U, Nasrullah S (2023) PoliWeet election prediction tool using tweets. Softw Impacts 17:100542. https://doi.org/10.1016/j.simpa.2023.100542
7. Mazza M, Avvenuti M, Cresci S, Tesconi M (2022) Investigating the difference between trolls, social bots, and humans on twitter. Comput Commun 196. https://doi.org/10.1016/j.comcom.2022.09.022
8. Olabanjo O, Wusu A, Afisi O, Asokere M, Padonu R, Olabanjo O, Ojo O, Folorunso O, Aribisala B, Mazzara M (2023) From twitter to Aso-Rock: a sentiment analysis framework for understanding Nigeria 2023 presidential election. Heliyon 9(5):e16085. PMID: 37215756; PMCID: PMC10199181. https://doi.org/10.1016/j.heliyon.2023.e16085
9. Shevtsov A, Tzagkarakis C, Antonakaki D, Ioannidis S (2022) Explainable machine learning pipeline for twitter bot detection during the 2020 US presidential elections. Softw Impacts 13:100333. https://doi.org/10.1016/j.simpa.2022.100333
10. Perera S, Karunanayaka K (2022) Sentiment analysis of social media data using fuzzy-rough set classifier for the prediction of the presidential election. In: 2022 2nd international conference on advanced research in computing (ICARC), Belihuloya, Sri Lanka, 2022, pp 188–193. https://doi.org/10.1109/ICARC54489.2022.9754173
11. Brito K, Adeodato PJL (2023). Machine learning for predicting elections in Latin America based on social media engagement and polls. Gov Inf Q 40(1):101782. https://doi.org/10.1016/j.giq.2022.101782
12. Ali H, Farman H, Yar H, Khan Z, Habib S, Ammar A (2021) Deep learning-based election results prediction using twitter activity. https://doi.org/10.21203/rs.3.rs-839553/v1
13. Bibi A, Abbasi WA, Aziz W, Khalil S, Uddin M, Iwendi C, Gadekallu T (2022) A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis. Pattern Recognit Lett 158. https://doi.org/10.1016/j.patrec.2022.04.004
14. Myilvahanan K, Yashas P, Pasha S, Ismail M, Tharun V (2023) A study on election prediction using machine learning techniques. In: 2023 3rd international conference on artificial intelligence and smart energy (ICAIS), Coimbatore, India, 2023, pp 1518–1520, https://doi.org/10.1109/ICAIS56108.2023.10073693
15. Cano-Marin E, Mora-Cantallops M, Sanchez-Alonso S (2023) Twitter as a predictive system: a systematic literature review. J Bus Res 157:113561. https://doi.org/10.1016/j.jbusres.2022.113561

Multi-Scale Decomposition, a Key Advanced Method for a Combination of Many Exposures of an Image Using Image Sequences

Yedlla Satyam, Guntoju Kalpana Devi, K. Srujan Raju, D. Sudha, E. N. V. Purna Chandra Rao, and Siva Skandha Sanagala

Abstract High dynamic range (HDR) photography is one example of how contemporary computational photography techniques have significantly advanced beyond the constraints of traditional digital cameras in handling situations with a wide scope of motion. It is difficult to capture all the nuances of real scenes in a single photograph because to the confined motion scope of commonly used image recorders. To decrease this issue, it may be possible to capture the scene using a collection of images taken at various exposure levels, which may then be combined via image fusion to create a useable image. A category of image fusion technique called multi-exposure image fusion, or MEF, combines a number of photographs of the same scene that were captured using various exposure settings. Recently, MEF methods have attracted a lot of attention because it is crucial to produce high dynamic range images. A recent study found that multi-scale structural patch decomposition-based MEF (MSPD-MEF), although losing detail, offers the best fusion quality and shortest runtime. In order to keep the fine features in the fused images in a particular course, we first incorporate edge-preserving components into our technique to look into this issue.

Y. Satyam
Department of Information Technology, CMR Technical Campus, Hyderabad, Telangana, India

G. K. Devi · K. Srujan Raju (✉)
Department of Computer Science and Engineering, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: ksrujanraju@gmail.com

D. Sudha · E. N. V. Purna Chandra Rao
Department of Electronics and Communication Engineering, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India
e-mail: dsudha@cmrcet.ac.in

S. S. Sanagala
Department of Computer Science and Engineering, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_4

Keywords Edge preserving · Multi-exposure image fusion · Quality · Patch decomposition

1 Introduction

In comparison with regular cameras, HDR photography may capture a wider range of brightness. Multi-exposure image fusion (MEF) is a common and simple method for producing HDR images. Furthermore, MEF has the capacity to enhance low-light images, eliminate haze, and pinpoint prominent spots. In order to be suitable for mobile devices, a good MEF technique ought to perform well in mutually static and dynamic scenarios while requiring the least amount of computational power.

Developed the cutting-edge technique known as structural patch decomposition multi-exposure image fusion. It combines many images captured at various exposure levels using structural patch decomposition to create an image by a high dynamic range and little noise. This method is particularly useful for capturing a wide range of brightness levels in high dynamic range pictures. Artifacts that can show up in the decomposed image as a result of SPD include halo effects and spurious edges. To reduce these artifacts, it can become more challenging to compute and necessitate more parameter tinkering. MSPD outperforms SPD by performing image decomposition at many scales and minimizing parameter dependence. It has drawbacks, though, including higher processing costs and a potential for over-smoothing, which can lead to a loss of detail in the fused image. To address the loss of details in the fused picture, we propose a technique known as the multi-scale decomposition approach, in which we can keep more detail information of the fused image by including edge-preserving components.

2 Literature Survey

The various techniques for fusing perceptual weights and the characteristics that are intended to be independent of exposure while calculating motion are highlighted in this section's summary of available MEF algorithms.

Li et al. [1] Fast multi-scale structural patch decomposition is a technique that decomposes large images fast by rapidly analyzing patches at different scales. It is commonly combined with other image processing techniques for object identification, feature extraction, and segmentation. However, it has drawbacks, such as the possibility of unreliable results from under or over segmentation, and it might not retain fine image details as it concentrates more on general structural data.

A structural patch decomposition technique was put out by Ma et al. [2] to combine several images of the same scene taken at varied exposure levels. The procedure partitions the images into patches and assigns weights to each piece based on structural information. Although the final fused picture keeps details in both bright and dark areas, the process can be computationally costly and could produce artifacts in regions with different exposure or noise levels.

Fang et al. [3] technique uses perceptual evaluation to assess the visual quality of composited images made from multiple exposures. Based on a variety of perceptual criteria, human observers assess the merged image's quality, providing a detailed examination of image quality that corresponds to human perception. The generalizability of this technique is constrained, and it can be time-consuming and subjective.

Hybrid learning, a technique to improve an image's brightness and contrast using multi-scale exposure fusion, was suggested by Zheng et al. [4]. It uses supervised and unsupervised learning strategies to fine-tune the fusion algorithm's settings, enhancing accuracy and robustness in the process. This method can be applied to a variety of industries, but it can take a while and specialized equipment or software to produce high-quality outcomes.

In order to rapidly combine multi-exposure images, Ma et al. [5] technique suggests using a deep convolutional neural network. High-speed photography and real-time video editing can both benefit from the method's improved brightness, contrast, and detail. However, to train and implement the model using this technique, you need a sizable training dataset as well as expensive hardware and software.

Cai et al. [6] suggests using a convolutional neural network (CNN) to instruct a SICE enhancer in the year 2018. How to create a training dataset of low-contrast and high-contrast image pairs intended for end-to-end CNN learning is a significant challenge. In order to achieve this, we create a sizable multi-exposure image data collection of 589 carefully chosen high-resolution multi-exposure sequence by 4413 photographs. Thirteen sample multi-exposure picture fusion and stack-based high dynamic range imaging techniques are utilized to produce the contrast-enhanced images intended for every sequence. The best photograph is then chosen as the reference image for each scene through subjective testing. Among the newly created dataset, it is easy to train a CNN as the SICE enhancer to increase the distinction of an underexposed or overexposed image.

In 2018, Yong et al. [7] we offer an efficient online background removal technique that can be reliably used to process real-world films with varying foreground and background elements. In contrast to other techniques that frequently model the foreground as Gaussian or Laplacian distributions, we model the foreground for every frame with a distinct mixture of Gaussians (MoG) distribution that is updated live frame by frame. In particular, our MoG model is regularized in every frame using the learned foreground/background information from earlier frames. Present results in a very stable, resilient, and adaptable online MoG model that can handle realistic foreground and background fluctuations. The suggested model might be expressed as a succinct probabilistic MAP model that the EM algorithm can easily solve.

According to Wang et al. [8] in 2016, a fundamentally essential but understudied problem in stereoscopic 3D imaging is the subjective and objective evaluation of the perceived eminence of depth information in symmetrically and asymmetrically deformed stereoscopic pictures. Here, we start by conducting a subjective test using the conventional complete grouping rating method, which is frequently employed in picture quality assessment studies in general. As a result of the significant influence that monocular cues and the spatial quality of images has on the depth quality ratings provided by individuals, we believe that this method is flawed because it is challenging to distinguish between the genuine assistance of stereoscopic signals to depth perception. To address this issue, we perform a fresh subjective inquiry in which, before applying various types and intensities of symmetric and asymmetric distortions, depth effect is synthesize at various depth levels.

By maximizing a brand-new objective quality metric dubbed the color MEF structural similarity (MEF-SSIMc) index, Ma et al. [9] developed a MEF strategy in 2018. The design idea we offer here differs greatly from those currently in use. We engage dynamically in the dimension of all pictures, looking for the image that optimizes MEF-SSIMc, as opposed to establishing an existing systematic computational architecture using MEF (for instance, multi-resolution transformation as well as domain fusion before performing image reconstruction). To be more specific, we start by strengthening and expanding the MEF-SSIM algorithm's application scope.

According to Ma et al. [10], the human visual structure is excellent in distinguish local blur of visual images in 2018. However, the essential mechanism is not unstated. Conventional models of blur, for instance those based on phase coherence loss at localized features and energy attenuation at high frequencies, have severe disadvantages. For instance, they have trouble telling flat from fuzzy areas. Now, we argue that high-level semantic information is crucial for accurately identify local blur.

Oh et al. [11] developed a novel high dynamic range (HDR) imaging technique in 2015 that makes use of rank minimization. When the intensity of each matching pixel is stacked together, the input low dynamic range (LDR) photographs taken with various disclosure times display a linear dependency and create a rank-1 matrix under the assumption that a camera responds linearly to scene radiance. In reality, the rank-1 structure of the LDR images is broken by misalignments brought on by camera movement, the presence of moving objects, saturations, and image noise. We describe a rank minimization approach that concurrently aligns LDR images and finds outliers for reliable HDR synthesis to solve these issues.

Multi-scale illumination fusion is a useful enhancement of images method for an elevated dynamic range (HDR) scene in 2017, claims Li et al. [12]. This paper proposes a unique multi-scale contact fusion method to combine low dynamic range (LDR) photographs that have been illuminated differently by utilizing a weighted guided picture filter for smoothing the Gaussian pyramidal of weight maps across each of the LDR shots. Without changing the relative brightness in the combined image, the suggested method better maintains information in the brightest and darkest parts of the HDR prospect. The multiple exposures of the photographs are also given a novel weighted structure tensor, which is then used to construct a feature removal component for the projected fusion approach. This allows users to edit the small features in the augmented image as they see fit. A straightforward single-image brightening solution is also created by means of the suggested multi-scale exposure fusion approach for both backlight and low-light imaging.

After delivering the image, SDR is finished when the image is recovered as a noise-free, interpolated image with no missing pixels [13–15] in order to change the missing image pixels at the receiver end, the Kriging interpolation procedure is implemented on the image. The radio system known as software-defined radio (SDR).

3 Proposed Model

Multi-scale decomposition using Laplacian pyramid.

At the outset of the algorithm, an essential and critical step involves decomposing the input images into multiple scales using a Gaussian pyramid. This step is instrumental in creating a multi-scale representation of the input images, allowing the algorithm to capture and preserve the intricate details of the images at various scales as seen the Fig. 1. The Gaussian pyramid works by smoothing and down-sampling the images to produce a series of increasingly lower resolution images. This process creates a hierarchical structure that reflects the varying levels of details in the input images.

After this initial decomposition step, the algorithm moves on to perform edge-preserving structural patch decomposition on each scale. The purpose of this decomposition is to separate the images into small, spatially coherent regions or patches, and to group these patches based on their structural similarity. This technique is particularly useful in preserving the fine details of the input images, while also reducing the impact of noise and artifacts that can often corrupt images.

Once the patches have been separated and grouped, the algorithm proceeds to perform exposure fusion on each scale separately. This step involves combining the

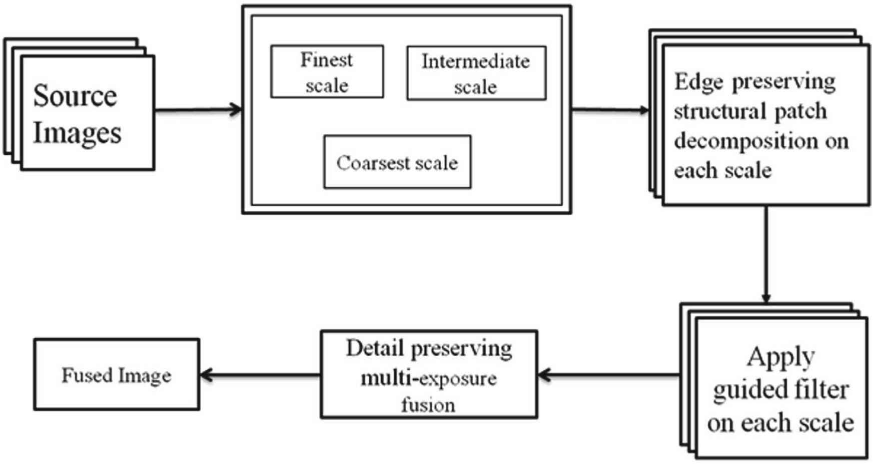


Fig. 1 Block diagram of projected technique

input images by selecting the best pixels from each image to create a single, unified output image for that specific scale. This process is designed to produce images that exhibit a higher dynamic range and greater visual appeal than the original input images.

In summary, the algorithm operates in a series of steps that involve decomposing the input images into multiple scales, performing edge-preserving structural patch decomposition, and finally performing exposure fusion to create a unified output image. The combination of these steps enables the algorithm to create images that exhibit a higher level of detail and clarity, while also reducing the impact of noise and artifacts that can often degrade image quality.

4 Results and Discussion

From below qualitative analysis, we can observe that the ultimate fused images that are obtained from the existing methods are less enhanced compared to the multi-scale decomposition method-based multi-exposure image fusion.

Fast multi-scale structural patch decomposition-based multi-exposure image fusion (MSPD-MEF) by [Hu Li, et al.] and structural patch decomposition-based multi-exposure image fusion (SPD-MEF) by [Kede Ma, et al.] are the two approaches that we compared.

As shown in Fig. 2, the final fused image we obtained retains more details with preserved edges than the images obtained by the previous existing methods.

5 Quantitative Analysis

From below quantitative analysis, we can see that the metric values of the previously existing methods are less compared to the edge-preserving, multi-scale decomposition method. Usually metrics are used to evaluate how will an image processing method performs in comparison.

The metrics included are:

Multi-scale structural similarity index (MEF_MS_SSIM): As shown the Fig. 3 in this, SSIM index between the MEF image and each source image in the source sequence at various scales is used to compute the MS-SSIM index. The final step is to calculate the overall MS-SSIM index by adding the separate SSIM indices from each scale using a weighted geometric mean.

The SSIM index is calculated as follows:

$$\mathrm{SSim}(m,n) = \frac{(2\mu_m\mu_n+\mathrm{D1})(2\mu_m\mu_n+\mathrm{D2})\left(\mu^2+\mu^2+D1\right)\left(\sigma^2+\sigma^2+\mathrm{D2}\right)}{mn}$$

wherever:

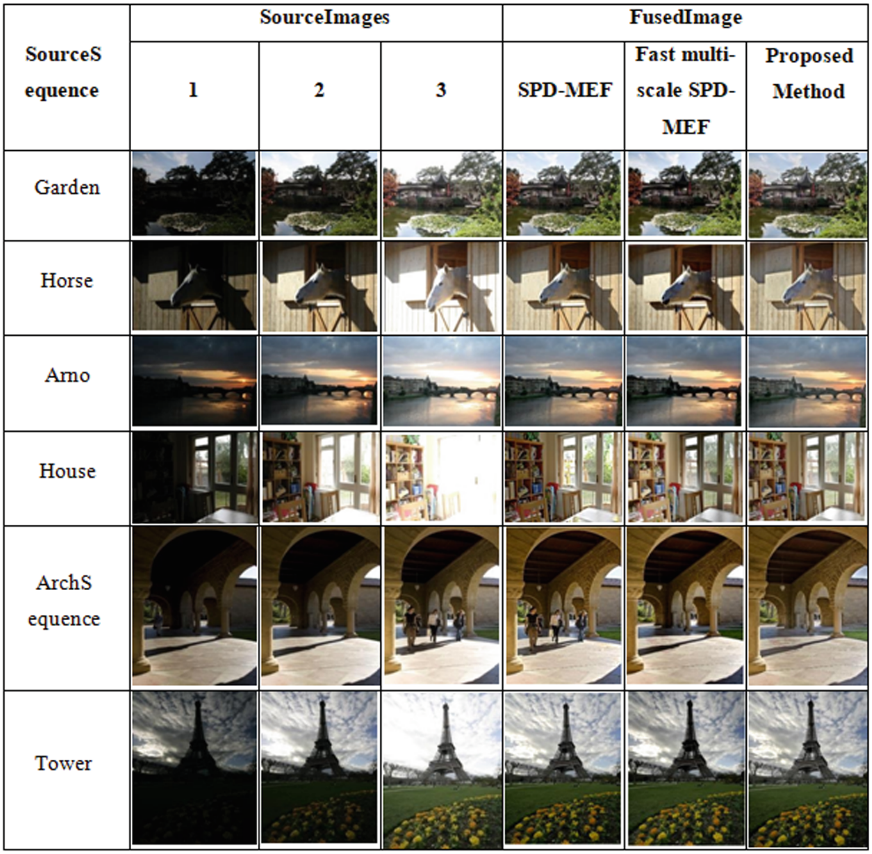
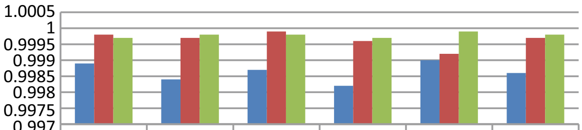


Fig. 2 Comparison of different datasets with existing methods and proposed method

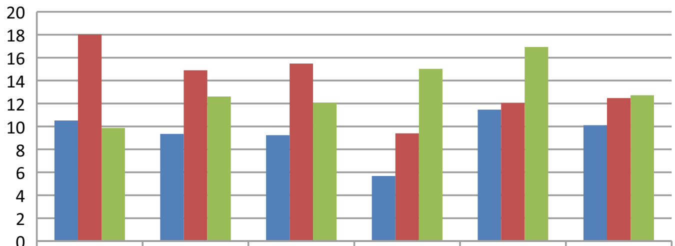


	Garden	Horse	Arno	House	Arch Sequence	Tower
SPD-MEF	0.9989	0.9984	0.9987	0.9982	0.999	0.9986
Fast Multi-scale SPD-MEF	0.9998	0.9997	0.9999	0.9996	0.9992	0.9997
Edge Preserving SPD-MEF	0.9997	0.9998	0.9998	0.9997	0.9999	0.9998

Fig. 3 Quantitative value for SSIM metric

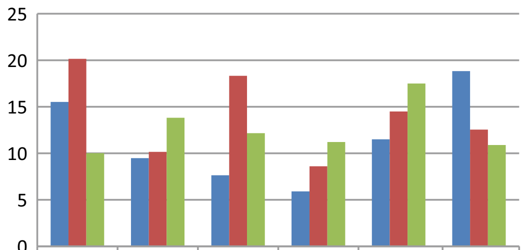
m and n are the two images being compared
μ_m and μ_n are the means of x and y correspondingly
σ_m and σ_n are the standard deviations of x and y correspondingly
σ_{mn} is the cross-covariance among x as well as y
D1 and D2 are constants added to avoid division by zero.
As shown Figs. 4, 5, and 6 are quantitative value for PSNR metric.
The next metric is Peak Signal-to-Noise Ratio (PSNR): PSNR is a widely used measure to assess an image's or video's quality.

$$\text{PSNR} = 10\ \log 10((\text{MAX2})/\text{MSE})$$



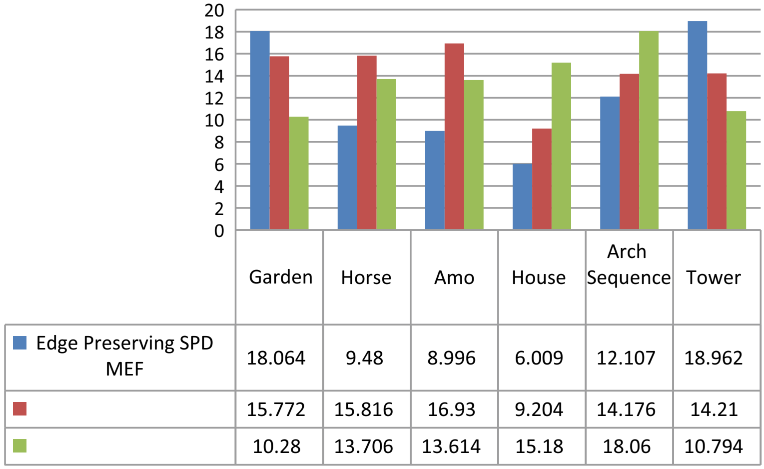
	Garden	Horse	Amo	House	Arch Sequence	Tower
SPD-MEF	10.52	9.349	9.236	5.68	11.462	10.115
	18.028	14.903	15.486	9.398	12.04	12.47
	9.862	12.605	12.06	15.024	16.925	12.72

Fig. 4 SPD-MEF quantitative value for PSNR metric



	Garden	Horse	Amo	House	Arch Sequence	Tower
Fast Multi Scale SPD MEF	15.524	9.48	7.64	5.898	11.5	18.834
	20.148	10.148	18.322	8.597	14.498	12.552
	10.012	13.825	12.163	11.204	17.498	10.892

Fig. 5 Fast multi-scale SPD-MEF values for PSNR metric



	Garden	Horse	Amo	House	Arch Sequence	Tower
Edge Preserving SPD MEF	18.064	9.48	8.996	6.009	12.107	18.962
	15.772	15.816	16.93	9.204	14.176	14.21
	10.28	13.706	13.614	15.18	18.06	10.794

Fig. 6 Edge preserving SPD-MEF values for PSNR metric

where

The maximum amount of pixels the image can contain is MAX. (For 8-bit images, for instance, 255).

The MSE, or mean squared error, measures the difference between the original and compressed images.

MSE is calculated as:

$$\mathrm{MSE} = 1/\mathrm{MX}\sum_{a=1}^{M}[I(a,b) - K(a,b)]2$$

where

M is the total number of pixels in the image.

$I(a, b)$ is the pixel value of the original image at position (a, b).

$K(a, b)$ is the pixel value of the compressed image at position (a, b).

Therefore, they have less noise or distortion than the original picture, images with higher PSNR values are generally of higher quality.

6 Conclusion

In this study, we used a multi-scale decomposition technique with edge preservation that allows for multi-exposure fusion, by incorporating edge-preserving factors that are safeguarded added particulars in the fused images. In comparison with the state-of-the-art methods, SPD-MEF and MSPD-MEF, our process can inherit all of their benefits, such as the suppression of halo artifacts, avoidance of the ghosting effect, and fastest running time, while also overcoming their weakness, namely detail loss. Therefore, the extensive experiments confirm that, when compared to various MEF methods, the current method can achieve the most cutting-edge visual quality intended for both static and dynamic scene settings. This strategy works better than the several image fusion techniques in conditions of mutually subjective and objective quality metrics. Additionally, the suggested approach is effective and adaptable, making it appropriate for use in real-world applications in industries like photography, computer vision, and medical imaging.

7 Future Scope

The technique used in this paper is intended for single-image multi-exposure merging. It would be intriguing to apply it to video processing. In order to do this, the algorithm may need to be modified to operate with temporal sequences of images, or novel methods for combining multiple frames into a single output may need to be developed.

References

1. Li H, Ma K, Yong H, Zhang L (2020) Fast multi-scale structural patch decomposition for multi-exposure image fusion. IEEE Trans Image Process 29:5805–5816
2. Ma K et al (2017) Robust multi-exposure image fusion: a structural patch decomposition approach. IEEE Trans Image Process 26(5):2519–2532
3. Fang Y et al (2020) Perceptual evaluation for multi-exposure image fusion of dynamic scenes. IEEE Trans Image Process 29:1127–1138
4. Zheng C et al (2020) Single image brightening via multi-scale exposure fusion with hybrid learning. IEEE Trans Circuits Syst Video Technol 31(4):1425–1435
5. Ma K, Duanmu Z, Zhu H, Fang Y, Wang Z (2020) Deep guided learning for fast multi-exposure image fusion. IEEE Trans Image Process 29:2808–2819
6. Cai J, Gu S, Zhang L (2018) Learning a deep single image contrast enhancer from multi-exposure images. IEEE Trans Image Process 27(4):2049–2062
7. Yong H, Meng D, Zuo W, Zhang L (2018) Robust online matrix factorization for dynamic background subtraction. IEEE Trans Pattern Anal Mach Intell 40(7):1726–1740
8. Wang J, Wang S, Ma K, Wang Z (2016) Perceptual depth quality in distorted stereoscopic images. IEEE Trans Image Process
9. Ma K, Duanmu Z, Yeganeh H, Wang Z (2018) Multi-exposure image fusion by optimizing astructural similarity index. IEEE Trans Comput Imaging 4(1):60–72
10. Ma K, Fu H, Liu T, Wang Z, Tao D (2018) Deep blur mapping: exploiting high-level semantics by deep neural networks. IEEE Trans Image Process Appear
11. Oh T-H, Lee J-Y, Tai Y-W, Kweon IS (2015) Robust high dynamic range imaging by rank minimization. IEEE Trans Pattern Anal Mach Intell 37(6):1219–1232
12. Li Z, Wei Z, Wen C, Zheng J (2017) Detail-enhanced multi-scale exposure fusion. IEEE Trans Image Process 26(3):1243–1252
13. Godavarthi B, Gopatoti A, Naik MC, Nalajala P (2018) Image processing for SDR applications to recreate lost image pixels/packets. J Fundam Appl Sci 10(6S):2826–38

14. Kirubakaran S, Nagarajan V, Dhayapulley SCK, Soubache ID, Pasupuleti SR, Kumar A, Rastogi R, Vasudevan S (2023) Hybrid deep learning-based grid-supportive renewable energy systems for maximizing power generation using optimum sizing. Electr Power Compon Syst. https://doi.org/10.1080/15325008.2023.2201249
15. Aelgani V, Gupta SK, Narayana VA (2023) Local agnostic interpretable model for diabetes prediction with explanations using XAI. Lecture notes in networks and systems, vol 606, pp 417–425. https://doi.org/10.1007/978-981-19-8563-8_40

Scrub Master: An IoT-Enabled Smart Cleaner

Puja Chavan, Shravani Bahirat, Vaishnavi Dalvi, Sakshi Darawade, Shravani Jagtap, and Sakshi Shakhawar

Abstract In twenty-first century, devices become an essential part of daily life. The increasing automation in daily lives is largely facilitated by the growing technology of the Internet of Things (IoT). Result of this the world experiencing a new era of convenience that is transforming daily routines. So this system is being introduced as revolutionary Smart Cleaner, utilizing advanced technology to efficiently and effectively enhance cleaning processes and elevate cleanliness standards in the automated world. In contrast to the existing system's limited functionality of smart brooms that solely perform sweeping tasks and target dust removal, the proposed system goes beyond these limitations by introducing an innovative approach where the cleaner efficiently wipes the floor using rotating mops, thereby significantly enhancing the overall cleaning effectiveness. Distinguishing itself further, this system incorporates an in-built water supply, setting it apart from conventional options. The ability to remotely control this smart mop through a user-friendly mobile application, coupled with its mobility facilitated by wheels, enables effortlessly navigation and efficient cleaning in various locations. Leveraging a Bluetooth module, users can effortlessly operate and supervise the mop's functionalities, making cleaning tasks more convenient and streamlined.

P. Chavan (✉) · S. Bahirat · V. Dalvi · S. Darawade · S. Jagtap · S. Shakhawar
Department of Multidisciplinary Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: puja.cholke@vit.edu

S. Bahirat
e-mail: shravani.bahirat22@vit.edu

V. Dalvi
e-mail: vaishnavi.dalvi22@vit.edu

S. Darawade
e-mail: sakshi.darawade22@vit.edu

S. Jagtap
e-mail: shravani.jagtap22@vit.edu

S. Shakhawar
e-mail: sakshi.shakhawar22@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_5

Keywords Bluetooth module · Effective cleaning · Internet of things (IOT) · Smart cleaner · Remote controlled

1 Introduction

In the recent years, the rapid advancement of Internet of Things (IoT) technology has revolutionized various industries, transforming conventional systems into smart and connected environments. One such domain that can greatly benefit from IoT integration is the field of cleaning and maintenance. Traditional cleaning methods often rely on manual labor and lack real-time monitoring and control, resulting in inefficiencies and suboptimal cleaning outcomes [1, 2]. The proposed IoT-based smart mop system consists of various components, including water pump, relay, actuators, connectivity modules, and microcontroller [3]. This IoT-based smart mop system utilizes Bluetooth connectivity for user control, allowing for seamless initiation of cleaning operations. It features an in-built water supply with a timed release mechanism, ensuring efficient water usage, while its exceptional agility makes it highly effective on flat surfaces. By employing IoT technology in the cleaning process, the smart mop system offers several advantages over traditional methods, and it eliminates the need for manual supervision and intervention, allowing cleaning operations to be performed autonomously [4]. The system's connectivity capabilities enable remote monitoring and control, allowing cleaning personnel or homeowners to manage the cleaning process from anywhere using a smartphone or web interface [5]. A standout feature of this smart mop is its in-built water supply, incorporating a time delay mechanism for optimal water usage. This innovative design ensures that water is dispensed efficiently, avoiding wastage and prolonging the mop's cleaning performance [6, 7]. In this paper, project will provide a comprehensive overview of smart cleaning technologies, their design, functionalities, and potential applications [8]. Paper discusses the current state of the art in this field and highlights the challenges and opportunities associated with smart cleaning solutions. It also provides a critical evaluation of the effectiveness of smart cleaners compared to traditional cleaning methods and explores the potential impact of these technologies on the environment, public health, and social well-being.

By exploring these aspects, paper aims to provide comprehensive insights into the capabilities, benefits, and potential challenges associated with this novel cleaning solution.

The remaining section is organized as: Sect. 2 reviews the existing smart cleaners. Section 3 represents the method used for execution of proposed system. Section 4 gives an idea about implementation and how the system to be controlled. Section 5.

2 Literature Review

Described the design and working of a floor cleaner robot which allows Bluetooth connectivity to any pairable device or component, and an Android Studio application program was built for it. Varsha et al. [1] the following are some papers that have been recently reviewed which are related to cleaning robots. Murdan et al. [2] proposed cleaning robot which can clean both dry and wet surfaces at the same time and it can be controlled by using an app which makes it more easy for users. However, the device created was a bit noisy. Anbumani et al. [3] aimed to design a noiseless vacuum floor cleaner robot with a water-level monitor that successfully cleans room corners. The robot has two modes of operation: wet mopping and hoover cleaning. There was no pre-calculated level of floor cleaning liquids that could be added to water specified in this approach. Mohammed et al. [4] developed a realistic and effective rubbish collection method for cleaning up waste from rivers, canals, and lakes. A type of conveyor is used to transport garbage from the river's surface to the specified region. Badamasi [5] explained various hardware components used in the Arduino board, the software used to program it, and how to construct and create your own projects. Also provided a few examples of Arduino projects. Jain et al. [6] created a smart broom using an ESP8266:Node-MCU and several sensors such as an accelerometer and a vibration sensor to read data from the ESP module. This broom keeps track of its own location. Also, keep an eye on the broom's operating time. This approach cannot be used for wet cleaning. Jarande et al. [7] provided an approach which uses IR sensors for obstacle detection. Along with cleaning, it also checks the battery status of the robot and accordingly sends the battery status to the MQTT server. This is beneficial for user as he/she can check the battery status from any place. Saleem et al. [8] provide a novel room cleaning robot that incorporates map storage and wall following capabilities, and is done by Arduino Uno. The efficiency of the robot is achieved by a combination of modern sensor assembly, intelligent algorithm, and a unique dust cleaner design that enables it to operate efficiently in a variety of situations. It is a bit expensive. Sridhar et al. [9] have built a floor cleaning system to function in a collision-free way, allowing it to cover the most possible cleaning area. The primary goal of this approach was to improve corner cleaning efficiency. However, a GSM module could be used which would send messages about the status of cleaning. Monika et al. [10] This presented paper consists of self-drive mode cleaner including spinning brushes and mopping. It uses ultrasonic sensor to avoid obstacles and changes its directions. It cleans the wet floor and sucks dirt. It is unable to reach edges and corners. RaveenaIshalavath et al. [11] utilized the Raspberry pi 3 model as main controller to control the cleaning robot. It communicates through Bluetooth. LCD displays in device which receives signals through every operation. It scans obstacles earlier than robot and to avoid collision when robot is in auto run feature. Manasa et al. [12] gave a detailed description of an RC car with a hoover attached. If an impediment is identified, an ultrasonic sensor attached to the front of the automobile is used to measure distance. Ahsan et al. [13] illustrated on an approach to control a robotic vehicle using internet. They overcame the constraints of wireless robots by using the Internet and a DTMF decoder IC and a preprogrammed microcontroller IC that controls the movement of the robotic vehicle based on the output of the DTMF decoder. Rissanen et al. [14] described the design approach and implementation details of an intelligent Bluetooth-enabled robot automobile. A review of the wireless robot car remote controlling system is also provided. Chaudhry et al. [15] proposed a system that focuses on the concept of how a robot can be controlled by the human voice. An Android app is employed as a conduit for transmitting human commands to the microcontroller in this system.

Kumar et al. [16] proposed an IoT-based Smart Floor Cleaner with Object Detection and Mapping capabilities. The smart floor cleaner utilizes Internet of Things (IoT) technology to enhance its functionality and efficiency. The key features of this intelligent cleaner include object detection, allowing it to identify and navigate around obstacles, and mapping capabilities, enabling it to create a map of the cleaning area for efficient coverage. The paper presents the design, implementation, and performance evaluation of the smart floor cleaner, showcasing its potential for effective and autonomous cleaning tasks.

Chavan et al. [17] introduced a Smart Floor Cleaner Robot using IoT technology. This intelligent cleaning robot leverages IoT to enhance its operational capabilities and provide seamless connectivity. The robot is designed to perform efficient cleaning tasks, aided by IoT-based control and monitoring systems. The paper details the development and implementation of the smart floor cleaner, along with its performance evaluation. The results demonstrate its effectiveness in autonomous cleaning and its potential to transform the cleaning process.

3 Methodology

The Scrub Master methodology is a best approach to cleaning and maintenance which borrows advanced technologies and intelligent systems to optimize the cleaning process. This methodology incorporates the use of Internet of Things (IoT) to revolutionize traditional cleaning practices and make them more efficient, cost-effective, and sustainable. Scrub Master is a Bluetooth controlled robot so to achieve this Bluetooth module HC-05 is used. Bluetooth module acts as a bridge between Android app and Arduino Uno. Commands for start, stop, right, left are provided on applications, and through Bluetooth module, Arduino receives the instructions and accordingly it performs activities.

The components included in Architecture of Scrub Master Robot are

1. Arduino Uno

The Arduino Uno is a popular and versatile microcontroller board that has gained immense popularity in the field of electronics and hobbyist projects. Developed by Arduino.cc, the Arduino Uno is based on the ATmega328P microcontroller, which provides a wide array of input and output pins for various connections. It offers 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog input pins, a 16 MHz quartz crystal, a USB connection for programming and power supply, an ICSP header for direct programming, and a reset button.

The board's simplicity and user-friendly interface make it an ideal choice for beginners and experienced electronics enthusiasts alike. It can be easily programmed using the Arduino IDE (Integrated Development Environment), which supports a simplified version of C++ language. The Arduino Uno can interact with various electronic components and sensors, making it an excellent platform for prototyping interactive projects, home automation systems, robotics, and more.

2. L293D Motor Driver Shield

The L293D Motor Driver Shield is a specialized add-on board designed to control motors with ease using Arduino and other compatible microcontrollers. Developed around the L293D IC, this shield provides an efficient and convenient solution for driving up to two DC motors or one stepper motor. It is widely used in robotics, automation, and various motor-driven projects.

The key feature of the L293D Motor Driver Shield is its ability to handle both direction and speed control of motors. The L293D IC is an H-bridge driver, which means it can control the rotation direction of the motors by toggling the input signals. This enables the motors to move forward, backward, or stop depending on the configuration of the input pins.

The shield conveniently plugs onto the Arduino board, making the connections straightforward and reducing the need for complex wiring. It features screw terminals for motor connections, making it easier to attach and detach motors securely. Additionally, it has built-in protection diodes that safeguard the circuitry against voltage spikes generated when the motors are turned off, enhancing the shield's reliability and protecting connected components.

3. Four Channel Relay

A four-channel relay is an electronic module that provides a convenient way to control multiple electrical devices or circuits using a single control signal. It consists of four individual relays, each capable of switching a separate circuit on or off in response to an external input. These relays operate as electrically controlled switches, allowing the isolation of low-voltage control signals from higher-voltage or higher-current loads.

The four-channel relay module typically interfaces with microcontrollers like Arduino or Raspberry Pi, making it a popular choice for home automation, robotics, and industrial applications. Each relay on the module is isolated, meaning the switching of one relay does not interfere with the others, providing a safe and reliable way to control different loads simultaneously.

4. Bluetooth Module

The Bluetooth module HC-05 is a widely used and popular wireless communication module that enables devices to establish Bluetooth connections and exchange data over short distances. It is based on the Bluetooth 2.0 standard and provides a simple and cost-effective solution for adding Bluetooth functionality to various electronic projects.

The HC-05 module operates as a Bluetooth Serial Port Profile (SPP) device, making it easy to use for serial communication. It uses the Universal Asynchronous Receiver/Transmitter (UART) interface to communicate with microcontrollers or other devices. This means that devices can communicate with the HC-05 module using standard serial communication commands, making it compatible with a wide range of microcontrollers, such as Arduino, Raspberry Pi, and others.

5. Water Pump

A water pump used for school projects is a simple and practical device that can be incorporated into various educational activities to demonstrate principles related to fluid dynamics, engineering, and physics. These projects not only enhance students' understanding of scientific concepts but also promote hands-on learning and critical thinking.

One common type of water pump used in school projects is the centrifugal pump. This type of pump uses an impeller to create a flow of water by centrifugal force. When the impeller rotates, it draws water into its center and propels it outward, creating pressure that moves the water through a pipe or hose. Centrifugal pumps are widely used in many applications, including pumping water in household appliances, irrigation systems, and industrial processes.

6. DC Motor

A DC motor (Direct Current motor) is a type of electric motor that converts electrical energy into mechanical rotational motion. It operates on the principle of Lorentz force, where a current-carrying conductor in a magnetic field experiences a force, causing it to rotate. DC motors are commonly used in various applications due to their simplicity, ease of control, and reliability.

The basic construction of a DC motor consists of a stationary part called the stator and a rotating part called the rotor. The stator contains field windings that create a magnetic field, typically using permanent magnets or electromagnets. The rotor, on the other hand, contains the armature, which is a set of windings connected to a shaft. When an electric current is supplied to the armature windings, it creates a magnetic field that interacts with the stator's magnetic field, generating a rotational force on the rotor. As a result, the shaft of the DC motor starts to rotate.

Steps to follow for designing the Scrub Master:

Step 1-Prepare Base

To begin with, you need to create a sturdy base for placing the parts. Start by cutting a piece of plywood measuring 12 $\times$ 12 inches. Next, drill two holes at the back of the base, ensuring that they are properly aligned and parallel to each other, to accommodate the motor clamps. Use screws to secure the clamps in place and then attach the motors to them. Finally, create the mops and attach them to the base, positioning them opposite to the motor clamps.

Step 2-Water Supply Mechanism

The water dispensing system includes a water pump with a 9 V power supply, which extracts water from the reservoir and dispenses it close to the mops on the floor. Start by positioning water pump, marking the installation points, drilling and securely fastening it in its place. Next, take an empty plastic bottle with a capacity of 600 ml. To obtain the desired portion, use a paper cutter to divide it in half and keep the lower segment. Attach the half-bottle to the robot base using warm glue. Obtain two sections of rubber tubing, with one connecting the inlet of the pump to the reservoir and the other linking the outlet of the pump to the floor. You can add straws to the pump's outlet later. To control the pump, use your smartphone just like other controls by turning it on/off.

Figure 1 shows the circuit diagram of a system, where all components are connected to the Arduino Uno board.

- The 100 rpm geared motors are connected to L293D motor driver shield with 9 V external supply and further it operated by Arduino Uno.
- The 75 rpm plastic geared motors are directly connected to the Arduino Uno which is used for mop rotation.
- The Bluetooth module HC-05 is used to provide remotely access of system which will be operate by the application

Connection of Arduino pins are outlined here (Table 1):

Next comes the Bluetooth module. Connections are (Table 2):

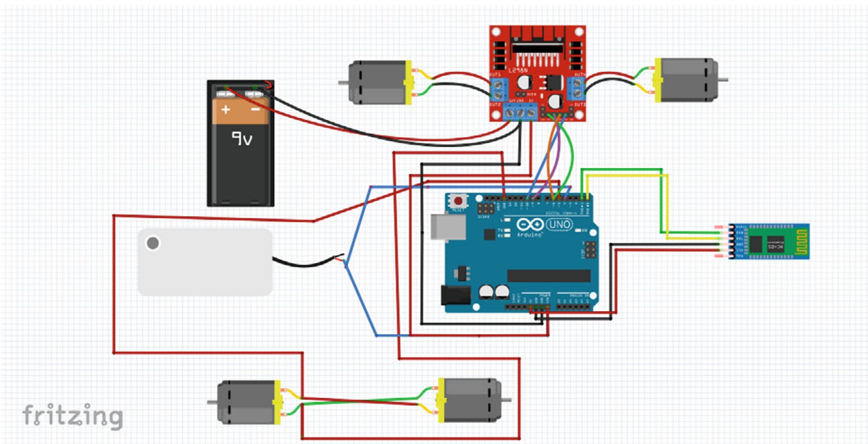


Fig. 1 Circuit diagram of system

Table 1 Connection from motor Arduino pins

Signals/voltage	Pins on Arduino
Signal 1	D8
Signal 2	D9
Signal 3	D11
Signal 4	D12
+ 5 V	+ 5 V
Gnd	Gnd
+ 12 V (motors will mov at this voltage)	Be connected to battery later

Table 2 Connection between bluetooth module and Arduino

Bluetooth module	Arduino pins
Vcc	+ 5 V
Gnd	GND
Rx	Tx
Tx	Rx

4 Implementation

In this advanced system, a smartphone serves as a remote control for a device. To enable this functionality, an Android application compatible with Bluetooth modules is utilized. The application receives instructions in text format, which essentially function as code for the microcontroller. These instructions are transmitted to the microcontroller via Bluetooth, with the HC-05 module serving as the receiver (Rx). The HC-05 module receives the code from the smartphone and then transmits it to the decode circuit for further processing.

Following are the steps to do so:

1. Download "Bluetooth Terminal HC-05" or any other app that supports Bluetooth communication.
2. Ensure that it is properly connected to your microcontroller or Arduino board.
3. Power on the module and put it into pairing mode. This involves holding the button on the module until the LED starts flashing rapidly.
4. On your smartphone, go to the Bluetooth settings and make sure if it is enabled.
5. Search for available devices and select the HC-05 module from the list.
6. When prompted for a password, enter "1234" (or any other password that you have set).
7. Once the pairing is successful open the Bluetooth terminal app you downloaded.
8. Select the paired module from the list of available devices.
9. You should now have a connection established between your smartphone.
10. You now use the app to send the command as mention in Table 3.

Table 3 Commands to perform activities

Action	Command
Forward Motor 1	‘M’
Stop Motor 1	‘m’
Forward Motor2	‘N’
Stop Motor2	‘n’
Turn Right	‘X’
Turn Left	‘Z’
Relay 1 ON	‘A’
Relay 2 ON	‘B’
Relay 3 ON	‘C’
Relay4 ON	‘D’
Relay 1 OFF	‘A’
Relay 2 OFF	‘B’
Relay 3OFF	‘C’
Relay4 OFF	‘D’

- Forward Motor 1: This action corresponds to the command ‘M’. When the system receives the command ‘M’, it initiates the forward motion of Motor 1.
- Stop Motor 1: This action corresponds to the command ‘m’. When the system receives the command ‘m’, it stops the operation of Motor 1.
- Forward Motor 2: This action corresponds to the command ‘N’. When the system receives the command ‘N’, it sets Motor 2 to move forward.
- Stop Motor 2: This action corresponds to the command ‘n’. When the system receives the command ‘n’, it halts the operation of Motor 2.
- Turn Right: This action corresponds to the command ‘X’. When the system receives the command ‘X’, it instructs the device or robot to turn right.
- Turn Left: This action corresponds to the command ‘Z’. When the system receives the command ‘Z’, it directs the device or robot to turn left.
- Relay 1 ON: This action corresponds to the command ‘A’. When the system receives the command ‘A’, it activates Relay 1, allowing current to flow through the corresponding circuit.
- Relay 2 ON: This action corresponds to the command ‘B’. When the system receives the command ‘B’, it activates Relay 2, enabling the associated circuit.
- Relay 3 ON: This action corresponds to the command ‘C’. When the system receives the command ‘C’, it turns on Relay 3, allowing current to flow through the respective circuit.
- Relay 4 ON: This action corresponds to the command ‘D’. When the system receives the command ‘D’, it activates Relay 4, enabling the associated circuit.
- Relay 1 OFF: This action corresponds to the command ‘E’. When the system receives the command ‘E’, it deactivates Relay 1, stopping the flow of current through the corresponding circuit.

- Relay 2 OFF: This action corresponds to the command ‘F’. When the system receives the command ‘F’, it turns off Relay 2, interrupting the flow of current through the respective circuit.
- Relay 3 OFF: This action corresponds to the command ‘G’. When the system receives the command ‘G’, it deactivates Relay 3, stopping the current flow in the associated circuit.
- Relay 4 OFF: This action corresponds to the command ‘H’. When the system receives the command ‘H’, it turns off Relay 4, interrupting the current flow through the corresponding circuit.

5 Results and Discussions

As shown in Fig. 2, the Scrub Master is designed using Arduino Uno, Motor Driver, Motors and connectivity features, it allows users to control and monitor the cleaning process remotely. System can control this system using the Bluetooth connectivity. The existing cleaning robot is specialized for particular surfaces or tasks, limiting their applications but the proposed cleaning robot is designed to be versatile, capable of cleaning a wide range of surfaces and environments, such as hardwood floors and tiles. By controlling it through Bluetooth, users can easily adjust its settings for specific cleaning requirements, making it suitable for various cleaning tasks.

Basically it offers a revolutionary approach to cleaning, leveraging advanced technology to enhance efficiency and users experience.

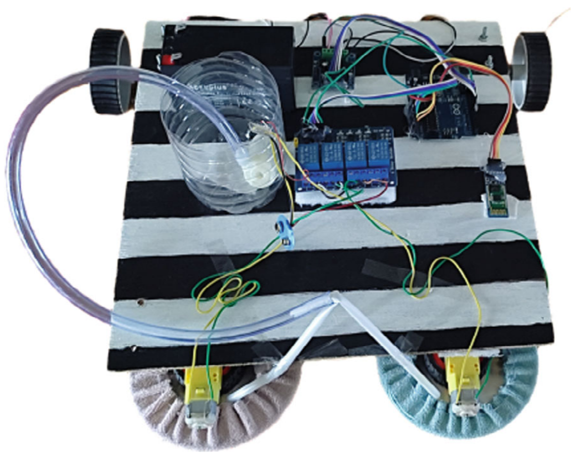


Fig. 2 Developed model of scrub master (top view)

6 Limitations

- As it is hardware and software, it needs time to time maintenance unless user will face some system breakdown
- As being a hardware robot technical issues may arrive.
- This floor cleaning robot is limited to clean floor stairs cannot be cleaned by this so it can be modified for cleaning of stairs.
- The water tank of a smart mop is usually limited, and it may need to be refilled multiple times during a cleaning session.

7 Conclusion

This paper illustrates how a robot can clean floor and make it easier and more efficient for humans by utilizing a wireless robotic system. This proposed system reduces the time and cost of labor. In the previous research papers, system had some drawbacks like colliding with objects, unable to clean wet surfaces, not able to clean edges and corners. But in this system, such drawbacks are overcome. In this research paper this system make avoid rugged surfaces and walls. This is one of the disadvantages of this research paper. Bluetooth module can make it little more user-friendly system.

8 Future Scope

The proposed system is convenient for users. But in the future, it is possible to add camera and voice assistant to enhance the functionality of this system. Setting a timer for a cleaning robot can save much time and effort of the user. As a future scope, advanced functionality such as rechargeable batteries, destroying disinfection on the floor. In the future by using AI and machine learning, the robot will be able to recognize various floor kinds such as hardwood, carpet, and tile. This information may be utilized to optimize the cleaning parameters of the robot, such as suction power, brush speed, and cleaning mode.

References

1. Varsha PH, Lavanya V, Meghana K, Rohan PS, Sneha R (2018) Sweepy—the smart floor cleaner. In: 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C), Bangalore, India, 2018, pp 124–126. https://doi.org/10.1109/ICDI3C.2018.00035
2. Murdan AP, Ramkissoon PK (2020) A smart autonomous floor cleaner with an android-based controller. In: 2020 3rd international conference on emerging trends in electrical, electronic

and communications engineering (ELECOM), Balaclava, Mauritius, 2020, pp 235–239. https://doi.org/10.1109/ELECOM49001.2020.9297006

3. Anbumani V, Geetha V, Renugha V, Praveenkumar V (2019) Development of ingenious floor cleaner using ARDUINO. Int J Recent Technol Eng 8(4):5848–5853
4. Mohammed MN, Al-Zubaidi S, Kamarul Bahrain SH, Zaenudin M, Abdullah MI (2020) Design and development of river cleaning robot using IoT technology. In: 2020 16th IEEE international colloquium on signal processing & its applications (CSPA), Langkawi, Malaysia, 2020, pp 84–87. https://doi.org/10.1109/CSPA48992.2020.9068718
5. Badamasi YA (2014) The working principle of an Arduino. In: 2014 11th international conference on electronics, computer and computation (ICECCO), Abuja, Nigeria, 2014, pp 1–4. https://doi.org/10.1109/ICECCO.2014.6997578
6. Jain V, Jasnani Y, Jota P, Solanki YS, Maheshwari A (2020) Smart broom. Int J Eng Adv Technol (IJEAT) 9(3s):2249–8958
7. Jarande PB, Murakar SP, Vast NS, Ubale NP, Saraf SS (2018) Robotic vacuum cleaner using Arduino with Wifi. In: 2018 Second international conference on inventive communication and computational technologies (ICICCT), Coimbatore, India, 2018, pp 1513–1517. https://doi.org/10.1109/ICICCT.2018.8473256
8. Saleem A, Iqbal A, Sabir A, Hussain A (2019) Design and implementation of an intelligent dust cleaner robot for uneven and nonstructural environment. In: 2019 2nd international conference on computing, mathematics and engineering technologies (iCoMET), Sukkur, Pakistan, 2019, pp 1–6. https://doi.org/10.1109/ICOMET.2019.8673450
9. Sridhar S, Himabindu R, Ashok Kumar P, Vasu R (2018) Design of IOT based vacuum cleaning and mopping robot. J Emerging Technol Innovative Res 5(4):520–525
10. Monika S, Aruna Manjusha K, Prasad SVS, Naresh B (2019) Design and implementation of smart floor cleaning robot using android app. Int J Innovative Technol Exploring Eng 8(4S2):250–252
11. RaveenaIshalavath M, Sandhyashree N, Rakshitha R, Arpitha Shankar (2020) Design and development of an automated floor cleaning robot for domestic application. Int J Eng Res Technol (IJERT) NCETESFT 8(14)
12. Manasa M, Vidyashree TS, Bindushree V, Rao S, Gowra PS (2021) Smart vacuum cleaner. Global Transitions Proc 2(2):553–558, ISSN 2666-285X
13. Ahsan SMT, Hassan MZ, Imam MT, Sadat KFMA, Hazari MR, Manzoor MLN (2014) Design and implementation of a robotic vehicle with real-time video feedback control via internet. In: 2014 International conference on electrical engineering and information & communication technology, Dhaka, Bangladesh, 2014, pp 1–6. https://doi.org/10.1109/ICEEICT.2014.6919047
14. Rissanen H, Mahonen J, Haataja K, Johansson M, Mielikainen J, Toivanen P (2009) Designing and implementing an intelligent bluetooth-enabled robot car. In: 2009 IFIP international conference on wireless and optical communications networks, Cairo, Egypt, 2009, pp 1–6. https://doi.org/10.1109/WOCN.2009.5010563
15. Chaudhry A, Batra M, Gupta P, Lamba S, Gupta S (2019) Arduino based voice controlled robot. In: 2019 International conference on computing, communication, and intelligent systems (ICCCIS), Greater Noida, India, 2019, pp 415–417. https://doi.org/10.1109/ICCCIS48478.2019.8974532
16. Kumar S, Mishra SK, Singh AK, Singh SB (2019) IoT-based smart floor cleaner with object detection and mapping. In: 2019 International conference on intelligent computing and control systems (ICICCS), 2019, pp 1–6. https://doi.org/10.1109/ICCONS.2019.8882099
17. Chavan DK, Borade SG, Sawarkar SV, Magar DB (2019) Smart floor cleaner robot using IoT. In: 2019 International conference on advanced communication control and computing technologies (ICACCCT), 2019, pp 1–6. https://doi.org/10.1109/ICACCCT.2019.8889752

Skin Disease Analysis Using Image Processing and Deep Learning

Rajeshwarrao Arabelli, Sriramoju Shruthi Vardhan, Nagelli Bhumika, Mekala Gopinath, Kancha Raghavendra, and Syed Musthak Ahmed

Abstract Cancer is a fatal condition brought on by the unchecked expansion of bodily cells. Every year, a considerable number of individuals die from cancer, which has been dubbed the most important public health issue. Any part of the human anatomy, which contains trillions of cells is susceptible to the development of cancer. Skin cancer, which starts in the skin's outermost layer, is one of the most common kinds of cancer. Previously, a range of imaging modalities and protein sequences were utilized in conjunction with machine learning techniques to identify skin cancer. Machine learning techniques have the drawback of requiring human-engineered qualities, which is a laborious and time-consuming operation. By enabling autonomous feature extraction, deep learning partially solved this issue. Several methods have been employed in this study to detect skin cancer using a public dataset. Cancer detection is a sensitive process that might go wrong if it is not carried out accurately and immediately. The capacity of every machine learning model to detect cancer has limitations. Individual learners' collective decisions should be more accurate than their individual decisions. The ensemble learning approach uses a range of learners to yield better outcomes. The ensemble learning approach uses a range of learners to yield better outcomes. Therefore, pooling student decisions on delicate topics like cancer diagnosis can increase prediction accuracy. In this research, we employ the Mobile Net, Xception, Efficient, Decision Tree, and Voting Classifier (DT + Gradient Boosting) models for recognition and categorization purposes. This work's experimental results provide strong support for its application in the diagnosis of various illnesses.

Keywords Deep learning · Computer version · Convolution neural network

R. Arabelli (✉) · S. S. Vardhan · N. Bhumika · M. Gopinath · K. Raghavendra
Department of Electronic and Communication Engineering, SR University, Warangal, India
e-mail: rajeshwarrao4322@gmail.com

S. M. Ahmed
Department of Electronic and Communication Engineering, Vagdevi Engineering College, Warangal, India
e-mail: musthak.ahmed@vecw.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_6

1 Introduction

The age of cells in the human animal generally follows a predestined model. The body's beginning, dynamic period, and passing need to happen aligned correctly capability appropriately. Disturbances to the normal request bring about the advancement of various infections. One of these infections is malignant growth. The human body, which might contain trillions of cells, can foster disease. Right when an individual has infection, certain cells begin to isolate fiercely, and these cells spread all through the body's tissues. Human cells ordinarily partition and increase to create new cells when the body needs them. Thus, the cell bites the dust, and new, sound cells have its spot. The exact and calculated instrument for cell demise is obliterated whenever malignant growth has spread. Thusly, the edge for cell irregularity and mischief increases definitively.

Cell endurance possibly happens when old cells bite the dust, and new cells are possibly created when required. The cells will keep on isolating uncontrolled on the off chance that they are not required, which could prompt the improvement of a neoplasm [1]. Disease makes both harmless and threatening growths structure. A large portion of developments made from cancer-causing cells are hurtful. Right when cells are perilous, they incite or pursue adjoining tissues. Threatening cancers partition, and a couple of harmful cells move by means of the blood or lymphatic framework to distant pieces of the body, shaping new growths that are distant from the first cancer. BENEDICT TUMORS, rather than harmful cancers, don't partition inside the body or inundate encompassing tissues. It is referenced that harmless cancers regularly fill in size. Nonetheless, while threatening cancers might recover after a medical procedure, harmless cancers can't come back after they are taken out. Although most harmless cancers in different pieces of the body don't represent a danger to life, the harmless growth in the cerebrum represents a huge danger to endurance [2, 3].

2 Literature Review

Advancements in machine learning (ML) techniques have simplified the analysis of skin diseases, particularly melanoma, which traditionally involved tedious and subjective manual examination. With computational resources and ML methods, the process of identifying skin lesions for melanoma detection has become more efficient and accessible.

This work examines methods for distinguishing healthy skin from diseased skin, aiming to advance robust models for automatically identifying skin conditions. Challenges include the complex nature of skin lesion photos and difficulties in distinguishing cutaneous cancers from epidermal images. The paper discusses preprocessing and segmentation techniques, recent research comparisons, and strategies for categorizing skin lesions across different types of skin conditions. Additionally, it addresses challenges in skin disease analysis using the ISIC 2018, 2019 datasets and evaluates classification methods with state-of-the-art AI algorithms [1]. Oncologists prioritize developing and discovering novel cancer treatments, with the National Cancer Institute (NCI) outlining its ongoing efforts. This includes describing specialists preparing for clinical trials and reporting on modified disease screening. The advantages and drawbacks of a molecular approach to drug development are discussed, emphasizing the importance of finding new cancer treatments despite differing opinions on the most effective methods [2].

This paper addresses the challenge of diagnosing dermoscopic skin cancer, particularly melanoma, which is aggressive if not detected early. It presents a machine learning technique for identifying melanoma using distinctive features extracted from dermoscopy images. The method demonstrates high accuracy, sensitivity, specificity, and AUC when evaluated on the PH^2 dataset, validating its effectiveness in melanoma detection [3]. The study proposes a method that combines different types of deep features and multi-scale convolutional neural networks to recognize various skin diseases more accurately. By merging deep features from different layers of neural networks and incorporating multi-scale analysis, the proposed method achieves better performance in identifying skin diseases compared to traditional methods [4].

By employing FCNs, which are specialized in image segmentation tasks, and leveraging the Jaccard distance metric, the proposed method achieves accurate and efficient skin lesion segmentation, aiding in medical diagnosis and treatment planning [5]. By training deep neural networks on a large dataset of skin lesion images, the proposed approach achieves accurate segmentation results, assisting dermatologists in identifying and analyzing suspicious lesions more effectively [6].

Through a systematic analysis of different deep learning architectures, training strategies, and datasets, the survey provides insights into the strengths and limitations of existing techniques for semantic segmentation, facilitating future research and development in this area [7]. By combining predictions from multiple EfficientNet models trained at different resolutions, the proposed approach achieves better performance in distinguishing between different types of skin lesions, offering potential benefits for dermatologists in clinical practice [8]. Through the incorporation of dilated convolutions and attention residual mechanism, the proposed approach effectively captures both local and global contextual information, enhancing the segmentation performance of skin lesions in medical images, which can assist dermatologists in diagnosis and treatment planning [9].

Zahra Waheed et al. in their 2017 paper present a streamlined machine learning technique aimed at diagnosing melanoma from dermoscopic images. The study likely delves into feature extraction methods to extract pertinent information from the images and employs a machine learning classifier to distinguish between melanoma and benign lesions. Their contribution lies in emphasizing efficiency within melanoma detection from dermoscopic images [10]. The research tackles imbalanced data in cancer prediction, particularly in breast and colon cancer cases. The authors propose a two-stage prediction method employing KNN and SVM algorithms, likely implementing techniques to handle imbalanced data. They compare the performance of these classifiers on various feature spaces. Their findings provide insights into selecting appropriate machine learning methods for imbalanced cancer prediction datasets [11]. Integration of support vector machine (SVM) with deep learning techniques has shown promise in improving the accuracy and robustness of melanoma detection systems. However, challenges such as class imbalance and dataset variability persist [12]. The paper introduces a method for automatic skin lesion segmentation in dermoscopic images using FCNs and the Jaccard distance loss [13]. This study explores using ResNets for automated melanoma recognition in dermoscopy images, leveraging their effectiveness in training deeper models and addressing gradient problems. Likely assessed on a public dataset, tfhe approach aims to improve melanoma detection accuracy. Their work advances deep learning for precise melanoma detection, potentially leading to early diagnosis and improved outcomes in dermatology [14]. The paper presents a new method for dermoscopic image segmentation using multistage FCNs, known for their suitability in image segmentation tasks. Their proposed multistage architecture cascades FCNs, potentially integrating techniques to enhance segmentation accuracy for skin lesions. Emphasizing improved accuracy over prior methods, crucial for skin cancer diagnosis, the study likely assesses its performance on a benchmark dataset and against existing techniques [15].

3 Methodology

Disease is a deadly disorder achieved by uncontrolled cell duplication in the body. Since it kills such countless individuals every year, cancer is the most serious issue for general well-being. Cancer can grow anyplace because of the human body's trillions of cells. Skin cancer is one of the most well-known types of cancer since it affects the skin's epidermis. Previously, using ML approaches, it was discovered that skin cancer had protein clusters and a range of imaging modalities. One drawback of ML techniques is the cycle of laborious and challenging requirements for human-designed qualities. The drawbacks of earlier techniques are that they are difficult and time-consuming.

4 Proposed Work

We employ the Mobilenet, Xception, Efficient Net, Decision Tree, and Voting Classifier (DT + Gradient Boosting) models for disclosure and depiction in this study. This study's exploratory outcomes support their application to the finding of different illnesses with persuading proof. Via preparing and assessing different models on the equivalent dataset, this study thinks about their presentation. In actuality, this connection is important and crucial since it will indicate which model performs best on this dataset and help us move forward with our investigation into the various model designs. The benefits of this research are substantial. Reliability and accuracy were guaranteed by the careful use of highly precise models. The experimental results reported in this study offer compelling evidence for the possible use of these techniques in the diagnosis of a range of diseases.

4.1 Functional Requirements

- Data gathering
- Pre-processing of data
- Instruction and assessment
- Modeling
- Predicting

4.2 Non-Functional Requirements

An idea building's distinctive trade name is based on a single non-functional requirement (NFR). They examine what unhelpful models—like comfort, safety, ease, openness, and differing rules—are essential to the new fashion structure's success. "How quickly does the page load?" is not a mandatory condition. In one instance, organizational systems that prohibit customers from meeting unrealistic models cannot satisfy customer requirements.

The explanation preparation might be limited or laid off across different active abundances using non-practical basics. The scene should load in three seconds, e.g., when there are 10,000 items linked to the internet quickly and there are experienced users. It is also evident that the portrayal of unrealistic essentialities is by no means essential as a supporting need. Figure 1 depicts a system architecture for a skin lesion classification process which consists of three stages.

- Necessities related to usability, sensibility, recoverability, security, data authenticity, limits, openness, scalability, interoperability, consistency, suitability, regulatory, and regular requirements.

  Image segmentation Segmented lesion Learning model Prediction model.

4.3 Modules

The modules listed below were used to complete this project:

- Using this module, we will import data into the system for data exploration. The purpose of this module is to read data for processing.

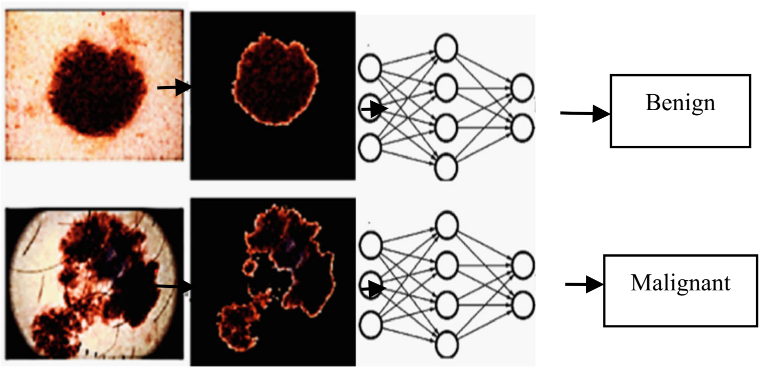


Fig. 1 System architecture

- Data splitting into train and test: This module will help separate the data into train and test.
- Build models with Voting Classifiers (DT + Gradient boosting), Efficient Net, Xception, Mobile Net, and Decision trees.
- By using this module, users will need to register and log in.
- User input: Information for forecasting will be available by using this module.
- The final forecast is shown in the forecast.

5 Implementation

5.1 MobileNet

A specific kind of neural network designed for embedded and mobile vision applications is called MobileNet. Its lightweight deep neural networks are made possible by its streamlined design, which lowers latency for embedded and mobile devices. By using depth-separable convolutional architectures, MobileNet improves efficiency without compromising performance in contrast to standard models. On the other hand, Xception is unique in that it is a 71-layer conventional neural network. Xception is still a potent tool even if it deviates from MobileNet's simple design; a trained version may be downloaded from the ImageNet database. It is noteworthy that this pre-trained network has exceptional adaptability, correctly categorizing a wide variety of items, including mice, keyboards, and animals. Conversely, EfficientNet presents a new method for designing and scaling neural networks. It equally scales depth, breadth, and resolution parameters by utilizing complex coefficients. Because of its ability to maintain optimal performance under a variety of computing restrictions, EfficientNet is a flexible option suitable for a broad range of applications. Figure 2 shows the steps involved in structural testing.

Types of structural testing.

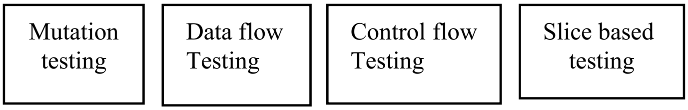
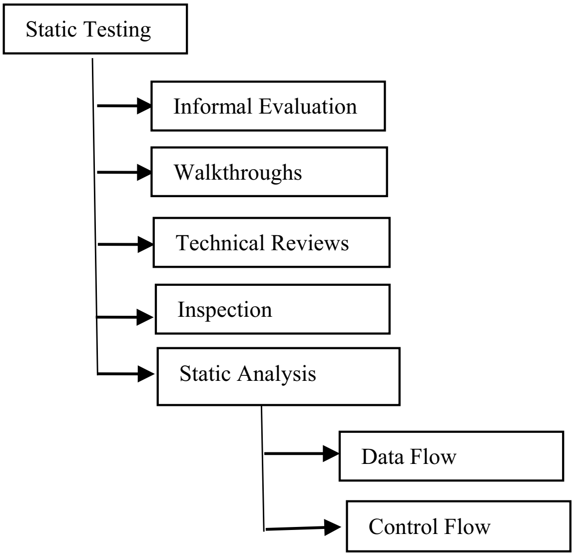


Fig. 2 Types of structural testing

Fig. 3 Structural testing



5.2 Decision Tree

The decision tree is a nonparametric supervised learning technique used in regression and classification applications. It is organized like a tree, with internal, leaf, branch, and root nodes. Figure 3 depicts the outlining of the various phases of structural testing. A voting classifier using gradient boosting plus DT: A Voting Classifier is a well-liked machine learning technique that Kaggle uses to boost model performance and climb the rank ladder. Voting Classifier has substantial disadvantages, yet it may be used to improve performance on real-world datasets.

6 Experimental Results

The development of a webpage titled "Skin Classification" is depicted in Figs. 4, 5, 6, 7, 8 and 9. Upon accessing the webpage, new users are redirected to the User Sign-up page (Fig. 5), which features fields for Username, E-mail, Mobile number, and Password entry. Conversely, returning users have the option to navigate directly to the User Sign-in page (Fig. 6) to input their login credentials. Following successful authentication, users are directed to the main page (Fig. 7) where they can upload images for classification. Subsequently, users proceed to upload the images to be classified (Fig. 8), with the results being displayed on the main page (Fig. 9). This interface facilitates the efficient classification of skin images, thereby contributing to dermatological diagnosis and analysis.

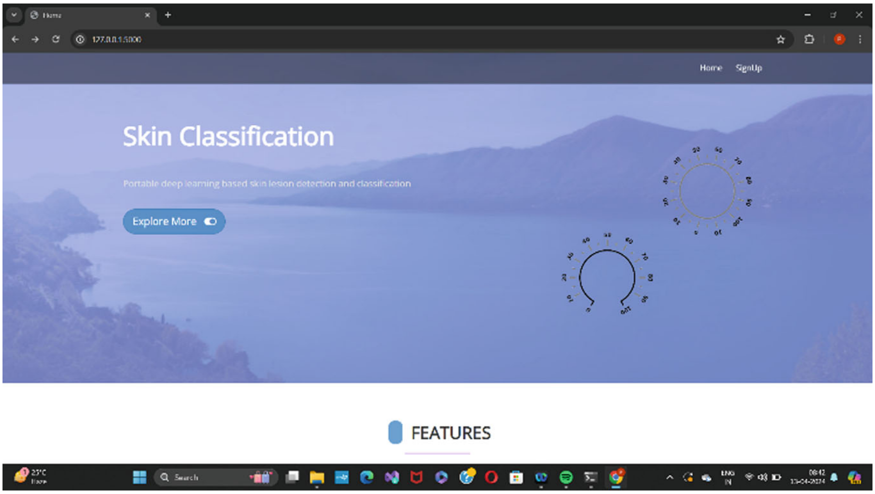


Fig.4 Home screen page

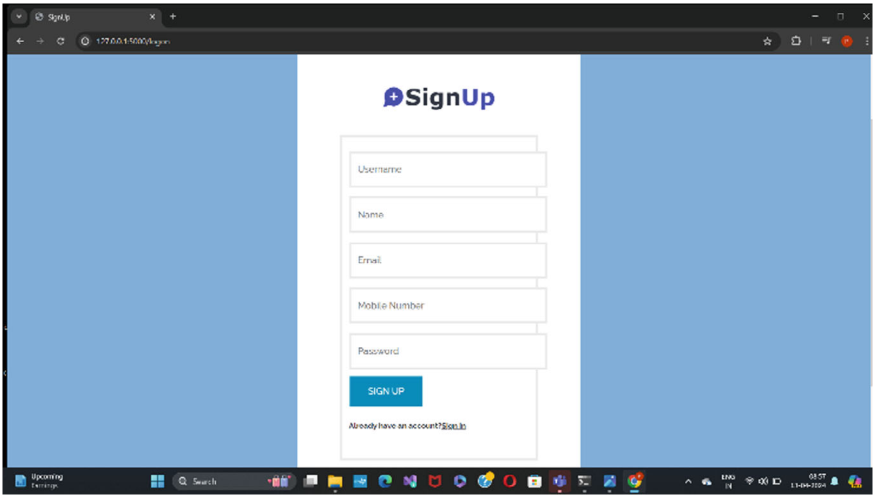


Fig. 5 User sign-up page

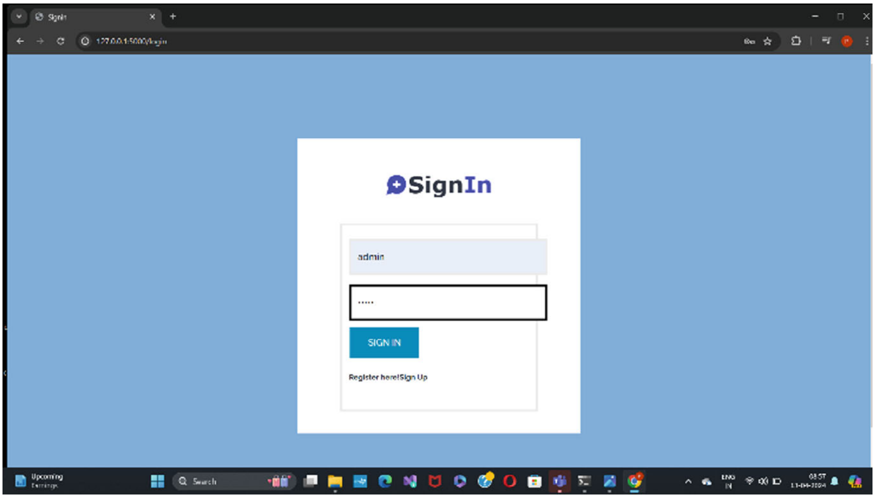


Fig. 6 User sign-in page

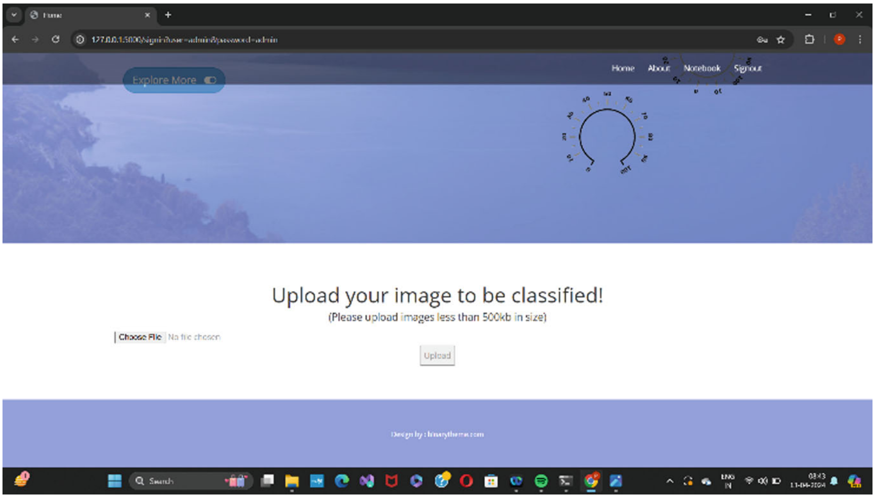


Fig. 7 Main page

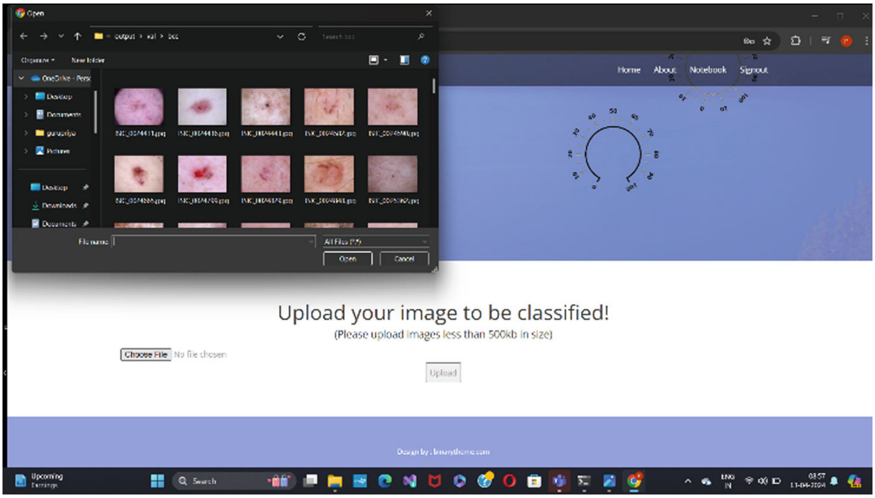


Fig. 8 User input page

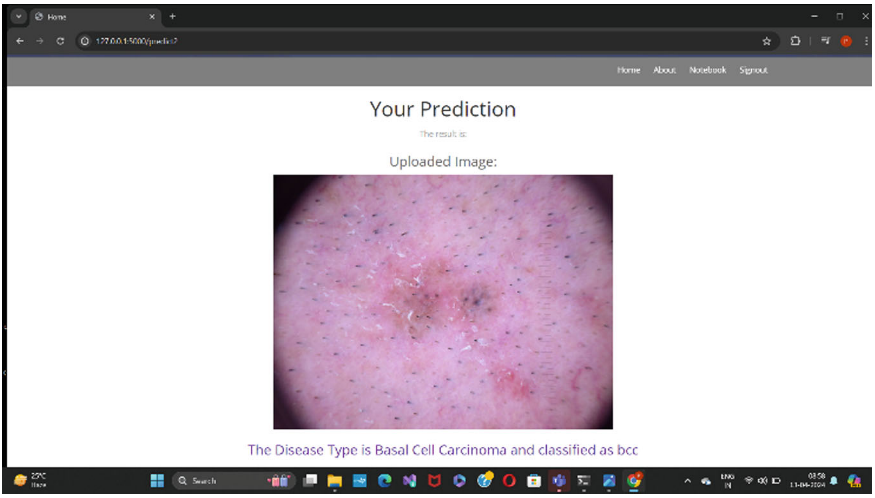


Fig. 9 Prediction result

7 Conclusion

The malignant lesion is the essential figure demise from skin cancer. Treatment might be conceivable assuming it is found in its beginning phases. In the composition, sickness distinguishing proof has been performed using deep learning procedures, and at this point, the sufficiency of individual understudies is confined. By consolidating the singular choices of different understudies, delicate points like malignant growth can further develop execution. A group model for distinguishing cutaneous disease was created in this review. It is made by present day deep learning models. The findings show that the suggested group achieved a mean accuracy of 93.5 percent with a characterization preparation season of 106 s. Future research will examine the feasibility of skin disease localization methodologies based on reinforcement learning.

Acknowledgements We would like to sincerely thank SR University for all of their help and advice over the course of the project. Their efforts have greatly improved this paper's quality.

References

1. Afroz A, Zia R, Garcia AO, Khan MU, Jilani U, Ahmed KM (2022) Skin lesion classification using machine learning approach: a survey. In: Proceedings global conference on wireless optical technologies (GCWOT), Feb 2022, pp 1–8
2. Grever M, Schepartz S, Chabner B (1992) The national cancer institute: cancer drug discovery and development program. Semin Oncol 19:622–638
3. Foundation MR (2016) Melanoma is the deadliest form of skin cancer. Available. https://melanoma.org/
4. Xie H, Lu Y, Cheng L, Yan Z, Zhou J (2019) Skin disease recognition using deep feature fusion and multi-scale convolutional neural networks. IEEE Trans Med Imaging
5. Yang D, Xu S, Zhou Z, Huang H (2018) Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. Comput Biol Med
6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) A deep learning approach for pigmented skin lesion segmentation. IEEE Trans Med Imaging
7. Al-Rizzo, Murugan SB, Chakravarthy JJK, Mohammed MA (2020) A survey on deep learning techniques for image and video semantic segmentation. Neural Comput Appl
8. Chen T, Kori P, Wadhawan K, Bernal L, Catanzaro B (2021) Skin lesion classification using ensembles of multi-resolution EfficientNet models. arXiv preprint, 2021
9. Zhang Y, Zhou Y, Xu Z, Miao C (2022) Automatic skin lesion segmentation using dilated fully convolutional network with attention residual mechanism. IEEE Access
10. Waheed Z, Waheed A, Zafar M, Riaz F (2017) An efficient machine learning approach for the detection of melanoma using dermoscopic images. In: Proceedings international conference on communication, computing and digital systems (C-CODE), Mar 2017, pp 316–319
11. Majid A, Ali S, Iqbal M, Kausar N (2014) Prediction of human breast and colon cancers from imbalanced data using nearest neighbor and support vector machines. Comput Methods Programs Biomed 113(3):792–808
12. Seeja R, Suresh A (2019) Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM). Asian Pac J Cancer Prev 20(5):1555
13. Yuan Y, Chao M, Lo Y-C (2017) Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Trans Med Imaging 36(9):1876–1886
14. Yu L, Chen H, Dou Q, Qin J, Heng P-A (2017) Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging 36(4):994–1004
15. Bi L, Kim J, Ahn E, Kumar A, Fulham M, Feng D (2017) Dermoscopic image segmentation via multistage fully convolutional networks. IEEE Trans Biomed Eng 64(9):2065–2074

An Efficient Soft Descrambling Procedure to Enable Soft-Source-Information-Combining for Wireless Networking

K. Shankar, Syed Javeed Basha, Vinit Kumar Gunjan, and D. Vishnuvardhan

Abstract This work explains the multi-stream networking type of systematic paradigm known as the soft-source-information-combining (SSIC) for the purpose to hold up wireless mode of Internet of Things (IoT)-based applications. This kind of methods require high amount of reliability to carry out the procedure in a technologically smooth manner. In the concerned SSIC networking, a source SSIC dispatcher is used which distributes duplicate packets among numerous streams that may be established across various physical wireless networks. The decoder makes soft information available in the event that a packet on a stream cannot be deciphered because of wireless interference and noise. In this work, the accessible soft information of duplicates is then combined by an aggregator to improve reliability. A significant challenge is determining how to decode scrambled soft bits from various streams in order to enable accurate SSIC. To address issue, soft descrambling (SD) methodology is provided to reduce bit error rate (BER) as well as packet error rate (PER) at SSIC's output. Reduced BER and PER is related to achieve SNR is beneficial for dependable wireless networking. As a result, this descrambling process is designed and developed to be simulated in order to obtain appropriate results using the MATLAB programming language.

K. Shankar
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India

S. J. Basha (✉)
Design Verification Engineer, SmartSoC Solutions Pvt. Ltd, Bengaluru, India
e-mail: munnu483@gmail.com; fahimaits@gmail.com

V. K. Gunjan
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, India

D. Vishnuvardhan
Department of Electronics and Communication Engineering, JNTUACE, JNTUA Ananthapuramu, Ananthapuramu, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_7

Keywords Packet error rate (PER) · Soft descrambling · Ultra reliable communications · Bit error rate (BER)

1 Introduction

Wireless communication is the most innovative and quickly growing technological sector in the communications sector. Wireless networks are crucial for facilitating communication. The modern human lifestyle is completely dependent on them, and as technology advances daily, new features are added. This may also help to ensure the reliability of wireless networks. Many IOT applications need high communication performance, and as wireless networks are truly error-prone, this could affect the reception data of destination. For the support of medical, audio visual, mobile industrial and Internet of Things (IoT) applications are with rigorous standards for service quality, wireless network dependability must be improved [1].

Extraction of source data during data transmission results in numerous errors and data losses because of inefficient extraction techniques. Due to the inability to extract data effectively in SSIC networking, utilizing typical descrambling techniques may result in data losses [2]. Soft descrambling technique has an ability to reduce data loss by taking soft information into account. If data packet is unable to provide source data, then using soft information of that packets helps to reproduce source data. Data losses are represented by bit and packet error rates, hence lowering these rates lowers data losses.

2 Literature Survey

Evaluating potential applied solutions for the high-reliability and low latency of wireless IoT networks, as well as numerous applications, fundamental performance constraints, and technical limitations. The MAC layer, physical layer, and network layer are discussed [3]. We address grant-free methods that have recently been presented for the MAC layer after first discussing techniques for managing spectrum and power resources that are optimal [4].

Innovating technology used in Oil and the Gas industry to reduce trading prices of them. The Internet of Things (IoT), which enables wireless concurrent data collecting, processing, and also analysis from a variety of equipment, processes, and activities to accomplish these goals, is at the vanguard of this digital transformation [5]. IoT-enabled solutions have a substantial impact on a variety of operations/applications in the upstream, midstream, and downstream sectors and present a range of options to boost socioeconomic advantages [6].

Techniques for reliability and the latency quantities are measures to being employed in Wi-Fi mesh network routing tree optimization technique [7]. WI-FI networks are designed to ensure the reliability and reduces the errors which makes a chaos in communication. We suggest a method for multichannel (MC) tasks. With this approach, bandwidth of signal is distributed among various mesh nodes in accordance with the anticipated traffic load that they can support [8].

In wireless networks, bit error occurs due to noise and interference may overcome by using coded and modulations [9]. The majority of the packets that are retransmitted in existing wireless mesh network protocols end up delivering that are squandering network capacity by being received repeatedly. Data packet recovery may get through some techniques [10].

3 Methodology

In SSIC networking, data can be separated into data packets at the transmitter side using various coding techniques. This is known as scrambling. Before transmission, a scrambler performs data stream manipulation. A descrambler is also included on the receiver side and can be used to reverse scramble data, turning data packets back into their original form. Conventionally, we follow some methods to perform descrambling. To reduce the error rate compared to the typical methods, we propose soft descrambling method.

4 Existing Method

Hard descrambling is a traditional method for doing the descrambling process. In this technique, typically, a decoder initially performs hard decoding by extracting the masking order from the preamble section of binary values. This is done after converting the Log-Likelihood Ratios (LLRs) of all the packets into binary values of 0 and 1. The ensuing payload's 1 or 0 values are subsequently decoded using the masking sequence (Fig. 1).

This method decodes each bit by treating it as either certainly 1 or 0 after taking a bits stream or block of bit sets from the threshold stage at the receiver. The received pulses are sampled, and their voltages are compared to their threshold values. A voltage is encoded as 1 if it is higher than the threshold value and as 0 otherwise. Regardless of how close of voltage values is to threshold, the decoding is performed. Since channel decoder basically does decoding by that the soft data might be lost, the standard descrambling is incongruous to combine the packets without fail. The masking sequence of hard decoder is frequently incorrect, resulting in inaccurate descrambling of the LLRs and incorrect pairing of the LLRs of the two packets. That can obtained by observing the bit and packet error rates.

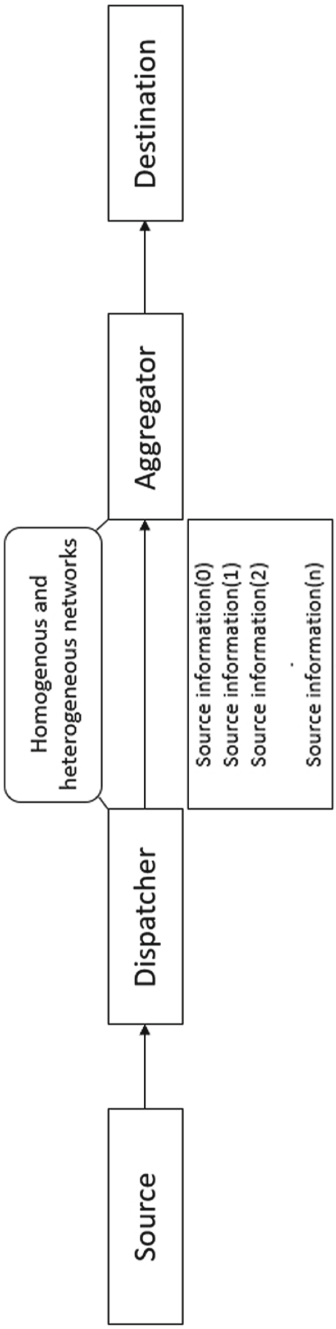


Fig. 1 SSIC networking

5 Proposed Method

Data losses happen while utilizing conventional techniques for descrambling because soft information is lost. We suggest a different method that can collect soft information in addition to the combining procedures because such packets cannot obtain information through that channel decoding. Generally, masking is to prevent long consecutive sequences of 1s or 0s, the source masks (XORs) the source bits with a random binary sequence before conducting channel coding. Different random masking sequences are employed within a NIC to XOR succeeding packets. Additionally, two NICs at the source may use various random sequences to mask copies of the same packet. The receiver must firstly decode mask of each duplicate packet to get recovery of the source bits that have been scrambled (Fig. 2).

The two NICs at the source could use various masking sequences to scramble the two packets. Prior to combining, the LLRs of resource bits output from the decoders of both channels must first be decoded using their respective scrambling masks. Although this approach is straightforward, it really decodes r0, r1,..., r6 in the same way as conventional HD; therefore, this is a hard r0, r1,..., r6 soft method. The only difference is that $x\ m$ obtained is still as soft information. In particular, if the hard bits are incorrectly decoded, the system will be unable to retrieve the right descrambled $\boldsymbol{x_m}$.

Next, two advanced SD variations are presented to acquire $x\ m$ is better estimated. There has been redundancy for both versions introduced to "r0, r1,..., r6" to allow for the presence of more than the seven pilot bits are employed to transport the data pertaining to r0, r1,..., r6.

That two variations are represented by HRSX and SRSX. Both of them contain a soft-source-combining work that is varied with the bits r and x nature either soft or hard. This variation is known as hard-r-soft-x (HRSX) because the scrambling bit is hard decoded in order to obtain the soft information. While waiting, it is important to remember that the naive technique is a variant of HRSX that does not use "redundant encoding". The second kind is known as soft-r-soft-x (SRSX) because soft decoding is used to create soft bits from the scrambling bit.

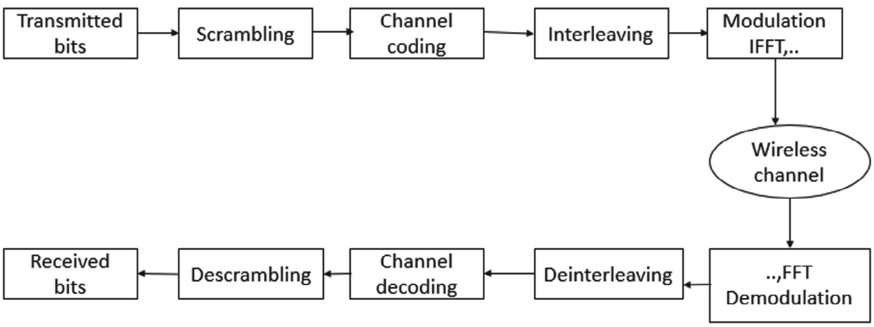


Fig. 2 Wi-Fi

Soft descrambling is a new technique for decoding (SD). First, SD enables the descrambling of scrambled LLRs of source bits without converting them to binary values. As a result, SD retains the soft data needed by SSIC. Second, the "soft" masking sequence is also produced by SD. In SSIC, accurate bit error rate (BER) and the packet error rate (PER) performance may result from decrypting the LLRs with the sequence of soft masking as compared to hard masking sequence.

6 Result

We ran MATLAB simulations in which the WLAN format waveform was created and transmitted across AWGN channel in order to verify the theoretical performance of SD. The total results are shown as graphs with each ratio's axis. We suggested a method called SD to increase reliability, and we wrote a piece of code in MATLAB to illustrate the results visually. Simulation may be useful when designing a system for a consistent network. The analysis of bit and packet error rate performances with the desired signal factor. Results are shown in logarithmic graphs taking x as SNR and y axis is different with a figures.

The simulation outcomes are shown in Fig. 3 which shows the performance of seed BER of the data streams by proceeds as y axis. In this graph, we compare the values of the y axis with the x axis for the traditional technique and the soft descrambling technique, these are indicated as dotted line and line with circles, respectively. In addition, the SD values are lower when compared to the conventional approach, and the error rate values steadily decline as the SNR values rise. The BER values for conventional, HRSX, and SRSX techniques are displayed in Fig. 4. Both variants of SD are slightly same values and it different from conventional. When compared to SRSX, HRSX has lower values at 9–12 (DB) SNR values, and both are adequate for descrambling to retain dependability.

Figure 5 shows the PER values as y axis and SNR values as x axis for all methods which shown in previous figure. Nearly all of the various method values in the graph are the same, and it doesn't indicate any differences between these approaches. In summary, SSIC with SD greatly increases communication reliability when soft information from many streams is present. Additionally, SRSX recommended when the SD is required in SSIC, since it still performs well even when L is tiny. For instance, SSIC with the SRSX may already achieve accurate BER and PER performance when $L = 16$, as shown in Fig. 3, and an additional increase in L is not required. Keep in mind that the SERVICE field of l IEEE 802.11 PPDU also comprises a 16 zero-length pilot bits. Since SRSX, $L = 16$ doesn't need any additional pilot bits above those provided by standard, it is best option for delivering SSIC over Wi-Fi networks.

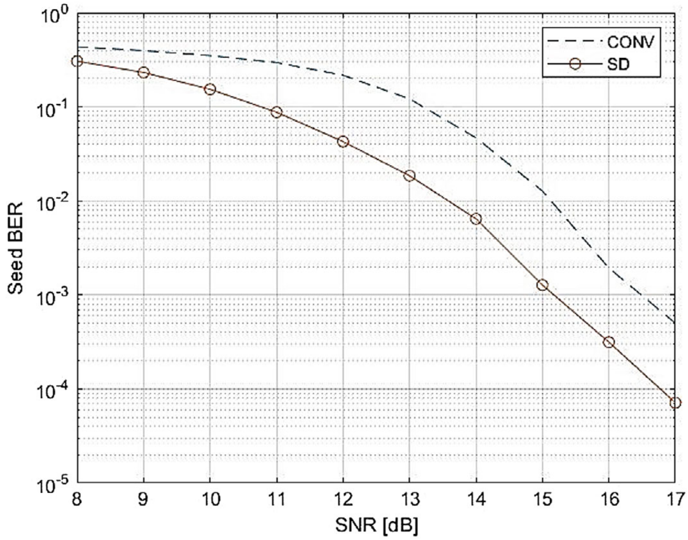


Fig. 3 $L = 16$

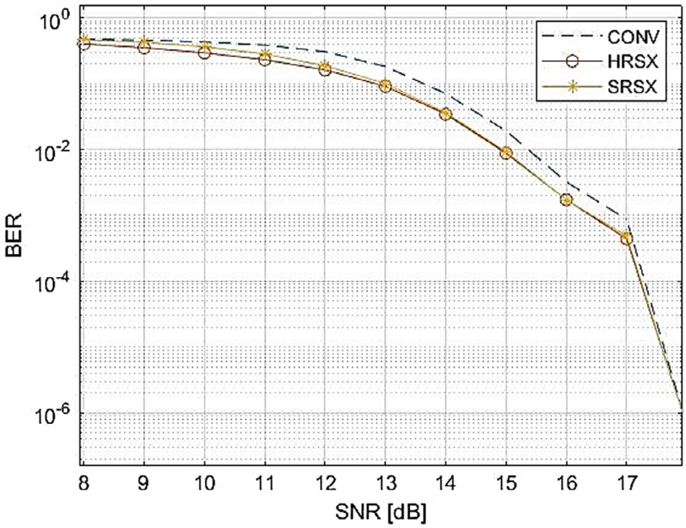


Fig. 4 $L = 16$

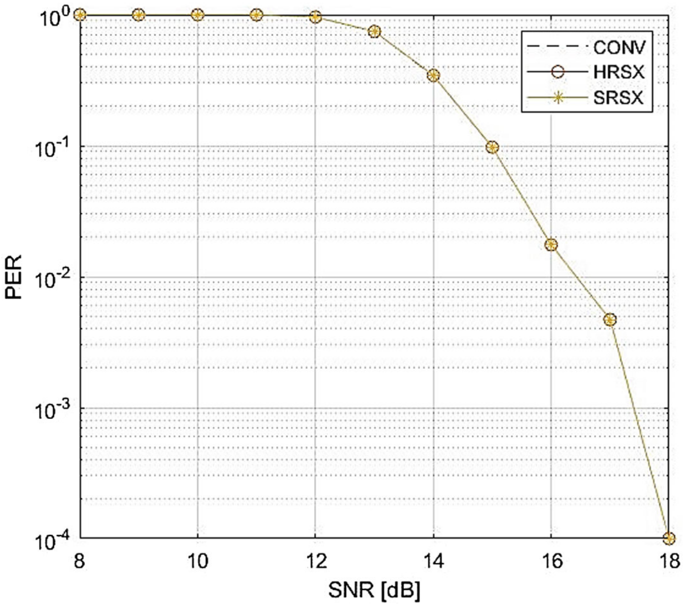


Fig. 5 $L = 16$

7 Conclusions

For the wireless communication to provide dependable network proposing a new technique to increase the extracting data capacity. In order to achieve extremely reliable communication, this research proposes that the soft-source-information-combining (SSIC) architecture that associates soft information of reproductions. Soft descrambling method in which the utilization of soft information to ensure the source data by combining as follows in this method. In this, the descrambler minimizes bit error rate and packet error rate w.r.t signal noise ratio (SNR). SD provides the sequence of soft masking needed to effectively performs decoding the soft information and reduce the data losses.

References

1. James LA, Wanasinghe TR, Mann GK, Gosine RG, De Silva O, Warrian PJ (2020) The internet of things in the oil and gas industry: a systematic review. IEEE Internet Things J 7(9):8654–8673
2. Karimullah S, Vishnuvardhan D, Bhaskar V (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Pers Commun 127:3041–3059. https://doi.org/10.1007/s11277-022-09909-2
3. Rampa V, Savazzi S, Spagnolini U (2014) Wireless cloud networks for the factory of things: connectivity modeling and layout design. IEEE Internet Things J 1(2):180–195

4. Karimullah S, Vishnu Vardhan D, Basha SJ (2020) Floorplanning for placement of modules in VLSI physical design using harmony search technique. In: Kumar A, Paprzycki M, Gunjan V (eds) ICDSMLA 2019. Lecture notes in electrical engineering, vol 601. Springer, Singapore. https://doi.org/10.1007/978-981-15-1420-3_197
5. Xiao M, Ma Z, Pang Z, Poor HV, Xiao Y, Vucetic B (2019) Highreliability and low-latency wireless communication for internet of things: challenges, fundamentals, and enabling technologies. IEEE Internet of Thing's J 6(5):7946–7970
6. Valenzano A, Scanzio S, Cena G (2016) Experimental evaluation of seamless redundancy applied to industrial Wi-Fi networks. IEEE Trans Ind Inf 13(2):856–865
7. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. Springerplus 5:507. https://doi.org/10.1186/s40064-016-2152-2
8. Karimullah S, Sai Sumanth Goud E, Lava Kumar Reddy K (2023) Spectral efficiency for multi-bit and blind medium estimation of DCO-OFDM used vehicular visible light communication. In: Kumar A, Senatore S, Gunjan VK (eds) ICDSMLA 2021. Lecture notes in electrical engineering, vol 947. Springer, Singapore. https://doi.org/10.1007/978-981-19-5936-3_83
9. Shaik F, Giri Prasad MN, Rao J, Abdul Rahim B, SomaSekhar A (2010) Medical image analysis of electron micrographs in diabetic patients using contrast enhancement. In: 2010 International conference on mechanical and electrical technology, Singapore, 2010, pp 482–485. https://doi.org/10.1109/ICMET.2010.5598408
10. Anu Priya D, Vishnuvardhan D, Mamatha G, Shaik F, Karimullah S (2023) VIP development of SPI controller for open-power processor-based fabless SOC. In: Kumar A, Ghinea G, Merugu S (eds) Proceedings of the 2nd international conference on cognitive and intelligent computing. ICCIC 2022. Cognitive science and technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_2

Design and Development of Fossil Image Segmentation System Using Hybrid Evolutionary Algorithms

S. Fahimuddin, Shaik Karimullah, Vinit Kumar Gunjan, P. Hari Obulesu, and D. Vishnuvardhan

Abstract Dinosaurs and other extinct animals are a part of our past, and by studying and analysing them, one may gain a clear understanding of what the planet was like millions of years ago. Even if the majority of them are extinct, palaeontology study may be done on their fossils or remnants. The importance of palaeontology to the modern and future worlds cannot be overstated. One may discover how creatures changed the physical world, as well as how climate change affected earlier species. It is also possible to better understand the concepts of extinction, evolution, and biodiversity. Scientists are hopeful that the environment will hold onto many dinosaur fossils and preserve and make pictures of them available. Applications for image segmentation in this field can be useful to palaeontologists. Applications in this area of picture segmentation might be useful to palaeontologists. This instrument may be used to determine the number of bones present in a sample and to learn the geometrical details of each section. In this research, image processing methods and various bio-inspired or metaheuristic algorithms are applied to pictures of dinosaur fossils. In order to filter additives depending on a range of parameters, the histogram equalization approach is utilized as a pre-processing stage. The contrast is altered using the Evolution Strategy histogram equalization. The input picture is quantized using the cultural method as well. The edge identification approach benefits from simulated annealing to effectively separate each piece. The first step consists of Particle Swarm Optimization in conjunction. An picture that has been segmented effectively.

S. Fahimuddin · S. Karimullah · P. H. Obulesu (✉)
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: p.hariobulesu@gmail.com

V. K. Gunjan
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, India

D. Vishnuvardhan
Department of Electronics and Communication Engineering, JNTUACE, JNTUA, Ananthapuramu, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_8

Keywords Dinosaurs · Histogram equalization · Fossils · Evolutionary algorithms · Paleontological

1 Introduction

The early 1800s saw the beginning of paleontological research. The English geologist William Smith first noted the value of using fossils to understand strata in 1815. Georges Cuvier, a French naturalist, began comparing the anatomy of living creatures to that of their prehistoric fossil relatives about the same period. By using the image segmentation technique, a digital image may be split up into several parts [1]. By either simplifying or modifying how it is portrayed, the objective is to make an image more intelligible and simple to study. Object identification is one of the many computer vision applications that makes use of the image segmentation approach. It is crucial to better comprehend how certain species became extinct in order to develop strategies to prevent it from happening again. It is crucial to learn about the prior Earthly events that have taken place. For millions of years, a wide variety of animals, including dinosaurs, have inhabited our world. While some of these dinosaurs, such as birds, are still alive today, the majority of them have perished as a result of asteroid strikes, natural disasters, and climate change. History has always been fascinating to humanity because it helps us understand in what way our descendants and other species survived, lived, and died at different times in history [2]. In the fossils that are still there, nature has managed to hold onto some remnants of this past. Here, dinosaur skeletons are examined using image processing techniques. Because it appears that this study mostly uses bio-inspired or metaheuristic algorithms variety of image processing tasks may be handled by them. These metaheuristic image processing techniques are more tempting to employ since they can easily outperform traditional methods. To create the final segmented image, we combine fundamental techniques with some metaheuristic image processing. Palaeontologists might analyse dinosaur remains more fully using the recommended method. Four dinosaur fossil images from distinct palaeontological periods are also used as test materials. On the same test pictures, the recommended technique is also related against K-means, Watershed, and Otsu algorithms, and its performance is evaluated using the same parameters for the three aforementioned performance metrics [3].

The importance of palaeontology in the present and the future cannot be overstated. Additionally, it is feasible to comprehend the ideas of extinction, evolutionary changes, and biodiversity better. We can discover both how creatures have changed the physical world and how past species have been influenced by climate change. Fossils provide valuable information on the evolution of life on Earth. They investigate the genesis of the species, how they overcame different challenges, how they adjusted to environmental changes, and if they developed or vanished [4].

It is clear that the K-means algorithm could separate small groupings of samples into different cluster sizes based on how close together they were in space. The same method might be applied to figure out the colour, intensity, or degree of grey for each pixel in an image. One of the best applications of K-means on photographs is that of Dhanachandra et al. Naturally, this method belongs to the field of cluster-based segmentation techniques. Watershed image segmentation technique is another example. This technique is based on the notion of the flow of water in valleys or rivers and is region-centric. The system's operation is dependent on its water level with the ground, as the height above the water acts as a boundary among each zone or portion. Otsu's thresholding technique segmentation is a renowned approach in the threshold-based segmentation area. It serves as a thresholding method for pixel intensity values, reducing inter-class fluctuation and dividing pixels into the foreground and background classes. The quantity of threshold values affects how many segments there are. Another interesting method for region-based segmentation is the use of Gabor filters. This method, also known as "image texture segmentation," is used by the system to try and separate the patterns of different objects. It is also fairly popular to segment digital images using algorithms that draw inspiration from biology. One of these instances is the segmentation of images using a genetic algorithm and K-means clustering [5].

2 Literature Survey

The region-based "watershed segmentation" method may be used with image morphology. There must be at least one marker, sometimes known as a "seed" point, inside of each object in the picture, including the background. The markers are either explicitly chosen by the operator or automatically created while accounting for the application-specific data of the objects using markers to accurately outline the limits by running the watershed over a certain background. It works by collecting specific backdrop and foreground information, running the watershed with markers, and determining the exact bounds. It costs a lot of money to construct a watershed. As a result, crop yield is increased. Watersheds do have certain drawbacks, such as the expensive cost of watershed building. Rainfall in the area should be adequate. A fall in the area required for a watershed construction should be adequate. Therefore, there aren't many locations where a watershed might be built. Watersheds require regular maintenance [6].

Equalization of the histogram in the evolutionary approach Algorithm for Evolutionary Strategies on 18 January 2022, Seyed Muhammad Hossein Mousavi published Histogram Equalization. These lines of code allow the Histogram Equalization function to apply the Evolution Strategy algorithm on the input parameter "target histogram" in order to fit the model in an evolutionary manner. The system uses several pre- and post-processing techniques that might be used as a picture-improving technique. Naturally, you may play with the parameters and use your own image to suit your preferences. This code can identify between veins and tissues in both low-quality and high-quality images, making it suitable for use in medicine. The method's benefit is that it is a rather straightforward technique that makes use of an invertible operator and adjusts to the input picture. If the histogram equalization function is known, the original histogram should theoretically be recoverable. This method has certain issues since it depends too much on the quantity of bins. The variable's maximum and lowest values are far too significant. It hinders the recognition of significant values. There is no evidence to justify the difference between continuous and discrete variables. It makes it difficult to compare distributions. It is challenging to do if you don't have all the knowledge in your memory [7].

Edge detection is a crucial technique for locating essential regions in the digital image that might be useful for feature extraction. In fact, edge detection makes it possible to crop off unnecessary areas of pictures. Although there are several edge detection techniques available, we advise employing a powerful evolutionary-based system to extract the crucial areas of the image. The system applies a number of pre- and post-processing techniques, such as filters and morphological operations, and it detects edges in the picture using a modified Ant Colony Optimization technique. The main goals are to calculate the system's performance and test it in various colour spaces. The research also uses colour images with depth data from Kinect V.2 in the validation phase to further understand the edge detection concept in depth data [8]. The system will then be compared to other well-known edge recognition systems under the same circumstances after being evaluated using seven picture quality evaluation criteria, including peak signal-to-noise ratio, mean squared error, structural similarity, and more (primarily related to edges). The durability of the proposed system is assessed using images with various noises added to them, such as Gaussian, Salt and Pepper, Poisson, and Speckle sounds. Additional calculations are necessary in order to produce the best edges possible, slightly increasing runtime computation. But this time has been significantly reduced by modern methods [9].

This technique is utilized for both machine learning-based along with image processing-based wound diagnostics to enhance performance. This study investigates the performance of three different image segmentation algorithms (Coye, Otsu's thresholding, and Grabcut) for examining the segmentation of skin lesion pictures. The Matthews Correlation Coefficient is used as a performance measure. Object detection is a technique that determines the quantity of objects in a segmented image. It is commonly employed alongside well-known metrics like Dice coefficient, accuracy, specificity, sensitivity, Jaccard index, precision, and Matthews correlation coefficient to assess the efficacy of segmentation. The Grabcut method yields optimal picture segmentation outcomes when these parameters are implemented. This study [10] showcases the importance and practicality of using the Matthews Correlation Coefficient score to evaluate the effectiveness of picture segmentation, specifically in the context of skin lesion segmentation.

3 Methods

3.1 Existing Method

Otsu's

The threshold for the thresholded black and white pixels is selected to minimize intra-class variation. Thresholding is a method that separates an intended object from the picture that serves as its background by giving each pixel a value of intensity T (threshold), which decides whether the pixel is thought of as being in the background or the object.

When the histogram has a bimodal distribution with a deep and sharp valley separating the two peaks, Otsu's method is successful. The Otsu approach performs poorly in environments with loud noise, tiny objects, uneven illumination, and larger intra-class variance than inter-class variance, much like all other global thresholding algorithms. For several situations, localized variations of the Otsu method have been developed. In addition, the mathematical foundation of Otsu's approach represents the image's histogram as a fusion of two normal distributions with the same variance and size. Similar to how statistical tests—which Otsu's technique is closely connected to—may still work well even if the underlying assumptions aren't totally accurate, Otsu's thresholding might still yield adequate findings even if these requirements aren't satisfied.

K-Means

It is used to keep the interesting subject and the backdrop apart algorithm without supervision. To distinguish the region of interest from the background, K-means clustering is performed. The iterative K-means method seeks to partition the dataset into K distinct, non-overlapping subgroups (clusters), each of which comprises just one group to which each data point belongs. The objective is to maximize the dissimilarity between clusters while ensuring the comparability of data points within each cluster. To minimize the sum of squared distances among the data points as well as the cluster centroid, that is the mean of all the data points within the cluster, the algorithm assigns the data points to the clusters in an ordered way. As the level of diversity within the cluster decreases, the uniformity (similarity) of the information points within that cluster increases. Each data point is assigned to a cluster according to its distance from a centroid in the centroid-based clustering technique K-means. The goal is to determine the Kth group in the dataset. The number of clusters (k) must be known in advance. It cannot handle erratic data or outliers. It is inappropriate to identify clusters with non-convex forms.

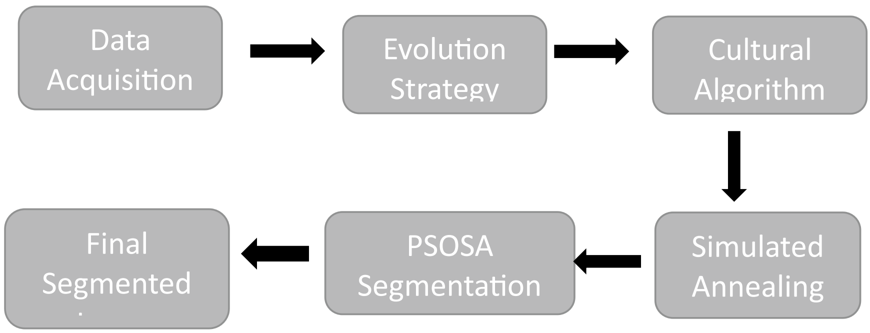
Watershed

Finding the precise bounds using a certain backdrop and the watershed method. It initially takes certain background and foreground information in order to run the watershed and determine the precise limits. Picture morphology is used in the region-based technique known as "watershed segmentation." At least one marker, often referred to as a "seed" point, must be present inside of each item in the picture, including the backdrop. The markers are either chosen manually or automatically while taking into account the application-specific information of the objects. Due to the high cost of development and the necessity for a lot of rainfall, a watershed's disadvantages include its large size and cost. Therefore, there aren't many instances locations where a watershed might be built. Watersheds need regular maintenance.

Genetic

Image analysis is used to solve difficult optimization problems. A genetic algorithm is a method for tackling both confined and unconstrained optimization problems that is based on natural selection, the mechanism that drives biological evolution. The mechanism through which qualities are passed on from parents to offspring is the subject of genetics research. The foundation for heredity is called as inheritance. The process through which qualities are handed down from one generation to the next is known to as this. Genetic testing might provide a limited amount of information about an inherited illness. A person's predisposition to acquire a problem, the severity of their symptoms, or whether the disorder will get worse over time are typically not predicted by the test.

3.2 Proposed Method



The suggested image segmentation technique consists of four biologically inspired image processing algorithms. In order to perform the histogram equalisation step, the system first grants the Evolution Strategy access to the raw image data. For the evolution strategy method to improve brightness, the objective histogram must fit the input image. The cultural algorithm uses the brightness image from the histogram equalisation in the subsequent phase of image size reduction. This step is essential for the rest of the procedure since it quantises the brightness image into a pre-set number of threshold levels. The quantitating image generated by the cultural algorithm is given to simulated annealing, which is utilized to precisely separate each component and detect edges. Simulated Annealing is coupled with Particle Swarm Optimization as a secondary method to divide the image into its constituent parts and account for the optimizer. The final results display segmentation images that are more precise and clearly defined.

4 Results and Analysis

In this paper, we mostly concentrated on the proposed PSOSA segmentation and Otsu's methods, which are widely used in many applications.

A PSOSA segmentation applies all evolutionary techniques, resulting in well-separated images. Comparison with traditional segmentation methods by Providian embryonic evolutionary segmented images and using them to improve image segmentation. It demonstrates each step of the suggested segmentation on an embryo sample. Also displays the outcomes of the five method comparisons on the Archaeopteryx test image. A comparison of three segmentation performance measures on four sample photos with various techniques (Figs. 1, 2, 3, 4, 5 and 6).

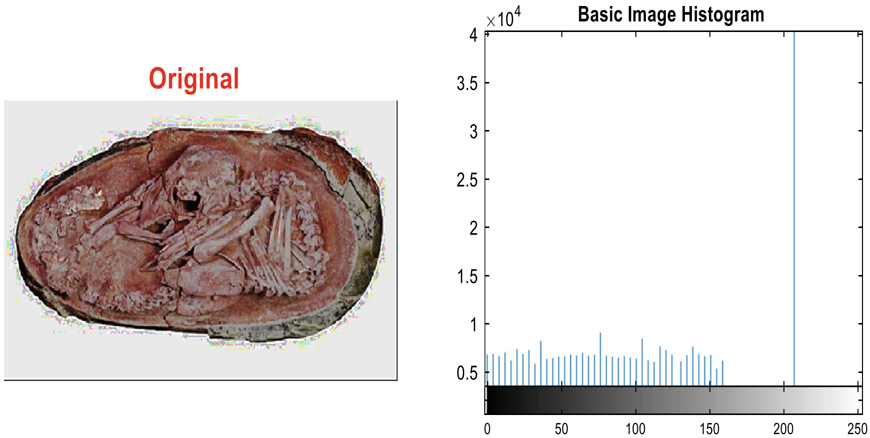


Fig. 1 Original image histogram

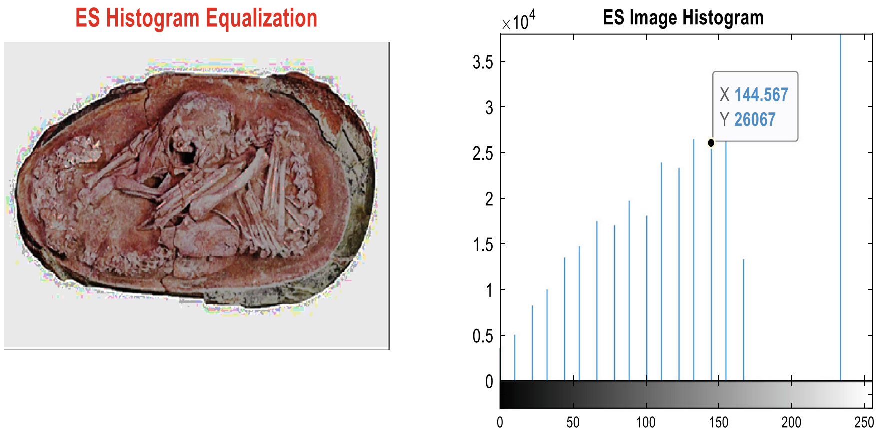


Fig. 2 Applying the evolution strategy equalization of histograms technique on the "Dilophosaurus" example image

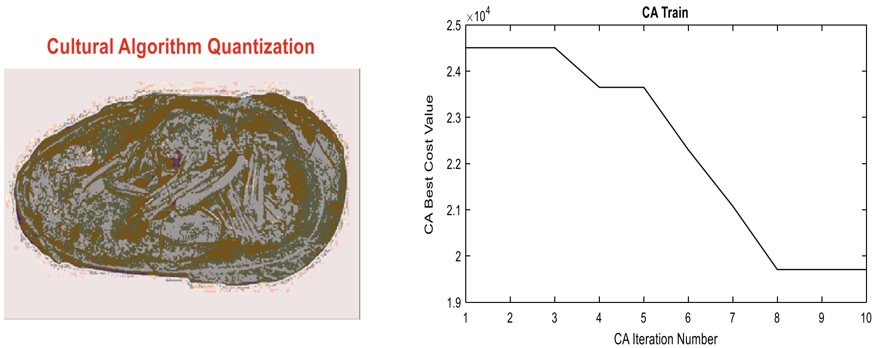


Fig. 3 Applying a cultural algorithm to quantize the image of the "Dilophosaurus" sample

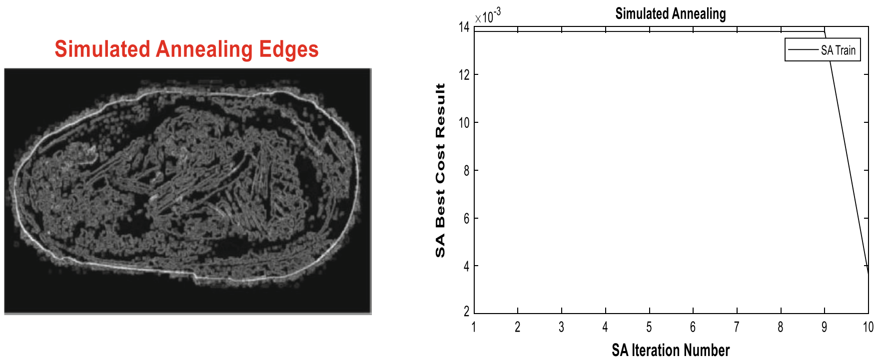
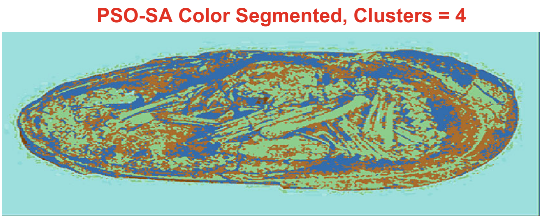


Fig. 4 Performing edge identification using the simulated annealing algorithm on the "Dilophosaurus" example image



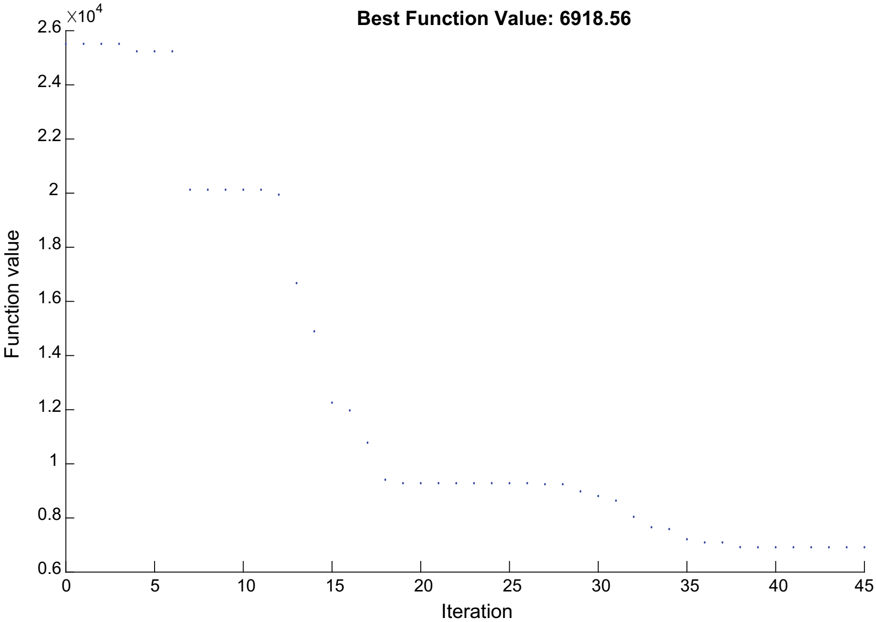


Fig. 5 Performing PSOSA image segmentation on the "Dilophosaurus" example image

Fig. 6 Proposed segmentation on "Embryo" image

5 Conclusion

The recommended technique can handle a variety of scenarios involving fossil images because it employs numerous bio-inspired image processing algorithms to create a final segmented image. The method produces a precise segmented image for palaeontology study; the only drawback is longer calculation times when compared to earlier methods. In general, an algorithm's parameters can be changed to make it run more quickly, or if the model is already available, it can be used in real time. However, traditional segmentation techniques could result in the desired outcomes for.

Despite the bulk of investigations, it makes sense for precise researchers to push some boundaries. Future research will entail utilizing this approach in fields other palaeontology. In order to obtain distinct or even better results, it is sometimes advocated to combine numerous bio-inspired algorithms.

References

1. Jusman Y et al (2021) Comparison of malaria parasite image segmentation algorithm using thresholding and watershed method. J Phys: Conf Ser. 1783:1
2. Mousavi SMH (2022) PSO-SA image segmentation. https://github.com/SeyedMuhammadHosseinMousavi/PSO-SA[1]Image-Segmentation. GitHub
3. Mousavi SM, Hossein VL, Prasath S (2019) Analysis of a robust edge detection system in different colour spaces using colour and depth images. Компьютерная оптика 43(4):632–646
4. Karimullah S, Lalitha M, Mounika, Mahesh Reddy S, Mahammad Waseem S, Irfan SM (2023) Performance Enhancement of wireless systems using hybrid RIS technique. In: Proceedings of the 4th international conference on data science, machine learning and applications. ICDSMLA 2022. Lecture notes in electrical engineering, vol 1038. Springer, Singapore. https://doi.org/10.1007/978-981-99-2058-7_17
5. Setiawan AW (2020) Image segmentation metrics in skin lesions: accuracy, sensitivity, specificity, dice coefficient, jaccard index, and matthews correlation coefficient. In: 2020 International conference on computer engineering, networking, and intelligent multimedia (CENIM). IEEE
6. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. Springerplus 5:507. https://doi.org/10.1186/s40064-016-2152-2
7. Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR In MIMO-OFDM using adaptive SLM and PTS technique. Int J Pure Appl Math 118(17):355–373. ISSN: 1311-8080 (printed version). ISSN: 1314-3395
8. Anu Priya D, Vishnuvardhan D, Mamatha G, Shaik F, Karimullah S (2023) VIP development of SPI controller for open-power processor-based fabless SOC. In: Kumar A, Ghinea G, Merugu S (eds) Proceedings of the 2nd international conference on cognitive and intelligent computing. ICCIC 2022. Cognitive science and technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_2
9. Oskouei AG, Hashemzadeh M (2022) CGFFCM: a color image segmentation method based on cluster-weight and feature-weight learning, pp 100228
10. Prasad PS, Pathak R, Gunjan VK, Ramana Rao HV (2020) Deep learning based representation for face recognition. In: Kumar A, Mozar S (eds) ICCCE 2019. Lecture notes in electrical engineering, vol 570. Springer, Singapore. https://doi.org/10.1007/978-981-13-8715-9_50

An Effective Clustering-Based Color Image Segmentation via Substantial Extracted Features

S. Fahimuddin, Shaik Karimullah, Mudassir Khan, and Vinit Kumar Gunjan

Abstract Image segmentation is the division of a picture into various categories. One aspect of an image that adds information is color. As computer technology has advanced, the importance of image processing techniques has increased in a wider range of applications. In the field of image processing, picture segmentation is a well-known subject as well as a hotspot and the subject of several methodologies. Several all-purpose algorithms and methods have been developed for image segmentation. Cluster-based methods capable of utilizing the key attributes gathered became essential when color image segmentation started to develop. Consequently, CGFFCM means cluster-weight and group-local feature-weight learning in Fuzzy C-Means is the cluster-based pigment image segmentation method reported in this paper. Automatic cluster weighting is applied in CGFFCM to reduce sensitivity to clustering initiation, while group-local feature weighting is used to enhance picture segmentation. Additionally, it uses the efficient of image characteristics which consists of few features from the three separate categories to increase the quality of picture segmentation. The technical language of MATLAB is to be used for the CGFFCM implementation.

Keywords Cluster-weighting · Feature-weighting · Clustering · Image segmentation

S. Fahimuddin · S. Karimullah
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India

M. Khan (✉)
Department of Computer Science at College of Science &Arts Tanumah, King Khalid University Abha, Abha, Saudi Arabia
e-mail: mudassirkhan12@gmail.com

V. K. Gunjan
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_9

1 Introduction

The image segmentation aims to divide the image into homogeneous and related sections. As compared to individual pixels, the regions in the segmented image have numerous unique characteristics, such as shape, texture, etc., [1]. Nearly all image segmentation methods that were previously established for grayscale images have also been used for segmenting color images. The processing of an objects' color image is frequently necessary for partitioning it from the background.

When a special angular metric is used, a segmentation method can be prone to shadows in the image. During the post-processing stage, the over-segmentation brought on by each of the two segmentation algorithms is removed, the highlights are taken out, and regions with related colors are combined. The issue of how the selected color space affects the segmentation outcomes is still up for debate. Since there is no one perfect color space, it is important to select one for color image segmentation that produces the best results for the purpose or image class [2].

Automatic rebound from failure is a clustered solution's key benefit. This method is typically less expensive and the sampling process of the approach uses fewer resources. Basically, the segmentation is used to reduce the noise in the image and to improve the quality of an image. So, taking example, this generation everyone is bound in capturing everything as an image. To make the image more clear and make it look as a good quality image, clustering process can be used [3].

2 Literature Review

Authors in [4] have explained about the feature weighting, which occurs during the clustering process locally and autonomously.

We also use cluster weighting to address the initialization difficulty. At each restart, the clusters' weights are determined by adding the intra-cluster distances of each member and the member's current feature weights. This method of cluster-weight calculation inhibits the formation of clusters with a high total of intra-cluster-weighted-feature distances (SIWD1). It hence permits the formation of high-quality clusters independent of the original centers. The main uses for picture segmentation are in data mining and other related fields like machine learning and machine vision. It contains a number of criteria that both lower and boost intra-cluster similarity. Somehow, these algorithms are able to reduce how sensitive the algorithms are to initialization. The primary benefits include a straightforward structure, simple implementation, and quick convergence. One of the drawbacks is that in some applications, they perform noticeably worse. The cost of computation is really high in general.

The S3MKIFC approach, a semi-supervised surrogate-assisted multiobjective kernel intuitionistic fuzzy clustering method, is recommended by the researchers in [5] to increase segmentation performance and processing time on colored images, image grouping in nonlinear systems, and suitable for supplying electricity to infrared image segmentation technology based on space distribution data. Semi-supervised kernel intuitionistic fuzzy function models are produced to help the Kringing model in forecasting the values of objectives functions rather than directly completing the time-consuming, costly object function computations. This approach saves time and yields good segmentation results on the current picture without the need of a background filter. The constraints include difficulty in obtaining reliable findings and the inability to handle data with complicated shapes.

In [6], the authors proposed a method where feature extraction stage comes first in the segmentation procedure. This CGFFCM clustering algorithm receives the retrieved characteristics. The ICA is used in conjunction with the clustering process to accurately determine the parameters needed by the clustering method. The weights of feature groups are automatically and appropriately determined using the ICA to enhance the CGFFCM. The main applications include image processing and transformation and pattern recognition. This process includes an efficient combination of the image feature to increase the segmentation quality. The advantage of clustering algorithms is providing reasonably good segments in a small amount of time. This method is more realistic. The initialization of primary clusters is highly important to this method's performance gives each image feature equal weight during a segmentation process, which results in inconsistent performance on the various image.

One of the most common visual features in image retrieval systems is the color feature were discussed in the paper [7]. Color can be extracted with ease and is independent of the size and orientation of the image. A user specifies the color of regions in a query image using a color-based image retrieval approach, and images that partially or completely better fit the user-specified query are then retrieved [8]. Color feature extraction methods can be divided into two types: color histograms and statistical methods [9]. The simple and most commonly used category to extract the color feature vector is color histograms feature extraction category. It can be thought of as a mass function of the image intensities and indicates the frequency of occurrences of each color in an image [10].

3 Methodology

3.1 Existing Method

k-Means Clustering Algorithm:

The k-means clustering technique is the one of the most well-liked and effective clustering methods. One of the simplest unassisted learning techniques used for clustering the image is k-means algorithm, where k represents the predetermined number of clusters. k-means clustering technique is explained in the following steps:

Step 1: Assume that there are k clusters.

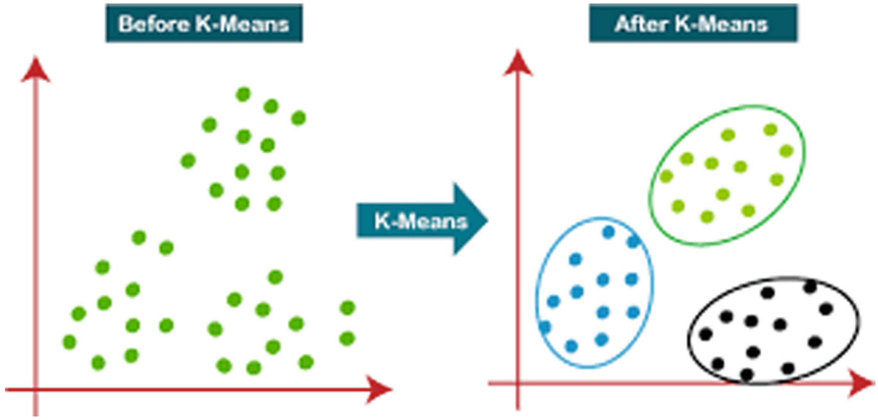


Fig. 1 K-means diagram

Step 2: Randomly initialize the clusters centers with values that fall within the dataset's range.

Step 3: Connect every sample into the dataset to one of the k clusters using the Euclidean distance (this depends on the shortest distance between the samples and the clusters centers).

Step 4: By averaging all the samples within each cluster, find the new centers of the k clusters.

Step 5: Stop; otherwise, proceed to Step 3 if none of the centers have moved.

The limitations of k-means clustering algorithm are, these are easy to put into practice, large data sets are scaled. As demonstrated in Fig. 1, this approach assures convergence and allows the centroids to warm up. This method easily adapts to new cases and expands to incorporate clusters of varying sizes and forms, including elliptical clusters.

3.2 Proposed Method

This proposed method as shown in Fig. 2 provides a novel approach that combines color, texture, and shape data to construct the feature vector in order to achieve high accuracy.

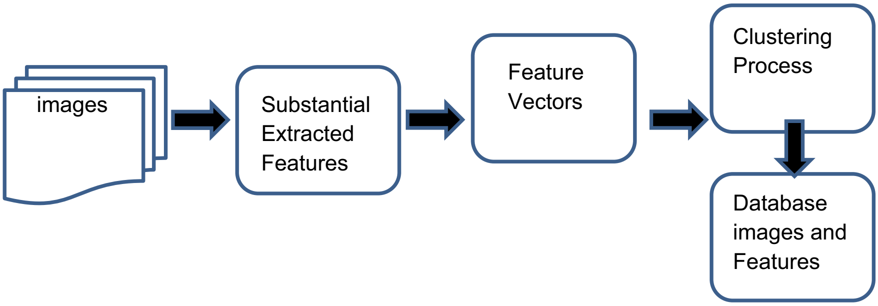


Fig. 2 Proposed system architecture

3.3 The steps for the proposed method are as follows:

Step 1: Repeat steps 2, 3, 4, and 5 for each image in the database.

Step 2: Using a color histogram, extract the color visual feature. To acquire the following color feature vector for each band, first apply color histogram to Hue, Saturation, and the Value.

Step 3: Using the Gabor filter, extract the texture's visual feature. Determine standard deviation (sd) and the mean (mn) of the magnitude of the already transformed coefficients that will be furtherly used to indicate the homogenous textural feature of the region. Each of the main three-color components Gabor measurements is calculated (RGB).

Step 4: Using this Fourier descriptors with brightness, extract the shape visual feature. The three-color components shape feature vector can be expressed, where bi stands for the Fourier converted coefficients, N for the number of border pixels, and I for the color band.

Step 5: Let us denote (RGB) F = {FTexture, FShape} as the feature vector for each color band.

Step 6: Execute the subsequent k-means Clustering algorithm steps.

(a) Enter the feature vector for every database image.
(b) Initialize the k clusters' centers at random with feature vectors from the images stored in the database.

4 Results and Analysis

In this research, we mainly focus on improved CGFFCM, which is widely used in various applications, for clustering-based color picture segmentation.

The enhanced CGFFCM utilizes an automated cluster-weighting technique to reduce sensitivity to cluster initialization, including a group-local feature distribution technique to improve picture segmentation. In order to improve the quality of the image, it is necessary to employ a proficient amalgamation of the eight characteristics derived from the three distinct categories. The three elements are local homogeneity, CIELAB color, and texture.

We run this software for the images which is available in a dataset. The difference between original image, ground truth image, and the segmented image using this CGFFCM technique is also mentioned in Figs. 3, 4 and 5 respectively.

Accuracy Score: The accuracy score is used to estimate the model's performance by calculating the ratio of total true positives to total true negatives across all made predictions. F-Measure Score: An estimate of a model's precision on a given dataset, whereas NMI Score is the Mutual Information (MI) score is normalized to produce the Normalized Mutual Information (NMI), which scales the outcomes from 0 (no mutual information) to 1 (perfect correlation).

The comparative analysis for proposed and existing methods is given in Table 1 and its corresponding plot is shown in Fig. 6. From the tabular column and graph, we can observe that the Accuracy Score, F-Measure Score, NMI score are high for the proposed method when compared to the existing methods. From this, we can say

Fig. 3 Original image

Fig. 4 Ground truth image

Fig. 5 The segmented image

that with the help of proposed method, we can get a good quality image with less noise.

Table 1 Comparative analysis

Methods	Accuracy score	F-measure score	NMI score
1. A novel fuzzy C-means clustering approach is proposed, which incorporates feature-weight and cluster-weight learning.	0.9764	0.9775	0.82421
2. Semisupervised surrogate segmentation based multiobjective kernel intuitionistic fuzzy clustering algorithm is proposed for color image analysis.	0.9655	0.9675	0.81540
3. Weighted clustering and localized grouping color picture segmentation is improved with the incorporation of feature-weight learning into the Fuzzy C-Means clustering technique.	0.9510	0.9440	0.81150
4. A method for segmenting color images using cluster-weight with feature-weight learning.	0.9650	0.9644	0.82110
Proposed method: An effective clustering-based color image segmentation via substantial extracted features	0.993206	0.996385	0.845040

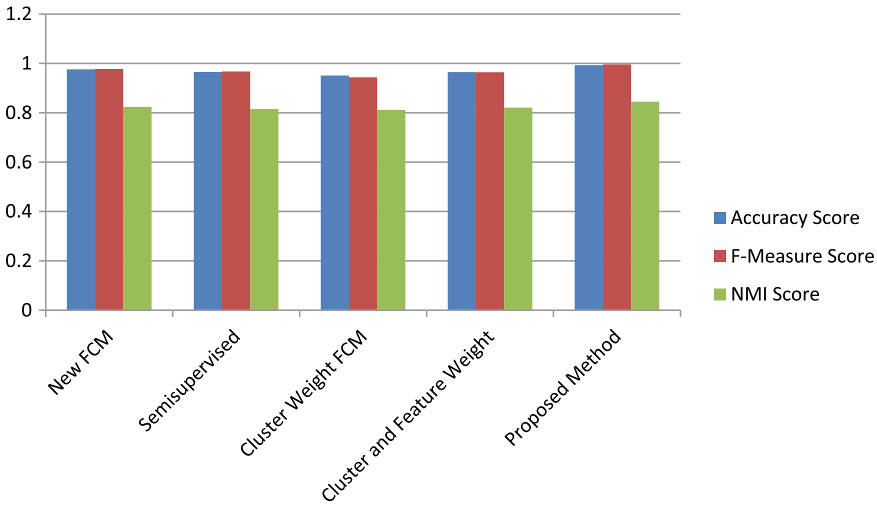


Fig. 6 Comparative analysis

5 Conclusion

Finally, it must be observed that there is probably no one best method for segmenting color images. The segmentation process—of which this is only the first step—clearly depends on the art of problem-solving. More complex segmentation approaches than those described in this study may be applied due to the enormous growth in processing capacity. Particularly crucial for the future advancement of color image processing are hybrid approaches, which use color, texture, and shape.

The experimental data presented in this study supports the claim that the CGFFCM outperforms state of the art which approaches for color image segmentation task. This source code included in this work has been specifically created to address the high-level difficulty in the clustering task and it is very efficient, flexible, simple to comprehend, and modular. It is beneficial for machine learning experts and also helpful for beginning and advanced researchers to experiment with new and best ideas. We can also adjust the method to handle various cluster-based jobs simply by making small modifications to the source code (such modifying the distance function).

On the subsequent stages of image processing, such as object detection and tracking, retrieval in image databases, etc., the quality of segmentation output has a significant effect. The authors intend to develop some modules, including a function for defining multiple distance measurements, in order to further enhance CGFFCM in the future.

References

1. M. Hashemzadeh, A. GolzariOskouei, N. Farajzadeh (2019) New fuzzy C-means clustering method based on feature-weight and cluster-weight learning. Appl Soft Comput 78:324–345. https://doi.org/10.1016/j.asoc.2019.02.038
2. Zhao F, Zeng Z, Liu H, Lan R, Fan J (2020) Semisupervised approach to surrogate assisted multiobjective kernel intuitionistic fuzzy clustering algorithm for color image segmentation. IEEE Trans Fuzzy Syst 28(6):1023–1034.https://doi.org/10.1109/TFUZZ.2020.2973121
3. Oskouei AG, Balafar MA, Motamed C (2021) FKMAWCW: categorical fuzzy k-modes clustering with automated attribute-weight and cluster-weight learning. Chaos Solitons Fractals 153:111494. https://doi.org/10.1016/j.chaos.2021.111494
4. Karimullah S, Lalitha M, Mounika S, Mahesh Reddy S, Mahammad Waseem S, Irfan SM (2023) Performance enhancement of wireless systems using hybrid RIS technique. In: Proceedings of the 4th international conference on data science, machine learning and applications ICDSMLA 2022. Lecture notes in electrical engineering, vol 1038. Springer, Singapore. https://doi.org/10.1007/978-981-99-2058-7_17
5. Mousavi SM, Hossein VL, Prasath S (2019) Analysis of a robust edge detection system in different colour spaces using colour and depth images. Компьютерная оптика 43(4):632–646
6. Karimullah S, Vishnuvardhan D (2020) Iterative analysis of optimization algotithms for placement and routing in asic design. In: Kumar A, Paprzycki M, Gunjan V (eds) ICDSMLA 2019. Lecture notes in electrical engineering, vol 601. Springer, Singapore. https://doi.org/10.1007/978-981-15-1420-3_199
7. Setiawan AW (2020) Image segmentation metrics in skin lesions: accuracy, sensitivity, specificity, dice coefficient, Jaccard index, and Matthews correlation coefficient. In: 2020 International conference on computer engineering, networking, and intelligent multimedia (CENIM). IEEE, 2020
8. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. SpringerPlus 5:507. https://doi.org/10.1186/s40064-016-2152-2
9. Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR In MIMO-OFDM using adaptive SLM And PTS technique. Int J Pure and Appl Math 118(17):355–373. ISSN: 1311-8080 (printed version); ISSN: 1314-3395
10. Prasad PS, Pathak R, Gunjan VK, Ramana Rao HV (2020) Deep learning based representation for face recognition. In: Kumar A, Mozar S (eds) ICCCE 2019. Lecture notes in electrical engineering, vol 570. Springer, Singapore. https://doi.org/10.1007/978-981-13-8715-9_50

Colour Image Segmentation Using Random Walker Approach

S. Fahimuddin, Shaik Karimullah, Vinit Kumar Gunjan, and D. Vishnuvardhan

Abstract Segmentation process is a necessary step more towards advanced image processing tasks such as image evaluation, pattern classification, and laptop vision. The methods used for segmentation are of great importance in many application areas of the image acquisition, because the outcomes of further stages of the image acquisition rely on the region-based quality. The literature has been developing several image segmentation techniques since the beginning. The segmentation algorithms were created largely for greyscale images since analysing a colour image takes much longer than processing a greyscale image. An eye can distinguish thousands of different colours and concentrations, but there are only 24 variations of grey, as has long been known. Despite the fact that most methods for colour image processing are derived on monochrome picture classification, colour image processing has grown more and more popular. Each cluster depicts a collection of pixels with comparable colour characteristics. There is no one colour mansion that will yield appropriate results with all sorts of photos because segmentation outcomes depend on the colour mansion used. Utilizing adequate Euclidean distance highlighted in a complementary colour scheme, the weights would've been determined. For quasi multi-label segmentation, we provide a novel local/global approach. The planned method does not include a thresholding post-process process and uses a completely distinct non-linear means of two terms. For the component of the stochastic process, the colour distances influence the evolution of the graph and the similarity among set of labelled pixels over time. Due to its connection to the distinctive Dirichlet shortcoming, we frequently think of stochastic aids organizations as a Laplacian-based multidimensional approach

S. Fahimuddin (✉) · S. Karimullah
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: fahimaits@gmail.com

V. K. Gunjan
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, India

D. Vishnuvardhan
Department of Electronics and Communication Engineering, JNTUACE, Ananthapuramu, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_10

with an extension to more broad information and some scientific validity. For semi-automatic segmentation, we tend to plan a slight enhancement to the random walker method. This is typically accomplished by adopting a novel concept of similarity and separation between components that makes use of a particular colour scheme and incorporates nearby pixels.

Keywords Segmentation · Random walk · Colour spacing · Pixels · Neighbours

1 Introduction

Image segmentation divides a picture into quasi groups of pixels. The sets that follow is described as supported visual features retrieved by native options and are refer to as sections (or segments or objects). The pixels must share some of these characteristics of uniformity or resemblance. Segmentation algorithms are categorized according to the following categories in the standard classification: Using algorithms that are pixel-based, edge-based, or region-based, segmentation can be carried out. While edge-based algorithms depend on the recognition of the boundaries existing in the image, pixel centred algorithms use specific pixel values as the sole data required to carry out segmentation [1]. Creating non-overlapping groups of pixels from an image is called segmenting it. The sets that follow are described as supported visual features retrieved by native options and are referred to as sections (or segments or objects). The first stage in high-level image processing activities like image processing, pattern matching, and laptop vision is picture segmentation. The results in succeeding steps of the picture acquisition based on the quality of the segmentation and the methods used for image segmentation are given high importance in many applications of the colour imaging process (the seeing and trailing, the retrieval in image databases, etc.). For instance, there is a great deal of scientific value in information blowouts from a tiny low set of seen instances towards the full image when using the Susceptible-Infectious-Recovered (SIR) concept to the task of image segmentation. The segmentation techniques were initially primarily intended for photos with a grey level [2]. This is because, despite the fact that colour information enables more extensive depiction of images and more precise segmentations, processing colour images takes machine times that are noticeably greater than those required for processing grey-level images. The strategies for segmenting monochrome images are backed by a number of principles, including edge detection, regional growth, and bar graph thresholding. These concepts are used in various colour image segmentation techniques and various colour models (such as RGB, La b, and HSV). These characteristics generally embed the regions' perceived complexity and their distinctive relationship. Similarly, so reducing the gap between both the computational division and also the one anticipated by the user [3]. One of the most common presumptions in the methodology for segmenting photographs is that homogeneous colours in an image relate to different clusters and, hence, important items within the image. To put it in other words, each cluster represents a group of pixels with complementary colour characteristics. No one colour area can achieve results which are appropriate with every type of image because the segmentation results depend on the selected colour region. As a result, various authors searched towards "the colour" but also "the colour" area which would address their particular colour image segmentation issue. The similarity of images in terms of colour and texture possibilities is communicated by placing more weight on just one set of vertices if we ordinarily think of design as just a network with observable pixels as its vertices. The weights will be determined using reasonable distance functions that are indicated with the appropriate colour region. Numerous research using this directed graph rephrased the segmentation problem to use the histogram-based goal operation. When the Merriman-Bence-Osher (MBO) methodology, also known as the unchanging application of warmness diffusion and thresholding, is discussed for binary image labelling requirements, we frequently demonstrate that the majority of those procedures call for other forms of energy, requiring the application of complex and time-consuming optimization techniques [4]

- The presence of a similarity metric or distance between pixels in a particular neighbourhood and pels using a particular colour, which supports both the accidental walker method and a simple bunch step.
- A combination of higher-than approaches that addresses the drawbacks those methods have when used alone.
- A transformed energy associated with random walker strategy that enhances the standard of the segmentation and recognizes solely the reduction of a non-linear operation.
- A strategy based on machine learning that modifies the Euclidian distance by changing the weights of colour separation as a pre-processing step for the images.

Other authors have thought about the potential for a changed energy associated with a random walker by applying some reasonably appropriate widths or some post-processing steps of the probability map generated with stochastic strategic process. We tend to imply that our methodology does not involve a thresholding post-process stage and instead uses a totally different quasi approach to two terms. The random walker separation algorithm is the new approach. Additionally, we'll talk about a new definition for quasi distance between pixels. It also focuses on numerical experiments. Here, we focus on the issue of discovering appropriate weights for such innovative colour distance that is mentioned. The case study's conclusions are given as a result [5].

Segmentation: The process of dividing a picture is thought of as segmentation. In segmentation the process is divided into either some part components or some objects. Whereas the process of the final image the independent is the most stringent within the segmentation. The time taken is additional within the segmentation.

For a correct technique, a strong segmentation takes a protracted time to prosperous resolution image of individual whole thing identification needs within the

Fig. 1 Example of image segmentation

complication. By segmentation, the image reduces complexity and analysis becomes easier. Once the segmentation processes is complete, it split and clusters the set of pixels from the image. Segmentation relies on separation and similarity. Edge detection happens by finding pixel boundary regions (Fig. 1).

2 Literature Review

In Laplacian coordinates for Seeded Image Segmentation [6], image segmentation with seeds helps to ensure the formation of a distinct output while being simple to utilize and conceptually understand. Furthermore, the formulation has an anisotropic action, which needs to keep pixels with comparable properties close together unintentionally enforcing huge jumps on the corners of picture areas to improve wearable on object borders.

Single Channel Image Enhancement (SCIE) of White Blood Cells Based on Virtual Hexagonal Filter (VHF) Designed over Square Trellis [7] WBC image quality and contrast enhancement were accomplished using only a virtual hexagon-shaped trellis (VHT)-based noise removal method. The virtual hexagonal filtration (VHF), with a dimension of three and a hexagonal framework, is constructed in this manner by incorporating features of true and square grid pixel resolution interpolation. This filtration is used to improve and adjust distinction through WBC ALL-IBD images.

In [8], improving user and digital tradition model interactive elements and information sharing by boosting computer image and vision computation inclusion with machine vision info from unmanned aerial systems was discussed. The abundance of images obtained through aerial image acquisition will also enable the use of Heritage Building Information Modelling (HBIM) structure in an ad hoc augmented reality (XR) project, allowing different types of people (experts, non-expert users, virtual tourists, and educators) and devices (mobiles, tablet devices, computers) to acquire details and data that are not visible from the ground.

An efficient position estimation of 3D fluorescent spherical beads in confocal microscopy is via Poisson denoising [9]. It first denoises each screen of the scanned quantity before estimating the centre and profile of the components' 2D crossings with the frames with a sets of data functional and a normalized balanced Least Squares fit. After that, geometrical properties are used to extract 3D coordinates from 2D data. The outcomes of the tests show that noise removal has a massive effect on physically based procedures, which depend heavily on the quality of an initial acquisition [10].

3 Methodology

3.1 Existed Method

K-Means Algorithm:

K-means clustering, an unsupervised learning approach, is used to tackle the clustering problems in computer vision or data science. We will learn about the K-means clustering approach in this chapter, including what it is, how it works, and how to use it in MATLAB. The K-means clustering technique for unsupervised learning separates the unlabelled dataset into different clusters. Here, K determines how many pre-defined clusters need to be created as part of the process; for instance, if $K = 2$, there will be two clusters, if $K = 3$, there will be three clusters, and so on. It gives us a simple way to classify the groupings in the unlabelled dataset our own that does not necessitate any training. We can also group the data into various clusters thanks to it. Since the algorithm is based on centroid, every cluster has a coordinate assigned to it. The main objective of this approach is to minimize the sum of the distances between each statistic and its associated clusters. The algorithm starts with an unlabelled dataset and separates it into K clusters. It then continues the procedure until it is unable to discover the optimal clusters. In this algorithm, K should have a known value.

The K-means clustering algorithm has two main purposes:

- Select the ideal value for K centre points or centroids via an iterative process.
- The nearest K centre is given to each data point. The pieces of data that are near a certain K centre group together to form a cluster.

As a result, each cluster stands out from the others and has some shared data points. Figure 2 explains the K-means clustering algorithm:

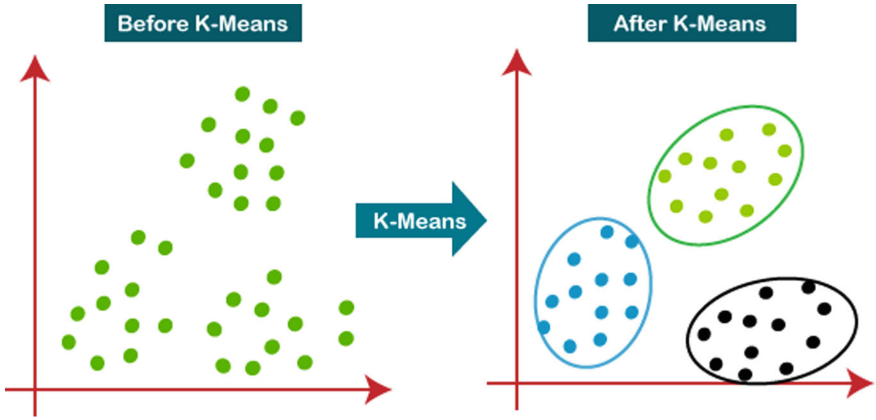


Fig. 2 K-means algorithm

3.2 Proposed Method

Random Walker (RaWa) Algorithm:

An undirected graph $G = (V, E)$ serves as the foundation of the RW, where V & E stand for sets of vertex & edges, respectively. The sets $V =$ Vm and U Vu can be created from the vertex set. The remaining pixels in the image are referred to as the "unlabelled vertices," or set Vu. One group is referred as the "marked vertices" (Vm) that are seeds that the user has designated as being a part of different objects. The pixels that are adjacent to one another in the image, such as the conventional 4- or 8-neighbourhood, make up the set of edges E. The edge connecting the vertexes vi and vj is represented by the symbol eij. Based on the difference in intensities between the two pixels, the function (vi, vj) can be used to indicate the weight of an edge.

The weight of the edges is between 0 and 1, with a weight near to 1 is pixels which is similar and 0 denotes dissimilar pixel. With aforementioned graph structure in mind, the RW method's basic concept is as follows. Each marked vertex of the VM is considered to correspond to one of the K potential regions (objects) that make up the image. If we take into consideration an edge with weight eij and endpoints vi and vj, the edges weight (vi, vj) (0, 1) could be seen as a measurement of the likelihood that random walk will pass from the one vertex to the next. If the vertexes are very close in colour, the random walk may move from vi to vj depending on the edge's weight, but if they are quite different, it is doubtful that it will do so. Given the aforementioned probabilities, the segmentation algorithm calculates the likelihood that a random walker departing from a given pixel will reach the kth item. This probability is denoted by the symbol xi^k. Following that, the image segmentation is carried out based on these probabilities. To be more precise, we designate any vertex vi as being a part of the kth region if xi$^k >$ xi^{-k} for any $\bar{\ }k = / = k$. Because they include a quick shift in colour or intensity, we see that the image's edges correlate to low transition probabilities. As a result, when conducting picture segmentation, this method tends to respect image edges. It was demonstrated in that by using the graph Laplacian matrix to solve a linear sparse system of equations, these probabilities may be computed analytically.

We look at how the vertices are categorized as two sets: where Vm, is user-marked vertex, and Vu, is unmarked vertex, with this results that the Vm U Vu $= V$ and Vm n Vu $=\varnothing$. Keep in mind that, regardless of their titles, all points mentioned in Vm are included.

4 Results and Analysis

The predicted image is provided in Fig. 3 and corresponding log and outcomes are provided in Figs. 4 and 5, respectively. The experimental outcomes demonstrate the superiority of the suggested strategy more traditional approaches like the K-means method and the random walk process is known also on measurement of distance in pixels. Moreover, RaWaCs did perform well in all colour formats, even if a particular colourspace (e.g. physiological pictures) may be a wiser alternative than others. Furthermore, the proposed procedure requires both classical RW and more expresses, like NRW and NLRW: RaWaCs is designed specifically for use with many labels, instead of just foreground and background extraction, though it excels at both.

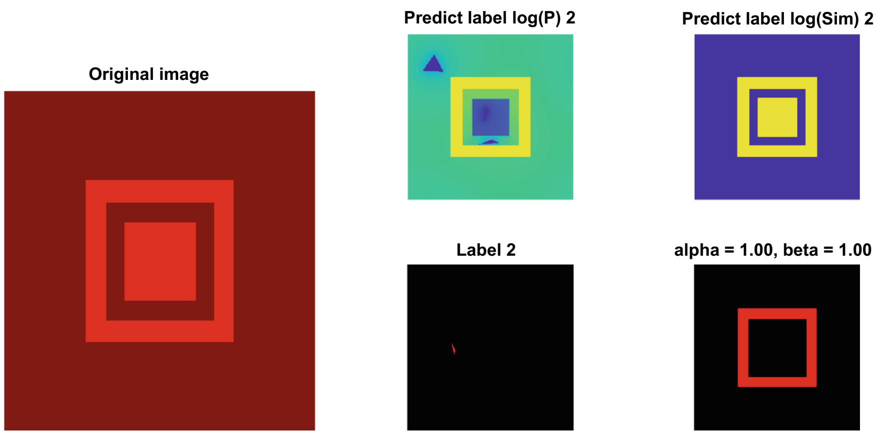


Fig. 3 Predicted image

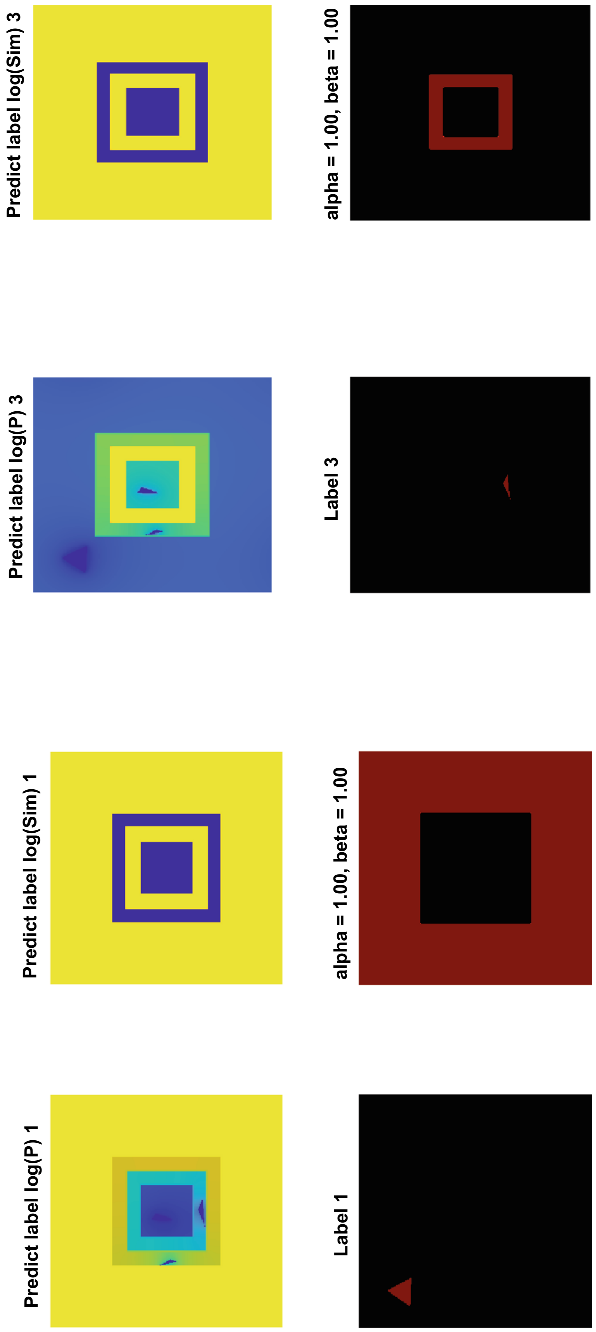


Fig. 4 Predicted image based on log values

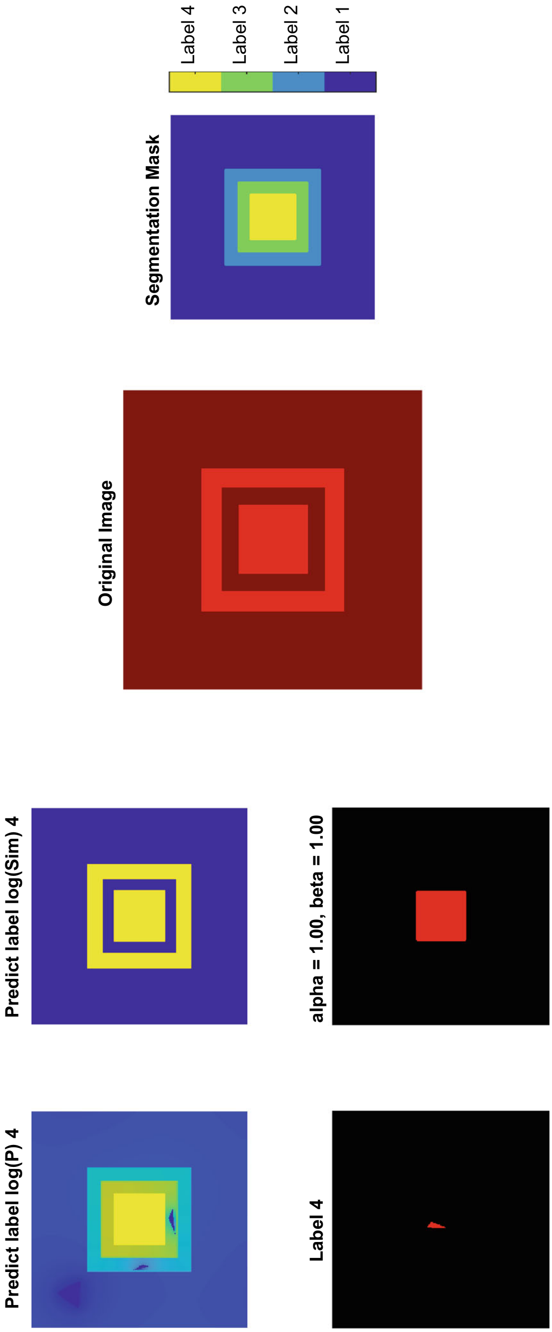


Fig. 5 Predicted image and its corresponding outcome

5 Conclusion

In this study, we suggested a quasi-segmentation improvement to the random walker method. This is done by using a separate concept of likeness and proximity among pixels in different colour spaces and between active pixels in a region. Then, in a manner similar to the traditional clustering procedure, a modified energy associated towards a random walker which incorporated the RW possibilities and a worldwide scale was taken into consideration. The costliest process involves estimating the distinct distance among pixels. In addition, we observe that these tasks can be carried out effectively in comparison, for instance, with a real-world deployment that uses graphics processing units (GPUs).

References

1. Casaca W, Gois JP, Batagelo HC, Taubin G, Nonato LG (2021) Laplacian coordinates: theory and methods for seeded image segmentation. IEEE Trans Pattern Anal Mach Intel 43:2665–2681. https://doi.org/10.1109/TPAMI.2020.2974475
2. Rasheed S, Raza M, Sharif M, Kadry S, Alharbi A (2022) Single channel image enhancement (SCIE) of white blood cells based on virtual hexagonal filter (VHF) designed over square trellis. J Pers Med 12:1232. https://doi.org/10.3390/jpm12081232
3. Banfi F, Mandelli A (2021) Computer vision meets image processing and UAS photogrammetric data integration: from HBIM to the extended reality project of Arco Della Pace in Milan and its decorative complexity. J Imaging 7:118. https://doi.org/10.3390/jimaging7070118
4. Benfenati A, Bonacci F, Bourouina T, Talbot H (2021) Efficient position estimation of 3D fluorescent spherical beads in confocal microscopy via poisson denoising. J Math Imaging Vis 63:56–72. https://doi.org/10.1007/s10851-020-00994-1
5. Karimullah S, Vishnuvardhan D (2020) Iterative analysis of optimization algotithms for placement and routing in asic design. In: Kumar A, Paprzycki M, Gunjan V (eds) ICDSMLA 2019. Lecture notes in electrical engineering, vol 601. Springer, Singapore. https://doi.org/10.1007/978-981-15-1420-3_199
6. Setiawan AW (2020) Image segmentation metrics in skin lesions: accuracy, sensitivity, specificity, dice coefficient, Jaccard index, and Matthews correlation coefficient. In: 2020 International conference on computer engineering, networking, and intelligent multimedia (CENIM). IEEE, 2020
7. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. SpringerPlus 5:507. https://doi.org/10.1186/s40064-016-2152-2
8. Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR In MIMO-OFDM using adaptive SLM And PTS technique. Int J Pure Appl Math 118(17):355–373. ISSN: 1311-8080 (printed version); ISSN: 1314-3395

9. Prasad PS, Pathak R, Gunjan VK, Ramana Rao HV (2020) Deep learning based representation for face recognition. In: Kumar A, Mozar S (eds) ICCCE 2019. Lecture notes in electrical engineering, vol 570. Springer, Singapore. https://doi.org/10.1007/978-981-13-8715-9_50
10. Karimullah S, Lalitha M, Mounika, Mahesh Reddy S, Mahammad Waseem S, Irfan SM (2023) Performance enhancement of wireless systems using hybrid RIS technique. In: Proceedings of the 4th international conference on data science, machine learning and applications ICDSMLA 2022. Lecture notes in electrical engineering, vol 1038. Springer, Singapore. https://doi.org/10.1007/978-981-99-2058-7_17

Network-Based Intrusion Detection by Using IOT Techniques

Nihal M. Pise, Girish R. Deshpande, Laxmi S. Naik, Sanjana D. Kamakar, Priyanka Patil, and Sanjeev Sannakki

Abstract The increased application of internet of things escalates the security concerns on these devices, when not handled properly would leave the system vulnerable to external threats. The current security mechanism requires high computation power and storage, they use technologies such as artificial intelligence, machine learning, and deep learning, moreover they cannot be implemented easily in all the environments. To facilitate this issue, we have proposed a network-based intrusion detection system on the internet of things device itself, that is simple to implement in any environment and requires less computation power and storage needs and has the ability to detect zero-day threats along with small number of false positive rate. The proposed system is developed using techniques such as deep packet inspection, packet sniffing, and regex string parsing along with the capability to block unauthorized client by dropping the incoming packet from that client. The proposed system uses a baseline profile which consists of valid operation of the internet of things device and will compare all the incoming traffic with this profile; moreover, it will determine whether to process the traffic normally or block it.

Keywords Internet of things · Security · Intrusion detection system · IOT security · Deep packet inspection

N. M. Pise (✉) · G. R. Deshpande · L. S. Naik · S. D. Kamakar · P. Patil · S. Sannakki
Gogte Institute of Technology, Belagavi, India
e-mail: nihalnaman25@gmail.com

G. R. Deshpande
e-mail: grdeshpande@git.edu

S. Sannakki
e-mail: sannakkisanjeev@git.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_11

1 Introduction

The recent advancement of the internet of things technology raises multiple applications in our surrounding. Using these devices without the proper security and monitoring tools would expose the network to external security threats. Because of the evolving nature of these devices and variety of applications, using the traditional security mechanism in some applications wouldn't be feasible. To overcome this issue, one workable solution would be to implement the threat detection tool on the internet of things device itself.

Because of the constraint on processing power and memory, such devices will have a limited set of functionality. We can achieve the above solution by developing an intrusion detection system that could alert the peer systems or users regarding a threat.

An intrusion is an unauthorized activity that will cause harm to the information system. An intrusion detection system is a software or hardware system that identifies these unauthorized activities [1]. We generally classify the intrusion detection systems into three types:

- Host-based intrusion detection system.
- Network-based intrusion detection system.
- Hybrid intrusion detection system.

The key difference between the above system is their placement in the host environment. The hybrid-based system is the combination of a network-based and host-based system. A network-based intrusion detection system is placed in an intermediate position, such that it can monitor all the incoming and outgoing network traffic. This kind of intrusion detection system is suitable for the proposed solution because we have the flexibility to analyze network traffic from both the ends.

The detection methodology of intrusion detection system is classified into following categories:

- Anomaly-based methodology
- Signature-based methodology
- State-full protocol analysis based methodology
- Hybrid-based methodology.

The anomaly-based approach is acceptable to the proposed solution because of the constraints imposed on the internet of things devices that would emphasize the devices will have a well-known and a few functionalities. In this technique, we would create a profile called baseline, which will consist of all the valid functionalities of the device. We will compare this profile with the incoming traffic. If the traffic matches the profile, then the data is processed normally, or else the system generates an alert and notifies the peer system or the users of a possible threat [2]. The profile comparison is done using deep packet inspection. In the deep packet inspection method, we will analyze the packet header and the payload of the packet [3].

The hardware requirements for the proposed solution would comprise:

- Raspberry Pi
- LCD
- Buzzer
- LED Sensor.

Raspberry Pi: It is a small computer that runs a unique operating system called “Raspbian.” This single-board system consists of multiple ports and pins that allow us to connect several internet of thing sensors, these pin in specific are called General Purpose Input Output Pin (GPIO) and we can use ports to connect keyboard, mouse, and monitor [4].
LCD: It is an electronic display module that can display 16 characters in each line, and it will include 2 lines on the display panel.
Buzzer: It is a magnetic speaker that produces an audio output on various input voltage frequencies.
LED Sensor: It is a diode that is doped especially for efficient light emission and is wrapped up in a transparent case.

2 Literature Survey

- **Smart TV: Features**

  1. Browsing the Web
  2. Online Streaming
  3. Smart Phone Connectivity
  4. Applications and Games.

- **Drawbacks**

  1. Smart TV Security and Privacy Risks are Real
  2. Internet Dependency
  3. Smart TVs have Inefficient Interfaces
  4. Smart TV Performance is Often Unreliable

     - Paper 1: A Survey on Intrusion Detection and Prevention Systems: Anomaly intrusion detection methodology, statistical anomaly detection, Knowledge/data-mining, and machine learning based [2].
     - Paper 2: Test Data Generation for State-full Network Protocol Fuzzing Using a Rule-Based State Machine: Fuzzing [5].
     - Paper 3: Survey of Intrusion Detection Systems: Techniques, datasets and challenges: Signature-based intrusion detection systems these are also known as Knowledge-based Detection or Misuse Detection [1].
     - Paper 4: Analysis of Encrypted Malicious Traffic: Deep packet inspection, Port-based methods [3].

3 Problem Definition

To design and develop a system that eliminates the exposure of the entire centralized system, that limits the attacker only with the access of the compromised device and discards all sorts of communications with the compromised device.

4 Existing System

The existing system is currently developed using technologies such as artificial intelligence, machine learning, and deep learning. Their accuracy rate depends on the training dataset they have used, and some datasets are specific to the protocol or the attack type, and they only can detect those protocol specific attacks. These systems are well suited for high-end computers and would devalue their performance when implemented in Raspberry Pi.

Disadvantage

- **A Deep Learning-Based Intrusion Detection System for MQTT Enabled IoT** [6]
  The drawback of this system is it requires more computational power and memory, when compared to implementing the system on Raspberry Pi.
- **Machine Learning-Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset)** [7]
  The accuracy of the system will depend on the training dataset, and the system cannot detect zero-day threats.
- **Generative Deep Learning to Detect Cyberattacks for the IoT-23 Dataset** [8]
  Requires high computation power and storage. Training the model would take a considerable amount of time and will totally depend on the computation power of the system.

5 Proposed System

The proposed system will require the following hardware components, such as a two Raspberry Pi, LCD, LED sensor, and a buzzer. The client script is executed on first Raspberry Pi, and it will send a humidity control command to the server system. The second Raspberry Pi will execute three scripts server, IDS, and intermediary. The server script will handle normal server functionality such as processing the data from the client, this script will communicate with the IDS script, which will perform deep packet inspection of the incoming traffic and compare it to the baseline profile. We will use the intermediary script for communicating with the android application and IDS script. Android application will alert and showcase the number of packets being processed, the number of valid and invalid packets, and the number of clients being blocked.

On the arrival of the incoming traffic, server and IDS scripts will perform their functionality simultaneously and the intermediary script will wait until it gets the data from the IDS script. The server script will wait for the response from the IDS script. If the traffic matches, the baseline profile IDS script will notify the server script to continue the execution normally and notify the intermediary script about the valid packet processed which in turn notifies the android application.

If an intrusion is detected, then IDS script will alert the server script and intermediary script, the packet being processed by the server script will be discarded the IDS script will add that media access control address to the violation list and increment the violation count of that client by one. If the client is already present in the violation list, only the violation count is incremented by one and the intermediary script will notify the android application about the invalid packet processed.

When the client reaches the violation count of three, the IDS script will add that client to the block list and drop all the incoming traffic from that client, and notify the intermediary script about the blocked client and power-ups the LED and buzzer along with the LCD with a message that showcasing blocked client's media access control address.

Advantages

- Requires less computational power and storage compared to the existing system.
- The proposed system has the ability to detect the zero-day threats.
- Proposed system can be easily implemented in any environment.
- Minimal power consumption compared to existing system and fewer false positive rates of intrusion.

6 Design Consideration

The following are the designs for the proposed system:

Figure 1 shows how different users/actors will interact with the proposed system, along with the relationship between the different blocks.

Figure 2 showcases the modular view of the proposed system and the association between the modules along with the user.

Figure 3 shows a sequence of interactions between different subsystems, along with a valid and invalid case of interaction. All the designs are developed using the Lucidchart and visual paradigm online tool.

7 Design Implementation

As shown in Fig. 2, the proposed system is divided into modules and each module has its own set of functionality as described in the following:

Client module: Allows the client to send Constrained Application Protocol (CoAP) packets to the server system, the packet consists of data, token, and the code, the

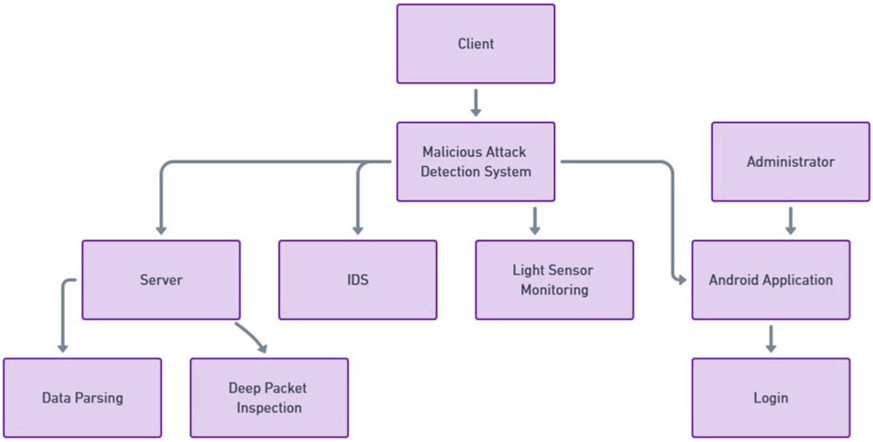


Fig. 1 Flow diagram

packet is constructed using aiocap. protocol module, where we have created context and assign the IP and port number, with additional feature such as encryption and decryption of the packet being sent and received.

Communication module: This module is responsible for transmitting the packet to its destination, here we use router as the communication module it will analyze the packet's media access control (MAC) address, IP address, and port number as well

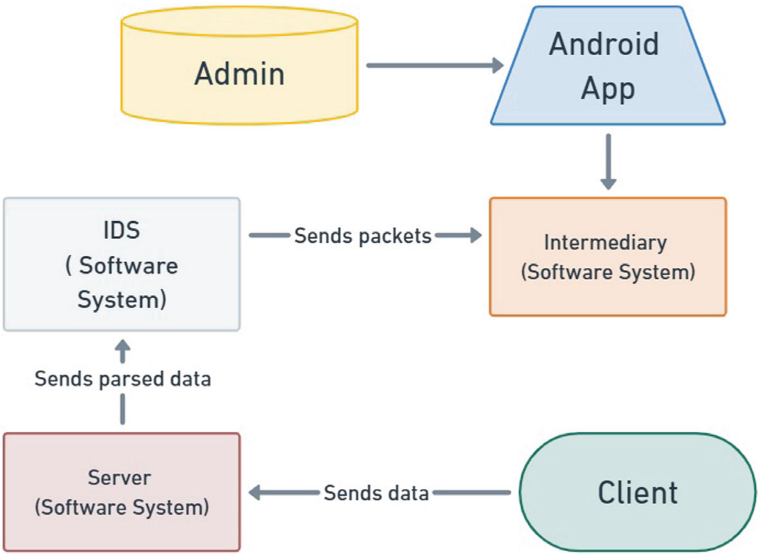


Fig. 2 System architecture

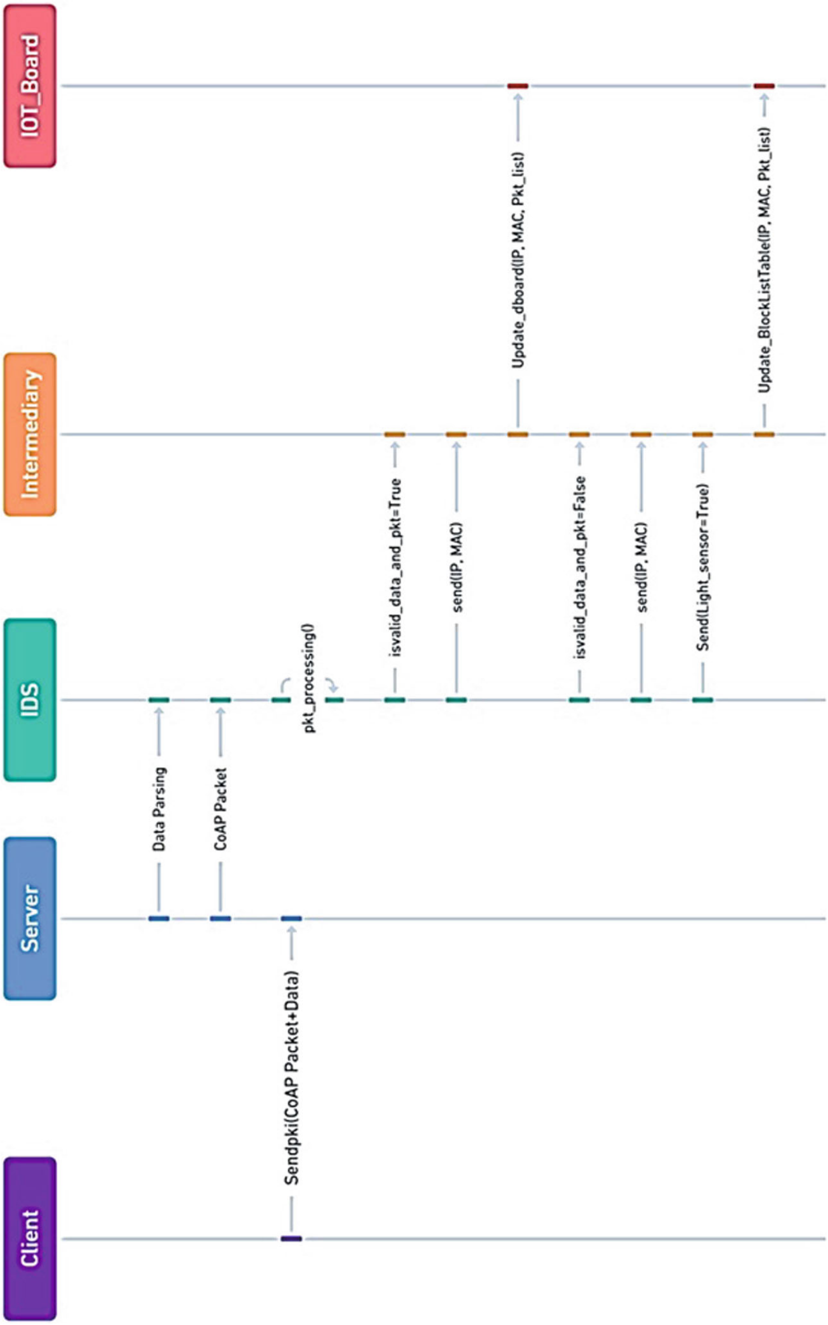


Fig. 3 Sequence diagram

as it will transmit the packet to appropriate destination, if the specified media access control address is not in the routers routing table then the packet is dropped.

Server module: It accepts the incoming packets and parse's the data, once the packet is received the packet data is decrypted using the decryption function and data parsing is done using regular expressions, for encryption and decryption, we use Chacha20Poly1305 in client and server.

IDS module: Used for packet inspection and classification of valid and invalid packet, if invalid packet is found it adds the details of that packet to the blocked list, for packet capture we use sniff() function of the scapy module that use's raw socket. And to drop the packet, we integrate our code with the iptables Linux build in firewall and this module is also responsible for powering up a LED, buzzer and LCD along with an alert message indicating the blocked client's media access control address.

Intermediary module: Used for integrating IDS and Android app, it consists of valid and blocked list of client along with packet capture rate and threat detection rate, and it uses stander internet protocol to transmit the data from the intermediary and the android application with a stander encryption–decryption policy for data confidentiality.

Android application module: Used to display the graphical view of packet capture and detection rate, along with valid and blocked list of clients.

The client module will communicate with the server module via the communication module, the intercommunication between server, IDS and intermediary module is done using sockets along with the security mechanism known as Chacha20Poly1305. The client, server, IDS, and intermediary module developed using the python language.

8 Results and Analysis

Here, we present the results of the proposed system:

Figures 4 and 5 show the normal execution of client and server program (Figs. 6, 7, 8, 9, 10, 11, 12 and 13).

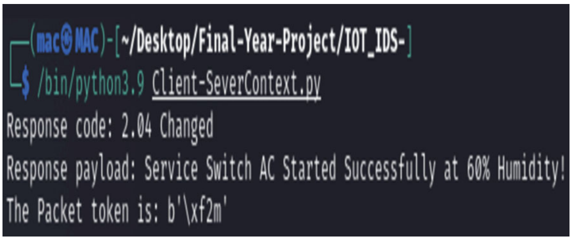


Fig. 4 Client program

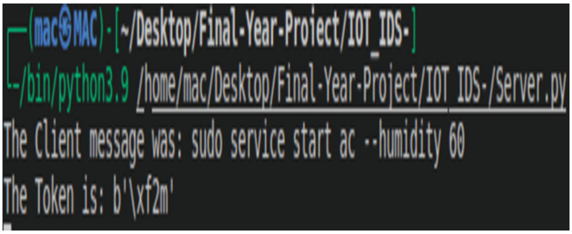


Fig. 5 Server program

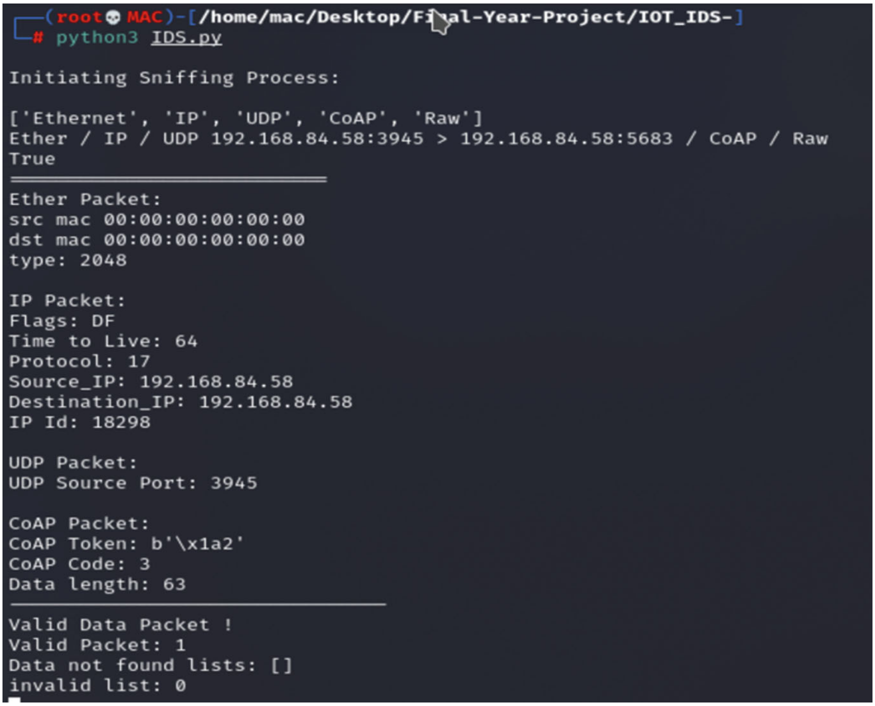


Fig. 6 IDS script processing normal packet

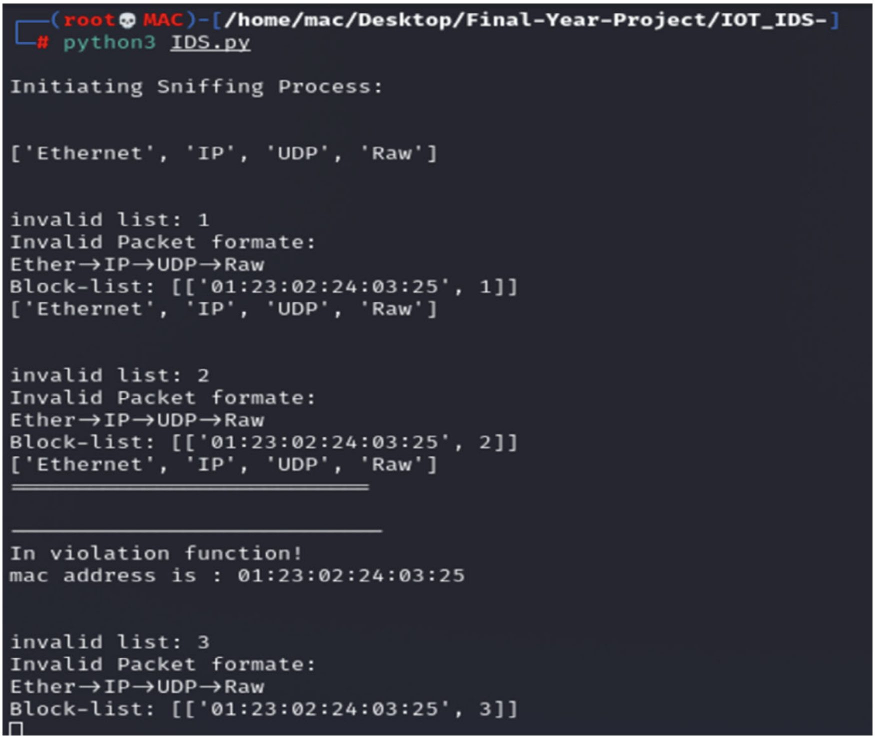


Fig. 7 IDS script processing invalid packets and adding client to the blocklist

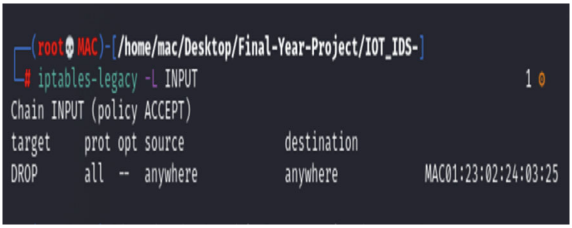


Fig. 8 IDS script added an invalid client to the iptables input chain and added a packet drop rule to that client

Fig. 9 User alert using SMS

Fig. 10 Initial view of the IOT kit will with all the necessary IOT component connected to the Raspberry Pi and proposed system running and executing normal packets

Fig. 11 Indication of intrusion from the Raspberry Pi on the LCD

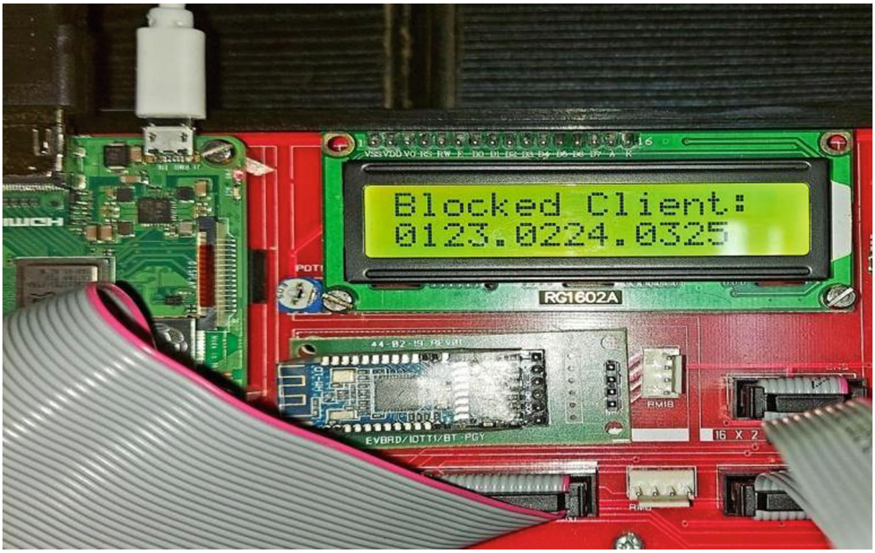


Fig. 12 Show casing the media access control address of the blocked client on the LCD in dot notation format

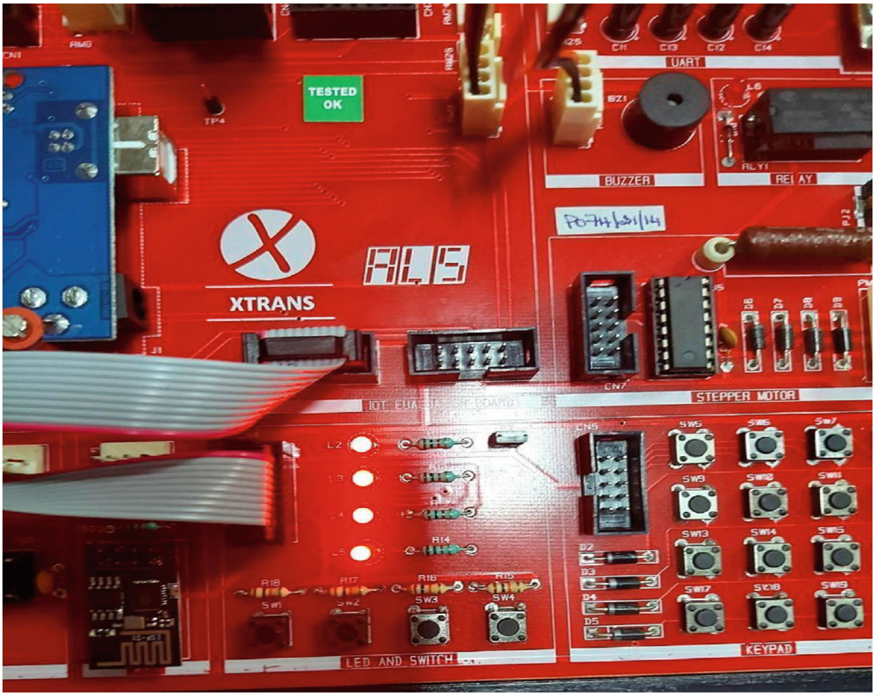


Fig. 13 Intrusion indication using buzzer and the LEDs

9 Conclusion

The goal of this work is to address the security concern of internet of things devices. By developing a network-based intrusion detection system on the internet of things, the device itself would allow us to easily implement the proposed system in any environment. The proposed system has the ability to categorize valid and invalid packets along with the feature of zero-day threat detection using a baseline profile that was developed using anomaly-based methodology, and packet classification is done using the deep packet inspection technique. The proposed system also has the capability to drop packets from invalid clients along with an alert triggering feature using LCD, LED, and buzzer with Raspberry Pi and an android application. Using the proposed system will provide us with an optimal intrusion detection rate with fewer false positives, with the additional benefits of low power consumption, low computation power, and storage requirements.

References

1. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J (2019) Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity 2(20)
2. Ghafir I, Husak M, Prenosil V (2014) A survey on intrusion detection and prevention systems. In: Student conference Zvůle
3. Shekhawat AS (2018) Analysis of encrypted malicious traffic. Master’s Projects 622
4. Deshpande GR, Sannakki S, Madi S, Kulkarni P, Madar A (2021) Advanced home automation by using Raspberry Pi. TURCOMAT 12(11):3031–3038
5. Ma R, Wang D, Hu C, Ji W, Xue J (2016) Test data generation for stateful network protocol fuzzing using a rule-based state machine. Tsinghua Sci Technol 21(3):352–360
6. Khan MA, Khan MA, Jan SU, Ahmad J, Jamal SS, Shah AA, Pitropakis N, Buchanan WJ (2021) A deep learning-based intrusion detection system for MQTT enabled IoT. Sensors
7. Hindy H, Bayne E, Bures M, Atkinson R, Tachtatzis C, Bellekens X (2020) Machine learning based IoT intrusion detection system: an MQTT case study (MQTT-IoT-IDS2020 dataset). In: Selected papers from the 12th International networking conference INC. Lecture notes in networks and systems, vol 180
8. Abdalgawad N, Kaddoura ASY, Zualkernan IA, Aloul F (2022) Generative deep learning to detect cyberattacks for the IoT-23 dataset. IEEE Access 10:6430–6441

A Survey Paper on Text Visualization Using Generative Adversarial Network

Aakanksha S. Choubey, Samta Gajbhiye, and Rajesh Tiwari

Abstract One of the primary uses of conditional generative models is the generation of visuals from text (natural languages). In addition to testing our conditional modeling and dimensional distribution capabilities at a theoretical level, text visualization offers a wide range of practical uses. Among the applications are photo editing and the generation of machine-aided content. Huge advances in generative adversarial neural networks have been accomplished in the past. Text visualization is one of the most intriguing findings achieved in the field of artificial intelligence in our century. The generative adversarial network for text to picture synthesis was unable to create correct and clear images in 2016. With technological advancements and model tweaks, it is now feasible to create clear and nearly totally correct pictures based on the description supplied. Visualizing a scene given a full description is a feat that humans can execute with little effort; nonetheless, it is a difficult activity that requires a mixture of numerous ideas defined in language so that they can be compared to how they seem in real life. In this research, we examine past work on picture synthesis from text descriptions in the context of improvements in generative adversarial networks (GANs), and we experiment with improved training strategies such as feature matching, smooth labeling, and mini-batch discriminating.

Keywords Convolutional neural networks · Generative adversarial networks (GAN) · Machine learning · Natural language processing · Neural nets · Object identification · Text visualization

A. S. Choubey · S. Gajbhiye
Department of CSE, Shri Shankarachrya Technical Campus, Bhilai, Chhattisgarh, India

R. Tiwari (✉)
Department of CSE, CMR Engineering College, Hyderabad, Telangana, India
e-mail: drrajeshtiwari20@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_12

1 Introduction

This section provides an overview of the research. The creation of an image based on a textual description has been shown to be a challenging issue in both the field of computer vision and the field of natural language processing. The extraction of descriptions and characteristics from an image has also been a challenging topic; this problem is the converse of the previous one. From a high level, this type of challenge is analogous to language translation problems. Text and pictures are two distinct "languages," in the same way that comparable semantics can be reduced to two languages. This challenge (image to text synthesis and vice versa) is special in that it is particularly multimodal in nature (involves more than one model to complete the synthesis and produce output). When translating a sentence such as "here is a handsome black man" into another language like German, there will only be a select number acceptable phrases that are equal to the one that was originally spoken [1]. There is a wide variety of possible outcomes that might occur whenever a person or even a piece of software attempts to generate an image that corresponds to the description that has been given. Multimodal behavior refers to problems in which many outputs correspond to multiple possible solutions; however, when the error associated with it is minimized to the point where the language is most often sequential, this problem may become easier to solve because the structure of the language is conditioned in such a way that the incoming generation of words depends on the words that were previously produced. In other words, the language is sequential. Multimodal behavior refers to problems in which many outputs correspond to multiple possible solutions; however, this problem may be easier when the error related with it is minimized to the point where the conversion of a picture into text is less complicated than the reverse process, which converts text into an image, and this is due to the need stated in the previous sentence. When it comes to image synthesis, the text may be put to a broad variety of uses, particularly when it is suited for commercial applications. For instance, rather of spending some hours and sometimes a no of days looking for comparable designed furniture, one would spend not as much of time specifying an item such as furniture that they want custom created. This would allow for more time to focus on the search.

GANs are effective models/methods. It helps train generative models. Multimodal learning challenges include learning a shared representation and anticipating missing data. Convolutional decoder networks have helped generator network modules of generative adversarial networks [2] synthesized pictures at several resolutions using a Laplacian pyramid of adversarial generators and discriminators. This technique produced eye-catching high-resolution pictures and may also be conditional on class labels for controlled production [3] employed a typical convolutional decoder but created a very effective and reliable architecture that included batch normalization to obtain impressive picture synthesis results. The primary difference between our work and the conditional GANs discussed above is that our model is based on text descriptions rather than class labels.

The generic adversarial networks may be used for a variety of reasons, but in our study, we employed them for text visualization in the form of image, which involves using text to generate a picture. To improve picture quality, we employed stack GANs, which created a low-resolution image in the first step. The most popular technique of training the generative model is to feed in both the picture and the text, which is then retained as a pair. As a result, if we provide text as an input, it will generate a similar image for display.

2 Literature Review

The study has supplied two phases of image synthesis from their written explanation. It is observed in the study report that the text description should contain the majority of the picture visual features. The image was generated using the information by the synthesizer and the more detailed and accurate the image description was, the more realistic the output image was. The continuous differentiable design generated was conditioned on extensive text descriptions to a certain extent than a single class label. The hybrid-level convolution neural networks were utilized to encode the text characteristics, which were subsequently used to condition the deep convolution generative model [4]. The picture generated convincing 64 by 64 graphics, but they never seemed real to humans. Rather than creating the photos in a single phase, the following author used two. Each step has a purpose in the generating process. First and second stages are the author's generation phases. The author utilized a generative model to build rough shapes for the items and the sampled noise vector from the previous distribution to generate a rough ideal backdrop. The initial stage only accepts the stage description [5]. The second step uses the first stage findings and the written description. The inputs detail the fundamental sketched shapes and increase image resolution. The author says the second-step drawings are more realistic and better. There is the detail of two steps.

In recent study [6], visuals were created from text captions using a vibration periodic auto encoder with the intention of painting the image in numerous phases, like DRAW [7]. The model can execute reasonable synthesis of wholly novel (unlikely for a person to create) material such as "a stop sign is flying in blue sky," indicating that it does more than just memories. While the findings are encouraging, the challenge is extremely difficult, and the generated images are not yet realistic, i.e., they can be mistaken for genuine. Our model can create esthetically credible pictures conditioned on text in many circumstances, and it is also unique in that our complete model is a GAN, rather than merely utilizing GAN for post-processing.

Generative adversarial networks (GANs) are made up of two components the generator and the discriminator behave like participants in a game. The concept is that generator will create a picture, and discriminator will try to determine whether or not the image created is the true image. However, the generator in this case attempts to deceive the discriminator.

3 Research Methodology

This section provides an overview of the methods and procedures used in this investigation, as well as the objectives that were deemed most suited for this study. It consists of the research objective, study design, ethical issues, and problems faced during the research process.

3.1 Research Purpose

This investigation is a theoretical analysis with the purpose of delivering succinct summaries of the findings obtained from a number of different GAN designs in addition to an experimental environment. In order to accomplish this goal, a comprehensive empirical and literary analysis of a large number of models and approaches is carried out. This study generalizes and covers a wide variety of topics pertaining to the field of adversarial networks. A descriptive research is carried out on the basis of the researcher's existing knowledge, which is then used to collaborate with the hypotheses. Additionally, the research questions that were posed are studied. It is important to take into account secondary data obtained from a variety of sources, including peer-reviewed papers, journals, case studies, and the conclusions and analyses that form the basis of the research.

3.2 Study Design

In this study, qualitative and research methods will be utilized to analyze the limitations and boundaries by conducting a thorough analysis of the most current academic publications and articles that have been subjected to peer review. In order to choose the most up-to-date papers in the subject matter that is being discussed, the researcher carried out a literature review that consisted of a survey as well as an analysis of academic journals sourced from both offline and online libraries that were already in existence. In addition, a case study analytic approach was utilized in order to explore the different factors. For the case studies, a new set of samples was used for the sampling process, and this time, objective sampling was utilized. The data analysis process involved functions that influenced qualitative and quantitative analysis, as well as procedures that influenced adversarial networks to create an experimental design, quantitative equations, and measurement variables In addition, the analysis of the data involved the creation of variables for the measurements. The analysis of the literature review provides the historical, present, and forthcoming patterns in GANs, which is significant for motivating further study since it provides the historical, current, and impending patterns in GANs. An empirical review and research were used to inform an offline, self-conducted study that contained study approaches based on a variety of various elements, and the findings are as follows: The research method utilized in this study is a combination of descriptive and correlation design. In order to provide evidence in support of the statements, modeling is done.

3.3 Ethical Considerations

The research was conducted in accordance with the ethical standards. To conduct the study, funding and permissions were obtained from all of the relevant organizations and governing bodies. The study that was examined for both the literature review and the empirical evaluation was found to meet the ethical requirements, and the results of this research were published in the public domain. In order to avoid accusations of plagiarism, each author of the literature review that was evaluated on the basis of secondary sources and for the purposes of analysis was recognized in the citation that was included within the text as well as in the references section.

3.4 Challenges Encountered

Several difficulties were faced during the research. Because it is a research based on both empirical data and an experimental setting, the theoretical modeling was compared at various places and on several variables to the various literature studies. Because there is little understanding of generative adversarial networks, it was challenging to obtain enough peer-reviewed publications for study.

4 Dataset and Preprocessing

4.1 Dataset

The dataset utilized was Oxford-102 flowers, which contains 8189 photos of flowers from various species. It comprises 102 classes, with 40–60 photos in each class and 10 written explanations for each image. We used 8000 photos for training in this investigation. For 300 epochs, this dataset was used to train the model.

4.2 Data Preparation

When the data was first collected and retrieved, it consisted of 8189 photos with varying resolutions and matching written descriptions. We utilized an NLTK tokenizer to turn textual sentences into words in order to normalize the textual data. These tokenized word lists were converted into an array of caption ids. The photos were loaded in order to be resized to the same dimensions. All training and testing photos were downsized to 128 × 128 resolution. The photos were transformed into arrays for training purposes, and both the language and images were fed into the model.

4.3 Generative Models Preprocessing

Generative models aim to solve the problem of converting text as a visual object like images. The ideal model for this transformation is generative adversarial networks. Brief background on generative models is presented. To illustrate, let's think of a set of n encoded images drawn at random as a vector of pixel values, $X = x_1, \ldots, x_n$, where P_{r} is an unknown sampling distribution where r is an actual image, any model that can learn to generate samples from a distribution. P_{g} while also estimating the sampling distribution P_{r} is considered a generative model. The model distribution P_{g} is the closest approximation to the true data distribution. By optimizing the log-likelihood function with respect to, say, P_{r} log (P_{g} (x)) [8] generative models may learn the distribution P_{g}. By taking fewer samples from further out and more from closer to X, maximum likelihood learning is analogous to concentrating more probability masses in the area surrounding X. Reducing the likelihood of a pullback assuming that P_{g} and P_{r} are densities, maximizing the Libeler divergence is the same as maximizing the log-likelihood function of the model. Due to the lack of needing to know the distribution of the population from which the sample was selected, this method of creating models is widely employed. With adequate data, the expectation may be roughly calculated using the weak law of large numbers. These days, GANs are the most popular model [9].

4.4 Stack Generative Adversarial Network

This paradigm is utilized in the natural language to picture conversion process. When a model like Stack GAN is fed picture descriptions, the model draws a low-resolution image in the first step by sketching the fundamental colors and rough outlines of the given items. Following that, it moves on to the next step, in which it takes the written description and the result from previous stage as input and creates a realistic and high-resolution image [10]. The enormous amount of photographs that potentially suit the specified description is one of the key obstacles this type of problem encounters. The generative adversarial network has recently improved as a consequence of the modifications made, and the outcomes obtained from sophisticated multimodal data synthesis and modeling are more promising than before. Despite the fact that text to image synthesis may produce plausible 64 × 64 graphics based on a text description, there are significant challenges related to the images that are formed, which are frequently abundant in object components and visual characteristics [11]. They might not grow eyes and beaks that are bright and colorful as birds do. The synthesis is unable to process photographs with a high resolution if further spatial clarifications of the topic picture are not given.

As was previously said, natural language to image conversion may be broken down into two basic sub-divisions that are easily accessible to the user by applying stacked GANs. These sub-divisions are stage one and stage two syntheses. In the initial stage of the process, the model learns to sketch primitive images by making use of the primary color. At this point, a vector of random noise that has been randomly selected from the prior distribution is used to construct the background region. The image that is created during the first phase is often warped as shown in Fig. 1, and it has a large number of faults; in the section on the adversarial network technique, we will cover the two phases in further detail [9]. After that, we go to step 2, which results in photographs of a superior quality and a greater likeness to the real thing. As was said before, the output of the first stage is used as an input for the second step of the synthesis, together with the picture text description [9]. At this stage, the primary emphasis is placed on making the photographs as comprehensive as is practically possible while also addressing any deficiencies. In the second step, we make use of all of the information that was disregarded in the first stage. This makes drawing the image easier because drawing a high-resolution image from scratch is really tough.

In two steps, stacked generative adversarial networks may generate realistic pictures from a given image text description as shown in Fig. 2. When compared to other models utilized in the text to picture conversion, our model outperforms others by about 20%. When extra spatial explanations are not supplied, the stack GAN may be employed in high-resolution picture production, but new models encounter difficulties when creating an image of 64 × 64 resolution level.

Fig. 1 The image output of the first stage (it is blurred and it's a rough output)

Fig. 2 The image output of the next stage is more clear and with better resolution (bitdepth)

4.5 Generative Adversarial Network

Most of the problems encountered by generative models may be remedied by employing the generative adversarial network. A model needed that could run in parallel and efficiently, yet the general level of graphical quality was low. In conclusion, there was a requirement for a model that could be modified according to the topology of the network that was producing samples as well as the loss function.

Since the loss it optimizes for does not include a biasness estimator, the P_r and P_g convergence equation is not biased [12]. However, these novel models have two key drawbacks: (1) they do not offer any signal of when convergence has taken place, and (2) they are unstable, especially during training. The stability of the generative adversarial network relies on a specific architecture and the thoughtful selection of key parameters (GAN). Nonetheless, it is not optimal if this is used in every kind of training [9]. The framework of generative adversarial is typically modeled as a competition between two players (the "generator" and the "discriminator/critic"). The purpose of this game is to see if the reviewer can be fooled into thinking that a computer-generated image is actually real. Once the picture is received, the discriminator will attempt to determine its authenticity [13]. As the game progresses, the two gain expertise in the given task, and their performance metric increases accordingly. In the end, they produce more photorealistic images [14]. In the given data set X, x_i are the number of samples that uniquely map to compact matrix x in the range [1, 1]. The critic is taught a (parametric) function of the form D_w: $x \rightarrow [0, 1]$. This algorithm takes an input image and returns a probability that represents the image's veracity [9]. Allowing J to represent the range of randomly chosen vectors J with a fixed; $P_z = N\ (0, 1)$.

Generator obtains the function G_r: $Z \rightarrow x$, which translates the random vector state to the random vector X state [15]. The visuals generated by the generator match to P_g X states. As a result, the generator may learn how to transfer pictures to a vector of noise.

Most of the time, the easiest approach to analyze and define the game is that it is a zero-sum game with G_r and D_w as the two players' tactics. Let v (D, G) denote the discriminator payoff function. The discriminator wants to enhance v, whereas the generator seeks to diminish it. To separate the true picture from the phony image, the value of D must be proportional to the value of V.

$$V(D, G) = \mathrm{EX} \sim P_\mathrm{r}\big[\log(D(x))\big] + \mathrm{EZ} \sim P_\mathrm{z}\big[\log(1 - D(G(z)))\big]$$

The role of discriminator is to force the generator to generate P_g which is closer to P_r.

4.6 Generative Models

To convert the GAN model to a conditional generative model, the vector C must be attached to discriminator and generator to create data [16]. It makes no difference which layer they are added to. As more inputs are supplied, the network will be able to trained and alter its settings accordingly.

4.7 Text Embedding

The text must first be victories before it can be utilized in the model—generative adversarial networks. Text embedding refers to the process of translating text to vector. In the text embedding calculations, the Char CNN-RNN encoder is employed. The encoder's job is to mapping between captions and images in common embedding space. Images that match are mapped to the inner product vectors with the highest inner product [17]. The CRNN is in charge of text description transformation, whereas the CNN is in charge of picture processing. Alternatively, a skip Thought vector model based just on language might be utilized. This model links sentences with similar meaning and grammar [18]. However, because it uses the matched pictures from the description, Char CNN-RNN is more effective in vision-related tasks. The embedding of the associated pictures is comparable to convolutional features. Because of this, they are visually prejudiced.

4.8 GAN Architectures

Early GAN designs comprised of fully connected neural networks; on the other hand, they were only effective with comparatively basic datasets such as MNIST and CIFAR-10. After that logical step was to employ convolutional neural networks (CNNs), which are significantly better suited for image recognition; unfortunately, training CNNs for GANs with the same capacity as those used in supervised learning tasks proved extremely challenging. DCGAN [17], the authors of which worked on a thorough search for different, was the first major advance in training GANs with CNNs. Convolutional architectures were studied, and some principles for constructing and training generators were developed [19–21].

GAN Designs—Fully connected neural networks made comprised the original generation of GAN designs, but they were only useful with small, straightforward datasets like MNIST and CIFAR-10 [22]. Convolutional neural networks (CNNs), which are far more suited for image recognition, were the logical next step. However, it turned out to be a very difficult challenge to train CNNs for GANs with the same capacity as those employed in supervised learning tasks. Following a thorough investigation of several convolutional architectures, the developers of DCGAN developed some recommendations for creating and honing generators and discriminators. They learn effective techniques for up sampling and down sampling using stride and fractionally stride convolutions, also known as deconvolutions or transposed convolutions, which ultimately leads to superior pictures. DCGAN [23] was the first big development in training GANs with CNNs [24]. Additionally, they suggest utilizing batch normalization, because it has an effect that helps to maintain stability throughout exercise. Finally, they recommend for adopting leaky ReLUs as the activation function since conventional ReLUs often have sparse gradients, which results in increased instability and, ultimately, superior picture quality. This is because leaky ReLUs have more gradients than conventional ReLUs. They also suggest employing batch normalization since it has an influence that is consistent throughout the training process. In conclusion, they propose utilizing leaky ReLUs as the activation function since traditional ReLUs often have sparse gradients, which results in increased instability. This is why they advocate using leaky ReLUs as the activation function [25, 26]. Due to the fact that the primary focus of our investigation is not on text embedding models, we will not go any deeper into this topic.

5 Conclusion

The outcomes and findings of the research are presented in the next section. It makes available a corpus of theoretical knowledge on the study issue, as well as a complete review of the literature. The generative adversarial model created the pictures in this article. The programs are written in Python 3.6 and include the GPU version of Tensor Flow. Tensor Flow is a Python open source library created by Google brain researchers. This library was designed to ensure that scientific and numerical computations were as efficient as feasible. Tensor Flow is a machine learning library that is frequently used. This is due to the fact that it provides both high-level and low-level APIs, which allows for faster iterations and better flexibility. It can do parallel calculation using modern graphics processing units, allowing it to handle operations effectively and efficiently. The proposed method decomposes the text visualization in two steps, first GAN sketches the object from given text descriptions. And then

GAN corrects the defects of first step results and adds more details, to achieve higher-resolution images with better quality. Extensive quantitative and qualitative results demonstrate the effectiveness of our proposed method.

References

1. Zhang H, Xu T, Li H, Zhang S, Huang X,Wang X, Metaxas D (2016) StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks 2(3):5. ArXiv preprint arXiv:1612.03242
2. Denton EL, Chintal S, Fergus R (2015) Deep generative image models using a Laplacian pyramid of adversarial networks. In: NIPS
3. Radford A, Merz L, Chintala S (2016) Unsupervised representation learning with deep convolution generative adversarial networks
4. Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolution generative adversarial networks. ArXiv preprint arXiv:1511.06434
5. St-Yves G, Naselaris T (2017) Decoding brain-like representations with a generative adversarial network. In: Conference on cognitive Com-600 mutational neuroscience
6. Mansimov E, Parisotto E, Ba JL, Salakhutdinov R (2016) Generating images from captions with attention. In: ICLR
7. Gregor K, Danihelka I, Graves A, Rezende D, Wierstra D (2015) Draw: a recurrent neural network for image generation. In: ICML
8. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Shi W (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR
9. Zhang H, Xu T, Li H, Zhang S, Wang X, Huang X, Metaxas D (2017) StackGAN++: realistic image synthesis with stacked generative adversarial networks. ArXiv preprint arXiv:1710.10916
10. Reed S, Akata Z, Yan X, Logeswaran L, Schiele B, Lee H (2016) Generative adversarial text to image synthesis. ArXiv preprint arXiv:1605.05396
11. Nie D, Trullo R, Lian J, Petitjean C, Ruan S,Wang Q, Shen D (2017) Medical image synthesis with context-aware generative adversarial networks. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham
12. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Bengio Y (2014) Generative adversarial nets. In: Advances in neural information processing systems, pp 2672–2680
13. Berthelot D, Schumm T, Metz L (2017) BEGAN: boundary equilibrium generative adversarial networks. ArXiv preprint arXiv:1703.10717
14. Kim T, Cha M, Kim H, Lee JK, Kim J (2017) Learning to discover cross-domain relations with generative adversarial networks. ArXiv preprint arXiv:1703.05192
15. Arjovsky M, Bottou L (2017) Towards principled methods for training generative adversarial networks. ArXiv preprint arXiv:1701.04862
16. Mirza M, Osindero S (2014) Conditional generative adversarial nets. ArXiv preprint arXiv:1411.1784
17. Johnson R, Zhang T (2015) Semi-supervised convolutional neural networks for text categorization via region embedding. In: Advances in neural information processing systems
18. Roweis ST, Saul LK (2010) Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500):2323–2326
19. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC (2017) Improved training of Wasser Stein GANs. In: Advances in neural information processing systems
20. Chen Z, Davis A, Dean J, Kudlur M (2016) TensorFlow: a system for large scale machine learning. In: Osdi, vol 16, pp 265–283

21. Behera AK, Tiwari R (2011) Analysis and design of software visualization tools for the behavior of object oriented programming. IJCA 13(1):975–8887
22. Park H, Waseem T, Ten WQI, Low YH (2021) Robustness evaluation of stacked generative adversarial networks using metamorphic testing. In: 2021 IEEE/ACM 6th International workshop on metamorphic testing (MET), 2-2 June 2021. https://doi.org/10.1109/MET52542.2021
23. Tan H, Liu X, Yin B, Li X (2023) DR-GAN: distribution regularization for text-to-image generation. IEEE Trans Neural Netw Learn Syst 34(12)
24. Peng J, Zhou Y, Sun X, Cao L, Wu Y, Huang F, Ji R (2022) Knowledge-driven generative adversarial network for text-to-image synthesis. IEEE Trans Multimedia 24
25. Zeeshan Khan Md, Jabeen S, Usman Ghani Khan Md, Saba T (2021) A realistic image generation of face from text description using the fully trained generative adversarial networks. https://doi.org/10.1109/ACCESS.2020.3015656
26. Khan M, Gajbhiye S, Tiwari R (2024) Fighting fake visual media: a study of current and emerging methods for detecting image and video tampering. In: Kumar A, Mozar S (eds) Proceedings of the 6th International conference on communications and cyber physical engineering, ICCCE 2024. Lecture notes in electrical engineering, vol 1096. Springer, Singapore, pp 545–556. ISBN: 978-981-99-7137-4, https://doi.org/10.1007/978-981-99-7137-4_54

English to Hindi Text Transliteration Using Deep Learning

Aparna Mete-Sawant, Aditya Kulkarni, Suketu Danke, Shravani Dhamne, and Raj Dharmale

Abstract In today's era, communication serves as a backbone of human interaction and plays a crucial role in various aspects of life. But the problem arises in approaching people from different linguistic communities and it also creates a language gap. Machine transliteration plays a vital role in reducing language barrier and enables effective communication between different scripts. In this paper, we have discussed various methods for English to Hindi Transliteration. We have implemented this project using Seq-2-Seq architecture with Long Short-Term Memory layers (LSTM), for encoding–decoding and to evaluate their performance-based factors such as accuracy, efficiency, and other performance factors. The proposed system is user friendly and provides the feature of incorporating with different platforms for better efficiency.

Keywords Machine transliteration · Long short-term memory · Natural language processing · Neural networks · Seq-2-Seq

A. Mete-Sawant
Assistant Professor, Department of Information Technology,
Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: aparna.mete@vit.edu

A. Kulkarni · S. Danke · S. Dhamne (✉) · R. Dharmale
Student, Department of Artificial Intelligence and Data Science,
Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: shravani.dhamne21@vit.edu

A. Kulkarni
e-mail: aditya.kulkarni21@vit.edu

S. Danke
e-mail: suketu.danke21@vit.edu

R. Dharmale
e-mail: raj.dharmale211@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_13

1 Introduction

The rapid growth of digital communication platforms necessitates the development of robust and accurate transliteration systems. Transliteration is a process of converting text from one writing system or script into another while preserving the pronunciation and phonetic characteristics of the original language. Transliteration between English and Hindi poses unique challenges due to the structural differences between the two languages.

English word: "University".
Transliteration in Hindi: यूनिवर्सिटी

The study of machine transliteration has been going on since 1960s. The earliest machine transliteration systems were rule-based and relied on manually defined transliteration rules. These systems mainly focused on specific language pairs. This system had limited features and could not handle complex linguistic patterns.

In the early 1990s, statistical approaches gained importance in machine transliteration. These methods used statistical models, such as hidden Markov models (HMMs) and finite-state transducers. This approach did improve accuracy but faced challenges in handling large data and out-of-vocabulary words.

Between (2000s–2010s) data-driven approaches became prevalent. Researchers started using large multilingual dictionaries to train transliteration models using statistical and machine learning methods. Some Machine learning algorithms were support vector machines (SVMs), conditional random fields (CRFs), and neural networks. Using this technique, it was possible to build real-time transliteration tools.

In the recent times, we mainly use deep learning and neural network-based methods such as convolution neural network and recurrent neural networks (RNNs), which includes various architecture which helps in building various real-life application. This method successfully builds the machine transliteration system with a good accuracy percentage.

2 Literature Survey

The paper by Ojaswi Srivastava had discussed the system which uses Convolutional Sequence-to-Sequence Neural machine translation framework and RNN-based NMT framework and between encoder and decoder it has used multi-hop attention mechanism which uses CNN with NMT framework [1].

Kshitij Tayal et al. (2020) discussed about the hybrid approach which is a combination of rule-based and data-driven methods. It uses a finite-state transducer to generate an initial set of candidate transliteration for a given input word. Then, a neural network is used to define the candidate transliteration and the model is trained using parallel corpus of English and Hindi text [2].

The paper by Shervin Malmasi and Mark Dras mentions the challenges faced in transliteration such as lack of orthographies, spelling variations, and language-specific issues. The paper also calculate various performance metrics and predict the result according to it [3].

Anurag Sharma et al. (2019) discussed the strength and weakness of rule-based and statistical methods. Deep learning models such as LSTM have been discussed and found that it has high accuracy as LSTM encoder/decoder approach requires a large number of sentence pairs [4].

The paper by Shruti Mathur and Varun Kumar Saxena discusses the hybrid approach for the transliteration. At first, the English words have been processed using rule-based approach which extracts individual homophones from the words and then it converts the English into its equivalent Hindi homophones using statistical approach and then eventually in the Hindi word [5].

Sandeep Saini et al. (2018) discussed about various machine translation technique and found out that Neural Machine Translation with a larger dataset and increasing more layers in the LSTM encoder/decoder could perform better and improve accuracy in comparison with other contemporary translation systems such as SMT and PBMT [6].

The paper by Baljeet Kaur Dhindsa suggests a method which transliterate words form English into Hindi language. It consists of two main models one CMU pronouncing dictionary of 122,230 words. Suppose a single word is not present in dictionary then it uses second model in which maximum five letters are required to generate a transliterated word [7].

P. Shalu et al. (2021) discussed and compared three Neural Machine Translation models: The first is Long Short-Term Memory and the second is Bidirectional LSTM with attention mechanism applied in three models, compares them, and states the better algorithm between LSTM and Bidirectional Long Short-Term Memory [8].

The paper by Jayashree Nair, Amrutha Krishnan K. discusses various technique for English to Hindi translation. The methods proposed are rule-based machine translation which first converts text into intermediate representation and then into a desired output. The other techniques uses word-by-word translation method, follows explicit linguistic rules, or learns translation pattern from parallel corpora [9].

Arshveer Kaur et al. (2018) proposed the *n*-gram-based approach to develop a tool for Punjabi to English Transliteration. The *n*-gram approach is a method commonly used in natural language processing (NLP) and text analysis. It involves breaking down text into contiguous sequences of n items, where an item can be a character, word, or even larger units such as phrases or sentences [10].

3 Methodology

The system uses Sequence-2-Sequence architecture with Long Short-Term Memory (LSTM) layer to convert English text into their corresponding Devanagari (Hindi) transliteration.

3.1 Dataset

The dataset used for the proposed system is Aksharantar dataset (English–Hindi pair) May 2022. The following dataset consist of about a million-language pair with high-quality transliteration. The dataset has two columns: The first column contained words in Latin script (English), and the second column contained words in Devanagari script (Hindi).

3.2 Tokenization

For any given input English sentence $\boldsymbol{S}_{\mathbf{Eng}}$ tokenization is performed to convert it into a sequence of character IDs:
Input Sentence: $\boldsymbol{S}_{\mathbf{Eng}}$ = "**Mai aam khaata hun**"
Tokenized Input Sequence (English): $T_{\text{Eng}} = [12,2,9,0,15,8,1,0,10,6,2,7,0,14,11,1,0,13,6,3,4]$

where

T_{Eng} represents the tokenized input sequence for the English sentence.
The integers in T_{Eng} correspond to character IDs.
0 is used for padding to ensure sequences are of equal length.

3.3 Sequence-to-Sequence Architecture

A sequence-to-sequence (Seq2seq) refers to a class of models used in natural language processing (NLP) and machine translation tasks which maps input sequence to output sequence, text summarization, and speech recognition. The model consists of two important components which are an encoder and a decoder.

In the context of the built model, the Seq2seq model utilizes two LSTM layers as the encoder and decoder components. It is a type of recurrent neural network. They contain memory cells that can maintain information over long periods of time, and they control all the information in the network and can store and retrieve in anytime depending upon the need of data (Fig. 1).

3.4 LSTM (Long Short-Term Memory)

Cell State—To transfer information throughout the sequence chain. The data is added or removed using gates.

Hidden State—It is meant to encode a kind of characterization of the most recent time step's data.

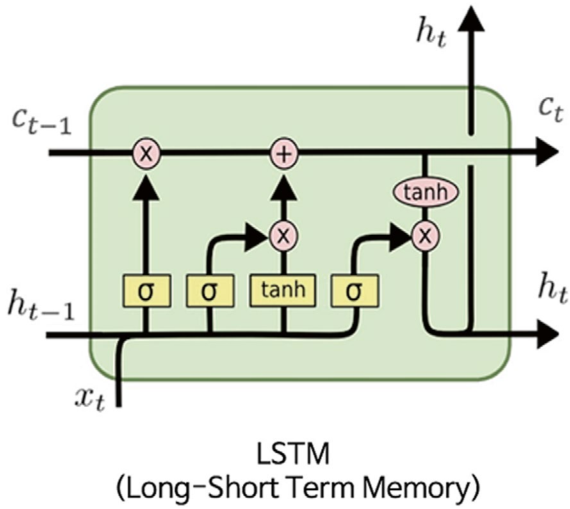


Fig. 1 Architecture of LSTM

Gates—It is meant to transfer data/information from one position to another. It is mainly of three types.

Forget Gates—Decides which information to keep or to throw away. The values vary between 0 and 1, closer to 0 means to throw away and closer to 1 means to keep the data.

Input Gates—This gate has two layers—Sigmoid Layer (transform value between 0 and 1 and Tanh Layer deals with the values between −1 and 1. The Sigmoid layer will decide which data to keep from the Tanh layer.

Output Gates—Decides what should be the next hidden state and transfers information to the Sigmoid layer.

Encoder: The encoder receives the input sequence, then each word in the sequence is passed through an embedding layer that maps the word to a dense vector representation. The embedded sequence is then fed into two LSTM layers in a sequential manner. The second LSTM layer which consists of final hidden state serves as an encoded representation of the input sequence.

The tokenized input $\boldsymbol{T}_{\mathbf{Eng}}$ sequence is passed through the encoder of the NMT model to obtain encoder states $\boldsymbol{S}_{\mathbf{H}}$ and $\boldsymbol{S}_{\mathbf{C}}$. The encoder consists of the following components:

Encoder Inputs: $\boldsymbol{E}_{\mathbf{In}} = \boldsymbol{T}_{\mathbf{Eng}}$
Encoder Masking: Masking is applied to handle variable-length sequences.
Encoder Embedding: Converts character IDs to dense vectors.

Encoder LSTM: Processes the embedded input sequence and produces encoder states.

$$S_{\mathbf{H}}, S_{\mathbf{C}} = \mathbf{LSTM}(E_{\mathbf{In}})$$

Decoder: The decoder takes the encoded representation which is generated by the encoder as input.

Similarly in this process also, each word in the target sequence is embedded using an embedding layer. The embedded sequence is then fed into two LSTM layers. At each time step, the decoder LSTM generates an output and updates its hidden state based on the previously generated output and the hidden state from the previous time step.

The decoder continues generating outputs until a maximum sequence length is reached (Fig. 2).

- **Data Preparation**: Three different datasets training, testing, and validation datasets are used. The English and the Devanagari words are tokenized separately. Then the English words are converted to padded sequences to ensure all sequences have the same length for training.

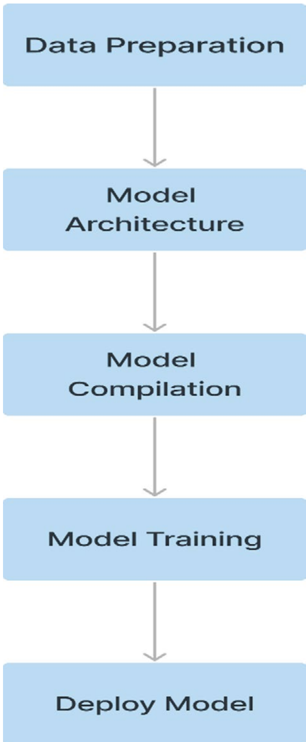


Fig. 2 Flowchart of proposed system

- **Model Architecture**: It sets up the input and output layers for the Seq2Seq model. Embedding layers are defined for both the encoder and decoder inputs, and LSTM layers are set up for the encoder and the decoder.
- **Model Compilation**: The model uses Adam optimizer for updating the model's parameters during the training process. The Adam-betas values were set to (0.9, 0.98).
- **Model Training**: The model is trained using the fit function, passing in the padded English and Devanagari sequences as input and target, respectively. The validation data is also provided to monitor the model's performance during training. The training process runs for a specified number of epochs, with a given batch size and validation split.
- **Deploy Model**: It is a visual representation of the proposed model which will be implemented using tools like MongoDB, flask server etc.

4 Result and Discussion

The proposed system designs an English to Hindi transliteration tool which provides a user-friendly platform. The system is efficient and the accuracy of the system comes out to be 86%.

5 Conclusion

The proposed system focusses on bridging the language gap and introducing a tool for English to Hindi Transliteration which enables effective communication between people from different linguistic communities. In the recent years, the transliteration has also been used in real-life application such as social media platforms. The system can be incorporated with different platforms to make it more efficient to work. The proposed system has got the accuracy of 86% and can be improved by implementing it with further coming architecture in the future.

Acknowledgements However, I would like to acknowledge the researchers, engineers, and experts who have contributed to the development of machine translation and transliteration system over the years and also to our guide Prof. Aparna Sawant for guiding us.

References

1. Srivastava O (2014) English to Hindi transliteration using machine learning. Int J Creative Res Thoughts (IJCRT)
2. Tayal K, Kulkarni V, Sundaram S (2020) A hybrid approach to English-Hindi transliteration using finite-state transducers and deep neural networks
3. Malmasi S, Dras M (2018) A survey of transliteration models. ACM Comput Surv

4. Sharma A (2019) A comprehensive survey of transliteration techniques for Indic scripts. J Intell Syst
5. Mathur S, Saxena VP (2014) A hybrid approach to English-Hindi name entity transliteration
6. Saini S, Sahula V (2018) Neural machine translation for English to Hindi. In: 4th International conference on information retrieval and knowledge management
7. Dhindsa BK, Sharma DV (2017) English to Hindi transliteration system using combination-based approach
8. Shalu P, Meera M. Neural machine translation for English to Hindi using GRU
9. Nair J, Krishnan KA, Deetha R (2016) An efficient English to Hindi machine translation system using hybrid mechanism. In: 2016 International conference on advances in computing, communication and informatics
10. Arshveer K, Goyal V (2018) Punjabi to English machine transliteration for proper nouns. In: 3rd International conference on internet of things smart innovation and usages
11. https://huggingface.co/datasets/ai4bharat/Aksharantar/tree/main

Machine Learning Models for Anomaly Detection in Data Mining

S. Dhanalakshmi, B. Pradeep, Balusupati Anil Kumar, Pirangi Vijay Kumar, Talari Swapna, and Rajendhar Reddy Gaddam

Abstract Anomaly detection is a critical task in data mining, aimed at identifying rare instances that deviate from the norm within a dataset. Anomaly detection has been greatly enhanced by machine learning models, which can understand intricate data patterns and correlations. This paper provides a comprehensive overview of machine learning models commonly used for anomaly detection in various domains. We discuss the principles underlying these models, their advantages and limitations, and their applications in real-world scenarios. Additionally, we highlight key considerations in model selection and evaluation for effective anomaly detection. By synthesizing existing research and discussing current trends, this paper aims to provide insights into the state of the art in machine learning-based anomaly detection.

Keywords Machine learning · Anomaly detection · Data mining · Classification · Clustering · Neural networks · Support vector machines

S. Dhanalakshmi (✉) · B. Pradeep · B. A. Kumar · P. V. Kumar · T. Swapna · R. R. Gaddam
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: rajamdhana@gmail.com

B. Pradeep
e-mail: pradeepb26@cmritonline.ac.in

B. A. Kumar
e-mail: anilkumarb@cmritonline.ac.in

P. V. Kumar
e-mail: vijaykumar@cmritonline.ac.in

T. Swapna
e-mail: swapnat@cmritonline.ac.in

R. R. Gaddam
e-mail: rajendarreddy.gaddam@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_14

1 Introduction

Anomaly detection in data mining plays a crucial role in identifying rare events or patterns that deviate significantly from the norm within a dataset. Machine learning models offer powerful tools for automating this process and have been increasingly adopted in various domains. One common approach is through supervised learning, where models are trained on labeled data to distinguish between normal and anomalous instances. Algorithms such as support vector machines (SVM), random forests, and neural networks [1] have been successfully applied in this context. However, obtaining labeled data can be challenging and may not always be feasible. Unsupervised learning techniques, on the other hand, are valuable when labeled data is scarce or unavailable. Clustering algorithms like K-means and DBSCAN [2], as well as density-based methods such as isolation forest and Gaussian mixture models, are commonly used for anomaly detection in this paradigm. These models identify outliers based on their deviation from the majority of the data points. Hybrid approaches combining both supervised and unsupervised methods are also prevalent, offering the advantages of both techniques. These models leverage the strengths of supervised learning for classification tasks while utilizing unsupervised methods to identify outliers in the absence of labeled data. Machine learning models provide effective solutions for anomaly detection in data mining, enabling the automatic identification of irregularities that may indicate fraudulent activities, system failures, or other significant events.

2 Literature Review

Anomaly detection, a long-standing method for identifying unusual components in data, has seen significant advancements over the years, with machine learning (ML) emerging as a pivotal technique in this domain. In their study, they conducted a systematic literature review (SLR) to analyze ML models used for anomaly detection across various applications. The review, spanning from 2000 to 2020, identified 290 research articles focusing on ML techniques for anomaly detection. They categorized the findings into four perspectives: applications of anomaly detection, ML techniques employed, performance metrics for ML models, and classification of anomaly detection. The review revealed 43 different applications, 29 distinct ML models, and 22 datasets used in experiments related to anomaly detection. Notably, unsupervised anomaly detection was found to be more prevalent among researchers. The study concluded that anomaly detection using ML models holds promise, offering numerous avenues for further research and provided recommendations and guidelines for future studies [3].

Intrusion detection has garnered widespread interest among researchers, yet persistent challenges remain, particularly in reducing false alerts during the identification of unknown attack patterns. Anomaly detection, a critical aspect of intrusion detection, focuses on detecting deviations from normal behavior that may indicate malicious activities. This paper provided an overview of research directions in applying both supervised and unsupervised methods for managing anomaly detection, addressing major theoretical issues and guiding researchers toward promising research directions [4].

With the exponential growth of data in today's world, the need for protecting data from attacks is paramount. Despite available techniques, vulnerabilities persist, leading to the emergence of data mining techniques for anomaly detection. This paper reviewed various data mining techniques used in anomaly detection to shed light on existing approaches and assist researchers in future endeavors in this field [5].

The proliferation of networked machines and the Internet has led to a surge in unauthorized activities, necessitating robust intrusion detection systems (IDSs). Anomaly detection systems (ADSs) play a crucial role in identifying malicious activities by monitoring deviations from normal behavior. This paper discussed the importance of ADSs in identifying attacks across various domains and highlighted the challenges and distinctions between different intrusion detection systems [6].

Anomalies, or rare events in datasets, are essential to detect as they may signify potential attacks or unusual occurrences. This paper provided a survey of different Such abnormalities may be located with the help of Anomaly Detection Techniques (ADT), which highlight the significance of selecting appropriate algorithms according to a number of criteria [7].

With the rise of IoT technologies, anomaly detection has become crucial in detecting suspicious events in data streams. This paper aimed to provide a comprehensive overview of state-of-the-art techniques and challenges in IoT data processing, addressing key issues such as data complexity, noise, and evolving characteristics [8].

Anomaly detection in building energy consumption is essential for identifying abnormal patterns and improving energy efficiency. This paper proposed a deep learning-based unsupervised anomaly detection framework and applied it to analyze energy data from residential houses, demonstrating its effectiveness in detecting abnormal energy consumption patterns [9].

Intrusion detection systems face vulnerabilities during data analysis and classification, requiring new analysis methods to monitor traffic effectively. This paper proposed a machine learning classifier for heterogeneous traffic and anomaly detection, aiming to improve intrusion detection accuracy [10].

Detecting anomalies in industrial processes is vital for spotting irregular activities and maintaining seamless operations. This study introduces a technique employing ML classification algorithms for real-time anomaly detection in industrial processes, showcasing its efficacy through a case study in the pharmaceutical industry [11].

Anomaly detection, a vital aspect in diverse applications, requires effective methods to learn complex structures from noisy data. A variety of anomaly detection algorithms based on deep learning were covered in this paper, including their uses, difficulties, and potential future developments [12].

Anomaly-based intrusion detection systems face challenges in accurately detecting attacks. This paper proposed a method for improving the detection accuracy of such systems using various machine learning algorithms, demonstrating their effectiveness in detecting anomalies using real-world datasets [13].

Anomalies pose threats to networks and require effective detection methods. This paper reviewed the challenges of anomaly detection in traditional and next-generation networks and discussed the implementation of machine learning in anomaly detection under different network contexts, providing insights into various machine learning models and their advantages [14].

3 Mathematical Model

Statistical Methods

Statistical methods for anomaly detection rely on assessing the likelihood of observing a particular data point given its distribution. Here are a few statistical techniques along with their equations:

Z-Score: The Z-score measures the number of standard deviations a data point is from the mean of the dataset. Anomalies are often defined as data points with a Z-score beyond a certain threshold.

$$Z = \frac{(X - \mu)}{\sigma}$$

where

- X is the data point,
- μ is the mean of the dataset,
- σ is the standard deviation of the dataset.

Grubbs' Test: Grubbs' test detects outliers in a univariate dataset by comparing the maximum or minimum value against the mean.

$$G = \frac{\left|X - \overline{X}\right|}{s}$$

where

- X is the data point being tested,
- $\overline{X}$ is the sample mean,
- s is the sample standard deviation.

The null hypothesis is that there are no outliers, and if the calculated G-value exceeds a critical value from a Grubbs' table for a chosen significance level and sample size, the null hypothesis is rejected.

Dixon's Q Test: Dixon's Q test identifies a single outlier in a univariate dataset by comparing the difference between the outlier and the nearest value to the range of the dataset.

$$Q = \frac{|X_2 - X_1|}{R}$$

where

- X_1 and X_2 are the values being compared,
- R is the range of the dataset.

If the calculated Q-value exceeds a critical value from a Dixon's Q table for a chosen significance level and sample size, the outlier is identified.

These statistical methods provide simple yet effective techniques for detecting anomalies in datasets, particularly when the underlying data distribution is well understood. However, they may not be as robust to complex data patterns as machine learning-based approaches.

Unsupervised Anomaly Detection Methods

Density-Based Methods

Density-based methods are another category of anomaly detection techniques that rely on the notion of data density to identify anomalies. One of the most well-known density-based anomaly detection algorithms is density-based spatial clustering of applications with noise (DBSCAN). Let us outline the mathematical model of DBSCAN:

Let $X = \{x_1, x_2, \ldots, x_n\}$ be the dataset, where each x_i is a data point in $\mathbb{R}^{\mathrm{d}}$.

- Core Point: A point x_i is considered a core point if there are at least MinPts points (including itself) within a distance ε (radius) of it.
- Border Point: A point x_i is considered a border point if it is not a core point but lies within the ε-neighborhood of some core point.
- Noise Point: A point x_i is considered a noise point if it is neither a core point nor a border point.

The DBSCAN algorithm operates as follows:

- Parameter Setup: Specify the parameters ε (radius) and MinPts.
- Density Reachability: For each point x_i, calculate the set $N_\varepsilon(x_i)$ containing all points within distance ε of x_i.
- Core Points Identification: Identify all core points x_i where $|N_\varepsilon(x_i)| \geq \mathrm{MinPts}$.
- Connected Components: Form connected components of core points based on density reachability.
- Cluster Formation: Assign each border point to a cluster of its nearest core point.

- Anomaly Detection: Any point that is not assigned to a cluster is considered an anomaly (noise point).

The mathematical representation of DBSCAN entails delineating core points, border points, and noise points by assessing the density of data points within a given radius ε. Subsequently, the algorithm detects clusters comprised of core points and incorporates border points into these clusters, treating noise points as outliers.

Clustering-Based Methods

Clustering-driven anomaly detection methods operate under the premise that anomalies exhibit distinct clustering patterns compared to the bulk of normal data points. A typical strategy involves employing clustering algorithms to segment the dataset into clusters and subsequently detecting outliers or anomalies by assessing their distance from cluster centroids or evaluating the density of points within the clusters.

Let us outline a mathematical model for a clustering-based anomaly detection algorithm using k-means clustering as an example:

- Initialization: Choose the number of clusters k and randomly initialize k cluster centroids $\mu_1, \mu_2, \ldots, \mu_k$.
- Assignment Step: For each data point x_i, calculate its distance to each cluster centroid and assign it to the adjacent cluster:

$$c_i = \arg\min_i \left\| x_i - \mu_j \right\|^2$$

Update Step: Recalculate the cluster centroids based on the mean of all data points assigned to each cluster:

$$\mu_j = \frac{1}{\left|S_j\right|} \sum_{x_i \in S_j} x_i$$

where S_j is the set of data points assigned to cluster j.

- Repeat: Repeat the assignment and update steps until convergence (when cluster assignments no longer change significantly or after a specified number of iterations).
- Anomaly Detection: After obtaining the clusters, anomalies can be detected based on their distance from the cluster centroids. Points that are farthest from their nearest cluster centroid or have low membership in any cluster can be considered anomalies.

$$\text{Anomaly Score } (x_i) = \min_j \left\| x_i - \mu_j \right\|$$

where μ_j is the centroid of the cluster to which x_i is assigned.

Thresholding: Anomalies are identified based on a threshold on the anomaly scores. Points with anomaly scores above the threshold are considered anomalies.

This is a simplified mathematical model for a clustering-based anomaly detection algorithm using k-means clustering. Other clustering algorithms, such as hierarchical clustering or DBSCAN, can also be used for anomaly detection with similar principles, where anomalies are identified based on their deviation from the cluster structures formed by normal data points.

Isolation Forest

Isolation forest is a tree-based anomaly detection algorithm that isolates anomalies by randomly selecting a feature and then randomly selecting a split value between the maximum and minimum values of that feature. The algorithm is based on the principle that anomalies are more likely to be isolated in the tree structure compared to normal data points.

Let us outline the mathematical model of the isolation forest algorithm:

Isolation Tree Construction:

- Randomly select a feature q and a split value p for the feature from the range of its values.
- Partition the data into two subsets based on the feature q and split value p.
- Repeat the above steps recursively until the tree reaches a maximum height or the number of data points in the subset becomes less than a predefined threshold.

Anomaly Score Calculation:

- For each data point x_i, calculate the path length $h(x_i)$ from the root to the leaf node in each isolation tree.
- The average path length $E(h(x_i))$ over all trees is calculated.

Anomaly Detection:

Anomalies are identified as points with shorter average path lengths, indicating they are easier to isolate in the tree structure compared to normal data points.

$$\text{Anomaly Score } (x_i) = 2^{-\frac{E(h(x_i))}{c(n)}}$$

where $c(n)$ is a normalization factor representing the average path length of unsuccessful search in a binary tree, given by:

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}$$

where $H(i)$ is the i-th harmonic number defined as

$$H(i) = \sum_{k=1}^{i} \frac{1}{k}.$$

Thresholding:

Anomalies are identified based on a predefined threshold on the anomaly scores. Points with anomaly scores above the threshold are considered anomalies.

This is a simplified mathematical model for the isolation forest algorithm, which isolates anomalies by constructing isolation trees and calculating the average path length of data points in the trees. Anomalies are then identified based on their shorter average path lengths compared to normal data points.

One-Class SVM

Sure, let us outline a simplified mathematical model for a common machine learning model used in anomaly detection: the one-class support vector machine (one-class SVM) [15]. This model is often used in situations where we have only one class of data (normal instances) and want to detect anomalies as instances that deviate significantly from the normal class.

Let us denote the feature vector of a data point as $x \in \mathbb{R}^n$, where n is the number of features.

The objective of the one-class SVM is to find a hyperplane that separates the normal data instances from the origin (center of mass) in the feature space with the maximum margin, while allowing a fraction of the normal instances to be on the wrong side of the hyperplane (i.e., the model tolerates a certain level of contamination).

The hyperplane can be represented by the equation:

$$w^{\mathrm{T}}x - b = 0$$

where w is the normal vector to the hyperplane, and b is the bias term.

In the case of one-class SVM, we introduce a slack variable $\xi_i \geq 0$ for each data point x_i to allow for some normal instances to be on the wrong side of the hyperplane. We also introduce a parameter ν $(0 < \nu \leq 1)$ to control the fraction of outliers (anomalies) allowed:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + \frac{1}{\nu m}\sum_{i=1}^{m}\xi_i - b$$

subject to:

$$w^{\mathrm{T}}x_i - b \geq 1 - \xi_i, \ \geq 0, \sum_{i=1}^{m}\xi_i \leq \nu m$$

The first term in the objective function represents the margin maximization, while the second term represents the penalty for allowing data points to be on the wrong side of the hyperplane. The parameter ν controls the trade-off between maximizing the margin and tolerating a fraction of outliers. This optimization problem can be solved using techniques like quadratic programming or convex optimization methods to find the optimal hyperplane parameters w and b. This is a simplified mathematical model for one-class SVM in the context of anomaly detection. Other machine learning models for anomaly detection may have different mathematical formulations, but they typically involve optimizing an objective function to find the optimal decision boundary or separation between normal instances and anomalies.

4 Semi-supervised Anomaly Detection Methods

Semi-supervised anomaly detection methods leverage a combination of labeled and unlabeled data to detect anomalies within a dataset. These techniques are particularly useful when labeled data is limited or expensive to obtain, as they can make use of a smaller amount of labeled data in conjunction with a larger amount of unlabeled data to improve detection performance [16]. Semi-supervised anomaly detection methods in data mining:

- Self-Training with Anomaly Scores: This method involves merging a supervised anomaly detection model with self-training. First, the model undergoes training on a limited labeled dataset. Subsequently, it evaluates unlabeled data points to assign anomaly scores. The instances with the highest confidence in classification are pseudo-labeled as either normal or anomalous and incorporated into the training set [17]. The model undergoes retraining using this enlarged dataset, and the cycle continues until convergence is achieved.
- Semi-Supervised Support Vector Machines (S3VM): S3VM extends traditional support vector machines (SVM) to incorporate unlabeled data during training. The labeled data is used to train a standard SVM, while the unlabeled data is incorporated into the optimization objective function to improve the decision boundary's generalization. This approach aims to find a decision boundary that maximizes the margin between normal instances and the decision boundary while minimizing the number of anomalies.
- Transductive Support Vector Machines (TSVM): TSVM is another extension of SVM that incorporates unlabeled data during training. Similar to S3VM, TSVM aims to find a decision boundary that maximizes the margin between labeled instances and the decision boundary while respecting the unlabeled data points. TSVM iteratively alternates between training on the labeled data and updating the decision boundary based on the unlabeled data until convergence [18].
- Generative Adversarial Networks (GANs) for Anomaly Detection: GANs can also be used for semi-supervised anomaly detection. In this approach, a GAN is trained on the normal instances to learn the data distribution. Then, anomalies are identified as instances that deviate significantly from this learned distribution. Additionally, GANs can be trained with both normal and anomalous instances, encouraging the generator to generate realistic instances while the discriminator learns to distinguish between normal and anomalous samples [19].

- Semi-Supervised Autoencoder-based Approaches: Autoencoder-driven techniques can be customized for semi-supervised anomaly detection through training on a mix of labeled and unlabeled data. This approach enables the model to reconstruct normal data instances while grasping the intrinsic data structure. Anomalies are then pinpointed by analyzing reconstruction errors, with instances exhibiting elevated errors flagged as anomalous.
- These semi-supervised anomaly detection methods leverage both labeled and unlabeled data to improve detection performance and are applicable in various domains where labeled data is limited or costly to obtain.

5 Supervised Anomaly Detection Methods

Supervised anomaly detection methods in data mining rely on labeled data, where anomalies are explicitly identified and labeled during the training phase [20]. These methods are effective when a sufficient amount of labeled anomaly data is available for model training. Here are some common supervised anomaly detection methods:

- Support Vector Machines (SVM): SVM is a popular supervised learning algorithm used for anomaly detection. In SVM-based anomaly detection, the algorithm learns a hyperplane that separates normal instances from anomalies in the feature space. Anomalies are instances that lie on the wrong side of the hyperplane or have a margin below a certain threshold.
- Random Forests: During training, random forests—an ensemble learning method—build many decision trees. In supervised anomaly detection with random forests, each decision tree is trained on a subset of the data, and anomalies are identified based on their frequency of occurrence in the leaves of the trees. Instances that are rare or have unusual feature values are classified as anomalies.
- Neural Networks: Deep learning techniques, such as neural networks, can also be used for supervised anomaly detection. In this approach, a neural network model is trained on labeled data to learn the normal behavior of the system. Anomalies are then identified as instances that deviate significantly from the learned normal behavior. Variants of neural networks, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), can be used for anomaly detection in different types of data, such as images, time series, and text.
- Ensemble Methods: Ensemble techniques amalgamate several base models to enhance predictive accuracy. In supervised anomaly detection, methods like bagging and boosting are employed to merge multiple anomaly detection models, such as SVMs or random forests, aiming for improved detection capability. By harnessing the distinctiveness of individual models, ensemble approaches efficiently discern anomalies within the dataset.
- One-Class Classification: One-class classification is a specialized form of supervised learning where the model is trained only on normal instances and learns to distinguish between normal and anomalous instances. One-class SVM, for example, is a variant of SVM that learns a boundary around normal instances in

the feature space. During testing, instances outside this boundary are classified as anomalies.

- Distance-based Methods: Distance-based methods measure the distance between data instances in the feature space and identify anomalies based on their proximity to normal instances. Examples include *k*-nearest neighbors (*k*-NN) and distance-based clustering algorithms like density-based spatial clustering of applications with noise (DBSCAN).

These supervised anomaly detection methods are effective when labeled anomaly data is available and can be applied in various domains, including fraud detection, network security, manufacturing, and health care.

Performance Metrics for Anomaly Detection

Performance metrics for anomaly detection evaluate the effectiveness of a model in identifying anomalies within a dataset. The accuracy with which the model can distinguish between typical and abnormal occurrences is measured by these criteria.

Some common performance metrics for anomaly detection include true positive rate (TPR) or recall, false positive rate (FPR), precision, F_1 score, area under the receiver operating characteristic curve (AUROC), area under the precision-recall curve (AUPRC), and accuracy.

True positive rate (TPR), alternatively referred to as recall or sensitivity, gauges the portion of actual anomalies accurately detected by the model. It is computed by dividing the count of true positives (accurately identified anomalies) by the sum of true positives and false negatives (anomalies erroneously categorized as normal). FPR measures the proportion of normal instances that are incorrectly classified as anomalies by the model. It is calculated as the ratio of false positives (normal instances incorrectly classified as anomalies) to the sum of false positives and true negatives (correctly classified normal instances).

Precision measures the proportion of true anomalies among the instances identified as anomalies by the model. It is calculated as the ratio of true positives to the sum of true positives and false positives.

The F_1 score is the harmonic mean of precision and recall, providing a balance between precision and recall, especially useful when there is an imbalance between the number of anomalies and normal instances in the dataset.

AUROC measures the model's ability to distinguish between anomalies and normal instances across various decision threshold levels. A higher AUROC indicates better discrimination ability of the model.

AUPRC measures the trade-off between precision and recall across different decision threshold levels. A higher AUPRC indicates better performance of the model in identifying anomalies while maintaining a high precision rate.

While accuracy is a common metric in classification tasks, it may not be suitable for evaluating anomaly detection models, especially in imbalanced datasets where the number of anomalies is much smaller than the number of normal instances. However, in some cases, accuracy can still provide useful insights into the overall performance of the model.

These performance metrics help assess the effectiveness of anomaly detection models and guide the selection of appropriate models and tuning parameters to optimize detection performance.

6 Conclusion and Future Work

The paper explores various machine learning techniques for anomaly detection, shedding light on their effectiveness and applicability in diverse data mining scenarios. The study investigates the strengths and limitations of algorithms such as isolation forests, one-class SVM, statistical methods, density-based methods, and clustering-based methods in identifying anomalies within datasets. Through rigorous experimentation and evaluation, the paper provides valuable insights into the performance metrics, scalability, and interpretability of these models.

For future work, the research suggests avenues for improvement and refinement of anomaly detection methods. It recommends exploring ensemble approaches that combine multiple algorithms to enhance overall detection accuracy. Additionally, the paper emphasizes the importance of addressing the interpretability challenges associated with complex models, encouraging further research into developing transparent and explainable anomaly detection solutions. Furthermore, the paper investigates the adaptability of these models to evolving data distributions and real-time applications, emphasizing the need for continuous evaluation and refinement. Overall, the paper lays the foundation for advancing anomaly detection in data mining through ongoing research and innovation in machine learning methodologies.

References

1. Boateng EY, Otoo J, Abaye DA (2020) Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: a review. J Data Anal Inf Process 8(4):341–357
2. Venkatkumar IA, Shardaben SJK (2016) Comparative study of data mining clustering algorithms. In: 2016 International conference on data science and engineering (ICDSE). IEEE, pp 1–7
3. Nassif AB, Talib MA, Nasir Q, Dakalbab FM (2021) Machine learning for anomaly detection: a systematic review. IEEE Access 9:78658–78700
4. Omar S, Ngadi A, Jebur HH (2013) Machine learning techniques for anomaly detection: an overview. Int J Comput Appl 79(2)
5. Agrawal S, Agrawal J (2015) Survey on anomaly detection using data mining techniques. Procedia Comput Sci 60:708–713
6. Narayana MS, Prasad BVVS, Srividhya A, Reddy KPR (2011) Data mining machine learning techniques—a study on abnormal anomaly detection system. Int J Comput Sci Telecommun 2(6)
7. Toshniwal A, Mahesh K, Jayashree R (2020) Overview of anomaly detection techniques in machine learning. In: 2020 4th International conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). IEEE, pp 808–815

8. Al-amri R, Murugesan RK, Man M, Abdulateef AF, Al-Sharafi MA, Alkahtani AA (2021) A review of machine learning and deep learning techniques for anomaly detection in IoT data. Appl Sci 11(12):5320
9. Xu C, Chen H (2020) A hybrid data mining approach for anomaly detection and evaluation in residential buildings energy data. Energy Build 215:109864
10. Guezzaz A, Asimi Y, Azrour M, Asimi A (2021) Mathematical validation of proposed machine learning classifier for heterogeneous traffic and anomaly detection. Big Data Min Anal 4(1):18–24
11. Quatrini E, Costantino F, Di Gravio G, Patriarca R (2020) Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities. J Manuf Syst 56:117–132
12. Wang R, Nie K, Wang T, Yang Y, Long B (2020) Deep learning for anomaly detection. In: Proceedings of the 13th International conference on web search and data mining, pp 894–896
13. Shin Y, Kim K (2020) Comparison of anomaly detection accuracy of host-based intrusion detection systems based on different machine learning algorithms. Int J Adv Comput Sci Appl 11(2)
14. Wang S, Balarezo JF, Kandeepan S, Al-Hourani A, Chavez KG, Rubinstein B (2021) Machine learning in network anomaly detection: a survey. IEEE Access 9:152379–152396
15. Hejazi M, Singh YP (2013) One-class support vector machines approach to anomaly detection. Appl Artif Intell 27(5):351–366
16. Song H, Jiang Z, Men A, Yang B (2017) A hybrid semi-supervised anomaly detection model for high-dimensional data. Comput Intell Neurosci 2017
17. Pang G, Yan C, Shen C, Hengel AVD, Bai X (2020) Self-trained deep ordinal regression for end-to-end video anomaly detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 12173–12182
18. Kondratovich E, Baskin II, Varnek A (2013) Transductive support vector machines: promising approach to model small and unbalanced datasets. Mol Inf 32(3):261–266
19. Li D, Chen D, Goh J, Ng SK (2018) Anomaly detection with generative adversarial networks for multivariate time series. arXiv preprint arXiv:1809.04758
20. Görnitz N, Kloft M, Rieck K, Brefeld U (2013) Toward supervised anomaly detection. J Artif Intell Res 46:235–262

Talkify: A Tool to Help People with Stuttering Condition

Aparna Sawant, Dhiraj Pawar, Parth Shethji, Mayur Patil, Richa Mahajan, and Vaishnavi Petkar

Abstract Stuttering is a global speech disorder that causes communication problems and emotional distress. In this paper, we introduce "Talkify", a cutting-edge deep learning tool that detects and analyzes stuttering patterns. Talkify uses a convolutional neural network (CNN) model trained on various stuttered speech samples to identify stuttering instances. Talkify uses spectrogram analysis to distinguish between stuttering events and fluent speech segments. The CNN model extracts relevant features from the spectrogram images to classify stuttering patterns. In this paper, we look at how Talkify can be used in speech therapy sessions and personal speech assessment tools as well as telehealth platforms. With Talkify, people with stuttering conditions can track their speech and track their progress in real time. Talkify provides a dependable, objective tool for stuttering pattern detection and analysis. Developing and implementing Talkify holds promise for improving communication support and quality of life for people with stuttering.

A. Sawant
Department of Information Technology, Vishwakarma Institute of Technology, Pune, Maharashtra, India
e-mail: aparna.mete@vit.edu

D. Pawar (✉) · P. Shethji · M. Patil · R. Mahajan · V. Petkar
Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Technology, Pune, Maharashtra, India
e-mail: dhiraj.pawar21@vit.edu

P. Shethji
e-mail: parth.shethji21@vit.edu

M. Patil
e-mail: mayur.patil21@vit.edu

R. Mahajan
e-mail: richa.mahajan21@vit.edu

V. Petkar
e-mail: vaishanavi.petkar21@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_15

Keywords Stuttering · Speech disorder · Talkify · Speech recognition · Natural language processing · Spectrogram

1 Introduction

Stuttering is a global issue of great concern given its wide occurrence and the resultant challenges in communication. Millions of individuals, amounting to one percent of the global population, are affected. The condition of stuttering can result in challenges with articulating thoughts smoothly, which can lead to feelings of frustration, anxiety in social situations, and low self-esteem. It has the potential to impact numerous areas of life, encompassing academic achievements, professional success, social connections, and overall well-being. Approximately 5% of young children are believed to experience a phase of stuttering during their early development. It is important to grasp the fact that not all children who stutter will continue to stutter as adults. Stuttering is often a temporary and natural stage experienced by young children, which typically resolves itself without requiring any intervention or treatment. On the other hand, certain children may undergo persistent stuttering which necessitates assessment and treatment from an expert in the field.

Stuttering is having different types as follows:

1. Developmental stuttering, a condition in children, can be effectively managed through early intervention and speech therapy, enabling them to improve their fluency and communication skills.
2. Neurogenic stuttering depends on the underlying neurological condition; it may involve teaching compensatory techniques to improve communication skills and manage speech interruptions.
3. Psychogenic stuttering, stemming from psychological factors, often necessitates psychotherapy or counseling for addressing underlying issues, making speech therapy less effective.
4. Persistent developmental stuttering, a condition with ongoing stammering in developmental years, typically requires speech therapy to collaboratively develop techniques for improving fluency and addressing emotional and social challenges related to stuttering.
5. Acquired stuttering depends on addressing the root cause, such as trauma or brain injury, to manage the associated speech difficulties.

It is crucial to keep in mind that the extent of a person's stuttering, along with other factors, and the effectiveness of speech therapy can all impact its overall usefulness. Speech-language pathologists are highly skilled professionals who possess the expertize to evaluate individuals and provide tailored therapies that cater to each client's specific needs and unique circumstances.

Automated stuttering treatment has made remarkable progress, especially in the realm of computer-based speech therapy. Nowadays, advanced AI algorithms are employed by virtual speech therapy systems to offer interactive exercises, real-time feedback, and continuous progress monitoring. Thanks to these automated solutions, individuals are now able to receive personalized assistance while practicing fluency skills remotely. AI systems have the capability to analyze speech samples automatically, identifying disfluencies and other acoustic characteristics related to stuttering. Automation streamlines the evaluation process and allows for tracking of treatment progress.

AI systems can provide objective evaluations of the severity of stuttering by analyzing speech patterns, thereby assisting in diagnostic and treatment planning. In addition, AI-powered platforms specializing in virtual speech therapy offer personalized and interactive exercises tailored to enhance fluency strategies and overcome speech challenges. Online environments provide individuals with a safe and motivating space to enhance and gain self-assurance in their communication skills.AI-based solutions have the potential to enhance the management and outcomes of individuals with stuttering issues through personalized assistance, immediate feedback, and constant availability of information and therapy, regardless of time and location.

2 Literature Review

According to the "Outcome of stuttering therapy on Egyptian school-aged children using the speak freely program", stuttering therapy utilizing the Speak Freely program has helped Egyptian school-aged children's fluency, communication skills, and general well-being. However, further research is required to increase the evidence base due to the sparse number of studies and methodological variances [1]. Future research should make use of bigger sample numbers, consistent outcome measurements, and long-term follow-ups to produce more reliable proof of the Speak Freely program's efficacy in this particular group.

The study on "Speech recognition and correction of a stuttered speech" is examined in this review of the literature. A complicated speech problem called stuttering is characterized by disturbances in the rhythm and fluency of speech output [2]. Technology advancements have resulted in the creation of voice recognition systems that can recognize and analyze speech patterns, potentially opening up new treatment and diagnosis options for stuttering. The study done so far shows intriguing results in the precise detection of speech patterns that stuttered and in the creation of systems that provide real-time feedback. The difficulties posed by individual variance and the incorporation of technology into therapeutic situations, however, call for further development.

The paper "FluentNet: End-to-end detection of speech disfluency with deep learning" describes a novel method for detecting voice disfluency [3]. With its end-to-end design and attention techniques, the suggested FluentNet model outperforms conventional approaches. The research provides insightful information on speech disfluencies identification, providing opportunities for enhanced speech processing systems and clinical language problem evaluations.

This paper "Optimization of stammering in speech recognition applications" offers a thorough examination of optimization methods for voice recognition initiatives that recognize stuttering [4]. The suggested methods seek to improve the effectiveness and usability of voice recognition systems for people who have stammering problems. The experimental findings show how successful the optimization strategies are and point to the room for growth in this field of study.

The paper "A tutorial on hidden Markov models and selected applications in speech recognition" describes, due to its capacity to describe sequential data with hidden states, hidden Markov models (HMMs) have found extensive usage in many domains, including voice recognition [5]. This paper gives a general overview of HMMs and voice recognition applications. Model structure, parameter estimation, and inference techniques are only a few of the key HMM ideas that are covered in the paper. Additionally, it explores a few uses of HMMs in voice recognition, including speech synthesis, language modeling, and acoustic modeling. In addition to outlining recent developments in HMM-based speech recognition systems, the paper suggests possible future research areas in this area.

This paper, "Support vector machine-based stuttering disfluency classification using GMM supervectors", offers a thorough overview of the Gaussian mixture models (GMM) supervector-based support vector machine (SVM)-based stuttering disfluency classification method. It discusses the fundamental ideas, approaches, and difficulties surrounding the categorization of stuttering disfluency [6]. The paper discusses the usefulness of the suggested strategies and evaluates their performance while highlighting the improvements made in SVM-based classification utilizing GMM supervectors.

The categorization of speech disfluencies—disturbances in the natural flow of speech—using multiclass support vector machines (SVM) and speech parameterization approaches is discussed in this paper [7]. The paper "Classification of speech disfluencies using speech parameterization techniques and multiclass SVM" offers a summary of the research in this field, stressing the several speech parameterization methods used and the potency of multiclass SVM in identifying speech disfluencies.

The "Study on automatic speech therapy system for patients" says use of automatic speech therapy devices to help individuals with speech impairments in their rehabilitation process has shown encouraging results [8]. The effectiveness of these systems, their components, assessment techniques, and the problems that need to be solved are all highlighted in this paper. It is a useful tool for academics, doctors, and programmers who are working to create and enhance autonomous speech therapy systems.

"Intelligent processing of stuttered speech" indicates improvements in screening, therapy, and support for those who stammer might be made with the use of intelligent speech processing [9]. The efficacy of intelligent processing approaches, their applications in voice recognition, natural language processing, machine learning, and augmented reality are highlighted in this paper, which offers a thorough overview of the status of the research in this field. Additionally, it shows the difficulties and potential paths for expanding technologically aided stuttering therapies.

The paper "Stuttered speech recognition using convolutional neural networks" focuses on how convolutional neural networks (CNNs) are used to identify speech that is stuttering [10]. This paper offers a summary of the relevant studies, emphasizing the creation and utility of CNNs for the identification of stuttering speech. It goes over the different approaches, data sets, and assessment criteria used in the investigations. It also examines the difficulties and potential approaches for enhancing stuttering speech recognition with CNNs.

The automated detection of stuttering in speech can be accomplished with the use of machine learning algorithms. The paper "Machine learning for stuttering identification: Review, challenges, and future directions" offers a thorough overview of the state of the art in this field of study, emphasizing the creation and use of machine learning models for stuttering diagnosis [11]. It covers methodology, features, assessment measures, and difficulties encountered. Additionally, it makes suggestions for future work on improving the detection of stuttering using machine learning, opening the door to better stutterer diagnosis, care, and support.

"Automatic speech recognition with stuttering speech removal using long short-term memory (LSTM)" ASR with stuttering speech removal utilizing LSTM networks has the potential to increase the accuracy and fluency of speech recognition for those who stammer [12]. The development and efficiency of LSTM-based models for the elimination of stuttering speech are highlighted in this paper, which offers a thorough overview of the status of the research in this field. It talks about the methodology, datasets, assessment criteria, and difficulties encountered. Additionally, it makes recommendations for future work on improving ASR through the use of LSTM networks to remove stuttering, opening the door for enhanced speech recognition and communication assistance for those who stammer.

"Stuttering detection using Atrous convolutional neural networks" has the potential to increase diagnostic precision and assistance for stutterers [13]. The progress and use of ACNN models for stuttering detection are highlighted in this paper, which offers a thorough overview of the status of the field today. The methodology, datasets, assessment measures, and difficulties are covered. It also makes recommendations for future work on improving stuttering detection using ACNNs, opening the door for better stutterer diagnosis and treatment planning.

"Introducing ECAPA-TDNN and Wav2Vec2.0 embeddings to stuttering detection" talks stuttering detection might be made more accurate and supportive to stutterers by including ECAPA-TDNN and Wav2Vec2.0 embeddings [14]. The creation and efficacy of these sophisticated models are highlighted in this paper, which offers a thorough overview of the status of research in this field. The methodology, datasets, assessment measures, and difficulties are covered. Furthermore, it offers future approaches for enhancing stuttering identification utilizing ECAPA-TDNN and Wav2Vec2.0 embeddings, opening the door for better stutterer diagnosis and treatment planning.

"KSoF: The Kassel State of Fluency Dataset—A Therapy Centered Dataset of Stuttering" as a stuttering-specific dataset, the Kassel State of Fluency Dataset (KSoF) is extremely valuable. A thorough description of the dataset, its creation, and its prospective uses are given in this paper [15]. It emphasizes the contributions and effects of the KSoF dataset on expanding our knowledge of stuttering and enhancing therapeutic results. It also discusses the dataset's restrictions and difficulties, and it makes recommendations for future development and application of the dataset in academic study and clinical practice.

3 Methodology

The first step of the method included the dataset "spe-28k" provided by Apple. This dataset consisted of podcast episodes from various sources, each episode lasting about an hour. To facilitate stuttering analysis and classification, the dataset was preprocessed by dividing the videos into shorter 3-s segments. In addition, clips were sampled at 16 kHz to improve audio processing.

To tag audio clips with different types of stuttering, the researchers used the tags and "spe-28k" dataset. This allowed them to assign appropriate labels to each 3-s clip based on the presence and characteristics of stuttering.

A preprocessing step was performed to improve sound quality and remove unwanted noise. We used the "noise reduction and a Python library for effectively muting audio clips. This denoising step was used to improve the accuracy of the subsequent analysis and classification steps.

After preprocessing steps, the 3-s audio clips were converted to waveforms (Figs. 1 and 3). An audio signal and its frequencies and amplitudes are plotted in a spectrogram (Fig. 2) to show how they change over time. It is a powerful tool for analyzing and visualizing audio data, as it provides in-depth information about audio qualities. The spectrogram provides a detailed overview of the frequencies present in the audio signal and their change over time. It is used to monitor stuttering patterns and abnormalities.

In the subsequent phase of the methodology, the spectrograms (Fig. 4) of the 3-s clips were categorized into two groups: stuttering and non-stuttering. In order to train a machine learning model, it was necessary to carry out this categorization process to prepare the data. In order to facilitate the learning process and differentiate between instances of stuttering and non-stuttering, the audio samples were classified into two distinct categories.

The classification challenge was addressed using a deep learning convolutional neural network (CNN) model. The use of CNNs is highly appropriate for the analysis of spectrograms due to their exceptional capabilities in image and pattern recognition tasks. In an effort to effectively identify and classify stuttering patterns (Fig. 5) occurring in podcast episodes, the researchers underwent training of the CNN model using classified spectrograms.

In the process of implementing the CNN model, numerous architectural elements were integrated to enhance its performance. In particular, the inclusion of max pooling and dropout layers was apparent. Max pooling is a technique used to decrease the dimensionality of feature maps, enabling the model to concentrate on crucial

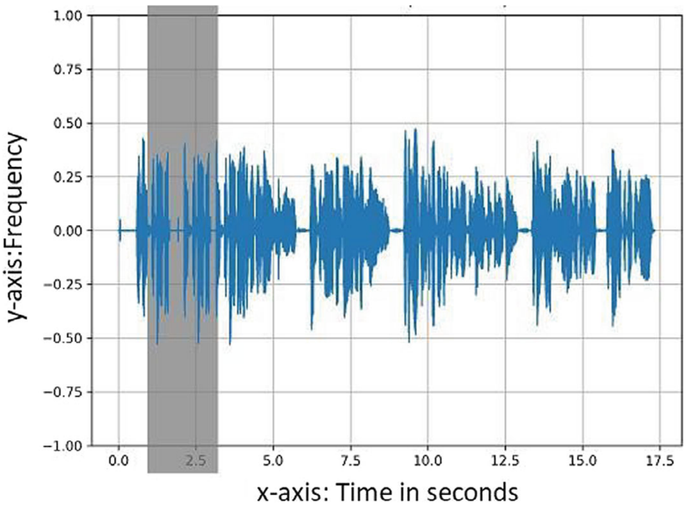


Fig. 1 Waveform

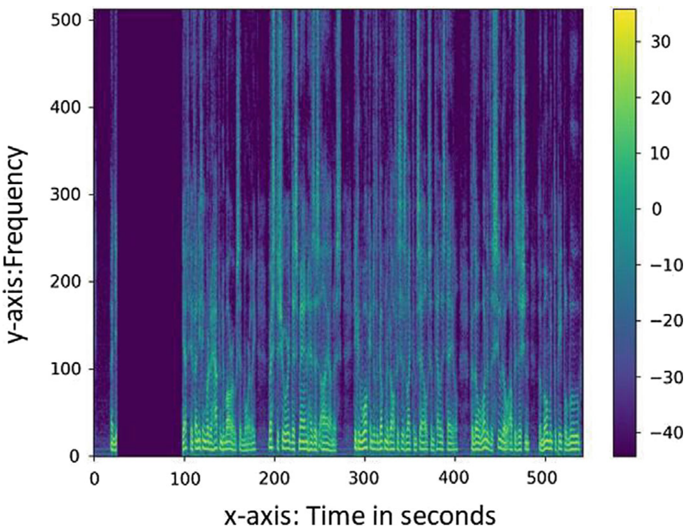


Fig. 2 Spectrogram

information. To address overfitting, dropout layers were introduced. These layers randomly eliminate units while training the model, thereby reducing reliance on specific features and facilitating generalization.

The training process involved passing the entire dataset through the CNN model for 10 epochs. In other words, the model underwent 10 cycles of going through the complete dataset. Following the completion of the training process, the model

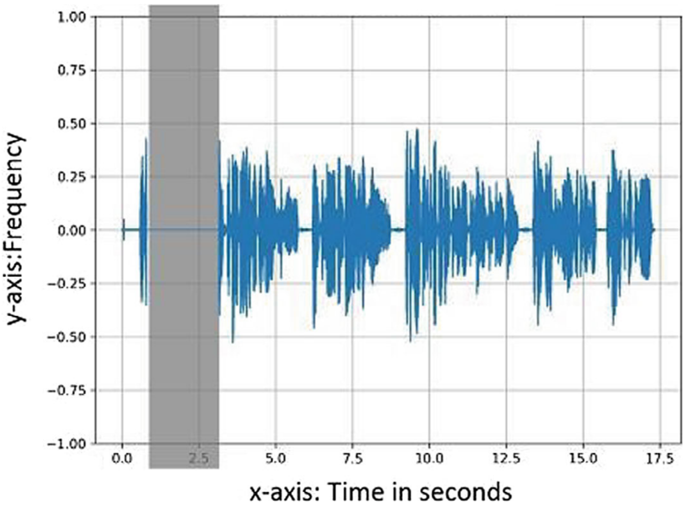


Fig. 3 Waveform

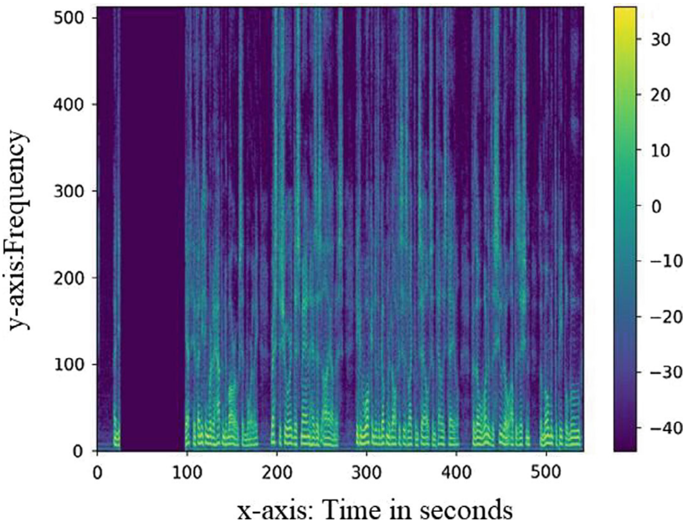


Fig. 4 Spectrogram

obtained an F_1 score and accuracy of 0.8. This outcome demonstrates its effectiveness in accurately distinguishing between instances of stuttering and non-stuttering (Fig. 6) throughout the podcast episodes (Fig. 7).

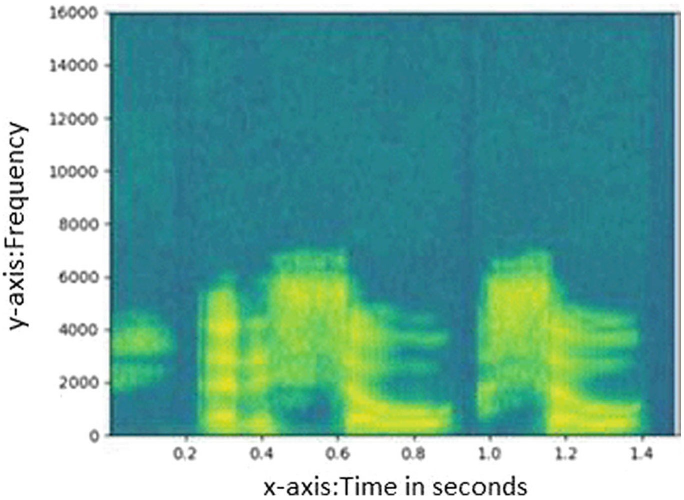


Fig. 5 Picture from stuttering folder

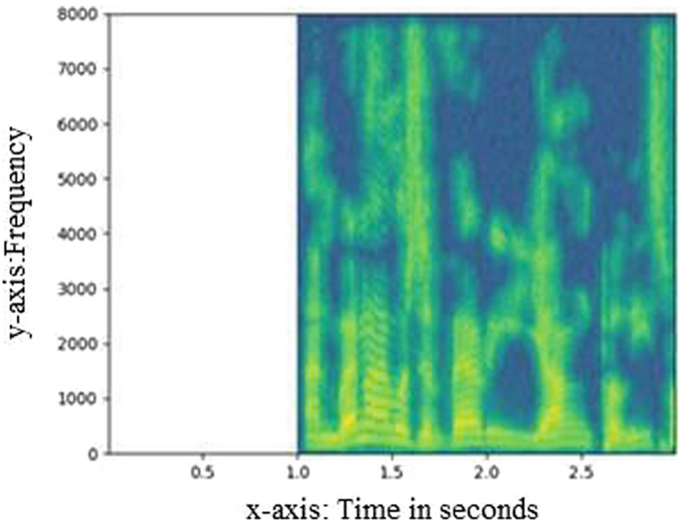


Fig. 6 Picture from non-stuttering folder

4 Results and Discussions

Talkify is a modern app designed to assist people who have trouble with their speech, like when they sometimes get stuck while talking. This app uses advanced technology to give users helpful tips and positive words right away, which makes them feel better and more self-assured in their communication skills.

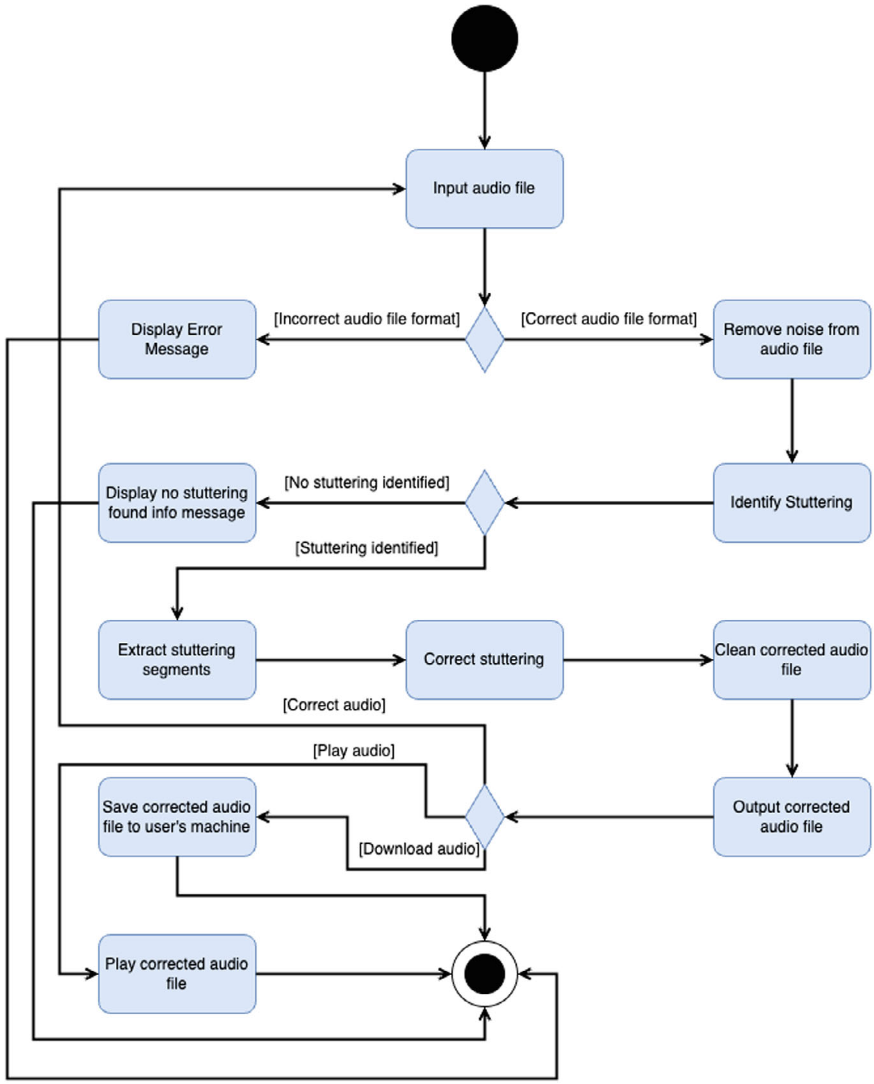


Fig. 7 Methodology

Think of epochs like how many times our model has learned from the data. Usually, as we train it more, the training accuracy goes up. But, after a while, the validation accuracy, shown in Fig. 8, might stop improving or even get worse. This happens because the model gets too good at the training data and struggles with new information. Accuracy is just the percentage of things our model gets right. So, a high accuracy means our model is doing well.

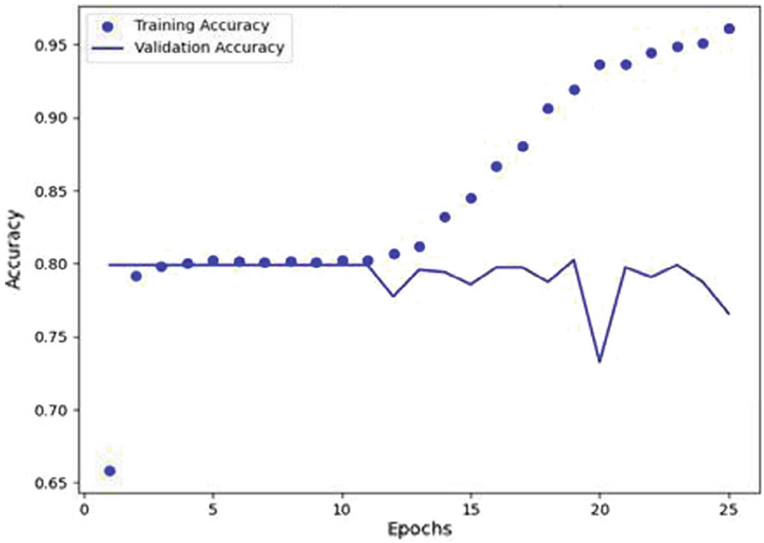


Fig. 8 Training and validation accuracy

We check how good our model is with something called the F_1 score, shown in Fig. 9. This score looks at both precision, which is how many times we correctly spot stuttering, and recall, which is how often we catch all the stuttering. If our F_1 score is high, it means our model is good at both. Just like before, epochs show how many times we have taught our model with data. Normally, as we teach it more, the training F_1 score gets better. But, after a while, the validation F_1 score might stop getting better or even get worse.

To figure out how well a stutter-detection model is doing, we can look at the training and validation loss graph. When these two lines are close together, it is a good sign. It means the model is working well with new data. However, if the training loss line (as shown in Fig. 10) is much lower than the validation loss line, it might mean the model is too focused on the training data and not as good with new information. This is called "overfitting", and it is not what we want.

When we calculate the precision score, which tells us how often the model predicts positive outcomes correctly, the training precision curve shows how well the model does on the training data. The validation precision curve, as shown in Fig. 11, tells us how well the model does on data it has not encountered before.

Think of the training and validation recall graph, like the one in Fig. 12, as a picture showing how good a model is at spotting stuttering as it learns. This graph usually has two lines: one for training recall and one for validation recall. The training recall line shows how well the model does on the training data it has been taught with, and the validation recall line shows how well it does with new data it has not seen before.

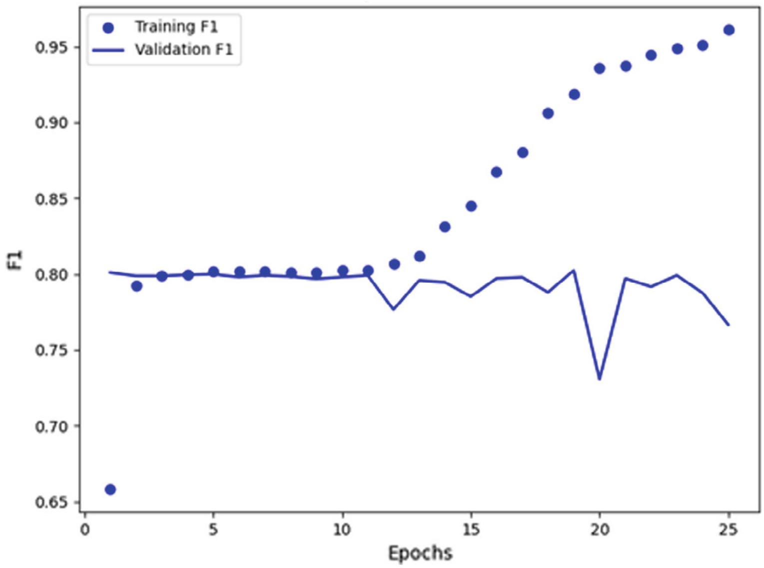


Fig. 9 Training and validation F_1

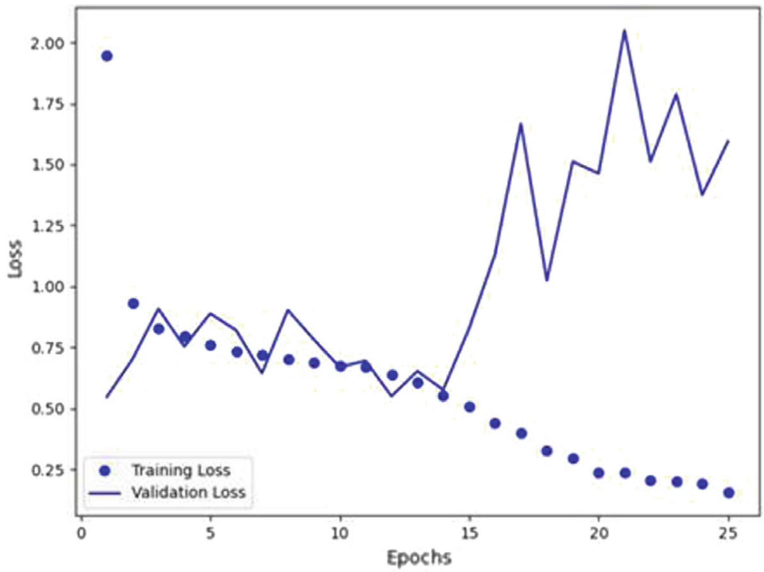


Fig. 10 Training and validation loss

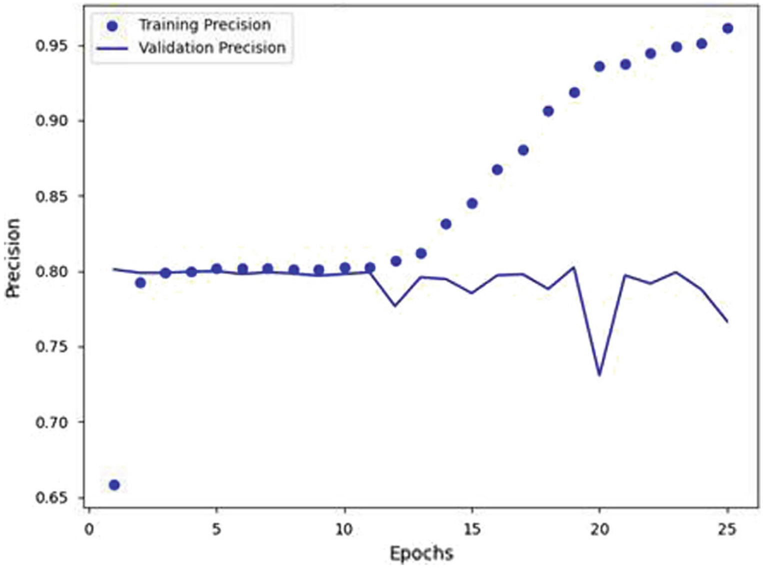


Fig. 11 Training and validation precision

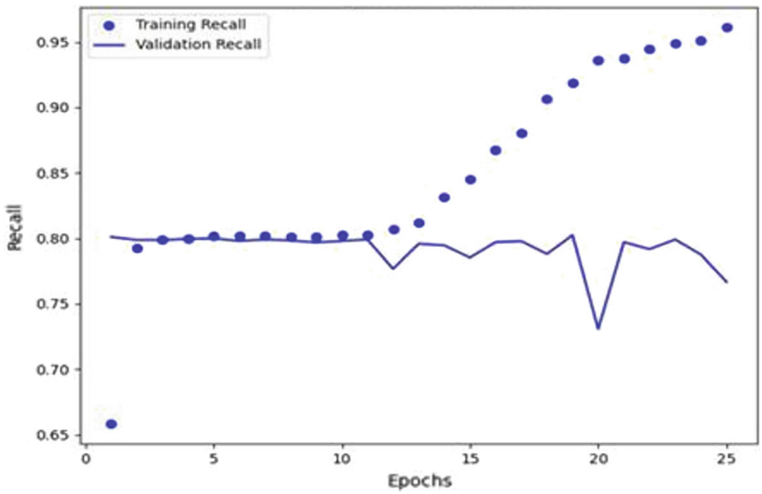


Fig. 12 Training and validation recall

Users of Talkify have shared that their speech has gotten a lot better thanks to the app. The app does this by using different techniques, like delayed auditory feedback, which has been proven to help reduce stuttering.

People who used Talkify mentioned feeling more self-assured. The app helps users manage their stuttering and reduces anxiety related to speaking.

People can gain a better understanding of how they speak and work on improving their fluency by looking at and analyzing their speech data using visual tools.

This tool adapts to how each person talks, what they prefer, and how they are progressing. It gives personalized advice that matches their unique goals and challenges.

In addition to traditional speech therapy, Talkify is often used as an additional tool. By integrating Talkify into therapy sessions, clinicians can better track patients' progress, provide targeted treatments, and increase the overall effectiveness of treatment for patients with stuttering issues.

While Talkify has shown positive results, it is important to know that each person might react differently to it. To make tools like Talkify even better, we need to keep researching and developing them, especially in the area of speech impairments.

Talkify has proven to be a helpful tool for people dealing with stuttering. It assists individuals in overcoming the challenges of stuttering by giving them personalized guidance, boosting their confidence, and improving their ability to speak fluently.

5 Future Scope

The Talkify project has a lot of future growth and integration potential. A stutter removal feature would be a big step forward. Talkify could use improved signal processing and machine learning algorithms for speech analysis and adaptation in real time to reduce or eliminate stuttering. This integration would be greatly improved supports communication and enables people who stutter to speak more fluently and confidently. Integration with popular voice assistants like Siri and other voices associated software would ensure a smooth user experience and wide availability. People stuttering, you can easily access Talkify features by voice orders. Speech analysis and aids would be more convenient for people with stuttering.

Talkify would also be more suitable for speech therapy and personal speech evaluation because it would be able to monitor and give feedback in real time. Integration with telehealth platforms would expand the reach and enable of Talkify speech language pathologists and therapists for remote monitoring and analysis of patient speech patterns and provide guidance and monitor progress. Last but not least, adding language support to the existing training suite would enable the use of Talkify more users around the world, increasing its global reach and improving their lives people who stutter around the world.

6 Conclusion

Talkify is a smart tool. It will change how we help people who have trouble speaking. It learned a lot from many people who stuttered when they talked. And guess what? It is really good at spotting and understanding stuttering. It can even tell when someone is stuttering or talking smoothly using some fancy tricks.

Talkify can do many useful things. It can work with speech therapy to check how well you are doing. It is also handy for checking how you talk and finding what makes you stutter. What is more, it can work with telehealth platforms, so you can get help with your speech over the Internet if you have a stutter.

Talkify represents a significant advancement in the quality of life for stutterers generally and in terms of communication support. Talkify gives people a tool for effective stuttering treatment and an objective, automated method to identify and analyze stuttering, empowering both individuals and professionals and involvement. Talkify's capabilities will continue to be expanded and improved through additional research and development. Talkify offers promising prospects for enhancing communication abilities and the general well-being of stuttering patients. It has the potential to revolutionize the treatment of stuttering.

References

1. Abo Ras YA, El-Maghraby RM, Madkour WM (2015) Outcome of stuttering therapy on Egyptian school-aged children using the speak freely program. Egypt J Otolaryngol
2. Dash A, Subramani N, Manjunath T, Yaragarala V, Tripathi S (2018) Speech recognition and correction of a stuttered speech. IEEE
3. Kourkounakis T, Hajavi A, Etemad A (2020) FluentNet: end-to-end detection of speech disfluency with deep learning. arXiv.org
4. Mishra N, Gupta A, Vathana D (2021) Optimizaton of stammering in speech recognition application. Int J Speech Technol
5. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. IEEE
6. Vinod DS, Mahesha P (2015) Support vector machine-based stuttering disfluency classification using GMM supervectors. Int J Grid Util Comput
7. Mahesha P, Vinod DS (2013) Classification of speech disfluencies using speech parameterization techniques and multiclass SVM. Springer, Singapore
8. Rao SB, Hegde S, Shetty S (2021) Study on automatic speech therapy system for patients. In: Advances in artificial intelligence and data engineering. Springer, Singapore
9. Czyzewski A, Kaczmarek A, Kostek B (2003) Intelligent processing of stuttered speech. J Intell Inf Syst
10. Bhushan SP, Vani HY, Shivkumar DK, Sreeraksha MR (2021) Stuttered speech recognition using convolutional neural networks. Int J Recent Technol Eng (IJRTE)
11. Sheikha SA, Sahidullaha Md, Hirschb F, Ounia S (2021) Machine learning for stuttering identification: review, challenges and future directions. arXiv.org
12. Girirajan S, Sangeetha R, Preethi T, Chinnappa A (2022) Automatic speech recognition with stuttering speech removal using long short-term memory (LSTM). Int J Recent Technol Eng (IJRTE)

13. Al-Banna AK, Edirisinghe E, Fang H (2022) Stuttering detection using atrous convolutional neural networks. IEEE
14. Sheikh SA, Sahidullah Md, Hirsch F, Ouni S (2022) Introducing ECAPA-TDNN and Wav2Vec2.0 embeddings to stuttering detection
15. Bayerl SP, von Gudenberg AW, Hönig F, Nöth E, Riedhammer K (2022) KSoF: the Kassel state of fluency dataset—a therapy centered dataset of stuttering

Assessment of a Significantly Better Aquatic Underwater Image Improvement Through the BBHE Algorithm

Swarna Venkatesh, K. Ranjith Reddy, Mahesh Kotha, B. Suresh Ram, M. Nagaraju Naik, and Vijaya Kumar Koppula

Abstract In this article, we present a useful technique for improving underwater photos that have been damaged by medium absorbing and scattering. Our solution uses a single image technique and lacks technological tools, understanding of the aquatic environment, or scene organization. It expands on the merging of two pictures that are directly derived from a description of the original deteriorated image that has been color-adjusted and white-balanced. In order to encourage the incorporation of edges along color contrast to the output image, both images to fused and their corresponding weight charts are created. Images taken in murky aquatic environments have weak contrast and limited clarity. Therefore, it is necessary to enhance the images obtained from aquatic research and examination efforts so that it may be processed to get the desired information. In this study, a brand-new technique for improving contrast dubbed "brightness-preserving bi-histogram equalization" (BBHE) is put forward. CLAHE significantly contributes to the enhancement of brightness in aquatic photos.

S. Venkatesh
Department of ECE, CMR Technical Campus, Hyderabad, Telangana, India

K. Ranjith Reddy (✉)
Department of CSE, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: cmrtc.paper@gmail.com

M. Kotha
Department of CSE (AI&ML), CMR Technical Campus, Hyderabad, Telangana, India

B. Suresh Ram · M. Nagaraju Naik
Department of ECE, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India
e-mail: bsureshram@cmrcet.ac.in

M. Nagaraju Naik
e-mail: nagarajunaik1976@cmrcet.ac.in

V. K. Koppula
Department of CSE, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India
e-mail: dean.academics@cmrcet.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_16

Keywords Scattering · Absorption · Color contrast · BBHE · CLAHE

1 Introduction

Another of the main subfields of digital image processing that finds application in many various domains is aquatic image processing. Such involves the observation of maritime environments. Additionally, it makes engineering piping inspection simpler. The physical characteristics of the undersea environment make underwater imaging a highly difficult field. Primarily having to do with how light is reflected and absorbed. Underwater photos lose contrast and degrade mostly as a result of poor visibility circumstances and factors including the absorption, contemplation, bending, and dispersion of light.

Due to suspended air particles, images taken in outdoor settings typically have poor visibility, which negatively impacts the pictures. Despite recent advancements in image dehazing techniques, more work needs to be done to enhance the images' visual quality [1]; it proposes a deep convolutional neural network (CNN) called PDR-Net for single image dehazing. PDR-Net consists of a haze removal subnetwork inspired by perception that reconstructs the latent dehazed image and a refinement subnetwork that further improves the contrast and color properties of the dehazed result through joint multi-term loss optimization. The advantages of the existing indoor and outdoor image dehazing training data are combined in our method.

2 Digital Image Processing

With the help of computer software, you may edit digital photographs thanks to the development of digital image processing. It is a branch of signal processing that concentrates mainly on visual data. The ability to perform various algorithms on a digital image as an input using digital image processing enables the stoner to provoke an affair. Depending on the request for a fair image, these techniques may change form image to image.

3 Image Enhancement

Mainly of the simplest sorts of image modification tools only alter an image's contrast or excellence or play with its grayscale or red–green–blue color patterns. Certain kinds of basic pollutants also permit converting color images to black and white, photograph-tone images, or the addition of visual elements.

4 Existing Methods

4.1 Histogram Equalization

Whenever the information of the image is expressed by contrast values, histogram equalization typically boosts the overall contrast of photographs. Another histogram can more evenly disperse the intensities. The most frequent values for intensity are properly spread out to achieve equalization of the histogram [2].

Images containing bright or dark background and centers can benefit from using this technique. The system's relatively easy design and invertible operators are two of its key advantages. Yet, if the equalization of the histogram procedure exists, the original histogram may also be restored. The magpie is a drawback of the method. It might reduce the amount of usable signal while enhancing the contrast because of noise in the background.

In order to learn about and comprehend the underwater world, underwater photographs are crucial. The dependability of underwater intelligent systems can be ensured by high-quality underwater photographs. Unfortunately, low contrast, color casts, blurring, low light, and uneven illumination are common in underwater photographs, which has a negative impact on how underwater information is perceived and processed [3]. Numerous strategies have been put forth to enhance the quality of underwater picture acquisition, particularly with the development of deep learning technologies.

Jiao [4] and Dimlo et al. [5] for resource detection, the quality of underwater photographs is a significant issue. The quality of underwater photographs can be impacted by light scattering and plankton in the water, though. This work proposes a novel method for underwater image restoration based on heat exchange optimization and non-convex, non-smooth variation.

4.2 Adaptive Histogram Equalization

Utilizing psychophysical bystander research, the use of adaptive histogram equalization (AHE) and brightness windowing were contrasted. Clinical CT (motorized tomographic) photographs of the sarcophagus have been shown to tolerant radiologists. The croakers were additionally shown the photos that had been reprocessed using AHE and brightness windowing. In some of the images, appropriate artificial lesions had been added. Their delicacy was evaluated when they were tasked to determine the likelihood that a specific image included the fake lesion. The findings of these tests reveal that there was indeed not a significant distinction between the two styles' abilities to display luminance contrasts for this specific individual task, and hence, further analysis of AHE utilizing supervised clinical trials is recommended.

5 Literature Review

The multi-scale adversarial networks for underwater image restoration developed by Bing Zheng [6] in 2019 depends on an adaptive underwater restoration of images system that makes use of cycle GAN and DCP. However, this representation is unable to provide a convincing image when the light source is not uniform.

A research study on differential gray-levels histogram equalization (DHE) and adaptive gray world (AGW) was conducted in 2018 by author Siaw-Lang [7]. Both of these approaches were effective. Utilizing the DHE methodology, which manipulates the intensity components of the raw underwater image, the AGW method removes the color cast from the image and enhances contrast. The undersea image's quality is considerably improved by this parallel arrangement. The DHE produces the best results for images with low contrast, but it may produce subpar results for images with high contrast.

Miao Yang's [3] deepwater underwater image reconstruction employing reflection-decomposition-based transmission map is expected to be completed in 2020. The calculation was centered on local backscattering lightening estimating and a dark channel reflection-illumination decomposition to improve clarity. On some things in the photos, however, it only generated mediocre quality restoration.

The output of the suggested technique is not the best, according to Qingliang Jiao's [4] underwater image restoration, which suggested a method for adaptively choosing regularization settings utilizing BRISQUE, FADE, and NIQE. The suggested approach can still be regarded as a superior underwater image restoration technique than other options when the restored images are combined.

In 2018, Codruta O's [8], a useful method to improve underwater photography, worsened as a result of medium dispersion and absorption. Our solution uses a single image and does not call for any specialist equipment or expertise in underwater circumstances or scene organization. It is based on the merging of two images that were produced by taking the original impaired image and applying color correction and white balance to it. To encourage the transfer of edges as well as color contrast within the output image, the two images to fusion and their related weight maps were determined. We also use a multiscale fusion approach to prevent the sharp weight map transitions from introducing artifacts in the low-frequency parts of the image that was reconstructed. Our in-depth analysis, both qualitative and quantitative, demonstrates that the upgraded photographs and videos are distinguished by increased global contrast, sharper edges, and better exposure of the dark areas.

Light always scattered and absorbed as it goes from the item to the camera in an underwater setting, according to Bowen Yao's [9] research from the year 2018; this has a significant impact on image quality. As a consequence, underwater images have less contrast and appear to be obscured by a haze. This work presents a modified dark channel prior (DCP)-based approach for image restoration. Initially the disparity among the blue and red channels is used to measure ambient light. Each of the three RGB channels' attenuations are subsequently determined individually. The remaining color distortion is then corrected using a color collection. The final studies show how effective this technique is at enhancing the clarity of underwater photos.

To improve the visibility of edge detail and the vibrancy of the output image, Miao Yang's [10] innovative underwater image restoration model has been shown for the year 2020. It depends on the dark channel's reflection-illumination breakdown and localized backscattering lighting prediction. Introducing statistical methods precondition to the offshore absorption coefficients creates an entirely new underwater picture restoration technique for turbid marine environments. It is demonstrated that the predicted transmitting map by lighting-reflection deconstruction offers greater clarity and color restoration over the other approaches, which divide the dark channel by the greatest backscattering light. The suggested approach outperforms modern algorithms on images obtained from offshore water that is distinguished by an elevated level of colored dispersed organic matter and total suspended material and is appropriate for quick underwater processing, according to a comprehensive qualitative and quantitative evaluation on hundreds of pictures taken underwater.

2020 will see Chongyi Li [11] arrive. Images and recordings of underwater sceneries become less visible due to wavelength-dependent extinction and dispersion. The quality of pattern recognition, visual comprehension, and essential extraction of features in underwater settings is impacted by the degraded underwater photos and movies. In this study, we present the UWCNN, an aquatic scene prior-based convolutional neural network (CNN) model for underwater picture improvement. The suggested UWCNN model directly recreates the clear hidden image beneath the water rather than guessing the underwater imaging model's parameters, taking employ of the recent underwater scene that may be employed to create training data for underwater image synthesis. Additionally, our UWCNN model may be quickly expanded to underwater films for frame-by-frame improvement depending on the lightweight network framework and efficient training data. We particularly provide underwater picture deterioration datasets that encompass a variety of water-related types and deterioration levels by fusing an underwater photography physical representation using optical characteristics of underwater sceneries. Then, using the appropriate training data, a lightweight CNN model is created for improving each type of underwater image.

According to Chunle Guo [1], 2020 considering recent advancements in image dehazing techniques, more work remains to be done to enhance the images' visual quality. In our research, we suggest a deep convolutional neural network (CNN) identified as PDR-Net for single image dehazing. PDR-Net consists of a haze elimination subnetwork prompted by perception that reassembles the latent dehazed photo and an enhancement subnetwork that further improves the brightness and color real estate of the dehazed result through joint multi-term reduction in optimization. The benefits of the current indoor and outdoor picture dehazing training data are combined in our approach, which, as contrasted to the prior methods, makes the suggested PDR-Net generalizable to diverse hazy photos and useful for enhancing the perceived quality of the dehazed outcomes.

6 Proposed Methodology

Matrix Laboratory is referred known as MATLAB. Deep reading and machine the ability to read and signal processing and communication, image and videotape processing, control systems, test and dimension, computational finance, and computational biology are just a few of the operations that may be performed in both academia and industry with MATLAB.

With the system of control in charge of issuing orders, administering, and controlling the characteristics of the other system, MATLAB offers control to the color systems and bias.

Wireless transmission: It is the most widely used phrase to describe the connection between two biases through wireless signals.

The impacts of the Internet: The network of automobiles, prejudice, house furnishings, and additional electronic bedded prejudice is what allows the prejudice to affect the information. Benefits of MATLAB: It provides the fastest IDE for calculating matrices and using direct algebra.

7 Block Diagram

BBHE: Most of the popular techniques for boosting the differentiation of provided images is histogram equalization (HE), which results in a performance image with a constant distribution depending on the slate conditions. It improves overall contrast by flattening and extending the picture's histogram's range of brightness. While a picture needs to be improved, HE can be used extensively, similar to medical image improvement. Although the "asked" histogram is flat, the average brightness of the proposition's relationship image remains in the center of the input means. In certain procedures whereby brilliance preservation is required, this characteristic is not desirable. To solve that issue, brightness-preserving bi-histogram equalization (BBHE) was originally suggested as shown in Figs. 1 and 2.

As shown in Fig. 3 to address the aforementioned issues with the standard histogram equalization, an additional version of the histogram equalization called the mean brightness-preserving bi-histogram equalization (BBHE) is developed. The main goal of the suggested approach is to preserve the image's mean brightness as well as saving brightness.

The image with brightness restoration is displayed below:

As shown in Fig. 4, the input image's histogram is first divided in half by its average by BBHE, resulting in two histogram ranges that do not overlap. It then separately equalizes the two sub-histograms. It was analyzed to show that if the input histogram displays a quasi-symmetrical dispersion around its mean, BBHE can retain its initial brightness to an extent.

As shown in Fig. 5, the minimum mean image is deliberated that is a MEAN 0 or black gray-level image.

Fig. 1 BBHE block diagram

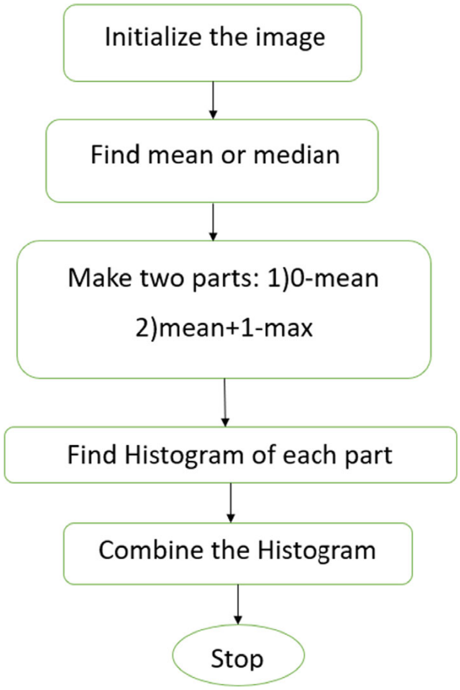


Fig. 2 CLAHE block diagram

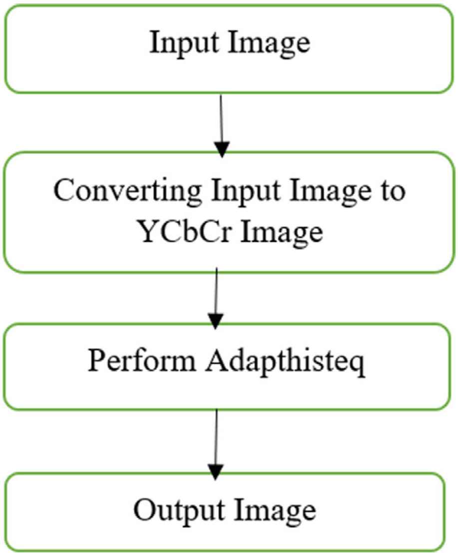


Fig. 3 Input image for BBHE technique

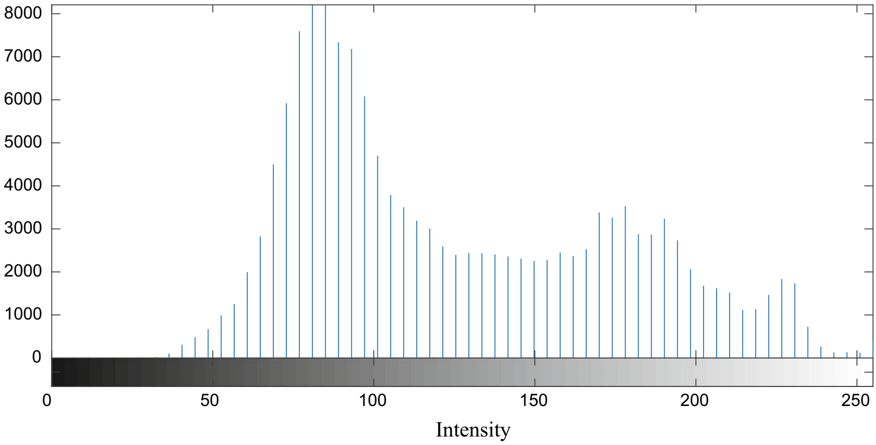


Fig. 4 Histogram of input BBHE image

Fig. 5 Minimum mean image

Fig. 6 Maximum mean image

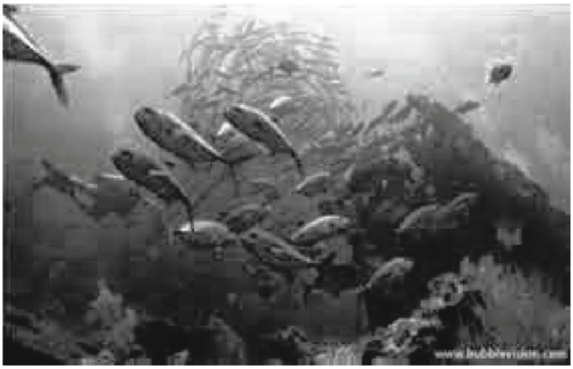


Fig. 7 Brightness preserved image

As shown in Fig. 6, the maximum image is deliberated with MEAN 256 or white gray-level image.

As shown in Fig. 7, one picture is created by combining the two histogram-equalized photos. Thus, it is evident that brilliance was successfully maintained.

The histogram of the beyond image is:

As shown in Fig. 8 may be seen in the output image's histogram, the histogram is evenly distributed among 0–250.

The PSNR, SNR, MSE, and SSIM values are calculated for input and output images.

The Peak-SNR value is 25.7917

The SNR value is 20.4330

The mean-squared error is 171.3614

The SSIM value is 0.9756.

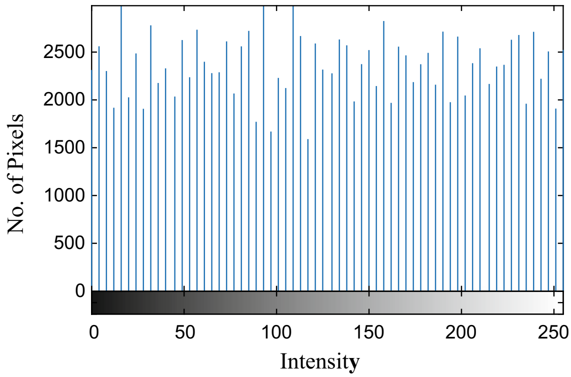


Fig. 8 Histogram of brightness preserved image

8 CLAHE

Although the global histogram equalization process amplifies noise in relatively homogeneous regions, it may worsen the focal loss issues in the dark regions of the image while solving the over-enhancement issue. Brightness limited adaptive histogram equalization (CLAHE) was suggested as a solution to this issue. CLAHE is a recognized block-based processing technique that uses basic histogram equalization to get around noise issues in the homogenous region of a picture. As shown in Fig. 9, the CLAHE technique is unique to regular HE in that it redistributes the image's brightness values using small portions of the image known penstocks. It does this by computing multiple histograms, each of which corresponds to a different section of the image.

The image under initializes the input picture:

As shown in Fig. 10, the input images have to go through RGB to YCbCr or gray-level image.

The histogram of the above image is:

Fig. 9 Input image for CLAHE method

Fig. 10 RGB to YCbCr converted image

As shown in Fig. 11, the YCbCr image, an adaptive equalization of the histogram, is carried out. By breaking the images into smaller data sections known as penstocks instead of processing the entire image and applying contrasting improvement, CLAHE eliminates over-amplification. Additionally, these pen stocks are returned to obtain the final contrast-enhanced image.

See in Fig. 12 the homogeneous images, AHE has a tendency to over amplify noise. This issue is solved by CLAHE by restricting noise modification.

The histogram of the above image is:

Employing the "adapthisteq" MATLAB functioning, the YCbCr picture is subjected to adaptive equalization of the histogram as shown in Fig. 13.

The PSNR, SNR, MSE, and SSIM values are calculated for input and output images.

Fig. 11 Histogram of YCbCr converted image

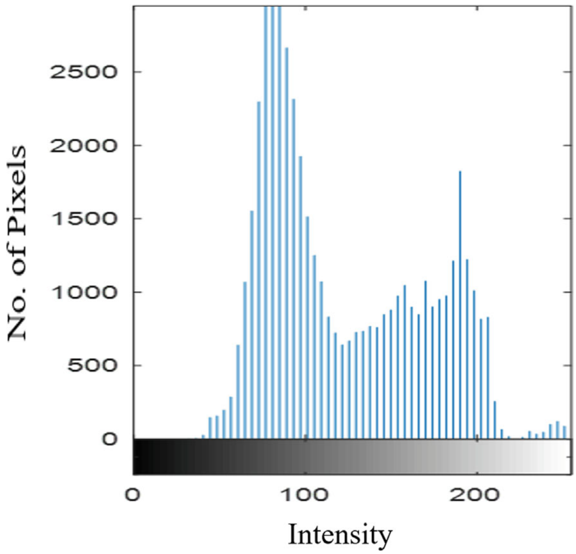


Fig. 12 Contrast-enhanced image

The peak-SNR value is 18.9768

The SNR value is 13.1329

The mean-squared error is 822.9966

The SSIM value is 0.7134.

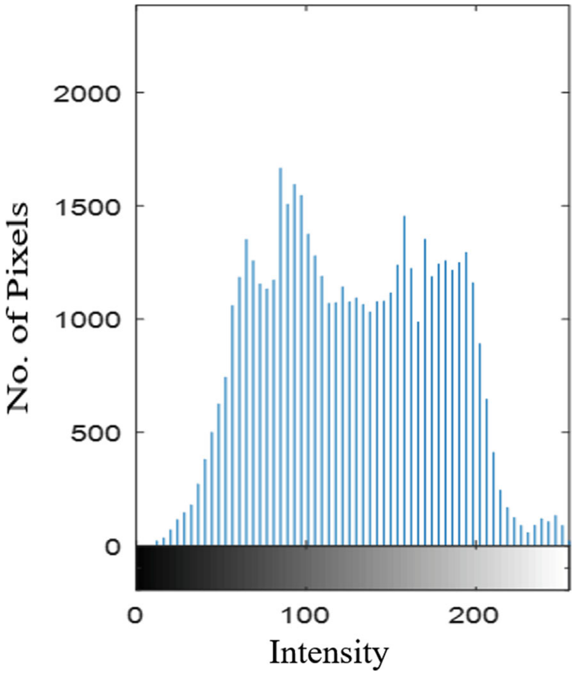


Fig. 13 Histogram of contrast enhanced image

9 Conclusion

BBHE and CLAHE were used for enhancing the underwater image's brightness and contrast to that, correspondingly. The image that is fuzzy, low contrast, low light, and foggy has been improved. The outcome we obtained is superior to HE. To address the aforementioned issues with the standard histogram equalization, an updated version of the histogram equalization called the average brightness-preserving bi-histogram equalization (BBHE) is developed. The main goal of the suggested approach is to maintain the image's average brightness despite preserving contrast. The global histogram equalization amplifies noise in fairly uniform parts, which causes the over-enhancement issue and point erosion issue in the darker parts of the image to persist. This can be fixed by the histogram equalization. Contrast limited adaptive histogram equalization (CLAHE) has been suggested as a solution to this issue.

10 Future Work

Although we have improved the photographs, the outcome is still unclear. Although there is significant noise, CLAHE does improve contrast. This is going to be addressed in subsequent work by altering the methodologies.

References

1. Guo C, Guo J, Han P, Fu H, Cong R (2020) Perception inspired single image dehazing network with refinement. IEEE
2. Bhaskar N, Ganashree TS, Patra RK (2023) Pulmonary lung nodule detection and classification through image enhancement and deep learning. Int J Biom 15(3–4):291–313. https://doi.org/10.1504/IJBM.2023.130637
3. Yang M, Hu K, Du Y, Wei Z, Sheng Z, Hu J (2020) Underwater image enhancement based on conditional generative adversarial network. IEEE
4. Jiao Q (2021) Underwater image restoration. IEEE
5. Dimlo UMF, Narasimharao J, Laxmaiah B, Rani DS, Kumar VN, Srinath E, Sandhyarani (2023) An improved blind deconvolution for restoration of blurred images using ringing removal processing. In: Reddy KA, Devi BR, George B, Raju KS, Sellathurai M (eds) Proceedings of fourth international conference on computer and communication technologies. Lecture notes in networks and systems, vol 606. Springer, Singapore. https://doi.org/10.1007/978-981-19-8563-8_34.nics, https://doi.org/10.37391/IJEER.100218
6. Zheng B (2019) Multi-scale adversarial network for underwater image restoration. IEEE
7. Siaw-Lang (2018) Differential gray-levels histogram equalization and adaptive gray world. IEEE
8. Ancuti CO, Ancuti C, De-Vleeschouwer C, Bekaert P (2018) Color balance and fusion for underwater image enhancement. IEEE
9. Yao B (2018) The dark channel prior method, as the direct application of DCP on the underwater image. IEEE
10. Yang M (2020) Offshore underwater image restoration using reflection decomposition based transmission map estimation. IEEE

11. Li C, Anwar S, Porikli F (2020) Underwater scene prior inspired deep underwater image and video enhancement. IEEE
12. Simon E (2018) Dehazing methods for underwater images. IEEE
13. Sharma (2019) Unique underwater enhancement using the discrete cosine transform. IEEE
14. Li Y et al (2019) Underwater degraded image. IEEE

Enhancing Mental Well-Being Through OpenBCI: An Intelligent Approach to Stress Measurement

Shrivatsa D. Perur and Harish H. Kenchannava

Abstract The COVID-19 pandemic has profoundly disrupted daily life, leaving enduring challenges in its wake, particularly in the realm of mental health. The pandemic's impact on individuals has been marked by substantial hardships and emotional distress, resulting in a spectrum of psychological strains. Consequently, there exists an imperative for research to address these mental health issues effectively. In 2013, Conor Russomanno and Joel Murphy introduced the OpenBCI, a revolutionary brain–computer interface (BCI) device characterized by its affordability, portability, and user-friendliness, representing a notable milestone in the field of brain computing. This study is dedicated to harnessing the potential of OpenBCI for the identification of adverse human emotions. The adoption of OpenBCI offers the promise of convenient stress measurement within the comfort of individuals' homes, fostering the development of cost-effective stress monitoring tools for neurological practitioners. This research delves into the challenges encountered during the initial phases of data collection and analysis, elucidating the methodologies employed to prepossess the acquired raw data. Furthermore, the study presents a comprehensive inventory of machine learning (ML) models deemed suitable for the task, offering insights into the outcomes achieved by these models when applied to the datasets.

Keywords Intelligent computing · Prediction of stress · Machine learning algorithms · Mental health · Personalized interventions

S. D. Perur (✉) · H. H. Kenchannava
Department of Information Science and Engineering, KLS Gogte Institute of Technology, Belagavi, Karnataka 590008, India
e-mail: perur35@gmail.com

Visveswaraya Technological University, Belagavi 590018, India

H. H. Kenchannava
e-mail: harishhk@git.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_17

1 Introduction

Today, stress is a constant in practically everyone's life. Stress affects people on a daily basis. Stress is described as "the body's non-specific reaction to any need for change." It could be a physical or emotional reaction to a certain circumstance. It has an impact on all living things, both psychologically and physically. Inevitably, there is not any way to completely escape stress, and therefore, we must learn to cope. The wise thing we are able to do is keep an eye on our stress levels and look for patterns so that we can manage stress [1]. Analysing our brain's EEG data is one method we might utilize to anticipate stress [2]. Although it might not seem like an easy alternative, technology is advancing and giving us strong solutions to every problem. One such gadget that is used to record a person's EEG is shown in Fig. 1; these gadgets are also known as brain computing interfaces.

The term "brain–computer interface" is an abbreviation. Simply described, it involves a brain and a computer working together. With the help of these gadgets, brain impulses can now control some exterior actions, including moving a cursor or a prosthetic limb. BCI enables a person with paralysis to write a book or operate a motorized wheelchair or prosthetic limb by using their thoughts alone. It does this by reading signals from a variety of neurons and utilizing computer chips and programs to interpret the signals into movement. One of these tools is made available by OpenBCI. It specializes in developing high-performance, reasonably priced brain–computer interfaces and biosensor technology. In more than 80 nations, gadgets have been employed as BCIs to operate devices and map neural activity by academics, producers, and enthusiasts. The OpenBCI Ultracortex Mark IV, an open-source, 3D printed headset, is the tool utilized in the experiment as depicted in Fig. 1. It has an 8-channel Cyton board for biosensing that can capture research-grade EEG as shown in Fig. 2.

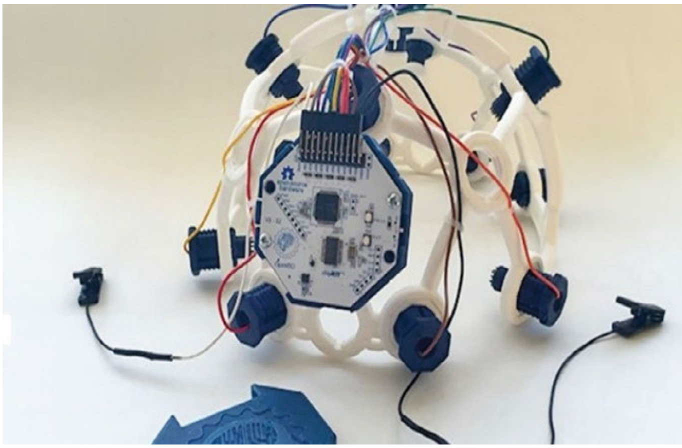


Fig. 1 3D printed headset, mounted with electrodes and the 8-channel Cyton board

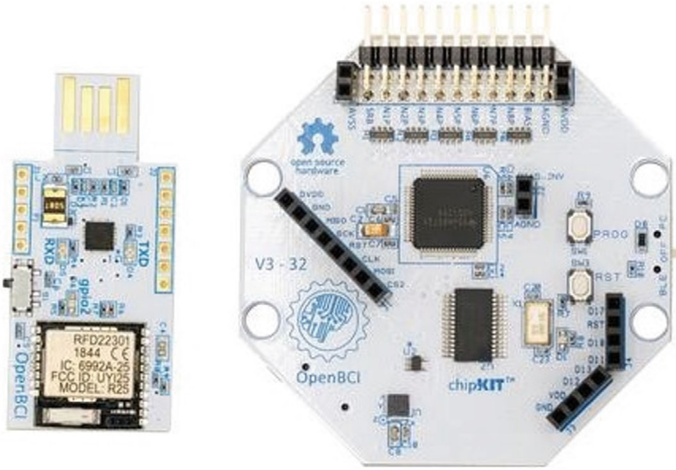


Fig. 2 Programmable dongle for Bluetooth connectivity and Cyton biosensing board 8 channel

2 Related Work

The survey aims to understand the benefits and outcomes associated with each technique, providing insights into their effectiveness for promoting relaxation, reducing stress, and improving mental well-being. By reviewing existing literature, this survey seeks to contribute to the perception of meditation practices and their potential impact on individuals' physical and mental health. Table 1 shows the related works in the domain of stress analysis.

3 Experimental Methodology

The most crucial and fundamental component of every data science effort is data. Both number and quality are crucial aspects of data. As a result, a lot of work and research went into creating an experiment to gather information for analysis. According to the device's computing and storage capabilities, the experiment was created. Figure 3 displays the experiment's whole flow diagram. The experiment was conducted in a quiet, well-ventilated space. The subject's values for blood pressure (BP), heart rate (HR), and EEG data, together with the subject's occupation, were the key parameters taken into account. Based on the subjects' ages and vocations, we divided them into three categories. The three age groups are 10–25, 26–40, and 41–60. Most commonly, these three age groups experience varying levels and forms of stress. For instance, those between the ages of 10 and 25 are in the process of establishing their schooling and job, and they frequently deal with peer pressure and academic stress. In a similar vein, those between the ages of 26 and 40 are

Table 1 Related works for analysis of stress

Paper	Methods	Conclusions	Results
[3]	EEG data acquisition under different conditions Preprocessing with band pass filter and wavelet transformation	10,159 students committed suicide in India in 2018 Ensemble classifier obtained highest classification accuracy of 87%	Experimental research done on stress among college students Ensemble classifier achieved highest classification accuracy of 87%
[4]	Biomedical signal analysis using EEG data Support vector machine (SVM) classifier for classification	EEG signals used to measure and monitor stress SVM classifier used to improve stress detection accuracy	Various stress detection models proposed using biomedical signals Support vector machine classifier improves stress detection accuracy
[5]	Support vector machine (SVM) Two-layer LSTM RNN deep learning classifier	Participants self-reported higher levels of mental stress following stress induction periods Post-stressor stress scores were significantly higher than pre-stressor scores	LSTM RNN classifier can identify mental stress with 93.27% accuracy EEG-based BCI shows promise for real-time stress assessment and mitigation
[6]	Preprocessing and feature extraction of EEG signals Support vector machine (SVM) classification method	EEG neurofeedback-based classification approach is efficient Highest classification accuracies: 92.48 and 88%	EEG neurofeedback-based classification is efficient for detecting anxiety disorder states The classification accuracy for healthy subjects is higher than for anxiety subjects
[7]	EEG signals recorded from 10 subjects Fractal dimension feature extracted from raw EEG	Investigation on impact of positive and negative emotions Average accuracy of 60% in emotion recognition	Emotions can be recognized from EEG signals Support vector machine (SVM) with RBF kernel achieved an average accuracy of 60%
[8]	Linear discriminant analysis (LDA) Support vector machine (SVM) *K*-nearest neighbors (k-NN) classifier	EEG features were analyzed from different cognitive states Power spectral density achieved the highest accuracy	EEG features were analyzed to control BCI devices Power spectral density achieved the highest accuracy
[9]	Sensors classification for stress analysis Use of behavioral and context data	Sensors classification for stress analysis Importance of using context and behavioral data	Non-obstructive wearable sensors and behavioral and context data will be used

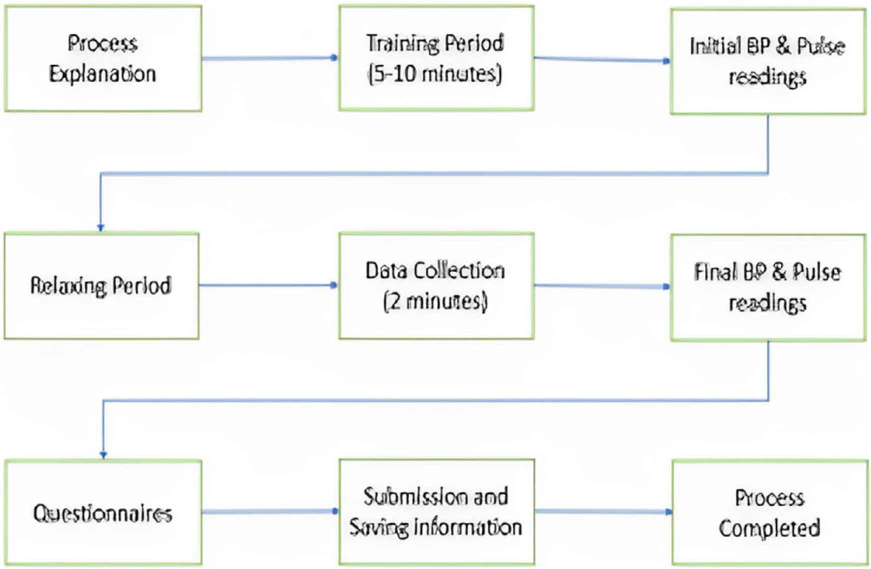


Fig. 3 Entire process of the data collection experiment

in their early professional stages and are under stress from their workloads, career development, etc., while those between the ages of 41 and 60 are in their seasoned professional stages and are typically coping with chronic stress, family pressure, etc. These three groups of people each have a unique style of handling stress. Students, professors, and other professionals make up the three profession-based groupings.

3.1 Data Collection

The research goals, details on the OpenBCI gadget, and other brain-related material were explained to the participants. This made it easier for the person to focus solely on the experiment. The Digital BP Monitor was used to record the initial blood pressure and heart rate data. The subject's head was fitted with the Ultracortex Mark IV OpenBCI headset. The subject was then required to complete a series of cognitive ability assessments, such as quizzes and puzzles, to record their EEG. This activity contained exercises like "find the difference," "spot the sentence error," "simple math problems," "locate the locations on the map," and others. This procedure was referred to as a training phase. This was carried out to increase the subjects' stress levels. Confirming that the sensors were in touch with the scalp and that the GUI was accurately displaying data retrieved from the Cyton board required simultaneous customization via the OpenBCI GUI screen. It took about five minutes to complete this task. Now it was instructed to the individual to rest by closing their eyes. At this point, data collection began. This procedure was referred to as the relaxing phase. Two minutes were spent gathering the data. The only information used for analysis was this. The subjects were instructed to reopen their eyes after two minutes, and the data collection was then halted. The final readings of blood pressure and heart rate were recorded. The subject's head was free of the headset. A PSS-10 survey was then presented to the subject. This questionnaire is a commonly used measure to evaluate a person's degree of stress. This was carried out with the support of psychiatrists and allows us to assess the subject's stress level using the perceived stress scale approach. Low stress, moderate stress, and high stress are the different categories of stress. These labels can be utilized as ML labels in the future to aid with model construction. The experiment was over after this. The experiment lasted 15 min on average for each subject.

4 Challenges Faced During Initial Phase

The purpose of the training phase was to increase the subject's degree of stress and to focus all of their attention on the experiment, improving the quality of the data collected during the relaxing phase. Stress can cause a variety of symptoms, such as forgetfulness, an elevated heart rate, swollen tongue, and mental confusion. The tasks should be created in a way that causes the subject to exhibit some of the symptoms listed above, even though our goal is not to predict the subject's psychological stress level. Some activities, such finding and memorising locations on a map, were also found on the famous National Geographic programme "brain games." Additionally, using timers, encouraging the participants for successful responses, and a little discouraging them for incorrect answers all contributed to the success of our training session. The subjects' curiosity and anxiety about their performance led to the anticipated behavioural shift [10].

To determine whether the data obtained with the Ultracortex Mark IV may be used for the study of brain states and whether the EEG data is sufficient to provide us the predicted model, a number of queries and conversations were conducted. The EEG data gathered possesses accuracy and can be utilized to identify different brain states, and it became obvious after extensive research and referencing OpenBCI literature. Neurologists and neurophysicists were consulted, and it was determined that the EEG had successfully distinguished stress from a state of rest with an accuracy of 85.55%, with an average classification rate of 90% utilizing Yule Walker algorithm[11]. Additionally, skin conductivity, heart rate variability, and EEG signals worked together to define stress with an accuracy of 84.1%. And finally, stress has been categorized with an accuracy of nearly 95% using blood pressure, conductance of the skin, heart rate variability, and EEG signals [11]. Therefore, it was evident that the model built using parameters would have a high level of accuracy. One of the classic tools used by psychiatrists to examine a subject's stress levels is the perceived stress scale (PSS) data. This knowledge made it possible to train the system more accurately and identify the information that was collected following the training phase. Additionally, adding variables like BP and heartbeat can help to achieve a considerably higher degree of accuracy.

For the 2-min data gathering experiment, the dataset had an average of close to 34,000 rows. The dataset measured at roughly 10 MB in size. The databases were over 300 MB in size and included data from 25 different patients. The problem with storing this huge quantity of data directly in online repositories is that they have a finite amount of storage space. The OpenBCI GUI programme should be able to be stored and shown on the laptop or desktop used for data gathering. In processing the GUI application, a computer with 8 GB RAM, a 2 TB HDD, and a 2 GB graphics card is sufficient. Additionally, care must be given when placing a headset on the subject's scalp because the spiky electrodes must properly contact the scalp and cannot be twisted more firmly while the electrodes are being adjusted.

Although there are many technologies present now to analyse digital signals, it is important to select the appropriate tool for the data. EEGLAB, a programme from MATLAB, includes tools for temporal analysis of frequencies and independent component analysis (ICA), as well as support for 20 distinct binary file formats. As a flexible language, Python has a wide range of packages, including scipy, dsptools/dsppy, pyopenbci, scikit-learn, etc. Numerous writers whose work contributed to the Python community in the area of DSP have created their own packages that meet the needs of their books and their readers; these are also accessible and usable worldwide. Similar to OpenBCI, BrainFlow is a library that may be used to gather, decode, and analyse EEG as well as other bio-signals. As the GUI was producing graphs using the data that had been cleansed in the backend by BrainFlow's APIs, we decided to use this library. The construction of a single data flow to directly analyse data from the board is aided by the fact that it not only has APIs for cleaning the data but additionally to gather information from the board and send it to a GUI and storage.

The raw, unfiltered data obtained from the OpenBCI GUI will be stored in a text file called txt. For the purpose of the following exploratory data analysis (EDA) activity, the data must therefore be cleansed. There were a number of preprocessing workflows accessible throughout the research phase; however, the methods were either incompatible with one another or inappropriate for the goal. Considering the history of each technique and by comprehending the terms, keywords, and concepts described in the techniques, a workflow was created, as shown in Fig. 4. Signals that are over the cutoff frequency are attenuated by the low-pass filter, while signals that are below the cutoff frequency are attenuated by the high-pass filter. Based on the documentation and code samples for BrainFlow, the setting values for the above filters were applied. Third-order low-pass and high-pass filters were used in order to obtain effective filtering values. Line noise at 50/60 Hz can also be eliminated with the notch filter. The notch filter is a narrow-band band-stop filter that frequently generates severe deformities of the underlying signal in frequency ranges outside of the notch band. The filter was used to verify the claim, and it was determined to be accurate. Hence, notch filter was removed from the list. The data is entirely filtered and cleaned after being converted from voltage to RMS voltage values and can then be used for the EDA process. The initial steep spike in the clean data's plot, which was later shown to be caused by the initialization of channels, was visible. By

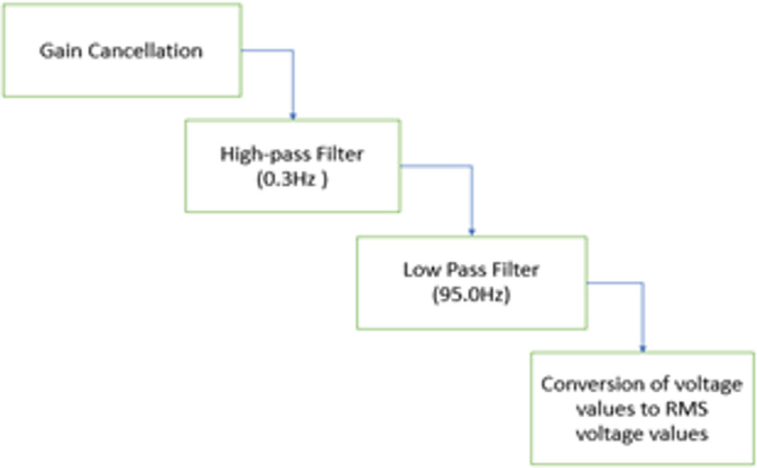


Fig. 4 Workflow of the data preprocessing

removing the first few rows from the dataset, this spike may be ignored because it has no bearing on the experimental statistical and exploratory data analysis.

EEG signal analysis is more difficult than signal analysis for other bio-signals like EMG or ECG. EEG tracks alterations in the brain's electrical activity throughout time [12]. The synaptic action between the neurons is what causes these electrical events. Under each electrode positioned on the scalp, there are hundreds to thousands of neurons. As a result, EEG analysis will be significantly more difficult because we cannot just analyse spikes per second or beats per minute. Additionally, the upper and lower limits of the EEG frequency range are hazy, and the ultra-fast and ultra-slow EEG components are not important for analysis. The two most popular EEG analysis methods involve averaging the data across numerous trials to detect an evoked potential and creating a spectrogram to track frequency variations over time. A complicated signal is broken down into its several frequency components to create a spectrogram, frequently using a Fourier transform.

The event reaction potential (ERP), which are electrical potentials connected to the presence of a stimulus or reaction, can be used to quantify quick and implicit cognitive activity, is one of the valuable features that can be derived from the EEG signals and utilized for EEG analysis. PSD—power spectral density—represents the power distribution of EEG signals in the frequency domain and is another significant feature that may be extracted from EEG signals. EEG data analysis using power spectral analysis is an established technique [13]. The PSD indicates the quantity of the specified band (α, β, γ, θ) [14] is present in the specified signal. The selection of EEG bands from the alpha, beta, gamma, delta, and theta bands that capture stress-related ERPs was the next step [15]. Each power band's frequency and amplitude are listed in Table 2. The alpha, beta, gamma, and theta bands were the ones that were important for our research. The beta and theta bands were the two most significant

Table 2 EEG signal frequency and amplitude ranges

EEG signal	Frequency (Hz)	Amplitude (μV)
Alpha	8–13	30–50
Beta	13–30	5–30
Gamma	> 30	< 5
Delta	0–4	20–200
Theta	4–8	> 20

ones. A discussion with a neuro expert clarified the choice of the electrodes and EEG bands. The estimation of PSD for the channels came next. In general, the most popular techniques for determining PSD from the frequency bands are Welch's method, FFT, and Burg's method [16]. The Welch's method is a modified FFT strategy that averages a tiny window's Fourier transform. So, in order to determine the PSD of the specific channel, we employed this method. The OpenBCI manual was used to set the parameters for the BrainFlow PSD API. The band powers of the necessary EEG bands were later determined using the PSD array that was produced following the estimation. For both channels, the band power ratios such as the α/β, β/γ, β/θ were computed. Other metrics were determined, including the SYS difference (difference between final and initial systolic blood pressure readings) and the DIA difference (difference between final and initial diastolic blood pressure readings). The difference in heart rate or pulse between the final and initial heart rate data was also computed. The PSS score was computed, and the labels associated with the PSS score were identified. These estimated parameters were transformed into a dataset, and the statistics of the datasets were visualized using several tools. Some of the data points were discovered to be incorrect since they had significant divergence from the other values; as a result, they were classified as outliers and eliminated from the dataset. According to the correlation matrix, the PSS score has a strong link with the DIA difference and a weak correlation with the beta/theta channel 1 ratio. When these two characteristics were visualized using a scatter plot, two groupings of data points were perfectly visible. The first one stands for low tension, while the second one is for moderate stress. Figure 5 shows the scatter plot for a few randomly chosen data points.

The next step after completing the EDA process was to experiment with the dataset using various machine learning models to predict the stress labels from the two recognized features.

5 Machine Learning

The experiment's main goal was to test the effectiveness of several machine learning algorithms using various models. Machine learning is a technique where we teach the model to analyse, comprehend, and predict precise results by feeding it data that is pertinent to our situation. Our task is made easier by the abundance of ML

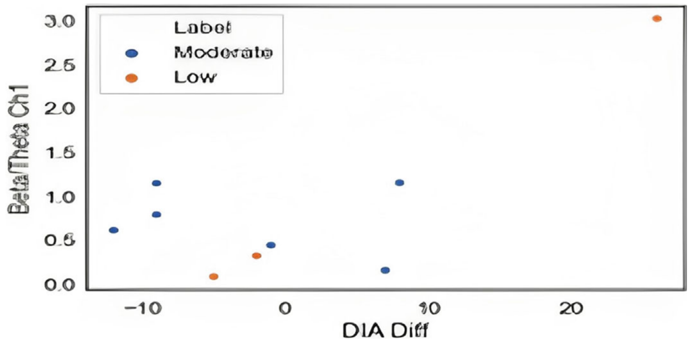


Fig. 5 Scatter plot showing the relation of stress labels with the two features

algorithms in the Python module "scikit-learn." With a ratio of 75:25, the dataset was divided into train and test sets. Different classification techniques, such as KNN, SVM, logistic regression, decision trees, and random forests, were used as a starting point for the training of the models, which were then tested for accuracy across 10 to 15 trials before being used to generate observations [17–19]. A supervised machine learning approach called *K*-nearest neighbours is employed for categorization based on the proximity of data points to a reference equation. Support vector machines, often known as SVMs, are supervised learning models that are effective for a variety of real-world issues. They can address both linear and nonlinear problems. Using this approach, the data is divided into classes by a line or hyperplane. The kernel SVM is a different variant or version of the SVM. Here, we can specify a parameter for the algorithm's kernel type. The accessible kernel types include sigmoid, radial basis function (RBF), polynomial, linear, and nonlinear. The type of curve that will be utilized to distinguish between the different kernel types will depend on the type of classes. Similarly, logistic regression is a supervised classification approach that divides the data into various groups using an equation. The decision tree model is a tree-structured classifier in which the branches reflect the judgements to be made, and the nodes typically represent the characteristics of the input. The random forest algorithm is an unsupervised machine learning technique that creates several decision trees and merges them all to produce more precise, accurate, and stable decisions as an extension to the decision tree classifier. The algorithms' outcomes were recorded and studied.

6 Results

The information of the performance of various ML models is provided in Table 3. With an average accuracy of 33%, the KNN algorithm did not perform as well as planned. The accuracy of other methods, such as the SVM, decision trees, and logistic regression, was consistently 67%. One thing to keep in mind is that the sample size

Table 3 Machine learning model accuracy

Algorithm	Accuracy (%)
SVM	67
K-SVM	69
KNN	33
Random Forest	94

and the random selection of the data may have an impact on how well these models function. However, the SVM and kernel SVM algorithms can do far better. Since the accuracy was constant throughout all trials and was unaffected by the dataset's sampling. The outcomes of the machine learning component give us a hint that if data is gathered from a big population, the algorithms' accuracy will be significantly improved, and the project as a whole will be a huge success. The experiment should be carried out in a method that allows for the collection of data from a larger number of subjects. A handful of the data rows were discovered to be imperfect throughout the EDA process and were subsequently removed from the dataset as outliers. The dataset's size was further decreased as a result.

7 Conclusions

We do know that technology has assisted us in completing many significant jobs that need for intense concentration, which humans frequently miss or are unable to handle. One such study used a method to estimate stress levels rather than depending on the subject's psychological information, which is frequently biased. This technique directly enables us to more precisely estimate human stress levels. With a few minor adjustments to the data collection experiment, the research can be utilized to anticipate other unfavourable feelings, such as anxiety, depression, tension, and so on. The data collection experiment should be set up in a way that allows for the analysis of other unfavourable feelings or mental states. Future work in the field can make use of this work. The documentation and BrainFlow APIs were used to prepare the data cleaning procedure, which can be easily translated into different languages, hence offering a simple solution for the complex and difficult process of EEG data preparation. Researchers may attempt to investigate novel spectrum analysis and EEG aspects in the future. The project provides a tool for parents to periodically check in on their child's mental health in addition to assisting neurospecialists and psychiatrists in performing fast analyses of individuals. Additionally, using BCI devices to forecast stress will be incredibly beneficial and simple for anyone to monitor from the comfort of their own homes.

References

1. Al-Shargie FM, Tang TB, Badrudddin N, Kiguchi M (2015) Mental stress quantification using EEG signals. In: International conference for innovation in biomedical engineering and life sciences, pp 15–19
2. Perez E, Cervantes I, Duran E, Bustamanate G, Dizon J, Chang Y, Lin H (2016) Feature extraction and signal processing of open-source brain-computer interface. In: 2nd Annual under-graduate research expo, University of Texas, Dallas
3. Rajendran VG, Jayalalitha S, Adalarasu K, Usha G (2022) Analysis and classification of stress among students using EEG as biomarker. ECS Trans. https://doi.org/10.1149/10701.1857ecst
4. Shaik FA, Raja AR (2023) Stress monitoring in humans using biomedical signal analysis. https://doi.org/10.1109/ICEEICT56924.2023.10156948
5. Sundaresan A, Penchina B, Cheong S, Grace V, Valero-Cabré A, Martel A, Martel A (2021) Evaluating deep learning EEG-based mental stress classification in adolescents with autism for breathing entrainment BCI. Brain Inform. https://doi.org/10.1186/S40708-021-00133-5
6. Chen C, Yu X, Belkacem AN, Lu L, Li P, Zhang Z, Wang X, Tan W, Gao Q, Shin D, Wang C, Sha S, Zhao X, Ming D (2021) EEG-based anxious states classification using affective BCI-based closed neurofeedback system. J Med Biol Eng. https://doi.org/10.1007/S40846-020-00596-7
7. Barjinder K, Dinesh S, Roy PP (2018) EEG based emotion classification mechanism in BCI. Procedia Comput Sci. https://doi.org/10.1016/J.PROCS.2018.05.087
8. Rashid M, Sulaiman N, Mustafa M, Khatun S, Bari BS (2018) The classification of EEG signal using different machine learning techniques for BCI application. Rev Informatica Teórica Aplicada. https://doi.org/10.1007/978-981-13-7780-8-17
9. Sysoev M, Kos A, Sedlar U, Pogacnik M (2014) Sensors classification for stress analysis: toward automatic stress recognition. https://doi.org/10.1109/IIKI.2014.31
10. Sharma R, Chopra K. EEG signal analysis and detection of stress using classification techniques. ISSN: 0252-2667
11. Niedermeyer E (1999) Normal EEG of the waking adult. Electroencephalogr Basic Principles Clin Appli Relat Fields 20(4):149–173
12. Read GL, Innis IJ (2017) Electroencephalogram (EEG). In: The International encyclopedia of communication research methods, pp 1–18
13. McFarland D, Mccane L, David SV, Wolpaw J (1997) Spatial filter selection for EEG-based communication. Electroencephalogr Clin NeuroPhysiol
14. Hossieni SA, Khalizadeh MA (2010) Emotional stress recognition system using EEG and psychophysiological signals. In: 2010 International conference on biomedical engineering and computer science, pp 1–6
15. Moreno MAR, Diaz-Padilla M, Valenzueka-Gómez KD, Vargas-Martínez A, Tudón-Martínez JC, Morales-Menendez R, Ramirez-Mendoza RA, Henríquez BLP, Lozoya-Santos J (2021) EEG-based tool for prediction of university students cognitive performance in the classroom. Brain Sci 11(6):698
16. Alam R, Zhao H, Goodwin A, Kavehei O, McEwan A (2020) Differences in power spectral densities and phase quantitites due to processing of EEG signals. Sensors 20(21):6285
17. Niha K, Aisha Banu W (2019) Brain signal processing: technologies, analysis and application. In: 2016 International conference on computational intelligence and computing research (ICCIC)
18. EEG Basics, CiMEC Wiki Page, 2020. ¡wiki.cimec.unitn.it/tiki-index.php?page=Data+pre-processing¿
19. BrainFlow: replicating the pre-processing steps from the GUI, OpenBCI Forum, MrKlean518, February 2020. ¡https://openbci.com/forum/index.php?p=/discussion/2478/brainflow-replicating-the-pre-processing-steps-from-the-gui¿

An Approach to the Detection of Brain Tumors by Processing the Magnetic Resonance Images

V. V. Hosallimath and A. S. Awati

Abstract A brain tumor is a lethal disease caused by the irregular growth of abnormal cells in the brain. Investigation of brain tumors with precise size, shape, and location is the most difficult task for medical experts, due to the noise inherited into the image during imaging. Also, investigating corrupted images due to instrumental errors is tedious and may lead to false detection of the tumor. To overcome the above-stated problems, an effective approach is made to detect the tumor by processing the MR images. The system initially pre-processes the acquired MR images and segments the regions using the manual threshold method, and the morphology of the tumor will be studied based on the shape, and texture-based parameters to identify the area, shape, and the region of tumor. This system efficiently detects the tumor on the brain tumor dataset and gives informative inputs about the tumor to the pathologist for diagnosis.

Keywords Magnetic resonance imaging · Morphology · Convex area · Convex hull · Entropy

1 Introduction

Irregular growth and the collection of abnormal cells in the body constitute a tumor. If such tumorous tissues are present in the brain, then they are called brain tumors. Tumors are of two types one is benign and another is malignant [1]. The cell division of benign tissues is quite slow, these tissues don't damage the neighboring tissues, and these types of tumors are so-called non-cancerous [2]. Benign tumors include fibroids in the uterus and lipomas in the skin. The tumor cells with high cell division and spread

V. V. Hosallimath (✉)
KLE Technological University, Hubballi, Karnataka, India
e-mail: hosallimathviru@gmail.com

A. S. Awati
KLS Gogte Institute of Technology, Belagavi, Karnataka, India
e-mail: asawati@git.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_18

rates are called malignant tumors. Examples of malignant tumors are carcinoma and sarcoma. According to an International Association of Cancer Registries (IACR) survey, around 28,000 brain tumor cases are diagnosed annually in India, and only 14% of these cases survive [3]. The German Brain Tumor Association (GBTA) dedicated the 8th of June as World Brain Tumor Day to raise awareness about the prevention, diagnosis, and treatment of the tumor. Magnetic Resonance Imaging (MRI) is an advanced method used for the acquisition of information from the internal organs of the body in the form of a 3D image. MR imaging is done by placing the human body in the MRI scanner, and the radio signals in the form of short pulses are sent to the interested organ of the body, these signals are absorbed by the nuclei of the hydrogen atoms. The nuclei intern creates slowly decaying radio signals, and the decaying speed of the radio signals varies, as the concentration of hydrogen atoms varies from one region to another in the organ. The time that is taken for the complete decaying of the radio signals from hydrogen atoms is considerably called the relaxation time. If the relaxation time of the radio signal at a particular region is less compared to other regions of the organ, then that region can be labeled as a tumor.

The brain tumor region can be detected and extracted from the MR image using an efficient technique called image processing [4]. In this technique, the MR images are initially resized and converted to grayscale, and then the grayscale images are filtered using the anisotropic filter for noise removal and also for improving the visual quality of textures in the image at steep angles [5]. The skull region in the image is stripped by using the erosion technique for better detection of tumors by the system. The pre-processed images are binarized, and the morphology of the tumors is studied by extracting features like tumor size, shape, and region. Entropy is the texture-based parameter used for better analysis of the tumor region. The paper is organized as follows, the introduction to the paper is given in Sect. 1. Section 2 explains the literature review, the methodology for designing the brain tumor detection system is explained in Sect. 3, and it is simplified in Sect. 4 using flowchart. Experimental results are discussed in Sect. 5, and Sect. 6 concludes the paper and its future work.

2 Literature Review

Kumar et al. [6] proposed a method to detect brain tumors using image processing. In his approach, pre-processing of the images was done by the wavelet-based method, then the images were filtered using the Gaussian filter, and finally, the features were extracted using the cluster-based segmentation method. SVM and SOM classification models were implemented to classify the benign or malignant type tumors. The proposed system was 97% accurate in detecting brain tumors.

Automated detection of brain tumors through magnetic resonance images using convolutional neural networks is the framework done by Gull et al. [4]. The framework was based on a Fully Convolutional Neural Network (FCNN) and transfer learning technique. The designed system initially pre-processed the MR images for noise removal and contrast improvements. The global thresholding technique was used to extract the tumor region, and the classification of the outputs was done by the GoogLeNet model. The proposed system achieved 97.5% accuracy in segmenting the images and achieved 97.31% accuracy in the classification of the tumor on the BRAT2020 dataset.

Banik et al. [5] proposed a very simple technique for detecting brain tumors using MR images. Filtering of the noise in the images was done by the Gaussian equation-based algorithm, and hybrid filtering was used to filter time-domain noises. A general segmentation technique was used to extract the brain tumor features such as shape, size, and region. By experimentation, the system was found to be efficient in detecting the tumor.

Maiti et al. [7] approached detecting brain tumors by processing the MR images. Firstly, the images were converted from GRB color components to HSV color components. The watershed segmentation method was implemented for the extraction of tumors. The canny edge detector detects the edges of the tumor and gives the actual region of the tumor in the brain. The results of the proposed algorithm were promising.

Dipu et al. [8] developed a deep learning-based brain tumor detection and classification model which used YOLO and the Library for the detection and classification of the tumor and its types, respectively. The YOLOv5 model achieved an accuracy of 85.95% in detecting the tumor and achieved an accuracy of 95.78% in the classification of the tumor types.

3 Methodology

The steps for designing the brain tumor detection system using image processing algorithms are explained using the block diagram, which is given in Fig. 1. According to the block diagram, initially the brain tumor MR images are pre-processed by resizing the images and converting the images to grayscale. Next, filter the images using the anisotropic filter and strip the skull part in the image using the erosion technique. Finally, extract the features by segmenting the images using the threshold method [8].

3.1 Image Acquisition

The brain MR images can be obtained by scanning the brain using a 1.5 T (Tesla) MRI scanner. The MR images dataset for testing the system was downloaded from the Kaggle.com website [8].

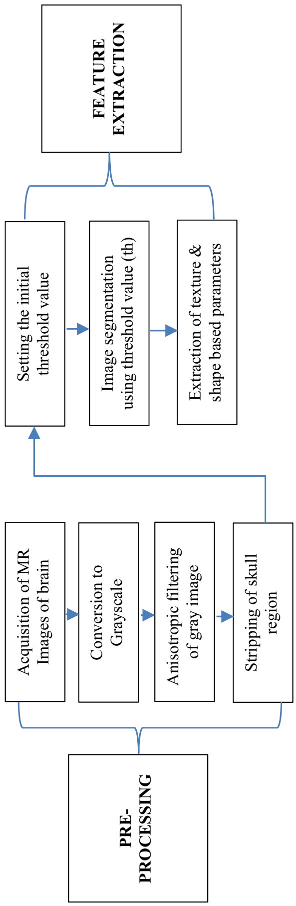


Fig. 1 System block diagram

3.2 Pre-processing

The MR images were pre-processed for feature extraction. Following are the pre-processing steps that were implemented for designing the brain tumor detection system [7].

3.2.1 Conversion of RGB to Grayscale

An MR image may look like a gray image but it contains shades of RGB colors. So, the MR images were converted to grayscale using the weighted method, in which the wavelength (weight) of the respective color is multiplied by the intensity of the color and finally the gray-scaled images were obtained [7].

3.2.2 Anisotropic Filtering

Anisotropic diffusion aims to filter the noise present in the image and gives a smooth image without ignoring the edges and minute structures in the image [8]. The basic expression of the anisotropic filter is given in Eq. 1.

$$\frac{\partial I(x, y, t)}{\partial t} = \mathrm{div}\big[g(||\nabla I(x, y, t)||)\nabla I(x, y, t)\big] \tag{1}$$

where t is the time parameter, $I\,(x, y, 0)$ is the input image, $\nabla I\,(x, y, t)$ is the gradient of the respective image at time t, and $g(.)$ is the conduction function. This function shows the maximum diffusion when $\log_{x\to 0} g(x) = 1$ within the uniform region and the diffusion stops across the edges of the image, if $\log_{x\to 0} g(x) = 0$. Two such functions are given in Eqs. 2 and 3.

$$g_1(x) = \exp\left[-\left(\frac{x}{k}\right)^2\right] \tag{2}$$

$$g_2(x) = \frac{1}{1 + \left(\frac{x}{k}\right)^2} \tag{3}$$

where k is the gradient threshold, which controls the rate of diffusion and provides a smooth threshold.

3.2.3 Skull Stripping

Erosion is the process of removal of the pixels on the object boundaries. In the MR images of the brain, the skull region should be stripped for a better analysis of the tumor region in the post-processing steps, and this can be done by eroding the image boundaries [5]. Stripping the skull region in the grayscale image is done by eroding the boundary using a disk-shaped structuring element. In general, the grayscale erosion of $A(x, y)$ by $B(x, y)$ is given in Eq. 4.

$$(A \ominus B)(x, y) = \min\left\{A\left(x + x', y + y'\right) - B\left(x', y'\right) | \left(x', y'\right) \in D_B\right\} \tag{4}$$

where D_B is the domain of structuring element B. For creating disc-shaped structuring elements use strel (disk, radius, n), where n is the number of lines that the structuring element uses to obtain an approximate disk shape. In Fig. 1, the original image is shown, and in Table 2, Y1 is the skull-stripped image.

3.2.4 Segmentation

Segmenting an image's background and foreground regions using a threshold value for extracting the features is called binarization. Binarization of MR images was done by a manual thresholding method [6]. The threshold value (th) is calculated by summing the initial threshold value with an average of the global maxima and global minima of image I. Mathematically the formula is given in Eq. 5.

$$\mathrm{th} = t_0 + ((\max(I(:)) + \min(I(:)))/2) \tag{5}$$

where t_0 is the initial threshold value, max ($I(:)$) and min ($I(:)$) are the global minima and global maxima. The shape of the tumor in the binary image varies as the initial threshold value varies, according to Eq. 5.

3.3 Feature Extraction

Morphology is the study of the size, shape, and texture of the organs in animals [9]. The morphology of the brain tumor was analyzed by extracting the texture and shape-based parameters such as entropy and area, respectively. These parameters are explained below.

3.3.1 Convex Area

Convex area is the number of white pixels in the convex image. An image that specifies the convex hull is called a convex image, and the small convex polygons that form regions are called convex hulls. The total number of white pixels in the convex hull constitutes the convex area [8].

Initially, all the convex hull areas in the segmented image are calculated. If the area of any convex hull in the image is greater than the experimental threshold area that is $T_a = 150$ mm [10], it is labeled as a tumor, and the convex hulls having areas less than T_a are ignored. Finally, we obtained the tumor-alone image.

3.3.2 Entropy

Entropy can be called the degree of randomness of white pixels in a binary image [5]. If the randomness of white pixels in the tumor-alone image is nearly equal to 1, that indicates the presence of the tumor in the image, and if it is nearly equal to 0, that indicates no tumor in the image. The mathematical formula for calculating the entropy of an image is given in Eq. 6.

$$\text{Entropy} = -\sum(p * \log 2(p)) \tag{6}$$

where p is the normalized histogram counts. The entropy values for different MR images of the brain are tabulated in Table 2.

4 System Flowchart

The methodology of the brain tumor detection system is simplified using a flowchart as shown in Fig. 2. The flowchart explains that the input images are pre-processed and segmented for feature extraction such as the area and entropy [11]. If the entropy is greater than or equal to 0.11 and Max_area greater than T_a the tumor is detected, else tumor is not detected.

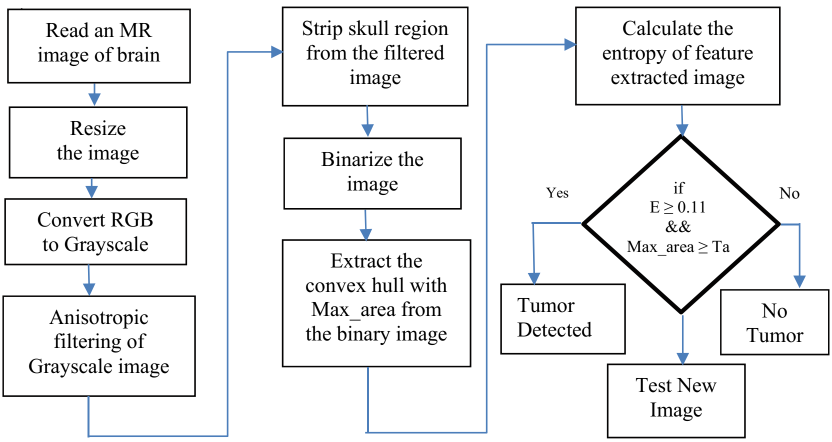


Fig. 2 System flowchart

5 Experimental Result Discussion

The results of the brain tumor detection system are tabulated in Table 1. The table contains the MR images of the brain, respective binary images, tumor-extracted images, the area of the tumor, and the entropy of the tumor-extracted image. Failure cases of the tumor detection system are tabulated as N1, N2, so on

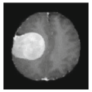
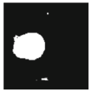
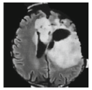
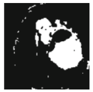
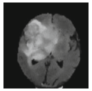
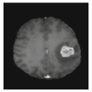
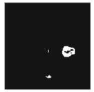
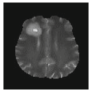
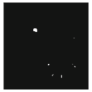
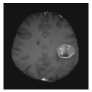
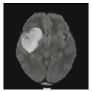
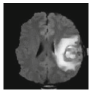
5.1 Experimental Results Discussion

The brain tumor detection system was efficient in detecting the tumor. The experiment was conducted in a DELL Inspiron RYZEN 5000 series machine, and the system was designed using the MATLAB R2022a tool. The brain MR images dataset was obtained from the Kaggle website, and the dataset had 255 images. All the images were pre-processed and features were extracted for the analysis of brain tumors. Considering image Y1 from Table 2, the system calculated the area of the tumor and the entropy of the tumor-alone image which are 1430 mm and 0.456, respectively. Since the Max_area of the convex hull is greater than the threshold area, and the value of the entropy is nearly equal to 1, the system concluded and displayed "Tumor Detected." The system was tested for image Y17, the area of the tumor calculated was 43 mm which is smaller than the threshold area, and the entropy of the tumor-alone image was 0.028 which is nearly equal to 0, then the system concluded and displayed "No Tumor Detected." The brain tumor detection system failed to detect and analyze the tumor due to poor image quality, instrumental errors while capturing the images, light-varying conditions, and poor processing of the images. The failure cases of the system are tabulated as N1, N2, and so on. Considering image N1, the system failed to detect the tumor due to poor stripping of the skull region from the image. The system also failed to analyze the tumor in image N5 due to poor input image quality, and the quality of the image may depend on the instrument and light conditions in the MRI room. In future, the system should be designed to overcome the above limitations.

6 Comparison Between Proposed and Existing Methodology

In an existing approach for the detection of brain tumors using MRI images [12], the system initially pre-processed the input images for the enhancement of the information then the features (tumors) were extracted by using the segmentation method. In this methodology, due to the lack of a noise filtration technique, the edges and regions of the images were not visible and such images lead to incorrect detection of the tumor region and shape. The skull region in the MR image is unwanted. During

Table 1 Output tabulation of brain tumor detection system

Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy	Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy
Y1				1430/ 0.456	Y15				2369/ 0.632
Y2				315/ 0.147	Y16				184/ 0.112
Y3			No tumor detected	24/0.017	Y17			No tumor detected	43/0.028
Y4				593/ 0.241	Y18				735/ 0.283

(continued)

Table 1 (continued)

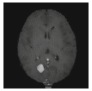
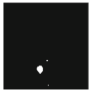
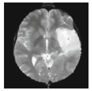
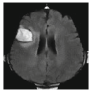
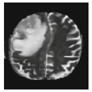
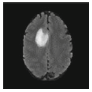
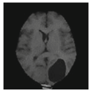
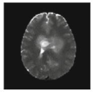
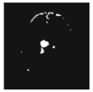
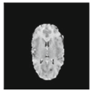
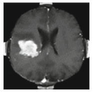
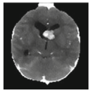
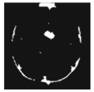
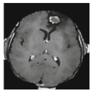
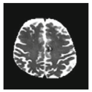
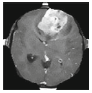
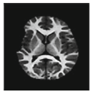
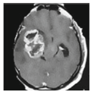
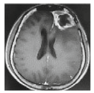
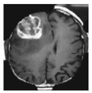
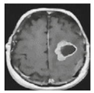
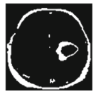
Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy	Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy
Y6			No tumor detected	73/0.044	Y19				1080/ 0.375
Y7				312/ 0.146	Y20				1231/ 0.411

Table 2 Output tabulation of brain tumor detection system

Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy	Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy
Y8				292/0.139	Y21			No tumor detected	12/0.009
Y9			No tumor detected	119/0.067	Y22				2320/ 0.624
Y10				483/0.206	N1				188/0.097
Y11			No tumor detected	59/0.037	N2				870/0.321
Y12				954/0.343	N3				844/0.312

(continued)

Table 2 (continued)

Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy	Sl. No.	MR image of the brain	Binary image	Tumor extracted image	Tumor area (mm)/ Entropy
Y13				499/0.211	N4				352/0.161
Y14				304/0.143	N5				1440/ 0.458

the feature extraction, there are chances that the system may recognize the skull region as the tumor, in the absence of the actual tumor in the brain. In our proposed methodology, the system filters the noise present in the image using the anisotropic filter, which gives a clear visualization of all the regions in the MR image. The skull region has been removed using the erosion technique so that the system processes only the brain region for the extraction of a tumor. And, entropy is a texture-based parameter that gives a good interpretation of the size and presence of the tumor.

7 Conclusion and Future Work

The experimental results of the brain tumor detection system were promising and effective. The system initially resized and converted the input images to grayscale, and the grayscale images were filtered using the anisotropic filter. The filtered images were segmented using the threshold method for feature extraction. The area and region of the tumor were detected and analyzed by a texture-based parameter called entropy. Tumor detection was affected by the angle of the capture of MR images, poor image quality due to instrument errors, and light-varying conditions during image capturing. Finally, we conclude that the proposed brain tumor detection system helps pathologists to analyze the tumor easily and provide better health advice to patients based on the technical outputs. In the future, the proposed system can be developed and modified to identify the different types of tumors, and stages of the tumor by implementing many other relative image processing parameters.

Acknowledgements This research is ethically supported by the Department of Electrical and Electronics Engineering, KLE Technological University. We acknowledge KLE Technological University for financial help in publishing this paper.

References

1. Malik M et al (2021) Comparison of brain tumor detection in MRI images using straightforward IP techniques. In: 2021 3rd International congress on HORA, Ankara, Turkey, pp 1–6. https://doi.org/10.1109/HORA52670.2021.9461328
2. Hashan AM et al (2021) MRI images using image processing techniques. In: 2021 4th International symposium on ISAMSR, BatuPahat, Malaysia, pp 24–28. https://doi.org/10.1109/ISAMSR53229.2021.9567799
3. Sethy PK et al (2022) Brain tumour localization using image processing techniques. In: 2022 Sixth international conference on I-SMAC, Dharan, Nepal, pp 753–756. https://doi.org/10.1109/I-SMAC55078.2022.9986502
4. Gull S et al (2021) Automated detection of brain tumor through magnetic resonance images using convolutional neural network. BioMed Res Int 2021, Article ID 3365043. https://doi.org/10.1155/2021/3365043

5. Banik R et al (2015) Automatic detection, extraction and mapping of brain tumor from MRI scanned images using frequency emphasis homomorphic and cascaded hybrid filtering techniques. In: 2015 ICEEICT, Savar, Bangladesh, pp 1–6. https://doi.org/10.1109/ICEEICT.2015.7307485
6. Kumar S et al (2019) Brain tumor detection using image processing. IJISA 11(1) (special issue). ISSN 0974-2255
7. Maiti I et al (2021) A new method for brain tumor segmentation based on watershed and edge detection algorithms in HSV color model. In: 2012 NCCCS, Durgapur, India, pp 1–5. https://doi.org/10.1109/NCCCS.2012.6413020
8. Dipu NM et al (2021) Deep learning based brain tumor detection and classification. In: 2021 International conference on CONIT, Hubli, India, pp 1–6. https://doi.org/10.1109/CONIT51480.2021.9498384
9. Zain Eldin H, Gamel SA, El-Kenawy EM, Alharbi AH, Khafaga DS, Ibrahim A, Talaat FM (2022) Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization. Bioengineering (Basel) 10(1):18. https://doi.org/10.3390/bioengineering10010018. PMID: 36671591; PMCID: PMC9854739
10. Sudharson M et al (2016) Brain tumor detection by image processing using MATLAB. Middle-East J Sci Res 24(S1):143–148. https://doi.org/10.5829/idosi.mejsr.2016.24.S1.30
11. Sravanthi N, Nagari S, Poreddy D, Siliveru R, Suwarna G, Sateesh N (2021) Brain tumor detection using image processing. Int J Sci Res Comput Sci Eng Inf Technol 348–352. https://doi.org/10.32628/CSEIT217384
12. Isselmou A, Zhang S, Xu G (2016) A novel approach for brain tumor detection using MRI images. J Biomed Sci Eng 9:44–52. https://doi.org/10.4236/jbise.2016.910B006

Review on Password-Based Door Locking System Using IOT

Kanegonda Ravi Chythanya, Koya Sampath Reddy, Manuka Pravalika, K. Jayapal, L. Nagesh, B. Ruchitha, and B. Mani Teja

Abstract To meet the new difficulties and present needs and provide positive results, the majority of systems are automated nowadays. Automated systems are more flexible and precise since there are fewer manual procedures. Therefore, every file prefers the automated control system, especially in electronics. The project objective is to create a unique system using mobile technology that can control the door just by entering the password as a message which opens/closes the door. This project also presents a low-cost security system that is widely used in our everyday life. It is designed to prevent unauthorized persons from opening/closing the door by entering the correct password. The password entry system will open the door lock once the user enters the right password, and the door lock will open. Traditional locks are heavy and not very protective as they can easily be broken down by some tools. Therefore, electronic locking systems are preferred over mechanical locks. The door is opened and closed by this method using Android technology.

Keywords Android technology · Telegram bot · Graphical client interface · Photodiodes · Microcontroller · Servo motors

1 Introduction

With the help of an Android device and a single keyword entered into an Android app, this extension is designed to open doors. Through a Bluetooth device, an Android OS-based device, such as a tablet or portable, is connected to the framework. The microcontroller, which has been modified with a Bluetooth device, is connected to the owner of the door must use a certain secret word to open and close the door. In

K. R. Chythanya (✉)
SR University, Warangal, India
e-mail: ravi.chythanya@sru.edu.in

K. S. Reddy · M. Pravalika · K. Jayapal · L. Nagesh · B. Ruchitha · B. Mani Teja
School of Computer Science and Artificial Intelligence, SR University, Warangal, India
e-mail: manukapravalika@sru.edu.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_19

an Android a screen-based program with a Graphical Client Interface. The goal of this project is to use an Android device to open a door by typing a single keyword into an Android app. A Bluetooth device attached to the microcontroller, which has been programmed by its owner with a unique secret word for opening and closing the door, connects an Android OS-based device, such as a portable device or tablet, to this system. A touch screen Android application featuring a Graphical Client Interface. When a vehicle approaches the door, the miniature scale controller sends a message to the owner's device, asking them to enter the password. When the Android mobile enters the password, it transmits that information to the microcontroller via Wi-Fi. When the microcontroller receives the stored password, it compares the two passwords and, if they match, sends the control signals to the hand-off driver. The transfer then uses the engine to carry out mechanical activities to open and close the door; however, for this project, the engine will be replaced with a servo engine for far better, far better, higher, stronger, improved, and enhanced utilization.

2 Literature Survey

Akshay Krishnadas Bhat [1] has expressed his inclination toward gifting an electronically controlled entrance way lockup framework that utilizes Arduino, an open-source microcontroller board capable of detecting, monitoring, storing, and controlling applications. This framework also employs an IoT-based log that monitors the entry and exit of customers.

Kailash Manger [2] aims to design a secure door lockup system. To accomplish this project, he has undertaken certain tasks such as planning the power supply for the entire electronic equipment, selecting the microcontroller, keypad, DC motor, and buzzer based on the requirements of the project.

Prof. A.Y. Prabhakar's [3] primary objective in this project is to provide safety at every common place, including homes and public areas. All data is stored in the database, and once the correct password is entered, the microcontroller will provide guidance to the servo engine, and the door will unlock. We all require a profound computerized innovation.

Saswat Kumar Das [4] aims to develop a combination electromechanical component that works together as a whole and can be easily installed in coordination with existing sliding or rolling doors of garages or shops worldwide. It can be operated with human effort, and the entry and exit will be controlled by an electromagnetic card or biometric sensor operated by a specified personnel's thumb.

The objective of Willians' [5] project is to create a low-cost system that keeps track of garage doors and sends data about their state to a receiver device installed in a practical location inside the user's home. This enables the customer to keep an eye on their garage doors without leaving the comfort of their home. The receiver device includes LEDs that make it simple to check the status of the garage door from a distance and a screen that shows system information. Because of the range of the technology, the receiver can be placed anywhere inside the home and still accurately report the state of the garage door. The purpose of this initiative is to improve home security through architecture.

3 Proposed Methodology

The digital pins of the ESP-32 are connected to the IR, ultrasonic, and actuator micro servo sensors in this project. The ESP32 is then wirelessly linked to a Telegram bot. To detect any motion of an approaching car, the ultrasonic sensor is constantly activated. The ultrasonic sensor promptly sends a message reading “Car has arrived” and requests a password when it notices a change in distance. The micro servo door will briefly open if the proper password is entered before automatically shutting. The sensors are connected to the ESP32 through connecting cables, and the circuitry is housed on a board. The password is communicated to the ESP32 after being input.

4 Experimentation

To perform this project, the software and hardware tools used are as follows.

4.1 Arduino IDE

Arduino is an open-source software library and breakout board for AVR microcontrollers. The Arduino IDE is the program utilized for coding and is available for download on the Arduino Web site. The physical board that stores and executes the uploaded code is referred to as the Arduino board. Both the software package and board are collectively known as Arduino (Fig. 1).

Showcases the Arduino IDE. Programs created in the Arduino IDE are referred to as sketches. The basic structure of an Arduino sketch comprises two mandatory functions, namely setup() and loop(). These functions are already present in a new sketch window upon opening the Arduino IDE. To introduce beginners to programming, the traditional ‘hello world’ problem is utilized to explain the basic necessities of function.

4.2 ESP32

ESP32 is a development board that comes with all the necessary features to create projects. The ESP32 is a powerful microcontroller that offers many advantages over the Arduino UNO and the ESP8266. It has a dual-core processor that allows for faster

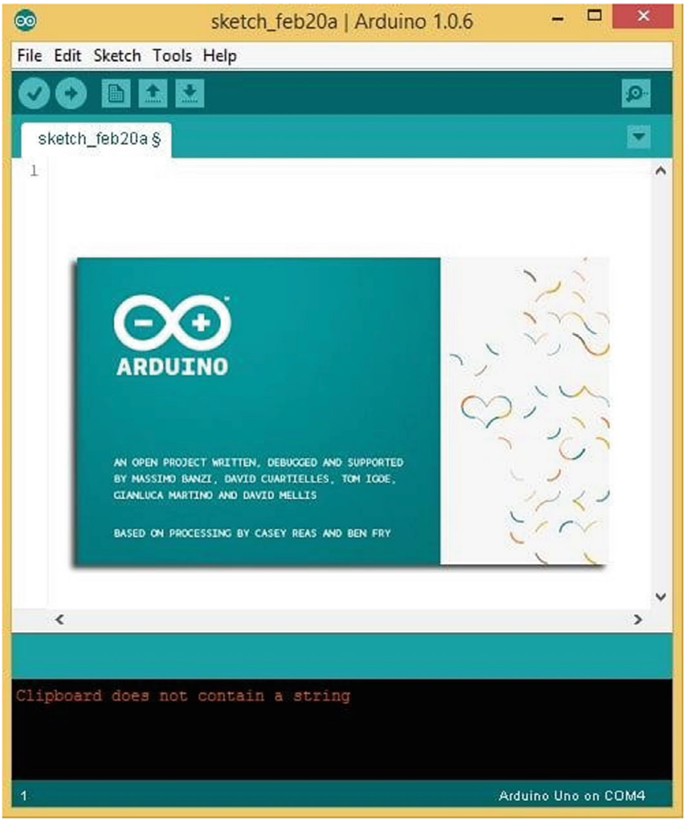


Fig. 1 Arduino IDE

and more efficient performance, as well as a better Wi-Fi module that supports higher speeds and longer range [6]. Additionally, Bluetooth 4.2 and Bluetooth low energy are supported by the ESP32, allowing for wireless connectivity with other gadgets and sensors. Additionally, the ESP32 contains built-in Hall effect sensors that can sense magnetic fields and touch-sensitive pins that can perform various purposes, such as waking up the device from deep sleep mode.

WROOM32 is a popular and versatile ESP32 microcontroller that we will explore and explain in this tutorial. It belongs to the ESP32-WROOM-32 family of microcontrollers, which are widely used and available in the market. Many ESP32-based development boards use WROOM32 as their core microcontroller, such as ESP32 Wemos Lolin32, ESP32-Core Board, and others. This tutorial can also help you if you are working with a different ESP32 microcontroller, because most of them have similar pins to WROOM32.

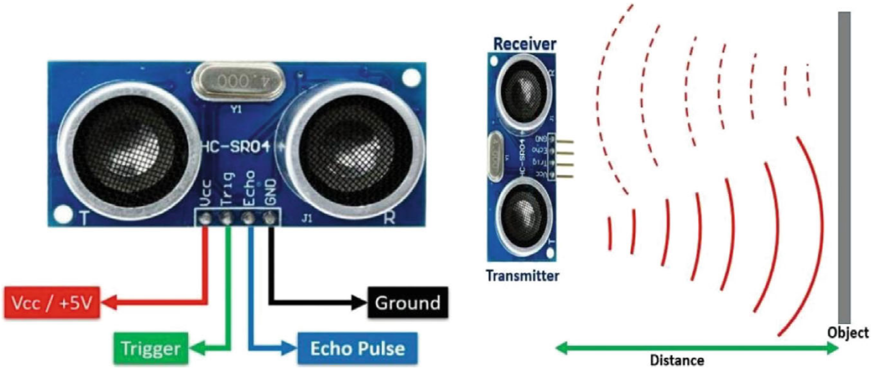


Fig. 2 Ultrasonic sensor

4.3 Ultrasonic Sensor

The ultrasonic sensors use sonar to calculate the distance to an object. There are some ways that explain how it works:

1. Trig pin sends out 40 kHz sound wave.
2. High frequency sound is used for ultrasound.
3. The sound is carried by the air. If it discovers an item, it returns to module.
4. The receiver converts the sound waves into electrical signals.
5. The signals are processed and displayed as an image (Fig. 2).

$$\text{Distance}(S) = \text{Speed}(v) * t$$

In this case, v controls the airborne ultrasonic wave speed. The speed of ultrasonic waves in air is 340 m/s, the same as the speed of sound.

4.4 IR Sensor

An IR sensor is a piece of technology that produces light to identify surrounding things. An IR sensor may detect movement in addition to measuring an object's heat.

Normally, all objects emit infrared heat radiation. Infrared sensors can detect these radiations even when the human eye cannot. Both the emitter and the detector are straightforward IR photodiodes [7]. When the photodiode's emission wavelengths match IR light, the photodiode can detect it. The output voltage and resistance of the photodiode will alter in direct proportion to the quantity of light received (Fig. 3).

An infrared source, an emission channel, a light-emitting component, infrared detectors or receivers, and signal processing are the five fundamental components of a conventional infrared detection system. Examples of infrared sources include

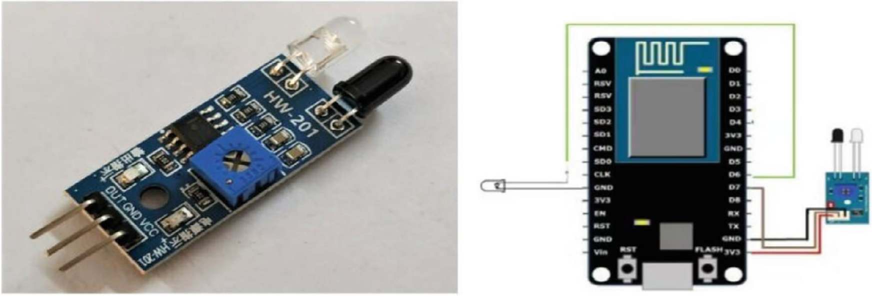


Fig. 3 IR sensor

lasers with infrared wavelengths and specific LED wavelengths [8]. Vacuum, atmosphere, and optical fibers are the three most typical infrared transmission medium types. Infrared light can be focused or the spectrum response limited using optical components.

4.5 Micro Servo

"Micro servomotor," also known as "SG90," refers to a little, light server motor with a lot of output power. A servo may perform tasks that more powerful servos cannot because it spins 180° (90 in each direction). Any servo hardware, software, or library can control these servos.

There are many different types of servomotors available on the market, each with a unique expertize and set of uses [9]. The following two words will determine the servomotor type that you choose for your project or system.

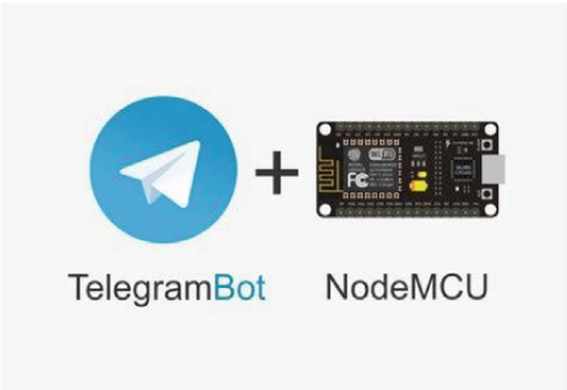
The operating voltage range for amateur servomotors is typically between 4.8 and 6.5 V; however, most of them run at + 5 V due to the potential for greater force. As a result of their gear arrangement, almost all amateur servomotors can only rotate between 0° and 180°, therefore be sure your project can withstand the half circle. If not, you can adjust the motor such that it generates a 360-degree circle or pick a different motor [10]. If your application calls for powerful, long-lasting motors, you can either use metal gears or just go with normal plastic gear because motor gears are prone to wear out (Fig. 4).

4.6 Telegram Bot

An instant messaging and voice over IP service based in the cloud is called Telegram Messenger. It is simple to set up on a PC, Mac, or Linux computer as well as an

Fig. 4 Microserver

Fig. 5 Telegram bot



Android or iPhone smartphone. There are no advertising, and it is free. Telegram enables the creation of chatbots that you can use.

We will make use of the Telegram Bot library, which gives the Telegram Bot API a simple user interface, to access Brian Lough's Universal Telegram bot (Fig. 5).

5 Conclusion

Using an ESP32 board, we will monitor the status of a door in this project. Whenever the door changes its state—open or closed, or depending on the status of a car arriving—you will receive a message on your Telegram account. No matter where you are, as long as your smartphone has Internet access, you will receive notifications.

The ESP32 board will be programmed using the Arduino IDE. The microcontroller in this project for a password-based door locking system is linked to Telegram, and messages from that app are used to open and close the door. If a password entered after one that is already stored matches, the door will temporarily unlock, after delaying the unlocking process for a predetermined amount of time.

References

1. Bhat AK, Kini SP (2018) Password enabled door locking system using arduino and IoT. Int J Eng Res Technol (IJERT) ICRTT 06(15)
2. Manger K, Maya Gurung H, Biswakarma S, Sharma S, Tamang K. Password based door locking system using arduino. Int Res J Eng Technol (IRJET) 08(08):813–816
3. Prabhakar AY, Oza SK, Shrivastava N, Srivastava P, Wadhwa G. Password based door lock system. Int Res J Eng Technol (IRJET) 06(02):1154–1157
4. Das SK, Tiwari DR, Pandey A, Sharma A, Aman A (2018) Automatic garage door opener. Glob J Res Anal 7(2)
5. Lehr T, Williams A (2016) Garage door security system
6. Mothe R, Tharun Reddy S, Chythanya KR, Supraja Reddy Y (2019) Challenges, open research issues and tools in bigdata analytics. Int J Recent Technol Eng (IJRTE) 8(2S11):2634–2641. ISSN: 2277-3878
7. Chythanya KR, Kumar KS, Rajesh M, Tharun Reddy S. Sensor cloud: a breakdown information on the utilization of wireless sensor network by means of cloud computing. Test J Eng Manage 82(1):13945–13954. ISSN: 0193-4120
8. Pramod Kumar P, Sagar K (2019) A relative survey on handover techniques in mobility management. IOP Conf Ser Mater Sci Eng 594:012027
9. Sandeep CH, Naresh Kumar S, Pramod Kumar P (2018) Security challenges and issues of the IoT system. Indian J Public Health Res Dev 11:748–753
10. Shaik MA, Sampath Kumar T, Praveen P, Vijayaprakash R (2019) Research on multi-agent experiment in clustering. Int J Recent Technol Eng (IJRTE) 8(1S4):1126–1129. ISSN: 2277-3878

Environmental Monitoring in Agriculture Using Wireless Sensor Networks (WSNs): Challenges and Applications

V. Indumathi, V. V. Bhavani, Tahseen Jahan, A. Hemalatha Reddy, Jalapala Sinjini, and M. Priyadharshini

Abstract With the increasing demand for efficient and sustainable agricultural practices, there is a growing interest in employing wireless sensor networks (WSNs) for environmental monitoring in agriculture. This review paper provides an overview of the current state of the art in utilising WSNs for agricultural monitoring applications. It explores the various sensor technologies, network architectures, and data processing techniques employed in these systems. Additionally, it discusses the challenges and opportunities associated with deploying WSNs in agricultural environments, including power management, data reliability, and scalability. Furthermore, the paper examines case studies and applications of WSNs in agricultural monitoring, highlighting their effectiveness in optimising resource usage, enhancing crop yield, and mitigating environmental impact. Through a comprehensive analysis, this review aims to provide insights into the advancements, limitations, and future directions of environmental monitoring in agriculture using WSNs.

Keywords WSN · Agriculture · Environmental monitoring · Sensor technologies

V. Indumathi (✉) · V. V. Bhavani · T. Jahan · A. Hemalatha Reddy · J. Sinjini · M. Priyadharshini
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: indu11mathi@cmritonline.ac.in

V. V. Bhavani
e-mail: bhavaniv@cmritonline.ac.in

T. Jahan
e-mail: tahseen048@cmritonline.ac.in

A. Hemalatha Reddy
e-mail: hemareddy@cmritonline.ac.in

J. Sinjini
e-mail: sinjini@cmritonline.ac.in

M. Priyadharshini
e-mail: priyadharshini@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_20

1 Introduction

The agricultural sector is only one of several that has seen rapid convergence thanks to recent developments in information and communication technology. The agricultural industry is anticipated to have a substantial increase in productivity and value due to this convergence and the incorporation of IT into agricultural technologies [1]. Nevertheless, in order to tap into this potential, there has to be omnipresent technology that is tailored to the agricultural sector. This technology should include middleware platforms, application services, routing protocols, and sensor hardware. Problems persist, especially in outdoor agricultural environments without adequate IT infrastructure, despite the fact that there have been encouraging instances of ubiquitous technology's use in agriculture, such as sensor nodes in vineyards and cattle farms. The majority of existing monitoring systems are designed for indoor environments, such as greenhouses, which means that data collected from these places must be manually validated. A server system for monitoring the agricultural environment using wireless sensor networks (WSN) is suggested as a solution to these problems. Outdoor environmental and soil sensors, CCTV cameras, and GPS modules would all contribute to this system's data set, which would then be available to farmers [2]. Data on light, temperature, humidity, wind speed, electrical conductivity, pH, and carbon dioxide levels are all part of the sent information. In addition, the system may be used in places with poor power infrastructure since it is intended to be powered by solar cells. The adoption of such a server system shows potential in empowering farmers to make choices based on data, which in turn might enhance field quality and productivity.

Components of WSNs

One finite resource that determines the lifetime of WSN nodes is the energy that powers them. For WSN networks, data transmission consumes the majority of the power. This needs the development of algorithms to transmit data from one node to another and ultimately to the gateway. The term for this type of transmission is "multi-hop transmission". The power consumption of a radio broadcast grows exponentially with distance. Nodes of a WSN can interact with one another and analyse, gather, and store data from sensors [3]. Wireless nodes can use a variety of transmission methods, including radio, laser, ultrasound, and infrared. The vast majority of WSN applications run on radio frequencies, often in the 433 MHz or 2.4 GHz range. In WSN networks, a single transceiver structure serves both the transmitting and receiving purposes. Transmission, reception, idle, and sleep are the fundamental modes of the transceiver. The battery life is significantly reduced while transitioning from sleep to data packet transmission mode, as each mode has its own particular power usage. Figure 1 depicts the architecture of a WSN node.

Figure 2 shows the five major components of a sensor node. A microprocessor receives data from the environment through one or more sensors. At its core, a wireless sensor node is a microprocessor. It processes all the information that comes in from the transceiver, sensors, or memory. Interacting with the natural world is the

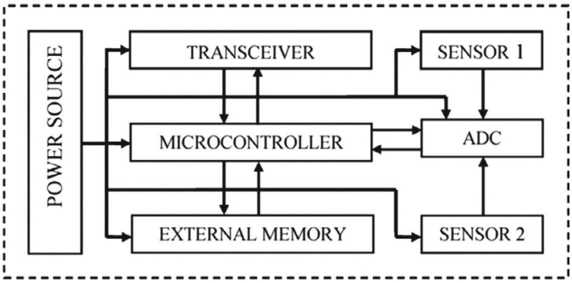


Fig. 1 WSN node architecture [3]

job of a transceiver. To transmit data wirelessly, it makes use of radio frequency (RF). In both directions, the transceiver may receive data from a microprocessor and transmit it wirelessly. The primary means of storing programmes and intermediate data received from sensors or the transceiver is a memory [4]. The sensor's intended use determines the memory's capacity. The energy for all components comes from the battery. Energy efficiency across the board is critical to guaranteeing a long enough network lifetime. Even while most sensors need a standard battery, there is preliminary work on developing battery-free sensors that use the same technology as passive RFID chips. It is common practice to disperse the sensor nodes throughout a sensor field. They may all gather information and send it on to the main hub. Internet or satellite communication between the base station and the task manager node is an option. Depending on the needs of the application, WSN has evolved with different designs and node distribution methodologies. The scientists zeroed down on five different deployment techniques for sensor networks: random, grid, group-based, and grid-group. One significant class of sensor network applications is environmental and habitat monitoring. The technological prerequisites for the construction of multifunctional, miniature sensor devices that can detect and observe physical events have been recently established by developments in low-power wireless network technology. There are several WSN examples in the literature that have been created for various applications, including environmental monitoring. An environmental monitoring system that utilises WSNs is conceptualised in Fig. 3.

Deploying WSNs, which consist of several disposable sensor nodes dispersed across a target region, provides an alternate method. When it comes to pollution monitoring, WSNs have a number of benefits, such as being easily deployable and saving time and money by not requiring power, wiring, or any of that associated infrastructure. Due of the energy consumption of petrol sensors, energy efficiency is a major consideration in these types of networks. Also, sensor nodes can run on batteries for months, if not years, without human intervention. The routing protocols and data-gathering algorithms have received special attention because of the potential differences between WSNs and conventional networks, and because energy consumption is a major factor in WSNs, researchers have created several protocols specifically for these networks [5]. The need to achieve optimal network performance with minimal energy consumption has spurred the recent proposal of several gathering systems for use in WSNs, where data collection is a crucial role. Clustering is

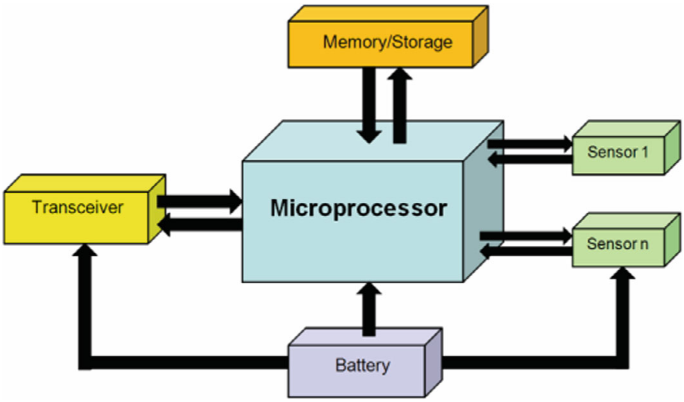


Fig. 2 Hardware components of a sensor node [4]

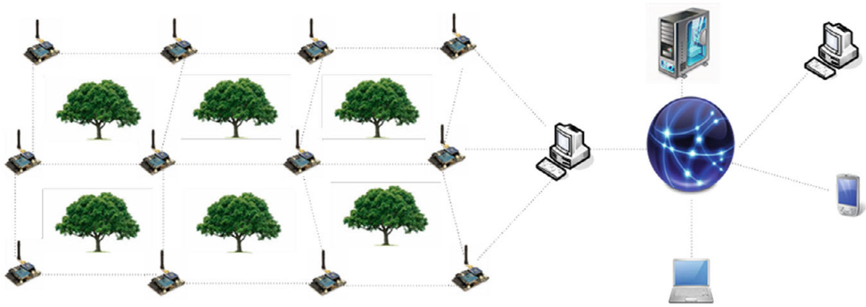


Fig. 3 Concept of real-time environmental monitoring using WSNs [5]

one of the ways to collect data from WSNs that uses the least amount of energy. One well-known method that allows a WSN to run very efficiently on energy is cluster-based protocols. To begin with, sensor nodes are grouped into groups called clusters. Then, each cluster chooses a node to act as its cluster head (CH), whose job it is to gather data from all the nodes in that cluster, compress it, and send it on to the sink node. Data transmission inside the network is reduced using cluster-based protocols. A number of clustering protocols, including LEACH, PEGASIS, LEACH-C, fuzzy C-means, and P-SEP, have been developed with the goal of extending the lifetime of the network. On the other hand, simulation experiments confirmed the efficacy of most of the previously used methods. Furthermore, certain methods quickly drain the energy of the nodes around the sink node since they need the transmission of several packets to them.

Wireless sensor network for precision agriculture application

The WSNs and, by extension, the whole system for controlling and monitoring the environment rely on sensors. It is easier to deploy wireless sensor technology in

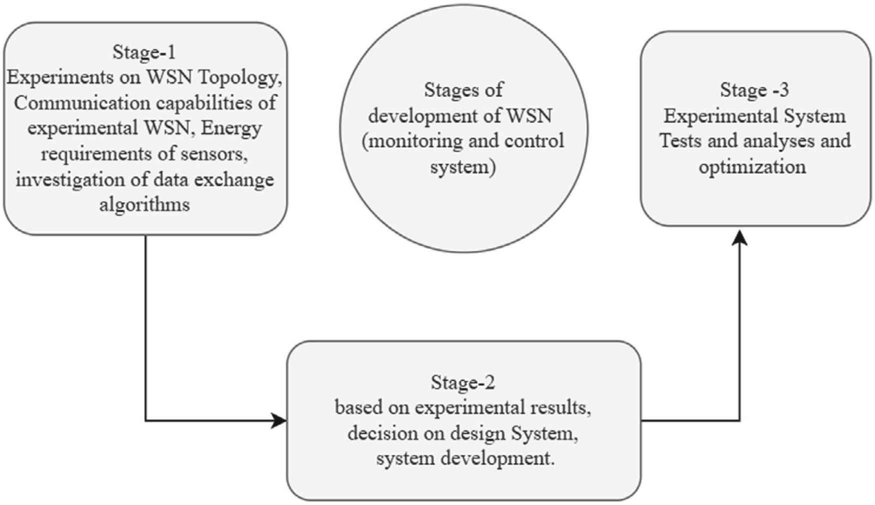


Fig. 4 Stages of development of WSN-based monitoring and control system

greenhouses than in outdoor applications since the crop conditions inside are mild and not heavily influenced by natural forces. The use of wireless sensor networks (WSNs) in precision agriculture entails placing nodes in green houses or gardens to collect data on environmental factors that impact the growth of crops [6]. A centralised base station receives the measured data wirelessly from the sensor nodes. Based on the decisions made by the base station, the nodes carry out the relevant tasks. Through WSNs, users may monitor product growth data and, with the help of an expert system, implement suitable management actions, such as controlling drip irrigation and fan facilities remotely, to enhance the product's microenvironment. Consequently, up-to-date precision agriculture relies on monitored data from large-scale sensor networks to determine the best course of action for controlling and adjusting environmental conditions. The goal is to maximise output yield while minimising resource use. To get exact control of the manufacturing setting, three things must be done: (1) keeping an eye on environmental conditions like temperature, humidity, and lighting since they have a significant impact on product quality and yield; (2) using this data to make decisions with optimisation help; and (3) putting the control mechanisms into action (Fig. 4).

2 Background

A server system for monitoring the agricultural environment was proposed by [2]. This system would use wireless sensor network (WSN) technology to collect data related to an outdoor agricultural production setting. The suggested agricultural environment monitoring server system included WSN-based sensors to collect data on the soil and outside environmental factors, CCTVs to capture images, and GPS modules to retrieve positions. In order to store the data, the agricultural environment monitoring server was used. This server had a sensor manager that dealt with data collected from WSN sensors, an image information manager that handled data collected from CCTVs, and a GPS manager that processed and provided producers with location information. A solar cell-based power source was also included into the server system to make it usable in agricultural environments with insufficient power infrastructure. This server system for monitoring the agricultural environment has the ability to enhance field quality and crop production by guiding farmers in making data-driven decisions. Even remotely, the device could keep tabs on data about the surroundings [2]. In [7], suggested a wireless sensor network design for a greenhouse growing vegetables; this would allow for more scientific cultivation while also reducing management expenses related to environmental monitoring. A wireless sensor network-based, practical, and inexpensive greenhouse monitoring system was developed to track critical environmental variables including lighting, humidity, and temperature in response to the characteristics of the greenhouse environment analysis. According to [8], many monitoring and tracking applications are driving the rising popularity of wireless sensor network (WSN) technology. Thanks to recent innovations in data processing and wireless sensor technologies, environmental management systems may now remotely monitor locations in real time. With the use of WSN, a plethora of environmental data, including as weather, air quality, plant health, and soil composition, could be tracked and managed, all of which had an impact on the conditions under which crops were grown. For precise depictions of present circumstances and risk assessments of the future, real-time sensor data collecting was used. Farmers might have a firm foundation to change strategy at any moment with the real-time data from the fields. A precision farming method takes into account individual variances and modifies management strategies appropriately rather than basing judgements on an idealised average situation that may not exist in the actual world. The software update downtime was minimised, and hardware could be reused thanks to modular design. It was possible to send the user a microcontroller module that was already constructed and programmed in case they needed a firmware update. The emergence of this subfield within precision agriculture is directly attributable to the recent proliferation of affordable wireless technologies. WSNs have the potential to outperform wired ones in several respects, including size, power, flexibility, affordability, and distributed intelligence, and they might function in a variety of situations. Multifunctional, low-power, and inexpensive sensor nodes have recently emerged as a result of innovations in wireless sensor networking. In addition to processing data, sensor nodes made it possible to sense the surroundings. Using data acquired from vineyard in-field sensors, this article outlined a framework for using WSN and a dynamic web application to track changes in environmental conditions, particularly microclimates. The tracking was meant to be a component of a bigger system that would include hardware and software that could gather data from the field even in the most challenging terrain. Data obtained from different monitoring stations in five countries was analysed and visualised using a database management system and related applications. Data from several sensor nodes, as well as sensors within the same node, as well as data from chosen sensors, may be compared and analysed. Data was received from the deployment locations successfully and could be analysed and streamed in real-time using web browsers. In [9], detailed the creation of a system for monitoring young sago palms in mist chambers using a wireless sensor network (WSN). Also covered was the functional system's impact on the measured environmental parameters. Temperature, relative humidity, and light intensity are the three environmental variables that the system was designed to track. A web portal connection and wireless network allowed for real-time readings of the monitored parameters. At 15-min intervals, the system communicated data from the node to the base over the Zigbee network, recording the environmental reading. After that, the GPRS gateway sent the information to a cloud storage platform. The mist chamber doors, which allow for natural ventilation, were controlled by linear actuators that were automatically activated by the threshold temperature. In addition, the system now has a smart alert system that, when activated, will send a message via SMS asking the user to do something. With just 2.3% packet loss rate (PLR), the created system reliably sent data. Due to the low monthly mobile data usage of 10 MB, the system was also able to run at a lesser cost thanks to the GPRS gateway. Finally, a WSN system was built for sago palm greenhouse mist chambers to capture, transmit, and display data in real time, allowing for environmental monitoring of the developing plants. A basis for smart farming was established by the created WSN system. According to [10], the use of wireless sensor networks (WSNs) has been accelerated in the agricultural sector due to the proliferation of Internet of things (IoT) devices that can detect changes in the environment and transmit that data wirelessly to farmers. The ability to monitor in real time via these WSNs enabled for smart decision-making to optimise yields while minimising costs. The context under consideration made the complex process of designing and implementing a WSN all the more difficult. In order to correlate obtained measurements, for instance, network synchronisation became necessary in these networks. A WSN that could support the sensing features of smart and precision agricultural applications was the primary emphasis of this work's design and implementation. In response to particular design needs prompted by the environment under consideration, a system was developed and put into operation. In order to provide time-correlated measurements using the sink node's clock as a reference, a straightforward synchronisation system was detailed. The usefulness of the suggested system in delivering a low-cost system that can get synchronised measurements was evaluated by installing it on an olive grove. An predicted modest discrepancy in the temporal correlation of the collected measures, largely due to serial transmission delays, was shown by the obtained findings, which demonstrated the system's overall efficacy and yielded a plethora of significant environmental parameters.

A valuable instrument that increased agricultural productivity was the monitoring of numerous interest criteria in a culture [11]. Several technologies might be used to accomplish crop monitoring in precision farming; however, the use of wireless sensor networks led to low-power installations, making them the most popular choice. Based on their findings, a new field monitoring system for agriculture might be created using sensors that could measure soil factors in addition to those that could measure air characteristics. A periodic hybrid routing method that is sensitive to the threshold for collecting environmental data was suggested in this chapter. By deploying sensor nodes using region-based cluster techniques, the suggested algorithm effectively covered the whole agricultural area. To further enhance energy efficiency in the WSN, we provide a clustering protocol that takes into account both residual energy and the distance between neighbouring nodes to get the ideal cluster head. Using metrics like energy usage, network longevity, and packet delivery, the simulation results demonstrated that the suggested routing method outperformed other popular algorithms. Significant environmental hazards, according to [12], include radioactive pollution, water waste, and air contamination. Sufficient oversight required in order to maintain a decent society while also guaranteeing economic sustainability. Thanks to advancements in the Internet of things and the development of improved detectors, environmental tracking has recently evolved into a smart environment monitoring (SEM) system. Significant accomplishments and scientific investigations on SEM, such as those pertaining to agricultural practices, radiation emissions, water management, and air quality control, were assessed in this context. This study put forward a framework for smart environment monitoring using wireless sensor networks and the Internet of things (WSN-IoT-SEC). We used SEM methods to partition the data, and then we looked at the detectors, ML models, and classifiers that were used for each goal. Based on the research's assessed outcomes and patterns, a systematic investigation was conducted, revealing crucial recommendations and the effect of the SEM analysis. Technological developments in mobile devices, the Internet of things (IoT), and wireless sensor networks have made it possible to regulate and monitor the environment intelligently, as the academics have described objectively. In the end, suitable WSN requirements were developed, and the idea of rigorous machine learning methodologies was put forward [12]. Wireless sensor networks (WSNs) have received a lot of attention from researchers in the field of networking in recent years [13], and they have several potential applications. The exponential expansion of WSN in the last few years has fallen short of fulfilling all requirements for environmental monitoring. In order to meet the demands of environmental monitoring, a wireless multimedia sensor network (WMSN) was necessary. This study details the use of WMSN for soil environmental monitoring in agriculture. When transmitting image, video, and scalar data, routing protocols with energy efficiency and quality-of-service (QoS) characteristics were essential for WMSN sensor node optimisation and data accuracy. This research not only identified pests but also suggested a priority-aware, energy-efficient quality-of-service data transfer method for use in agricultural soil environments. It made optimal use of the available bandwidth and data rate to reduce energy consumption while transmitting data. Experimental findings showed that the recommended algorithm outperformed competing approaches in terms of latency and energy usage in a performance study that used a Java simulation of the proposed system [13].

According to [14], worked in agriculture. Consequently, agricultural efforts had to prioritise improving crop quality and quantity in order to boost total cultivation output. Farmers had to put in a lot of time and energy using the old ways, which meant their crops were slower to mature. Overall agricultural productivity was negatively impacted since it was difficult to forecast weather patterns and pinpoint exactly where weeds, insects, etc., were present, necessitating prompt treatment. Therefore, it was necessary to automate it, and smart agriculture and precision agriculture (PA) offered the means to achieve this goal. One area that saw extensive use of wireless sensor networks (WSNs) was precision agriculture. These networks comprised of numerous sensors dispersed across the field that could detect and record changes in environmental factors like temperature, humidity, soil moisture, soil PH value, precipitation, water level, and so on. The goal was to improve crop yield, quality, quantity, and profitability. Machine learning methods have the potential to revolutionise precision agriculture by enhancing crop growth, streamlining crop cultivation, and cultivating optimal conditions for increased yield with little human intervention. With the use of wireless sensor networks, this article shed light on a number of machine learning approaches utilised in precision agriculture. A new level of Internet of things (IoT) technology has recently emerged in the agricultural and farming business [15], which is the precise network connection of sensors. Potentially useful for far-flung agricultural operations is the Internet of things (IoT) long distance network based on cloud computing and wireless sensor networks (WSNs). This study presents a design for an extensible wireless sensor network that may be used for remote farming and agricultural monitoring and control via the Internet of things. Water resource irrigation and efficient use of water resources were important management aspects of precision agriculture and farming (PAF). Applying WSN technology via the Internet of things might lead to more efficient water irrigation control. In order to boost farmers' output, PAF processed the efficient communication of different wireless sensors utilising the Internet of things (IoT). They have examined the WSN architecture with an eye on maximising throughput while minimising delay, achieving a high signal-to-noise ratio (SNR), minimising mean square error, and expanding coverage. The testing findings demonstrated that the suggested approach outperformed the status quo of traditional Internet of things (IoT) farming and agribusiness. In [16], listed air pollution, water contamination, and radiation contamination as three of the most significant environmental threats. In order for the globe to attain sustainable development while preserving a healthy society, appropriate monitoring was required. As the Internet of things (IoT) and new sensors have advanced, traditional environmental monitoring has evolved into a smart environment monitoring (SEM) system. In this context, the current publication set out to conduct a critical evaluation of significant contributions and research investigations on SEM, which included the monitoring of agricultural systems, radiation pollution, water quality, and air quality. The review was organised according to the tasks that used SEM methodologies. For each task, the sensors, machine learning approaches, and classification algorithms utilised were further examined. The comprehensive review came before the in-depth analysis, which, based on the outcomes of the debate and the patterns in the studied study, had proposed important suggestions and implications of SEM research. Recent developments in sensor technology, the Internet of things (IoT), and machine learning have transformed environmental monitoring into a very intelligent system, which the authors have thoroughly investigated. At long last, a framework for reliable machine learning techniques, denoising algorithms, and WSN standard development was proposed.

Several studies have been carried out over the last few decades to create wireless sensor networks (WSN) for a broad range of uses, including health care, environmental monitoring, agricultural, and military applications, among many others. Many research problems in this area of technology are now being addressed by academics and researchers. Problems that researchers encountered included developing energy-efficient data routing protocols, techniques for deployment, network topologies, and low-cost, low-power motes. While much work had been done in each of the aforementioned domains separately, there was a dearth of literature or research papers that addressed studies spanning many areas at once. Only a small fraction of the available literature describes the whole process of designing, developing, and deploying a WSN; the rest focuses on simulation work. From the initial concept to the accompanying applications, this research covered every facet of WSN design in great depth via a literature review. One of the most important applications for livelihood, crop production or the agricultural domain, was the subject of this article despite the fact that WSNs were being employed for monitoring purposes by many other applications [17] (Table 1).

3 Methodology

Designing a mathematical model for environmental monitoring in agriculture using wireless sensor networks (WSNs) involves capturing various parameters related to the agricultural environment. The model should consider factors such as soil moisture, temperature, humidity, light intensity, and possibly other relevant variables. Below is a simplified example of a mathematical model for monitoring soil moisture using WSNs.

Let us denote the following variables:

$M(t)$—Soil moisture, $I(t)$—Irrigation input, $E(t)$—Evaporation, $R(t)$—Rainfall, $S(t)$—Sensor reading at time t.

The soil moisture dynamics can be represented by the following equation:

$$M(t+1) = M(t) + I(t) - E(t) + R(t)$$

This equation represents the change in soil moisture over time, considering irrigation input, evaporation, and rainfall.

The sensor reading, $S(t)$, is related to the actual soil moisture, $M(t)$, with some measurement error ($N(t)$):

$$S(t) = M(t) + N(t)$$

Here, $N(t)$ represents the measurement noise.

Table1 Comparative review

References	Technology used	Area of research	Methodology	Conclusion
[2]	Wireless sensor network	Agricultural environment monitoring system	Environmental and soil, image data collection and location data. Implementation of solar cell-based power supply	Proposed an agricultural environment monitoring server system utilising WSN technology for outdoor agricultural production environment
[7]	Wireless sensor network (WSN)	Greenhouse environmental monitoring system	Design of a practical and low-cost greenhouse monitoring system based on WSN technology for key environmental parameters	Proposed a low-cost greenhouse monitoring system using WSN technology to monitor key environmental parameters
[8]	Wireless sensor network (WSN)	Environmental management system	Real-time monitoring and management of environmental data. Precision farming approach for accurate decision-making based on real-time sensor data	Proposed a framework WSN and a dynamic web application for real-time environmental monitoring
[9]	Wireless sensor network	Environmental monitoring system for sago palms	Monitoring of temperature, humidity, and light intensity in mist chambers. Cloud storage through GPRS. Linear actuators for mist chamber control	Developed a WSN system for greenhouse to monitor plant growing environment in real time
[10]	Wireless sensor network (WSN)	Precision agriculture monitoring system	Design and implementation of a WSN for smart and precision agriculture. Evaluation in an olive grove	Designed and installed a WSN for smart and precision agriculture applications, focusing on low-cost synchronised measurements

(continued)

Table1 (continued)

References	Technology used	Area of research	Methodology	Conclusion
[11]	Wireless sensor network (WSN)	Agricultural field monitoring system	Deployment of sensor nodes based on region-based clustering. Clustering protocol for optimal cluster head and improved energy efficiency	Proposed a hybrid routing algorithm for WSN-based agricultural monitoring, showing superior performance in simulations
[12]	Wireless sensor network (WSN)	Smart environment monitoring system	Proposal of WSN-IoT-SEC framework for SEM. Systematic study on SEM techniques, detectors, machine learning models, and classifiers	Proposed a WSN-IoT integrated system for smart environment monitoring, discussing the application of various ML techniques
[13]	Wireless multimedia sensor network	Agriculture soil environment monitoring system	Proposed priority-aware energy-efficient QoS-based data transmission for soil and detection of pests	Priority-aware energy-efficient QoS-based data transmission for agriculture soil and environment monitoring using WMSN
[14]	Wireless sensor network (WSN)	Precision agriculture and intelligent farming	Overview of machine learning techniques for precision agriculture using WSN	Provided insight into various ML techniques used for precision agriculture using WSNs to enhance crop productivity
[15]	Wireless sensor network (WSN)	Remote area agriculture and farming monitoring	Proposed scalable WSN architecture for monitoring and control in PAF. Analysis based on throughput maximisation, latency minimisation, etc.	Proposed a scalable WSN architecture for agriculture and farming in remote areas, emphasising water resource management

(continued)

Table1 (continued)

References	Technology used	Area of research	Methodology	Conclusion
[16]	Internet of things (IoT), cloud computing, WSN	Smart environment monitoring system	Critical review of SEM research, analysis of sensor technology, IoT, and machine learning methods. Proposed framework for robust machine learning methods and denoising methods	Reviewed and analysed SEM methods and their applications in monitoring air quality, water quality, radiation pollution, etc.
[17]	Wireless sensor network	Wide variety of applications including agriculture	Literature review on WSN design, deployment strategies, network topologies, and energy-efficient data routing protocols	Presented a comprehensive literature review of various aspects of WSN design and its applications in agriculture and environment

The wireless sensor network may also have constraints or limitations, such as battery life, communication range, and data transmission delays. These factors can be incorporated into the model as well, depending on the specific characteristics of the WSN.

4 Conclusion and Future Work

This paper has highlighted the significant potential of wireless sensor networks (WSNs) in revolutionising environmental monitoring within agriculture. Through a comprehensive analysis of existing literature, we have identified the numerous advantages offered by WSNs, including real-time data collection, cost-effectiveness, and scalability. Additionally, the paper has outlined various applications of WSNs in agriculture, such as soil moisture monitoring, crop health assessment, and environmental parameter tracking. These applications not only enable farmers to make informed decisions but also contribute to sustainable agricultural practices by optimising resource utilisation and minimising environmental impact. Moreover, the review has emphasised the importance of addressing challenges such as energy efficiency, data security, and interoperability to fully realise the potential of WSNs in agricultural monitoring.

Moving forward, several avenues for future research and development in the field of environmental monitoring using WSNs have been identified. Firstly, there is a need for the continued advancement of sensor technologies to enhance the accuracy, reliability, and durability of WSNs deployed in agricultural environments. This includes the development of sensors capable of measuring additional parameters relevant to agriculture, such as nutrient levels and pest infestations. Secondly, efforts should be directed towards improving the energy efficiency of WSNs through the exploration of alternative power sources and energy harvesting techniques. Additionally, research focusing on the integration of WSNs with other emerging technologies, such as Internet of things (IoT) platforms and artificial intelligence, holds promise for enhancing the functionality and intelligence of agricultural monitoring systems. Furthermore, there is a growing need for the development of standardised protocols and frameworks to facilitate interoperability among heterogeneous WSNs, enabling seamless data exchange and integration across different agricultural applications. Lastly, addressing concerns related to data security and privacy will be crucial to fostering widespread adoption of WSNs in agriculture, necessitating the development of robust encryption and authentication mechanisms. Overall, the future work outlined in this review paper underscores the importance of continued innovation and collaboration across multidisciplinary fields to unlock the full potential of WSNs in environmental monitoring within agriculture.

References

1. Rehman A, Jingdong L, Khatoon R, Hussain I, Iqbal MS (2016) Modern agricultural technology adoption its importance, role and usage for the improvement of agriculture. Life Sci J 14(2):70–74
2. Hwang J, Shin C, Yoe H (2010) Study on an agricultural environment monitoring server system using wireless sensor networks. Sensors 10(12):11189–11211
3. Coca E, Popa V, Buta G (2009) Wireless sensor network nodes performance measurements and RSSI evaluation. In: 2009 15th International symposium for design and technology of electronics packages (SIITME). IEEE, pp 105–112
4. Vujic DA (2015) Wireless sensor networks applications in aircraft structural health monitoring. J Appl Eng Sci 13(2)
5. Alhmiedat T (2017) Low-power environmental monitoring system for zigbee wireless sensor network. KSII Trans Internet Inf Syst 11(10)
6. Jawad HM, Nordin R, Gharghan SK, Jawad AM, Ismail M (2017) Energy-efficient wireless sensor networks for precision agriculture: a review. Sensors 17(8):1781
7. Srbinovska M, Gavrovski C, Dimcev V, Krkoleva A, Borozan V (2015) Environmental parameters monitoring in precision agriculture using wireless sensor networks. J Clean Prod 88:297–307
8. Ghobakhlou A, Zandi S, Sallis P (2011) Development of environmental monitoring system with wireless sensor networks
9. Kho EP, Chua SND, Lim SF, Lau LC, Gani MTN (2022) Development of young sago palm environmental monitoring system with wireless sensor networks. Comput Electron Agric 193:106723
10. Zervopoulos A, Tsipis A, Alvanou AG, Bezas K, Papamichail A, Vergis S, Stylidou A, Tsoumanis G, Komianos V, Koufoudakis G, Oikonomou K (2020) Wireless sensor network synchronization for precision agriculture applications. Agriculture 10(3):89
11. Khelifi F (2020) Monitoring system based in wireless sensor network for precision agriculture. In: Internet of things (IoT) concepts and applications, pp 461–472
12. Yu Q, Xiong F, Wang Y (2022) Integration of wireless sensor network and IoT for smart environment monitoring system. J Interconnection Netw 22(supp 02):2143010

13. Manjula E (2021) An agricultural soil environment monitoring using wireless multimedia sensor network. Turk J Comput Math Educ (TURCOMAT) 12(10):5533–5540
14. Goel S, Guleria K, Panda SN (2022) Machine learning techniques for precision agriculture using wireless sensor networks. ECS Trans 107(1):9229
15. Sanjeevi P, Prasanna S, Siva Kumar B, Gunasekaran G, Alagiri I, Vijay Anand R (2020) Precision agriculture and farming using Internet of Things based on wireless sensor network. Trans Emerg Telecommun Technol 31(12):e3978
16. Ullo SL, Sinha GR (2020) Advances in smart environment monitoring systems using IoT and sensors. Sensors 20(11):3113
17. Kumar P, Reddy SRN (2020) Wireless sensor networks: a review of motes, wireless technologies, routing algorithms and static deployment strategies for agriculture applications. CSI Trans ICT 8(3):331–345

An Efficient Parallel Barcode Recognition System

Chidanand S. Kusur, M. S. Shirdhonkar, and M. M. Bachalapur

Abstract In computer vision and image processing, object detection and recognition is the process of detecting specific objects from input digital images or videos. It has a wide range of applications like surveillance systems, automated vehicle systems, machine inspection, content-based image retrieval, face detection, pedestrian detection, recognizing fire from forest, etc. In our proposed work, a novel time efficient parallel approach has been applied for library stock verification, i.e., detect and recognize barcodes pasted on the cover page from book images. Barcodes are recognized in parallel by mapping the pool of workers for the barcode detection. The experimental results show that the time complexity of both sequential and parallel execution is lesser than the time required for the manual scan of barcodes from book images.

Keywords Object detection and recognition · Barcode recognition · Parallelism and time complexity

1 Introduction

In today's era, barcodes play a crucial role in various aspects of business and daily life due to their numerous advantages and benefits, i.e., for the efficient inventory management, faster and more accurate data entry, enhanced supply chain visibility,

C. S. Kusur (✉) · M. S. Shirdhonkar · M. M. Bachalapur
Department of Computer Science and Engineering, B.L.D.E.A's V. P. Dr. P. G. H. College of Engineering and Technology, Vijayapur, Karnataka, India
e-mail: cs.kusur@bldeacet.ac.in

Visvesvaraya Technological University, Belagavi 590018, India

M. S. Shirdhonkar
e-mail: cs.malakappa@bldeacet.ac.in

M. M. Bachalapur
e-mail: bachalapur@bldeacet.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_21

and helps to track products from the manufacturer to the end consumer. It helps to improve customer service for the better data analytics, etc. [1]. A barcode is a method of representing data in a visual and machine-readable form. It consists of a series of parallel lines or geometric patterns with varying widths and spaces, which can be scanned and interpreted by optical scanners or barcode readers. We find two types of barcodes, namely 1D and 2D barcodes. Universal Product Code (UPC) and European Article Number (EAN) are the most commonly used 1D barcodes, and quick response (QR) codes are the examples for 2D barcodes [2]. We find UPC barcodes on retail products that typically consists of 12 numeric digits and is used for inventory management, pricing, and point-of-sale transactions, and QR codes can store more data than traditional barcodes and can represent alphanumeric characters. They are often used in marketing, advertising, ticketing, and mobile payment applications and in many applications.

In our proposed work, both the sequential and parallel techniques are applied for the recognition of barcode from images of books present in a rack for the annual library stock verification. A book barcode is a scannable version of the International Standard Book Number (ISBN). Each book barcode contains a unique set of shaded lines that stores the information about a book such as a book accession number, i.e., linked to other details of a book like its title, author, edition, publisher, and price integrated with library automation software.

The conventional method used for the library stock verification requires more number of human resources for the manual scanning of barcodes from each book from a rack by using handheld scanner, i.e., to get total count of books present in a library, which is more time-consuming processes.

Most of the image processing algorithms are based on random access machine (RAM) model that executes the instructions sequentially by using a single CPU. These sequential algorithms are not well suited for today's modern multi-core systems and heterogeneous architecture of a machine. Therefore, there is a need to design and develop parallel algorithms that execute instructions based on the parallel random access machine model (PRAM) to reduce the time complexity of computationally intensive image processing steps which can be achieved by using multi-core CPUs with other processors like GPGPU, DSP, FPGA, etc.

The paper is organized into four sections including introduction section. In Sect. 2, we discuss the literature review and its summary. In Sect. 3, we discuss the proposed methodology for the barcode recognition. In Sect. 4, we discuss the experimental setup and results. Finally, paper has been concluded.

2 Literature Survey

The history of barcode begins since 1940, Silver and Woodland [3] proposed technique to create a system to recognize product information for the automated data collection and identification.

In the 2006, Rajendiran and Bhushan [4] described the use of barcode technology and its recognition for the atomization of library stock verification, i.e., they concluded that the barcode technology is reliable and easy to use that helps to improve accuracy and accelerates the time required for library stock verification and helps to eliminate the typing errors and delay in manual stock verification process.

In the 2010, Yahyanejad and Strom [5] proposed barcode deblurring algorithm to remove the translator motion from a blurred barcode image. The algorithm works on the bimodal characteristics of barcode image histograms that construct a simple target function that measures how similar a deconvoluted image is to a barcode. Authors used a collection of a 138 barcode images with varying camera blur and found increase in the recognition rate from 32 to 65%.

In the 2017, Hansen et al. [6] proposed deep learning-based detector of You Only Look Once (YOLO) for the purpose of detecting both 1D and QR barcodes in a fast and reliable way from images from the dataset of Muenster BarcodeDB and achieved the detection rate of 0.991.

In the 2019, Islam et al. [7] proposed a model for the multiple barcode detection from an image in business system that detects 1D barcodes such as EAN-13 barcodes and Code-128 barcodes that helps to detect multiple barcodes at once from multiple products that saves huge time of individual scan. Authors used Zamberletti algorithm using OpenCV and ZBar library to detect and decode the barcodes from the input image.

In the 2021, Zhang et al. [8] proposed the fast barcode detection based on ThinYOLOv4model that aims to resolve the barcode detection problem for datasets with different shortcomings, including images with blur, reflection, and high density problem and found the significant improvement in the processing speed of barcode detection. A barcode dataset consisting of 16,545 images were used for the experiment and obtained the proposed method achieves a recall rate of 93.8% on the provided dataset.

In the 2021, Quenum et al. [9] proposed a new approach for the detection of barcode, i.e., semantic segmentation to achieve a fast and accurate detection of barcodes of various scales in ultra high-resolution (UHR) images has long been a challenging problem in computer vision due to the varying scales of the targeted objects. The concept of pipeline was used with a modified region proposal network (RPN) on images of size greater than 10 k $\times$ 10 k and a newly proposed Y-Net segmentation network for barcode detection from images. The proposed system has a latency of 16 ms, which is 2:5_ faster than YOLOv4 and 5:9_ faster than mask R-CNN.

In the 2021, Huo et al. [10] proposed improved artificial intelligence algorithms for the recognition of QR codes from distorted QR image due to uneven background fluctuations, inadequate illuminations, and distortions due to the improper image acquisition method. They proposed improved adaptive median filter algorithm and a QR code distortion correction method based on backpropagation (BP) neural networks. This combination of artificial intelligence algorithms is capable to identification of QR codes that are more severely distorted, especially those printed on objects that are prone to wrinkles. The proposed method provides a significant improvement of 14%.

In the 2021, Thao Do et al. [11] proposed a model for decoding underexposed, occluded, blurry, wrinkled, and rotated barcode in wilder conditions by using deep convolutional neural network that runs on portable devices. Researchers proposed a special modification of inference based on the feature of having checksum and test-time augmentation, named smart inference (SI), in the prediction phase of a trained model. It boosts accuracy up to 95.85% and reduces the false prediction for trained models.

In the 2022, Wudhikarn et al. [12] reviewed the latest advances and progress in barcode recognition using deep learning (DL) techniques that provides better accuracy and speed for the larger dataset of barcodes, and the proposed DL model needs to have multitask learning capabilities to deal with different types of barcodes with varying size and the ability to learn from real-time data. The researchers concluded that the future DL-based barcode recognition techniques should be improved to meet actual industrial requirements or real-life usages. It should be boosted to ensure a 100% success rate.

In the 2022, Dai and Hsieh [13] proposed new model using a deep learning classifier based on YOLACT object detection network for the recognition of discount stickers and barcodes from products. Researchers improved the training performance of deep learning model by using a novel data augmentation approach to generate various data that are close to the real-time scene. The data augmentation consists of two steps, i.e., to individually augment discount stickers and barcodes with different status such as distortion, blur, and complex background and to paste the generated discount stickers or barcodes or both to complex backgrounds and then augment images with various real lighting conditions.

The conclusion of literature review is that there is a need of both sequential and parallel code for the detection and recognition of barcodes from images to get details associated with it instead of manual scan using handheld barcode scanner. Hence, we proposed both sequential and parallel programming techniques for the barcode recognition from images.

3 Proposed Methodology

In the proposed study, we used two programming techniques called sequential and parallel for barcode recognition from images. They are as follows.

3.1 Sequential Approach for Barcode Recognition

The sequential approach makes the use of a single CPU for the barcode detection as shown in Algorithm 3.1. The core steps involved in the sequential algorithm are as follows.

1. The first step is to define empty *list* in Python to store recognized book's accession number after decoding barcode from images.
2. Next step is to pass the all **.jpeg* images in a sequence from a current folder to user-defined function *BarcodeReader(img)* that makes the use of *decode(img)* method from *pyzbar* module form *OpenCV* library for the detection of barcode.
3. If found, decoded accession number will be appended into the globally declared *list* data type.
4. The total elapsed time for the recognition of each barcode is recorded in microseconds (10^{-6})
5. Finally, extracted accession numbers from *list* data type are exported to MS Excel sheet for further use for the library stock verification.

Algorithm 3.1 Sequential algorithm for barcode recognition

```
SEQUENTIAL ALGORITHMBarcodeRecognition(img)
//Sequential Algorithm for the barcode recognition from set of images
//Input: Multiple captured *.jpg files are stored in a folder
//Output: Codes list data structure containing recognized barcodes
Begin
Set: path = "C:\\Users\\Hp\\...\\Python311\\ME35IMAGES\\*.jpg"
Set: startTime = datetime.datetime.now()
For file in path
BarcodeReader(img)
End For
Set: endTime = datetime.datetime.now()
Output: Print "Time taken for sequential execution in microseconds",endTime-startTime
Output: Excel Sheet = codes
End
codes = []# defining empty list to store records of recognized barcodes from a book
BarcodeReader(img)
img = cv2.imread('C:\\Users\\Hp\\...ME35IMAGES\\%s'%(image))
detectedBarcodes = decode(img)
if not detectedBarcodes:
print("Barcode not detected !")
return
else:
for barcode in detectedBarcodes:
print("Barcode Detected !")
if barcode.data! = "":
if barcode.data not in codes:
codes.append(barcode.data)
return
```

3.2 Parallel Approach for Barcode Recognition

The parallel approach makes the use of *multiprocessing* module from Python for the parallel execution of user-defined function named *BarcodeReader(img)* on multiple images from working folder. The proposed parallel algorithm as shown in Algorithm 3.2 makes the use of single instruction multiple data (SIMD) technique and pool of workers for the parallel execution of code. The core steps involved in the proposed parallel algorithm are as follows.

1. A pool of workers is used for the execution of user-defined function named *BarcodeReader(img)*in parallel on multiple images using 4 cores at once; i.e., *pool.map()* function distributes the tasks among the worker processes.
2. Finally, decoded barcodes are recorded in the globally shared *list* data structure and exported to MS Excel sheet for further use for the library stock verification.

Algorithm3.2 Parallel algorithm for barcode recognition

PARALLEL ALGORITHM*BarcodeRecognition(img)*
//Sequential Algorithm for the barcode recognition from set of images
//Input: Multiple captured *.jpg files are stored in a folder
//Output: Codes list data structure containing recognized barcodes

(continued)

(continued)

Algorithm3.2 Parallel algorithm for barcode recognition

```
Begin
Set: path = "C:\\Users\\Hp\\...\\Python311\\ME35IMAGES\\*.jpg"
Set: list_image = os.listdir(path)
Set: workers = os.cpu_count()
Set: startTime = datetime.datetime.now()
with Pool(workers) as p:
outputs = p.map(BarcodeReader, list_image)
endTime = datetime.datetime.now()
Output: print("Time Taken in microseconds (Parallel Execution):",
endTime - startTime)
# pseudocode to convert from nested to flat list
flatList = []
for element in outputs:
if type(element) is list:
for item in element:
flatList.append(item)
else:
flatList.append(element)
## pseudcode to remove duplicate items from list
flatList = list(set(flatList))
#flatList.remove(None)#pseudocode to remove partially extracted barcode from a book
for i in flatList:
if(len(i) > = 5):
continue
else:
flatList.remove(i)
print(flatList)
Output: Excel Sheet = codes
End
codes = []
Barcode Reader(img)
img = cv2.imread('C:\\Users\\Hp\\...ME35IMAGES\\%s'%(image))
detectedBarcodes = decode(img)
if not detectedBarcodes:
print("Barcode not detected !")
return
else:
for barcode in detectedBarcodes:
print("Barcode Detected !")
if barcode.data! = "":
if barcode.data not in codes:
codes.append(barcode.data)
return
```

Some of the built-in functions used in the proposed parallel algorithms are as follows.

os.cpu_count(): A function from the *os* module that returns the number of multi-cores in the system, i.e., 4 cores in our system. ***workers = os.cpu_count().***

Pool Class: The pool class in multiprocessing helps to handle number of processes to be executed. It helps to run multiple jobs per process. The ***multiprocessing*** module and the ***multiprocessing.Process*** class provide process-based concurrency. It is used to create a pool of worker processes.

pool.map(): This function has been used to distribute the tasks among the worker processes and to get the decoded barcode from image.

4 Experimental Setup and Results

The experimental setup for the proposed study is as follows.

4.1 Experimental Setup

The dataset used for the experiment includes the barcode images of 515 books out of 92,000 books from a central library. The books from five book racks titled ME34, ME35, ME32M CV89, and ME33 are used for the experiment of barcode recognition. The sample racks of book images are shown in Fig. 1.

4.2 Experimental Results

We listed number of rack-wise ground truth of barcodes present and number of barcodes recognized by both sequential and parallel approach with their accuracy of barcode recognition in the following Table 1. The overall accuracy of barcode detection and recognition for counting books from all the racks is 95.94%, i.e., 483 books were detected out of 515 by both the sequential and parallel algorithms, i.e., we obtained the overall accuracy of barcode recognition as 95.94%.

We can see the accuracy of 86.47% which is lesser than 100%, i.e., 13.53% of accuracy is lost for the rack ME34 because of varying thickness of books present in it with incorrect paste of barcodes and blurred barcodes. Which was intensely considered for the detection to check the accuracy of proposed methods. Similarly, we lost 6.77% of accuracy for the detection of books from rack ME32 because of smaller thickness of books present in the racks and distortion in the barcodes. Whereas in the remaining racks named ME35, CV89, and ME33, we got 100% count because of larger thickness of books and clear barcode images present on books. The rack-wise true count of barcodes and number of detected barcodes has shown graphically with its accuracy in the following Fig. 2.

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

Fig. 1 Sample racks of books

Table 1 Overall accuracy of barcode detection and recognition

Rack	Ground truth barcodes	Correctly detected barcodes	Accuracy (%)
ME34	170	147	86.47
ME35	135	135	100.00
ME32	133	124	93.23
CV89	51	51	100.00
ME33	26	26	100.00
Total	515	483	95.94

4.2.1 Analysis of Time Complexity for Barcode Recognition

The experimental results show that time required for the barcode detection and count of books is reduced by the proposed system. We observed the following two improvements by the proposed system.

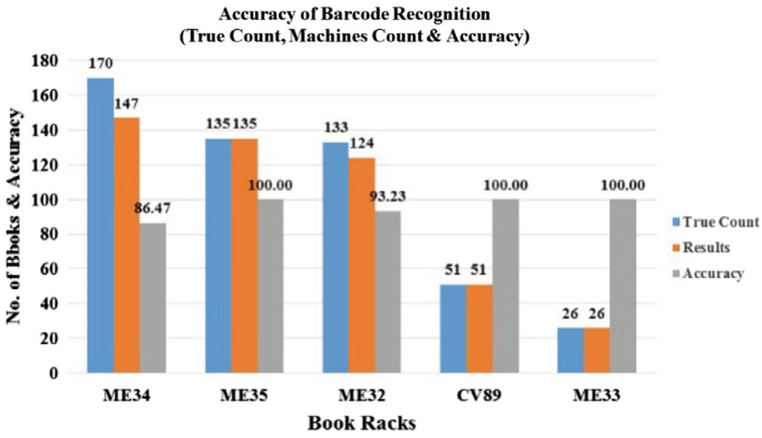


Fig. 2 True count versus system count for both sequential and parallel

- Time taken by both the sequential and parallel approach are lesser than the time required for manual scan of barcodes from books using handheld scanner.
- Time taken by parallel approach is lesser than the time taken by sequential approach

The actual speedup has been calculated by using the following Eq. (1) to find the performance difference between the proposed sequential and parallel algorithms.

$$\text{Speedup} = \frac{t_s}{t_p} \tag{1}$$

where t_s is execution time on single processor, and t_p is execution time on multiprocessor.

In our experiments, for 4 cores the actual speedup gained is 10.09/4.73 which is equivalent to 2.13 for the rack named ME34. Similarly, we calculated the speedup for remaining rack of books and listed in Table 2.

In addition to actual speedup, we calculated the maximum speedup that can be achieved for 4 cores using Amdahl's law as shown in Eq. (2).

Table 2 Time complexity and speedup

Rack	No. of books	t_s	t_p	Actual speedup	Overall speedup by Amdahl's law	Performance difference
ME34	170	10.09	4.73	2.13	2.12	+ 0.01
ME35	135	7.66	2.92	2.62	2.12	+ 0.50
ME32	133	6.91	3.4	2.03	2.12	– 0.09
CV89	51	2.82	2.12	1.33	2.12	– 0.79
ME33	26	1.10	1.74	0.63	2.12	– 1.49

$$\text{Speedup}(f, n) = \frac{1}{(1 - f) + \frac{f}{n}} \tag{2}$$

where f = parallelizable fraction of a program and n = number of processors

In our case, 50% portion of code has been parallelized by using 4 cores, i.e., $f = 0.50$ and $n = 4$. So, the maximum speedup that can be achieved is 2.12.We observed that the actual speedup gained by our experiments is very much closer to the maximum speedup that can achieved on 4 cores as shown in Table 2.

We obtained positive performance gains for the barcode recognition of books from ME34, ME35, and ME32 racks because of more number of books present in the racks. Whereas, for the last two racks named CV89 and ME33 got negative performance gain because of less number of books present in the racks.

The time taken for the barcode recognition from each rack of books has been shown in the following Fig. 3. We observed that the time taken by parallel execution is lesser than the time taken for sequential execution.

The time complexity of proposed sequential programming technique is $O(n)$, i.e., the basic operation is calling user-defined function named *BarcodeReader(image)* to pass N images in total from the current folder. So, we calculated the total count for the number of times the basic operation is executed. The mathematical analysis is as follows.

$$C_{(n)} = \sum_{i=1}^{n} \text{BarcodeReader(img)}$$

$$C_{(n)} = \sum_{i=1}^{n} 1$$

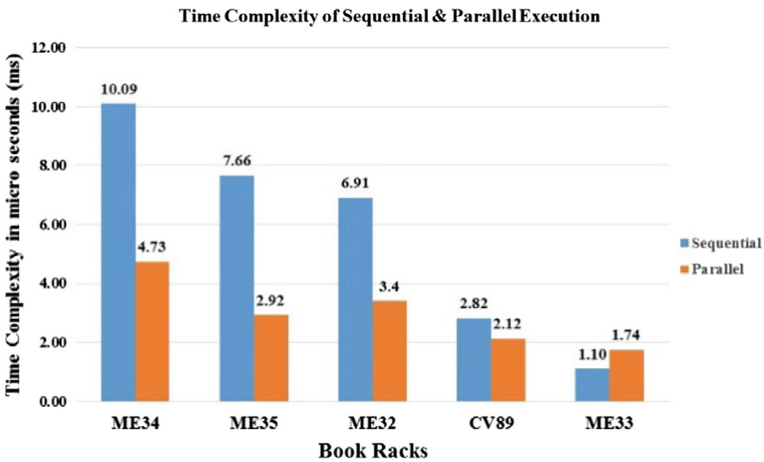


Fig. 3 Time complexity for the barcode recognition

$$C_{(n)} = n - 1 \in (n)$$

$\therefore$Time complexity of sequential algorithm is $O(n)$.

Similarly, we calculated the time complexity of proposed parallel programming technique as $O(\log_4 n)$. Here, the user-defined function *BarcodeReader(img)*is executed in parallel by 4 multi-cores at once. Hence, the total count of basic operation is calculated as follows.

$$C_{(n)} = \log_4 n$$

$\therefore$Time complexity of parallel algorithm is $O(\log_4 n)$.

Herewith, we conclude that time complexity of parallel execution is $O(\log_4 n)$ which is better than the time complexity of sequential execution, i.e., $O(n)$for the barcode recognition from book images.

5 Conclusion

The experimental results shown that time required for barcode recognition from images will be faster for both sequential and parallel execution of code in comparison with physical barcode recognition from books using barcode reader. This work can be further extended to extract barcodes from video frames of recorded video of the rack of books by extracting a key frame from a video clip.

Acknowledgements I want to express my sincere gratitude to all the central library staff of BLDEA's Vachana Pitamaha Dr. P.G. Halakatti College of Engineering and Technology, Vijayapur, Karnataka, for their continuous guidelines and support to do experiments for the recognition of barcodes from textbooks.

References

1. McCathie L (2004) The advantages and disadvantages of barcodes and radio frequency identification in supply chain management. Honours Theses, Faculty of Engineering & Information Sciences 9
2. ByteScout Team of Writers (2003) What information does a barcode hold? https://bytescout.com/blog/barcode-information.html
3. Silver B, Woodland NJ (1949) Classifying apparatus and method. Patent Filed on October 20
4. Rajendiran P, Bhushan I (2006) Automated library stock verification with barcode and LibSys. DESIDOC Bull Inf Technol 26(5):17–21
5. Yahyanejad S, Strom J (2010) Removing motion blur from barcode images. In: IEEE computer society conference on computer vision and pattern recognition—workshops. San Francisco, CA, USA, pp 41–46

6. Hansen DK, Nasrollahi K, Rasmussen CB, Moeslund TB (2017) Real-time barcode detection and classification using deep learning. In: Proceedings of the 9th international joint conference on computational intelligence (IJCCI). Funchal, Madeira, Portugal, pp 321–327
7. Islam A, Sharmin M, Sakib N (2019) A study on multiple barcode detection from an image in business system. Int J Comput Appl 181(37):30–37
8. Zhang L, Sui Y, Zhu F, Zhu M, He B, Deng Z (2021) Fast barcode detection method based on ThinYOLOv4. In: Cognitive systems and signal processing, Zhuhai, China. Springer, Singapore, pp 41–55
9. Quenum J, Wang K, Zakhor A (2021) Fast, accurate barcode detection in ultra high-resolution images. In: IEEE international conference on image processing (ICIP), pp 1019–1023
10. Huo L, Zhu J, Singh PK (2021) Pavlovich and Pljonkin Anton. Research on QR image code recognition system based on artificial intelligence algorithm. J. Intell. Syst. 30(1):855–867
11. Do T, Tolcha Y, Jun T, Kim D (2021) Smart inference for multidigit convolutional neural network based barcode decoding. In: International conference on pattern recognition (ICPR). Milan, Italy, pp 3019–3026
12. Wudhikarn R, Charoenkwan P, Malang K (2022) Deep learning in barcode recognition: a systematic literature review. IEEE Access 10:8049–8072
13. Dai X, Hsieh Y (2022) Image-based discount and barcode recognition system. In: 12th International conference on artificial intelligence, soft computing and applications. Sydney, Australia, pp 01–07

Predictive Analytics in the Healthcare Industry: Machine Learning for Early Diagnosis

Rajendhar Reddy Gaddam, S. Dhanalakshmi, B. Pradeep, Balusupati Anil Kumar, Pirangi Vijay Kumar, and Talari Swapna

Abstract When it comes to early diagnosis in particular, predictive analytics—made possible in large part by machine learning—has the potential to radically alter the healthcare system. Predictive analytics has many uses in healthcare, but this review paper mainly discusses how it might help with early diagnosis. Using the possible advantages of using machine learning algorithms for predictive modeling, we address the significance of early diagnosis in enhancing patient outcomes and decreasing healthcare expenses. Early diagnosis across many medical problems is investigated using a variety of machine learning techniques, such as supervised, unsupervised, and deep learning approaches. We also look at the difficulties of integrating predictive analytics into preexisting clinical workflows, issues with data protection, and the interpretability of models. We conclude by outlining potential avenues for further study and development in this dynamic area, with an emphasis on the importance of data scientists, healthcare providers, and legislators working together to fully utilize predictive analytics for healthcare early diagnosis.

Keywords Machine learning · Healthcare · Early diagnosis · Predictive analytics

---

R. R. Gaddam (✉) · S. Dhanalakshmi · B. Pradeep · B. A. Kumar · P. V. Kumar · T. Swapna
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: rajendarreddy.gaddam@cmritonline.ac.in

B. Pradeep
e-mail: pradeepb26@cmritonline.ac.in

B. A. Kumar
e-mail: anilkumarb@cmritonline.ac.in

P. V. Kumar
e-mail: vijaykumar@cmritonline.ac.in

T. Swapna
e-mail: swapnat@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_22

1 Introduction

Machine learning methods are used in predictive analytics within the healthcare business [1] to facilitate early illness detection. Gathering various types of healthcare data, such as patient records [2], test findings, and imaging data, from different sources is the first step. In preparation for analysis, this data is cleaned, has missing values handled, and normalized by preprocessing. Feature selection considers things like correlation and domain expertise to find the variables that are important for prediction.

The data is prepared and machine learning models like logistic regression, neural networks, decision trees, etc., are trained on it. After dividing the data into a training and validation set, hyperparameter tuning methods like cross-validation are used to train the model. To find out how well the trained models work, we measure them using things like $F1$-score, recall, accuracy, and precision.

A. Predictive Analytics

Predictive analytics is a subfield of advanced analytics that aims to foretell future occurrences or actions by combining past data with statistical algorithms and machine learning methods [3]. It entails looking at both recent and past data to see what might happen or what trends might emerge.

The goal is to identify patterns and relationships within the data that can be used to make predictions and inform decision-making.

The concept of predictive analytics involves several key components:

- Data Collection: Gathering relevant data from various sources, including structured and unstructured data, such as databases, sensor data, social media, and more.
- Data Preprocessing: Cleaning, transforming, and preparing the data for analysis by removing inconsistencies, handling missing values, and standardizing formats.
- Model Building: Develop statistical models or machine learning algorithms using historical data to predict future outcomes. This step involves selecting appropriate algorithms, training the models, and optimizing their performance.
- Validation and Testing: Evaluate the accuracy and reliability of predictive models using validation techniques such as cross-validation and testing them on unseen data.
- Deployment: Implement the predictive models into operational systems or decision-making processes to generate insights and make predictions in real-time.

B. Types of Predictive Analytics Techniques

- Regression Analysis: Uses a set of predictor variables to make a prediction about a continuous outcome variable.

- Classification Analysis: Predicts categorical outcomes by assigning observations to predefined categories or classes.
- Time-Series Analysis: Predicts future values based on past observations of a time-dependent variable.
- Machine Learning Algorithms: Techniques such as decision trees, random forests, support vector machines, neural networks, and ensemble methods can be used for predictive modeling.
- Clustering Analysis: Identifies natural groupings or clusters within data and predicts the cluster membership of new observations.

C. Predictive Analytics Has Numerous Applications in Healthcare, Including:

- Disease Prediction and Diagnosis: Predictive models can be used to classify individuals at risk of developing diseases or conditions based on their demographic, genetic, and routine factors. They can also aid in diagnosing diseases by analyzing symptoms and medical history data.
- Hospital Readmission Prediction: Predictive analytics may assist healthcare practitioners in identifying patients at risk of readmission to hospitals, allowing for preemptive treatments to reduce readmissions and enhance patient outcomes.
- Personalized Treatment Plans: Predictive models can analyze patient data to tailor treatment plans and interventions based on individual characteristics, preferences, and predicted outcomes.
- Healthcare Resource Management: Predictive analytics can optimize healthcare resource allocation by forecasting patient demand, predicting equipment maintenance needs, and optimizing staff schedules.
- Drug Discovery and Development: Predictive analytics can aid in drug discovery by analyzing molecular data to identify potential drug candidates and predicting their efficacy and safety profiles.

2 Literature Review

A. Predictive Analytics in Healthcare

This research presents a model for improved diabetes categorization based on glucose, age, gender, blood pressure, sugar, red blood cells, hemoglobin, blood urea, and a few external diabetes markers. There are 250 variations in our dataset, and each one has 16 distinct characteristics. For this forecast, we have used random trees, support vector machine, and logistic regression. A total of three models—logistic regression (94.5% accuracy), support vector machine (96.5% accuracy), and random tree (97.5% accuracy)—were trained using data submitted via tenfold cross-validation [4].

Predictive analytics, which use machine learning approaches to uncover improved decision-making, is a hot topic of study right now. The best possible results may be predicted from health factors using big data analysis, which opens exciting new possibilities for the future of healthcare. Their methodology for disease prediction was the big data predictive analytics model with the naive Bayes technique (BPA-NB). With the premise of feature independence, it offers probabilistic classification based on Bayes' theorem. When dealing with massive datasets, the naive Bayes method works well, particularly with big data. Heart disease data extracted from the UCI machine learning repository is used to train the naive Bayes algorithm. The next step was for it to use the test data to make classification predictions. The findings show that when it comes to predicting the illness rate, the suggested BPA-NB scheme offers a much greater accuracy of around 97.12%. To get substantial insight from healthcare data, the suggested BPA-NB system used Hadoop–Spark, a big data computing platform. Predicting the future health state of various individuals is the goal of the studies. To estimate the health parameters needed for classification, the training dataset is used. To predict their future health, the data demonstrate early illness identification [5].

The research on several machine learning prediction methods and tools is presented in this publication. In addition to discussing its prominent position in the healthcare business, this article provides a peek of machine learning's applicability in many disciplines [1].

Predicting a patient's chance of getting an illness or condition has shown potential using machine learning algorithms. Patients' prognoses and quality of life are enhanced when chronic illnesses including diabetes, cancer, and cardiovascular disease are detected at an early stage. They conclude that ML algorithms might change the game when it comes to early illness detection and prediction, but that further study is required to overcome the obstacles of using these algorithms in clinical settings [6].

This work is to provide a thorough overview of the current state of healthcare prediction using machine learning and deep learning techniques, and to highlight the inherent challenges associated with applying these techniques to this field [7].

B. Applications of Machine Learning in Early Diagnosis

The research [8] suggested a two-layer model that utilized machine learning methods, based on the approach used by dementia support centers for early dementia diagnosis. Information had been gathered from individuals who underwent dementia screening at the Gangbuk-Gu dementia center in the Republic of Korea between 2008 and 2013. The suggested methodology began by dividing MMSE-KC results into normal and abnormal categories. Dementia and moderate cognitive impairment were classified using CERAD-K data in the second stage. Utilizing precision, recall, and F-measure, the algorithms' performance had been contrasted with that of naive Bayes, Bayes network, begging, logistic regression, random forest, support vector machine (SVM), and multilayer perceptron (MLP). When looking at the F-measure values for normal, moderate cognitive impairment (MCI), and dementia, the MLP had come out on top with 0.97 for normal, while the SVM seemed to have come out on top with 0.739 for MCI and dementia. The suggested early diagnosis strategy for dementia might have simplified the diagnostic process while reducing the time and financial load.

In [9] stated that effective therapies for neurodegenerative illnesses had been desperately needed on a global scale. In this paper, they had gone over some of the ways in which machine learning might have helped with things like early diagnosis, medical picture interpretation, and the creation of novel treatments. Automated extraction of actionable insights from numerous high-dimensional data sources had been a common thread throughout machine learning's various uses; each source had offered a unique perspective on illness.

Recent research by [10] had indicated that despite improvements in speech processing that had made prognostic evaluation of PD patients easier, there had still been no objective way to diagnose the disease early in a clinical context. The purpose of this research had been to examine the potential of a new hybrid AI-based classifier to improve PD early diagnosis. The proposed feature-driven algorithm (MLP-LSVM) had shown promise for early PD diagnosis in clinical settings, with results showing a 100% area under the receiver operating characteristic curve and overall classification accuracy, and relatively faster convergence.

According to [11], the use of machine learning to analyze the spectra of exhaled breath for the purpose of early illness detection had been explored in their study. A quantum cascade laser and a Herriot astigmatic gas cell had made up the experimental setup, which might have been adjusted to operate within a wavelength range of 5.4–12.8 μm. Biomarkers and their mixes had been identified using principal component analysis and a shallow convolutional neural network. For optical channel lengths up to 6 m and signal-to-noise ratios below 3, a minimum detectable concentration of acetone and ethanol at the sub-ppm level might have been achieved. A lower detection limit for the same signal-to-noise ratio in the measured spectrum had been produced by neural networks compared to statistical approaches.

In the study by [12], several areas of medicine had been making extensive use of machine learning. More accurate information on the early warning signs of a wide range of illnesses had been available because of developments in medical technology. Degeneration of brain cells had ultimately resulted in memory loss in Alzheimer's disease, a chronic neurological disorder. Neuroimaging tests such as CT, MRI, PET, EEG, and others had been recommended for patients with cognitive mental disorders including forgetfulness and confusion, as well as other symptoms like behavioral and psychological issues. The paper's goal had been to apply machine learning algorithms to analyze neuroimaging data in order to predict early-stage Alzheimer's disease.

C. Challenges and Limitations in Implementing Predictive Analytics in Health

The study conducted by [13] provided a comprehensive overview of medical informatics machine learning and predictive analytics approaches in that chapter. Starting with a review of the existing literature, they went on to define important tasks, identify unresolved research difficulties, and examine the present status of predictive modeling in diagnostic medicine. This was in line with the conventional hierarchy of supervised, unsupervised, and RL. Afterward, they had a look at the latest studies on differentiable computing techniques including deep learning as well as semi-supervised, active, and transfer learning. Deep neural networks with typical applications in computational medicine, such as generative adversarial models for class imbalance correction in differential diagnosis and anomaly detection, recurrent neural networks for time series, and convolutional neural networks for image analysis.

As stated by [14], while resource-intensive causal inference was often the focus of clinical research, predictive analytics' potential with ever-growing big data sources had been mostly untapped. Big data made basic prediction much simpler, without relying on causal inference. The use of big data in this more straightforward way could improve emergency treatment. As basic heuristics for risk categorization, predictive analytics had long been useful in emergency medicine. Typically, these tools adhered to a consistent procedure: minimal requirements, simple computation, and separate validation with different populations. Although simplicity was very desirable in a prediction tool, it was no longer required due to technological advancements. Data science methods with an abundance of possible input factors found in electronic medical records might improve emergency treatment by making clinical predictions. Large data sets and reduced resource needs for comparing each clinical encounter to the previous one might allow for more accurate stratification of patients' risks, which would improve clinical decision-making and the operations of health systems. Early in the clinical encounter, when choices involving resources were being made and diagnostic and prognostic uncertainty were high, predictive analytics seemed to be the most valuable.

According to [15], updates to clinical decision support systems and predictive modeling had resulted from the integration of real-time analytics into existing electronic health records. The capacity of these systems to impact patient outcomes, however, remained uncertain at that time. Despite nurses making up the bulk of the healthcare workforce, very little was known about how clinical decision support influenced their choices and the results for patients. Pressure ulcers, failure to rescue, falls, and infections were four nursing-sensitive patient outcomes that were investigated in a scoping literature review that looked at the effects of clinical decision support systems that included healthcare predictive analytics.

The study conducted by [16] stated that the incorporation and integration of large datasets from many databases about the environmental and molecular causes of illness into analytical frameworks was a primary objective of precision medicine. This would enable the creation of personalized, context-dependent methods of diagnosis and treatment. To that end, analytical models of complicated diseases could be constructed using AI and ML techniques with the goal of predicting individual health states and outcomes. Because people varied greatly in terms of both genetic susceptibility and social and environmental factors, these models needed to be able to account for that. The ability to effectively handle, display, and integrate massive information in both organized and unstructured forms was essential for computational approaches to medicine. All of this had to be accomplished inside a single analytical framework, preferably while adhering to various degrees of secrecy.

D. Current Trends and Developments in the Field

The study conducted by [17] discussed how there had been a lot of interest in using deep learning techniques to assist in computer-assisted illness detection from medical data. They explained that it had been a medical domain challenge to make illness predictions using healthcare data that supported meaningful information in huge quantities about individuals with different conditions. In order to estimate the likelihood of an illness and devise a strategy to mitigate its impact on patients or eliminate it completely, disease risk prediction had been essential. They pointed out several advantages to illness prediction, including the ability to diagnose diseases at an early stage, reduce morbidity, and avoid death. In their research, they discussed how the healthcare industry might have used machine learning to forecast the ratio of illness detection.

The study conducted by [18] noted that global healthcare systems had been under constant strain from an increasing and an aging population, not to mention rising expenditures. Because of this problem, scientists throughout the globe had been trying to find ways to maximize healthcare spending by making use of digital resources that were dedicated to health and wellness. They explained that the study of how to give computers the ability to learn, reason, and make decisions similar to humans was known as artificial intelligence (AI). They pointed out that new developments in artificial intelligence, together with the widespread use of sensor-based technology like smartphones and wearables, had heralded a golden age in medical treatment. They provided an outline of popular ML algorithms to readers, along with an introduction to the ML train-test process. They concluded by outlining obstacles that must have been overcome in order to fully harness the power of ML in common digital health contexts.

In [19], discussed how recent advances in big data analytics and the Internet of Things (IoT) had opened up exciting new avenues for healthcare and biomedical technology. They explained that this database could have been used to generate structured information for remote diagnostics. They also covered more recent developments in robotics-based telemedicine that made use of AI techniques. They mentioned that AI-based telemedicine offered limitless assistance for healthcare improvement due to its adaptability and versatility. They provided a quick overview of the IoRT and its significance in relation to the development of telerobotic surgery. Additionally, they described wearable gadgets that were then on the market and aimed at biomedical and healthcare applications. These devices could gather data, analyze it using established procedures, and then use artificial intelligence to forecast health problems. They mentioned applications of these technologies that had recently emerged in contemporary biomedical research and the healthcare system, such as personalized medication design and targeted drug delivery.

According to [20], they explained that advancements in artificial intelligence (AI) for medical purposes had been driven by the proliferation of medical software and hardware applications and the digitalization of health-related data. They pointed out that this development pointed the way for the future of artificial intelligence in healthcare while also presenting new possibilities and obstacles. The purpose of their study had been to examine the present landscape of artificial intelligence (AI) in healthcare, including its potential, obstacles, and real-world effects. They focused their overview on the most current changes and future trends during the last five years. They talked about potential opportunities and where they could have gone from there based on these developments. They mentioned that new healthcare methods that included AI had been made possible by technological advancements, but obstacles had to be overcome. They stated that they had been actively working on solutions for multi-modal data integration, protecting model security, federated learning, model bias, and striking a balance between quantitative algorithm performance and qualitative model interpretability.

3 Methodology

Creating a mathematical model for predictive analytics in healthcare, particularly for early diagnosis using machine learning, involves several interconnected components. First, we need to represent the input data accurately. This includes gathering relevant features such as patient demographics, medical history, and test results. Let X denote the input features, where each feature is represented by x_i within the feature vector $X = (x_1, x_2, ..., x_n)$.

Next, define the parameters of machine learning model as θ. These parameters capture the relationships between the input features and the output prediction. The output prediction, denoted by Y, is a function of the input features and the model parameters, $Y = f(X; \theta)$ The goal of the predictive model is to accurately predict certain outcomes, such as the presence or absence of a medical condition, based on the input data to train the model. This is typically defined using a cost function $J(\theta)$, which measures how well the model performs on the training data. The objective during training is to minimize this cost function by adjusting the model parameters θ iteratively using optimization techniques such as gradient descent.

Choosing the right machine learning algorithm is crucial and depends on various factors including the complexity of the data, interpretability of the model, and performance requirements. Common algorithms used in healthcare predictive analytics include logistic regression, support vector machines, decision trees, and neural networks.

During the training process, the model is trained using a dataset comprising input–output pairs $\{(X(1), Y(1)), (X(2), Y(2)), ..., (X(m), Y(m))\}$, where m is the number of training examples. The model parameters θ are adjusted iteratively to minimize the cost function $J(\theta)$ on this dataset.

After training, the model's performance is assessed against a separate validation or test dataset to determine predicted accuracy, sensitivity, specificity, and other relevant parameters. Once confirmed, the trained model may be used to generate predictions on previously unknown data. Early diagnosis in healthcare involves using the trained model to estimate the probability of a certain medical condition based on input data. This allows for prompt intervention or treatment, possibly improving patient outcomes and lowering healthcare expenditures. As a result, the mathematical model provides a core framework for using predictive analytics in healthcare for early detection utilizing machine learning approaches.

A. Implementation Flow Model

Predictive analytics in the healthcare industry, specifically focusing on machine learning for early diagnosis, involves several steps.

Gather relevant healthcare data (e.g., patient records, laboratory results, and imaging data) and include data from various sources such as hospitals, clinics, and wearable devices.

Clean the data to remove errors and inconsistencies. Handle missing values and outliers. After that, we normalize or standardize the data for data preprocessing.

Identify relevant features (variables) for prediction, and use techniques like correlation analysis, feature importance, or domain knowledge for feature selection.

Choose appropriate machine learning algorithms for early diagnosis (e.g., logistic regression (LR), decision trees (DT), support vector machines, and neural networks).

Train model using splitting the data into training and validation sets, and train the selected models on the training data. Use techniques like cross-validation for hyperparameter tuning.

Evaluate the trained models using appropriate metrics (e.g., accuracy, precision, recall, $F1$-score, and ROC-AUC) and compare the performance of different models.

Deploy the best-performing model in a real-world healthcare setting and integrate the model into existing healthcare systems or workflows. To maintain the deployed model up to date, continuously assess its performance and retrain it with new data on a regular basis. Update the model as required depending on healthcare experts' comments and fresh research results (Fig. 1).

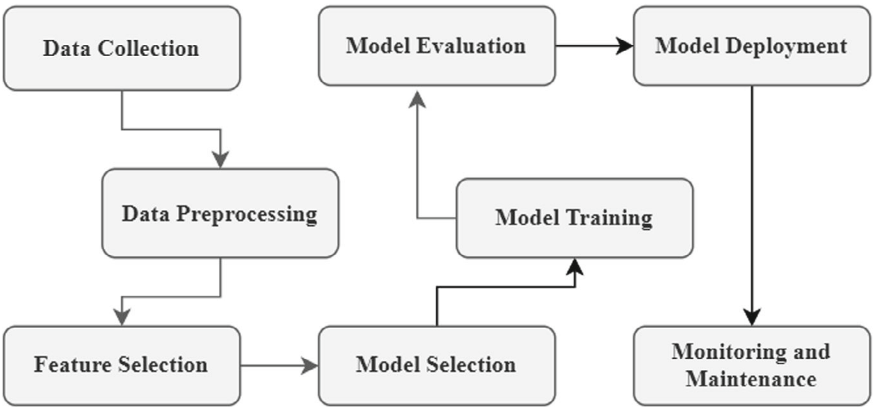


Fig. 1 Implementation flow model

4 Early Diagnosis in Healthcare

Early diagnosis in healthcare is crucial for improving patient outcomes across various medical conditions [21]. Here is a breakdown of its importance, challenges, and the role of predictive analytics in facilitating early diagnosis:

5 Importance of Early Diagnosis for Patient Outcomes:

- Better Treatment Outcomes: Timely intervention made possible by early identification frequently results in better treatment outcomes. Early diagnosis may greatly improve the prognosis for many diseases, including infectious disorders and cancer.
- Stop Disease Progression: By diagnosing illnesses early on, rates of morbidity and death can be decreased as well as the likelihood that they will proceed to more severe and perhaps irreversible phases.
- Improved Quality of Life: Patients with earlier diagnoses are able to better manage their diseases, which improves overall quality of life and symptom control.
- Lower Healthcare Costs: Because early illness treatment may include fewer resources and actions, it is frequently less expensive than managing later disease treatment.

6 Challenges in Early Diagnosis

The difficulty in identifying certain diseases in asymptomatic patients at an early stage is partly due to the lack of screening programs. It becomes considerably more difficult to detect these diseases before symptoms appear in the absence of established screening techniques. As a result, people might not become aware of their illness until it reaches a more severe level.

The fluctuating nature of some diseases' symptoms [22], especially in the early stages, is another confounding element. These illnesses can have nebulous or nonspecific symptoms that are simple to ignore or misread. Because of this, diagnosis is frequently postponed until the illness has progressed, which reduces the efficacy of treatment and may also lower the patient's prospects of recovery.

The difficulties in early detection are further exacerbated by diagnostic ambiguity. The diagnosis process may become more difficult in some situations when diagnostic tests produce unclear or inconclusive results. This uncertainty makes it difficult to determine the underlying origin of the symptoms, which is especially problematic when many illnesses present with similar clinical presentations.

A key factor in early diagnosis is also access to healthcare services. An individual's capacity to obtain prompt diagnostic services can be impacted by several circumstances, including geographic location, socioeconomic status, and the presence of healthcare infrastructure. Some people may have delays in diagnosis because of disparities in access to healthcare, which exacerbates the difficulties involved in early disease identification and treatment.

7 Role of Predictive Analytics in Facilitating Early Diagnosis

In the healthcare industry, predictive analytics [1] is essential because it makes risk stratification possible, which is the process of identifying people who have a higher chance of contracting specific diseases. Analyzing a variety of variables, including clinical history, genetic predispositions, and demography, allows for this. Healthcare practitioners might potentially prevent or mitigate the progression of diseases by implementing focused screening programs and early therapies after identifying high-risk people.

Furthermore, early warning system creation in healthcare contexts is made easier by predictive analytics. These systems make use of sophisticated algorithms to continuously monitor patient data and identify minute alterations that can point to the beginning or development of a disease. Early warning systems provide proactive actions by instantly alerting healthcare providers to these changes, lowering the risk of unfavorable outcomes and enhancing patient care.

Predictive analytics is not only helpful for early intervention but also a useful tool for healthcare decision support. Utilizing patient-specific data, predictive models can predict possible complications, treatment responses, and the course of the disease. These models improve the quality of care and patient outcomes by giving healthcare providers relevant insights to help them make better clinical decisions.

Predictive analytics is also essential to the advancement of personalized medicine. Through the analysis of extensive datasets comprising a variety of patient data, such as genetic profiles, clinical records, and treatment outcomes, predictive analytics facilitates the creation of customized treatment regimens that are specific to each patient's unique attributes. These individualized strategies maximize the beneficial effects of treatment while reducing side effects, which eventually improves patient outcomes and satisfaction.

Overall, risk classification, early warning system development, clinical decision support, and personalized medicine advancement are all made possible by predictive analytics, which has the potential to revolutionize the healthcare industry. Predictive analytics will become more crucial to improving patient care and outcomes as healthcare organizations continue to adopt data-driven approaches.

8 Machine Learning in Healthcare—Challenges

There are many possible uses for machine learning in healthcare, but it also brings its own unique set of difficulties. Training and validating machine learning models efficiently requires high-quality data, which presents a substantial challenge [23]. The predicted accuracy of the models is reduced without trustworthy data, which affects their usefulness in actual clinical situations.

The problem of restricted data availability, especially when it comes to uncommon diseases or ailments, makes this problem even more complex. Progress in building reliable predicting algorithms for such scenarios is hindered by the scarcity of relevant data, which restricts the capacity to train models efficiently.

In addition, healthcare records often contain missing data, which makes training and predicting models much more challenging. Accurately evaluating patient conditions and outcomes becomes more complicated when there is little data, which also weakens the models.

Feature engineering, the process of discovering and extracting relevant characteristics from data, is another essential component. Determining which components are most important to include in the model architecture is a challenging and time-consuming task that necessitates domain knowledge.

Another significant obstacle to healthcare machine learning projects is protecting patients' personal information and data. Strict safeguards must be put in place to guarantee the privacy and authenticity of patient data as it moves through the machine learning process to comply with legislation such as HIPAA.

Healthcare data, which includes a wide variety of sources including medical pictures and time-series data, is both complicated and high dimensional, adding another layer of difficulty. It takes specialized knowledge and computing resources to design and train models that can efficiently handle such complex data types.

Another important issue in healthcare is the interpretability of machine learning models, especially deep learning models. In order to build confidence among healthcare practitioners, algorithms must be visible and easy to understand, as the opaqueness of these models' decision-making processes might have serious consequences for patient care.

Lastly, there are a number of challenges associated with deploying and integrating machine learning models into clinical practice. Among these, there is a widespread lack of confidence in algorithmic systems among healthcare providers and stakeholders, in addition to technological difficulties, a lack of standardization, limited resources, concerns about data privacy and security, and limited acceptance.

Ultimately, if machine learning is truly going to transform healthcare, it must first overcome these complex obstacles. Only then will it be able to improve patient care and clinical results to their maximum potential.

9 Conclusion and Future Work

We investigated the use of predictive analytics, especially machine learning, in the healthcare business, with an emphasis on early detection. Machine learning algorithms in healthcare have demonstrated encouraging outcomes in terms of improving early diagnosis, prognosis, and treatment planning for a wide range of medical disorders. Machine learning algorithms can detect patterns and trends in massive datasets such as electronic health records, medical imaging, genetic data, and wearable sensor data that human physicians may not be aware of. Predictive analytics-enabled early diagnosis improves patient outcomes while also lowering healthcare expenditures and constraints on healthcare systems.

References

1. Nithya B, Ilango V (2017) Predictive analytics in health care using machine learning tools and techniques. In: 2017 International conference on intelligent computing and control systems (ICICCS). IEEE, pp 492–499
2. Agniel D, Kohane IS, Weber GM (2018) Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. Bmj 361
3. Rehman TU, Mahmud MS, Chang YK, Jin J, Shin J (2019) Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Comput Electron Agric 156:585–605
4. Hassan MM, Billah MAM, Rahman MM, Zaman S, Shakil MMH, Angon JH (2021) Early predictive analytics in healthcare for diabetes prediction using machine learning approach. In: 2021 12th International conference on computing communication and networking technologies (ICCCNT). IEEE, pp 01–05
5. Venkatesh R, Balasubramanian C, Kaliappan M (2019) Development of big data predictive analytics model for disease prediction using machine learning technique. J Med Syst 43:1–8
6. Siddiq M (2022) Use of machine learning to predict patient developing a disease or condition for early diagnose. Int J Multidiscip Sci Arts 1(1)
7. Badawy M, Ramadan N, Hefny HA (2023) Healthcare predictive analytics using machine learning and deep learning techniques: a survey. J Electr Syst Inf Technol 10(1):40
8. So A, Hooshyar D, Park KW, Lim HS (2017) Early diagnosis of dementia from clinical data by machine learning techniques. Appl Sci 7(7):651
9. Myszczynska MA, Ojamies PN, Lacoste AM, Neil D, Saffari A, Mead R, Ferraiuolo L (2020) Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat Rev Neurol 16(8):440–456
10. Parisi L, RaviChandran N, Manaog ML (2018) Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Syst Appl 110:182–190
11. Fufurin IL, Golyak IS, Anfimov DR, Tabalina AS, Kareva ER, Morozov AN, Demkin PP (2020) Machine learning applications for spectral analysis of human exhaled breath for early diagnosis of diseases. In: Optics in health care and biomedical optics X, vol 11553. SPIE, pp 113–122
12. Khan A, Usman M (2015) Early diagnosis of Alzheimer's disease using machine learning techniques: a review paper. In: 2015 7th International joint conference on knowledge discovery, knowledge engineering and knowledge management (IC3K), Vol. 1. IEEE, pp 380–387
13. Lamba D, Hsu WH, Alsadhan M (2021) Predictive analytics and machine learning for medical informatics: a survey of tasks and techniques. In: Machine learning, big data, and IoT for medical informatics Academic Press, pp 1–35

14. Janke AT, Overbeek DL, Kocher KE, Levy PD (2016) Exploring the potential of predictive analytics and big data in emergency care. Ann Emerg Med 67(2):227–236
15. Jeffery AD (2015) Methodological challenges in examining the impact of healthcare predictive analytics on nursing-sensitive patient outcomes. CIN: Comput Inform Nurs 33(6):258–264
16. Martinez-Garcia M, Hernández-Lemus E (2022) Data integration challenges for machine learning in precision medicine. Front Med 8:784455. Li W, Chai Y, Khan F, Jan SRU, Verma S, Menon VG, Li X (2021) A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system. Mob Netw Appl 26:234–252
17. Rathore DK, Mannepalli PK (2021) Recent trends in machine learning for health care sector. Int J Innov Res Technol Manage 5(2)
18. Cummins N, Ren Z, Mallol-Ragolta A, Schuller B (2020) Machine learning in digital health, recent trends, and ongoing challenges. In: Artificial intelligence in precision health. Academic Press, pp 121–148
19. Banerjee A, Chakraborty C, Kumar A, Biswas D (2020) Emerging trends in IoT and big data analytics for biomedical and health care technologies. In: Handbook of data science approaches for biomedical engineering, pp 121–152
20. Hinton G, Obermeyer Z, Emanuel E, LeCun Y, Bengio Y, Hinton G, Bekhet L (2019) AI in health: state of the art, challenges, and future directions. Yearb Med Inform 28(01):016–026
21. National Academies of Sciences, Engineering, and Medicine (2015) Improving diagnosis in health care. National Academies Press
22. Fernández-de-Las-Peñas C (2022) Are patients exhibiting post-Coronavirus Disease (COVID) symptoms at 12 months the same at 5 or 9 months? The fluctuating nature of Post-COVID. Clin Infect Dis 75(1):e1208–e1208
23. Alanazi A (2022) Using machine learning for healthcare challenges and opportunities. Inform Med Unlocked 30:100924

Generation of Opinionated Abstractive Summaries from the Knowledge Graph Using Transfer Learning with CNN

T. Priyanka, A. Mary Sowjanya, and A. Deepthi

Abstract The amount of data generated through online has increased rapidly due to the emergence of the Internet and the rapid progress of social media. This accumulated data offers unprecedented opportunity to different organizations for analyzing the view or opinion available as reviews toward any product, article, or any other entity, which aids in efficient development of business strategies as well as decision making. However, this humongous data available in an unstructured manner and in natural language, poses a great challenge for extracting any significant information in text. Thus, to obtain the sentiment of the users as well as concise representation of the text, sentiment rating prediction and abstractive summarization is essential. Hence, a novel method for extracting the abstractive summary from reviews using CNN-transfer learning is devised for predicting sentiment rating using a hybrid deep learning network comprising deep neuro fuzzy network (DNFN) and deep maxout network (DMN), for estimating the sentiment rating. Further, the techniques are evaluated for their effectiveness in terms of precision, recall, f-measure, and ROUGE rouge measure. The developed model showed superior results compared to existing methods for abstractive summarization.

Keywords Sentiment prediction · Abstractive summarization · Deep learning · Transfer learning · BERT tokenization

T. Priyanka (✉)
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, KL (Deemed to be University), Hyderabad, Telangana, India
e-mail: priyanka.t@klh.edu.in

A. M. Sowjanya
Department of Computer Science and Systems Engineering, Andhra University College of Engineering (A), Visakhapatnam, Andhra Pradesh, India

A. Deepthi
Department of Computer Science and Engineering (AIML), Narayanamma Institute of Technology and Sciences, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_23

1 Introduction

With modern advances in Web usage, feature-based opinion summary is becoming crucial. Almost thousands of individuals all over the world post comments for every product purchased online. Consequently, reading such a large number opinions is a tedious task. However, appealing to a tiny proportion of them might result in a skewed judgment. As a result, sentiment classification, summarizing, and subjectivity analysis became crucial. Sentiment analysis is used to determine the positivity or negative of a sentiment, while opinion mining only looks for statements that express a view upon anything. A summary is a means to convey a significant amount of data in a few words while yet keeping it meaningful and relevant. Similar to that, opinion summarizing provides an overview for numerous opinionated statements. The method of opinion summary that offers concise but crucial information incorporating summaries of several aspects of the intended brand is feature-based sentiment summarization. These have grown to become a more vital and frequently desired type of summary because it concentrates on various aspects rather than providing broad information about a product. It consists of three key specific phases: opinion forecasting, feature/aspect and opinion categorization, and summary creation. The summary generated must have sentences, which preserve the order of important notions in the original document, without altering the meaning of text and the sentences should be consistent and non-redundant [1]. As the abstractive summarization is a complex process, it can be effectively solved by using deep learning methods. Deep learning techniques also find application in various NLP processes as they comprises of various data processing layers with the ability to learn higher level data representations [2]. Several deep learning methods, like sequence-to-sequence model, gated recurrent unit (GRU), recurrent neural network (RNN), long short-term memory (LSTM), convolutional neural network (CNN), etc., have been developed for performing abstractive summarization [3, 4].

Extracting a summary from a text document using CNN-transfer learning is devised. Initially, the aspect terms and tokens are extracted from the input review data, which are then subjected to CNN-transfer learning for abstractive summarization. Here, the CNN is utilized with the architecture of AlexNet as well as the hyperparameters from VGG-16.

2 Literature Review

Several approaches have been put forth for performing abstractive summarization and sentiment rating prediction, among which a few are detailed here. Xu et al. [5] presented a rating-boosted abstractive review summarization (RARS) for generating a personalized summary. Here, the product and user preferences are learned from the review by using a neural review-level attention model, and these were utilized for producing a summary based on the saliency scores by the personalized decoder. This technique was highly effective in extracting the significant details of input text that are consistent with the reference summary; however, it suffers from high computational complexity.

Tan et al. [6] developed a unique hierarchical encoder–decoder framework with a graph-based attention method and a hierarchy beam search method to produce multi-sentence summaries. However, there is a paucity of training examples and the neuro-abstractive approach on the multi-document summarizing problem.

Zhang et al. [7] proposed a CNN-based documentation summarizing system which is suggested to simultaneously acquire text features and carried out sentence grading. The underlying CNN model has been modified to convert the ranking problem into a regression procedure for ranking sentences. Employed for a variety of document summarizing tasks using pre-trained word embeddings because it does not require for any previous knowledge.

Melnyk et al. [8] created Grapher, a multi-stage knowledge graph creation technique that partitions the whole graph generation. Using a pre-trained learning algorithm, the nodes are initially produced from the text document. The output graph is then built using the generated edge features from the node features.

Suhara et al. [9] developed Opinion Digest as a fusion of ABSA and seq2seq models that does not necessitate any training on gold-standard summary. The approach trains a transformer system to recreate the actual reviews from such extractions after extracting opinion terms from reviews using an ABSA model. When it came time to summarize, integrated the sentences from different reviews and chose the most favored ones.

Wu et al. [1] developed convolutional neural network (CNN)-driven opinion summarization method and Ortony–Clore–Collins (OCC) paradigm for Chinese microblogging services via sentiment classification rules set. A rule-based sentiment transfer technique provided by the OCC model's with eight behavioral categories as the fundamental feelings that arise during crisis situations.

Badgujar et al. [10] provided a graph-based method to produce abstractive summaries. By using the sentiment infusing technique, a graph is created, with every node including details about the word's positioning in the text, its sentence's location, and the estimated POS tag.

3 Model Developed: Abstractive Summarization Using CNN-Transfer Learning

A different technique is developed for performing abstractive summarization by using CNN-transfer learning. Initially, the input review [2] data is acquired from datasets subjected to bidirectional encoder representations from transformers (BERT) tokenization [3], the input sentence is divided into individual words or tokens. The tokens are then forwarded to the aspect term extraction (ATE) phase for extracting the specific aspects from the input data. The aspect terms along with the tokens are then fed to the graph generation module for generating graphs based on the frequency of the word occurrences. From the graph, the co-occurrence matrix is formed, which is provided as the input to the CNN-transfer learning [4], which in turn generates another co-occurrence matrix. The obtained co-occurrence matrix is converted back into a graph from which the abstractive summarized output is attained. Here, CNN is utilized with the architecture of AlexNet [5] as well as the hyperparameters from VGG-16 [11]. A hybrid deep learning model is used to predict the sentiment rating score via weighted correlation method.

3.1 Identifying the Sentiment of Review Entities

Initially after acquiring the reviews from the dataset and preprocessing the data, the sentences are forwarded to BERT to perform the tokenization process. Later identification of prominent features via TF-IDF, BoW, etc., helpful to diminish the dimensionality. Then, the aspect term extraction phase has multiple steps: categorizing the pairs of sentences and identifying the features observed globally and local context features using the BERT shared layer. Using the head pooling the polarity of each aspect is discovered, by identifying the corresponding positions in the given sequence [12]. These extracted features are interlinked together and forwarded to a deep learning network composed of deep neuro fuzzy network and deep maxout network, to get the overall sentiment.

3.2 Hybrid Deep Learning for Predicting Sentiment Rating

The feature map is applied to both the DNFN and the DMN, which constitutes the hybrid deep learning network. To generate the estimated rating, the output from both deep learning systems is integrated using weighted correlation as shown in Fig. 1.

3.2.1 Deep Neuro Fuzzy Network (DNFN)

The DNFN has a number of layers, including input, output, and several hidden levels. The defuzzification step forms the output layer, although the input layer makes use of many input variables in addition to the system's fuzzified values. Defuzzification, normalizing, and rule layers constitute the hidden layers.

The system's axioms and implications signify the key simple hyperparameters, and they both have an impact on fuzzification and defuzzification chores. To build rules, this network employs a fuzzy system of inference.

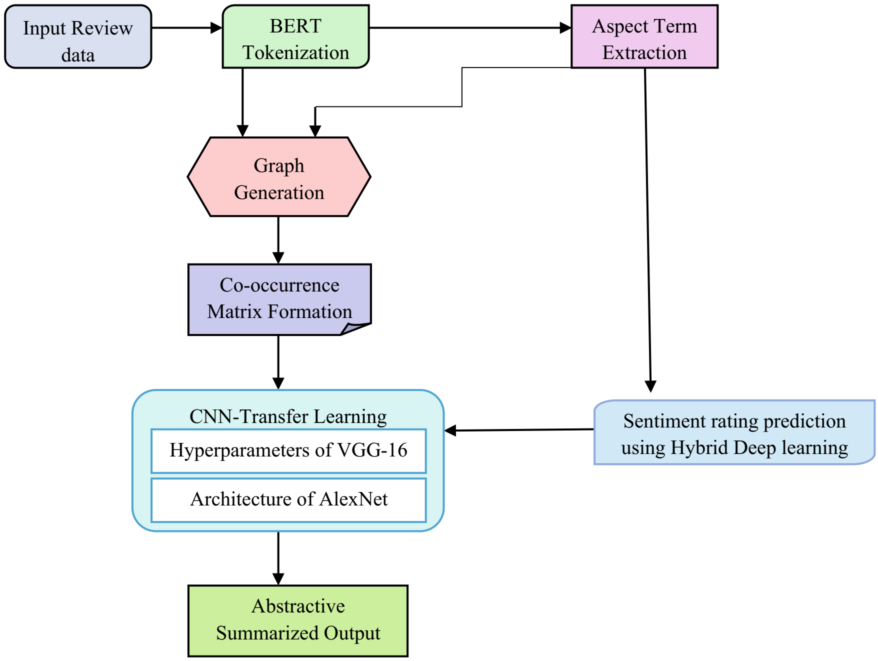


Fig. 1 Architecture of the developed CNN-transfer learning for abstractive summarization

3.2.2 Deep Maxout Network (DMN)

Numerous maxout layers interconnected progressively comprise the multi-layer DMN architecture. The DMN can be used successfully in circumstances with a restricted lexicon and provides the advantage of greater effectiveness in a context with dwindling resources. The DMN's maxout layer includes maxout functions that are employed to trigger the hidden activations. The layers also include hidden subunits in non-overlapping subgroups.

3.2.3 Integration Using Weight Correlation Coefficient

The weighted correlation step, in which both outputs are merged to estimate the overall sentiment score, is on the outputs of DNFN and DMN. The weighted coefficient of correlation is applied in this context to generate the sentimental score. At this stage, the final sentiment score was anticipated.

3.3 *Abstractive Summarization*

To generate abstractive summaries, graph-based structure is most utilized generally. The text is displayed as a knowledge graph, which aids in collecting redundant information [13]. The phrases are altered such that they comply with the fundamentals of convolutional neural networks, such as the adoption of the matrix form as a depiction of the text representations in CNN structure.

This section details the process of abstractive generation, wherein the aspect terms T_a as well as the tokenized data Y_a are utilized for creating the summary. Here, abstractive summary is generated by using the CNN-transfer learning [14], wherein the CNN is trained and used along with the pre-trained models like AlexNet and VGG-16. The process is detailed in the ensuing parts.

3.3.1 Graph Generation

The primary goal of the graphical method is to extract the most significant statements from sources. By creating the article's semantic structured graph and then employing the properties of both the document and the graphs to create the text summary leveraging co-occurrence matrix, it is possible to extract the most crucial phrases from the original content [15]. The aspect terms T_a as well as the tokenized data Y_a are subjected to the graph generation phase, wherein a graph is generated based on the occurrence of the words.

For example, consider the three sentences given below; Consider a window size of 1, and the neighborhood of the words has to be determined. For the word "I", the neighborhood terms are "enjoy" and "like", which appear once and twice, respectively. For the word "enjoy", the neighborhood terms are "I" (1) and "flying" (1). The numbers in the bracket indicate the number of times the corresponding words appear in the document. For "flying", the neighborhood term is "enjoy" (1). For the word "like", the terms are "I" (2), "NLP" (1), and "Deep" (1). Similarly, the neighborhood term for "NLP" is "like" (1), for "Deep", the terms are "like" (1), "learning" (1), and for "learning", the neighborhood is "Deep" (1).

3.3.2 Co-occurrence Matrix Formation

The weight among words can indeed be obtained from the co-occurrence matrix, and this weight will be utilized to illustrate the significance of the connections between both the nodes, i.e., the greater the weight, the stronger the connections between the vertices [16]. A weighted, directional graph with the words operating as the vertices and the score on the edges serving as the weights can be defined using a co-occurrence matrix. The co-occurrence matrix is used to depict the text as a weighted and strongly connected graph, with a weight assigned to each edge denoting the intensity of the links between the different phrase pairs in the document. This graph displays every connection while also taking into consideration the text's context and semantic structure. For generating the co-occurrence matrix, the graph G1 generated from the aspect terms and tokens is considered as the input. Let us assume that the words "I", "like", "flying", and "enjoy" are aspect terms (AST). The co-occurrence matrix is generated by considering the following equation,

$$A_{ij} = \frac{1}{2}\left(b_{ij} + c_{ij}\right) \tag{1}$$

where b_{ij} is the frequency of occurrence of the ith and jth word and c_{ij} is given by,

$$c_{ij} = \begin{cases} 1 & \text{if } i\&j \text{ are AST}, \ \ i \neq j \\ 0.5 & \text{if } i \text{ or } j \text{ is AST} \\ 0 & \text{if } i \text{ or } j \text{ are not AST} \end{cases} \tag{2}$$

The value of co-occurrence matrix entries is computed using Eq. (1).

The co-occurrence matrix thus generated is then fed as input to CNN-transfer learning, which performs abstractive summarization.

3.3.3 CNN-Transfer Learning for Abstractive Summarization

The co-occurrence matrix A is fed as input to the CNN-transfer learning network, wherein the CNN is utilized with the hyperparameters of VGG-16 and architecture of AlexNet. CNN-transfer learning offers the advantage of using the information from the pre-trained models devised for a particular issue to solve another related issue, thereby minimizing the time required for training. Moreover, the CNN-transfer learning has the capability to effectively reduce the false positives, while performing classification. Further, the issues concerning limited samples as well as disproportionately distributed data are resolved successfully. Initially, the CNN-transfer learning is trained by considering the hyperparameters of VGG-16 and architecture of AlexNet, and then, abstractive summarization is carried out in the testing phase as shown in Fig. 2.

The summary is produced by adopting the shortest way that begins with the initial sentence of the original document and concludes with the concluding sentence after the graph has indeed been generated. After removing duplicates and making the competition between phrases more clear, abstractive summarization increases sentence coherence. The abstractive summarized output thus obtained is "I like NLP, Deep learning".

In the testing phase, the co-occurrence matrix is directly applied to the CNN, as it is already trained using the hyperparameters of VGG-16 and AlexNet. The output of the CNN will be the abstractive summarized output.

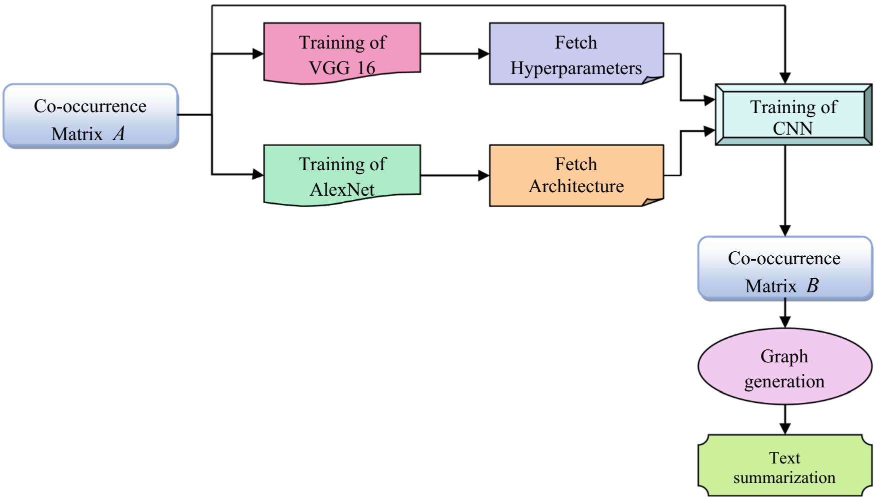


Fig. 2 Block diagram of the CNN-transfer learning

4 Sample System Generated Summary

Table 1 presents the proposed system generated output summary for the review input data. The summarized output generated for the review data which contains plain text review, ratings, user, as well as product information is given in Table 1

Table 1 System generated summary

Review data	Summarized output
Chocolate bar was ordered for an affordable rate and was delivered on time. As I couldn't drive my car to visit a mall, it was indeed ideal for me	Chocolate bar was purchased for a fair cost and arrived quickly. A massive variety of almost everything our family can enjoy. The lollipop has no taste and is simply scarlet
What else can I say, it's caramel. A wide choice of everything our family will love. With a seven family, stuff acts quickly. Perfect diversity. Take 3–5 and much more sweets, cookies, and Hershey's peanut butter	
The lollipop has almost no flavor and aroma and is merely scarlet. Just chewy and plain. Never again would I purchase them	

5 Dataset and Evaluation Metrics

The evaluation of the techniques is carried out using the summarizing text with Amazon Reviews. The dataset comprises reviews regarding fine foods available on Amazon. A total of 568,454 food reviews. The dataset contains ten attributes, such as Id, ProductId, UserId, ProfileName, Time, summary, text, score, Helpfulness Denomination and Helpfulness Numerator. The proposed CNN-transfer learning for summarization is evaluated for its effectiveness using parameters, like precision, recall, as well as, $F1$-score and rouge.

6 Conclusion

The reviews available online provide extraordinary opportunities to business organizations and researchers for understanding the perception of the user toward any article or product. However, the huge information available poses a great challenge in analysis, thus requiring effective techniques for extracting the relevant information and predicting the sentiments. Here, an efficient scheme for performing abstractive summarization using CNN-transfer learning is devised. Here, BERT tokenization is utilized for splitting the input review into tokens, from which aspect terms are extracted. The aspect terms are used for estimating sentiment rating. From the tokens, and aspect terms, CNN-transfer learning approach extracts the opinionated summary.

References

1. Wu P et al (2020) Social media opinion summarization using emotion cognition and convolutional neural networks. Int J Inf Manage 51:101978
2. Amazon Reviews dataset available at https://www.kaggle.com/currie32/summarizing-text-with-amazon-reviews/data. Accessed on June 2022
3. Devlin J, Chang MW, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint: arXiv:1810.04805
4. Shi Z, Hao H, Zhao M, Feng Y, He L, Wang Y, Suzuki K (2019) A deep CNN based transfer learning method for false positive reduction. Multimedia Tools Appl 78(1):1017–1033
5. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25
6. Tan J, Wan X, Xiao J (2017) Abstractive document summarization with a graph-based attentional neural model. In: Proceedings of the 55th annual meeting of the association for computational linguistics (Volume 1: Long Papers)
7. Zhang Y, Meng JE, Pratama M (2016) Extractive document summarization based on convolutional neural networks. In: IECON 2016–42nd annual conference of the IEEE industrial electronics society. IEEE
8. Melnyk I, Dognin P, Das P (2022) Knowledge graph generation from text. arXiv preprint: arXiv:2211.10511
9. Suhara Y et al (2020) OpinionDigest: a simple framework for opinion summarization. arXiv preprint: arXiv:2005.01901

10. Badgujar C, Jethani V, Ghorpade T (2018) Abstractive summarization using graph based methods. In: 2018 Second international conference on inventive communication and computational technologies (ICICCT). IEEE
11. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint: arXiv:1409.1556
12. Al-Radaideh QA, Bataineh DQ (2018) A hybrid approach for Arabic text summarization using domain knowledge and genetic algorithms. Cogn Comput 10(4):651–669
13. Suleiman D, Awajan A (2020) Deep learning based abstractive text summarization: approaches, datasets, evaluation measures, and challenges. Math Probl Eng
14. Wang L, Yao J, Tao Y, Zhong L, Liu W, Du Q (2018) A reinforced topic-aware convolutional sequence-to-sequence model for abstractive text summarization. arXiv preprint: arXiv:1805.03616
15. Song S, Huang H, Ruan T (2019) Abstractive text summarization using LSTM-CNN based deep learning. Multimedia Tools Appl 78(1):857–875
16. Xu H, Liu H, Zhang W, Jiao P, Wang W (2021) Rating-boosted abstractive review summarization with neural personalized generation. Knowl-Based Syst 218:106858

Enhancing Diabetic Retinopathy Classification Using Deep Learning with Data Augmentation Techniques

K. Deepthi, M. S. Josephine, Suresh Betam, and Bathula Prasanna Kumar

Abstract People with diabetes can develop diabetic retinopathy (DR), an eye disease that is a severe issue that affects more than a quarter of people with diabetes. The importance of early DR identification has increased since diabetes prevalence, particularly in developing countries, is on the rise. The prevention of DR begins with early detection and treatment. Detecting the DR condition is a time-consuming and labor-intensive process requiring an expert to analyze the images of the patient's retina. It makes slow up in diagnosis and treatment. However, automated examination of retinal images could significantly increase the effectiveness and coverage of DR screening programs. In light of this, we propose two predictive models—transfer learning and CNN—that incorporate the Alex model. The study shows that deep learning has great potential, and the EfficientNet transfer learning model achieves an impressive 0.80 accuracy with excellent recall, precision, and F-score metrics. While achieving a respectable 0.7347 accuracy, the CNN with Alex variation must be higher than its counterpart. The experimental outcomes conclusively demonstrate that the proposed approach surpasses traditional methods in utility and effectiveness. Our study enhances the field by revealing the significance of deep learning and transfer learning in DR severity prediction, ushering in a more effective era of diagnosis.

K. Deepthi (✉)
Department of Computer Science & Engineering, Dr. MGR Educational and Research Institute, Chennai, India
e-mail: deepthi.it@anurag.edu.in

Department of IT, Anurag University, Hyderabad, Telangana, India

M. S. Josephine
Department of Computer Applications, Dr. MGR Educational and Research Institute, Chennai, India

S. Betam
Department of Computer Science and Engineering, Vikas Group of Institutions, Vijayawada, Andhra Pradesh, India

B. P. Kumar
Department of Computer Science and Engineering, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_24

Keywords Diabetic,·Retinopathy classification · CNN · Data augmentation technique · Transfer learning

1 Introduction

DR is a diabetes-related eye condition with severe consequences if left untreated [1], and it is essential to take a comprehensive approach to manage the condition. It arises when high blood sugar levels damage the retina's blood vessels, leading to vision problems and even blindness in developed countries, particularly among working-age people, if not appropriately managed. DR is a condition that can have severe consequences for those who suffer from it. Unfortunately, the number of cases is expected to rise significantly in the coming years, which is a significant concern. It is responsible for roughly 2.6% of cases of blindness worldwide [2]. Early detection and timely treatment are essential for managing this progressive disease, and the number and types of lesions present in fundus images typically determine the severity. It is critical to emphasize that abnormalities in crucial components of the retina, such as optic discs, macula, and blood vessels, may be signs of underlying eye disorders. Furthermore, proliferative and non-proliferative diabetic retinopathy are the two main types of DR, according to traditional classifications [3].

The expansion of DR is attributed to a lack of blood sugar control and an unhealthy lifestyle, but it remains asymptomatic initially. However, as it progresses, patients start experiencing symptoms such as blurred vision, floaters, distortions, and a decline in visual acuity. Research involving some patients found that 23.86% had DR and 25.8% had already acquired the severe PDR stage [4]. Automated methods have significantly transformed the field of DR detection. They are budget-friendly, time-saving, and offer better accuracy than manual diagnosis [5].

The manual misdiagnosis system is risky in healthcare and effectively treating patients. With the integration of computer vision and AI breakthroughs, the healthcare industry has entered a new era where medical images can be interpreted more precisely, and data analysis capabilities are enhanced. Deep learning (DL) [6] and transfer learning (TL) [2] algorithms have been established to be very productive in navigating the complexities of medical imaging, leading to noteworthy advancements in the accuracy of analysis and classification. Ensuring everyone has equal access to these cutting-edge technologies is essential, as this is a crucial step toward achieving healthcare equity and fairness.

Our article is divided into different sections, which will help us to achieve our objective. In Sect. 2, we will review existing work, their result, methods, and findings. As we move on to Sect. 3, we will briefly explain the methodology information used in our work. Finally, we will conclude our article with a summary and a detailed analysis of our research findings in Sect. 4.

2 Related Work

Gangwar and Ravi [6] have demonstrated the potential of a data-driven DL algorithm to automate DR screening. By training the algorithm on a dataset of 75,137 fundus images from diabetic patients, they achieved impressive results, with 0.97 AUC, 94% sensitivity, and 98% specificity on fivefold cross-validation with reference to the evaluation metrics discussed by Scholar [11]. External validation on the MESSIDOR 2 and E-Ophtha databases produced AUC scores of 0.94 and 0.95, respectively. The algorithm also generated informative abnormality heat maps, which can aid in timely and accurate diagnoses, preventing vision loss.

Gargeya and Leng [7] employed DL to develop an algorithm for detecting DR and macular edema in retinal images, a critical challenge in medical imaging. The algorithm was trained on a vast dataset of 128,175 images graded by ophthalmologists, and it exhibited impressive performance. It acquired an AUC of 0.991 for EyePACS1 and 0.990 for Messidor-2. At a specificity-focused operating point, it demonstrated a sensitivity of 90.3% and 87.0% for EyePACS-1 and Messidor-2, respectively, with 98.1% and 98.5% specificities. At a sensitivity-focused point, sensitivity increased to 97.5% and 96.1%, with specificities of 93.4% and 93.9%. These outcomes highlight the algorithm's potential as a valuable diabetic retinopathy screening tool, which necessitates further research for clinical integration and assessment of its impact on patient care compared to current standards. Researchers developed the source-free transfer learning method for detecting diabetic retinopathy in a recent study [8]. This approach uses unannotated retinal pictures and depends only on a source model during training. It is divided into two primary modules: one for creating retinal pictures in the target style using target data and the source model, and the other for improving the classification model using these images. With a source model from EyePACS, the studies on the APTOS 2019 dataset produced encouraging results, illustrating the potential of SFTL in identifying diabetic retinopathy. The SFTL model achieved an accuracy of 91.2%, a sensitivity of 0.951, and a specificity of 0.858. These findings highlight the competitiveness of the SFTL model compared to state-of-the-art supervised learning methods, underscoring its potential as an effective tool for referable DR detection.

Scholars [9] introduced a segment-based learning approach that simultaneously learns classifiers and features from data, resulting in significant improvements in DR image recognition and lesion detection, and adopted a pre-trained CNN to obtain segment-level diabetic retinopathy estimation (DRE) and used diabetic retinopathy mapping (DRM) to integrate these segment-level DREs for classification. This approach was effective in handling irregular DR lesions. On the Kaggle dataset, the method achieved an impressive area under the ROC curve of 0.963 and sensitivity and specificity of 96.37% each in high-specificity and high-sensitivity settings, surpassing existing models. These findings highlight the potential of the segment-based learning approach for robust DR detection and its superiority over current methods.

3 Methodology

This section examines our proposed methods for detecting DR employing advanced DL models. We aim to revolutionize the earlier detection of DR by leveraging the power of artificial intelligence and cutting-edge image analysis. We have developed a well-organized pipeline that outlines the procedures for the automated examination of retinal images. Figure 1 presents a graphical representation of the overall steps in our work. In the subsequent sections, we will describe the research step in detail.

3.1 Dataset

The dataset was collected from the Kaggle (APTOS 2019). The dataset comprises 3662 meticulously collected retina images, evaluated using fundus photography techniques. Additionally, it contains 35,126 retina images that have been labeled by experienced clinicians employing the same standardized scale as the primary dataset. This curated dataset has the potential to significantly enhance the accuracy of detecting and managing DR, thereby propelling the field of medical image analysis forward.

3.2 Data Preprocessing

Based on our goal to improve the generalization capabilities of our model while preventing overfitting, we implemented a comprehensive data augmentation and preprocessing pipeline for our dataset. We introduced various modifications to our training data through data augmentation techniques like rotation and flipping, which expanded our dataset's diversity and provided our model with enhanced adaptability to real-world scenarios. We applied early stopping to prevent overfitting during training. Combining these techniques ensured that our model learns from the provided data effectively, enabling it to generalize with remarkable proficiency to previously unseen instances. As a result, our model consistently delivers superior performance.

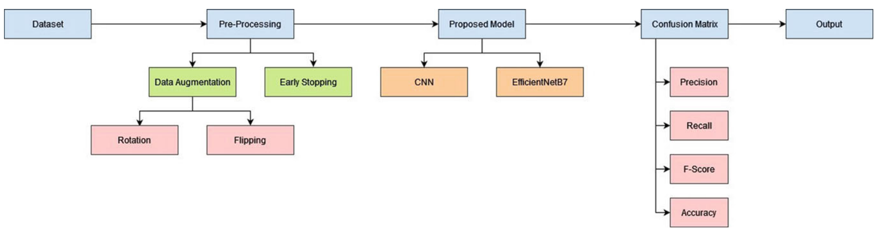


Fig. 1 Graphical representation of our methodology

3.3 Proposed Model

CNN: A crucial task in medical image analysis is the categorization of diabetic retinopathy, and the suggested convolutional neural network (CNN) architecture is specifically designed for this task. The model uses several convolutional layers to extract features from 28 × 28 pixel grayscale input images. The model's capacity to learn complicated patterns and minimize overfitting is improved by the first layers, which comprise max pooling and dropout layers, followed by two double convolution operations with 32 filters each. Two further convolutional layers with 64 filters are added after that to improve the feature extraction process even more. Then, max pooling with modified strides and dropout layers is utilized again. The flattened output is then fed into a 256-unit layer that is densely linked and uses the rectified linear unit (ReLU) activation function. A dropout layer with a rate of 0.5 is added to decrease overfitting. Then, a final output layer with five units and the softmax activation function is used for multi-class classification. Convolutional, pooling, and dropout layers are thoughtfully integrated into this carefully planned architecture. This exhibits a balanced trade-off between model complexity and generalization performance, making it well-suited for research on diabetic retinopathy and fundus image analysis. The proposed CNN architecture is visualized in Fig. 2.

The parameters employed during the training of this architecture are illustrated in Table 1.

Transfer Learning: In this study, the EfficientNetb7 architecture was employed. It is a convolutional neural network derived from the EfficientNetb7 model for image

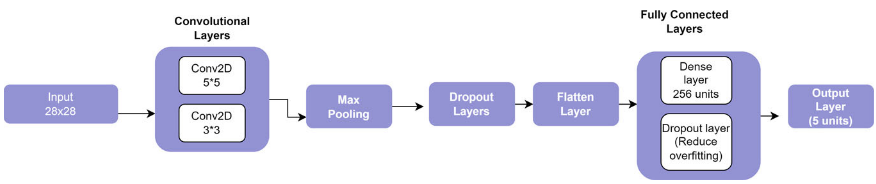


Fig. 2 Architecture of proposed CNN model

Table 1 Parameter and values of CNN architecture

Parameter	Value
Learning rate	0.001
Patience	3
Verbose	1
Factor	0.5
Min lr	0.00001
Epochs	30
Batch size	100

Table 2 Parameter and values of Efficientnetb7

Parameter	Value
FL	CE
Epochs	16
ES	6
ROL(η)	1 × 10 − 3
SOS	6
GOS	0.6

classification tasks. The architecture comprises several layers, each with a specific role in image data processing. The first layer is the stem convolutional layer, which uses a convolution operation with filters and spatial dimensions of the input data. A batch normalization phase follows this to maintain consistent activation distributions during training. The backbone of the EfficientNetb7 architecture consists of Mobile Inverted Bottleneck Blocks (MBConvBlocks). These blocks utilize a depth-wise convolution to apply a convolution to each input channel. Together, all of these capture fine details in the data. To further refine the high-dimensional representation of the data, a head convolutional layer applies a 1 × 1 convolution. Then, a global average pooling (GAP) layer aggregates spatial data to produce a condensed 1 × 1 summary of spatial dimensions. A dropout layer is also added to stop overfitting during training. This architecture's last layer is the completely linked layer. This layer maps the elegant features to the specific output space for this task, consisting of five distinct classes. The architecture's efficiency is achieved through the strategic use of depth-wise separable convolutions and the integration of SE operations, significantly reducing computational demands without compromising performance. The model employs pre-trained weights during initialization to expedite the training process. During inference mode, computational resources are conserved by selectively freezing particular layers and adaptively configuring the final fully connected layer. For more information on the parameters employed during the training of this architecture, please refer to Table 2. A visual representation of the model's architecture can also be found in Fig. 3.

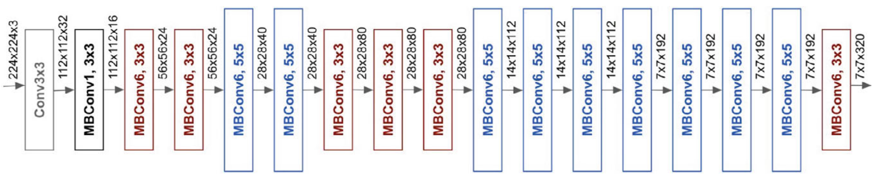


Fig. 3 Architecture of proposed Efficientnetb7 model

4 Result and Discussion

Through our comparative analysis between EfficientNet and CNN, we have better understood their capabilities in deep learning and image recognition. It is worth noting that EfficientNet consistently outperforms CNN across precision, recall, *F*-score, and accuracy metrics. With a precision of 78%, it is particularly effective in minimizing false positives, while its recall rate of 81% showcases its ability to capture positive instances accurately. The harmonic mean or *F*-score, 76%, underscores a balanced blend of precision and recall. Moreover, EfficientNet's overall accuracy rate of 81% solidifies its superiority in making correct predictions throughout the dataset. These findings highlight the potency of newer architectures like EfficientNet, offering promising prospects for various computer vision tasks. However, when deciding between models, task-specific requirements and computational resources should be considered, bearing in mind the transformative potential of advanced architectures in the ever-evolving landscape of deep learning.

These findings suggest that the EfficientNet model distinguishes individuals with diabetic retinopathy, making it a precious tool for early opinion. These results have essential counteraccusations for clinical practice, pressing the eventuality of advanced ML algorithms, particularly the EfficientNet, in enhancing the delicacy of diabetic retinopathy opinion. The model's exceptional precision and recall values and outstanding delicacy demonstrate its capacity to contribute significantly to early and precise discovery, eventually perfecting patient issues and conserving vision. Scholars [10] have clearly explained the formulas for evaluation matrices. Table 3 comprehensively overviews our models' evaluation criteria.

Figures 4 and 5 visually represent our two models' training and confirmation results. These graphical delineations offer a deeper understanding of how the models performed during the training and confirmation phases. They illuminate the models' capacity to effectively learn from the training data and their capability to generalize

Table 3 Performance metrics of the two proposed models

Model	Precision (%)	Recall (%)	*F*-score (%)	Accuracy (%)
EfficientNet	78	81	76	81
CNN	73	73	73	74

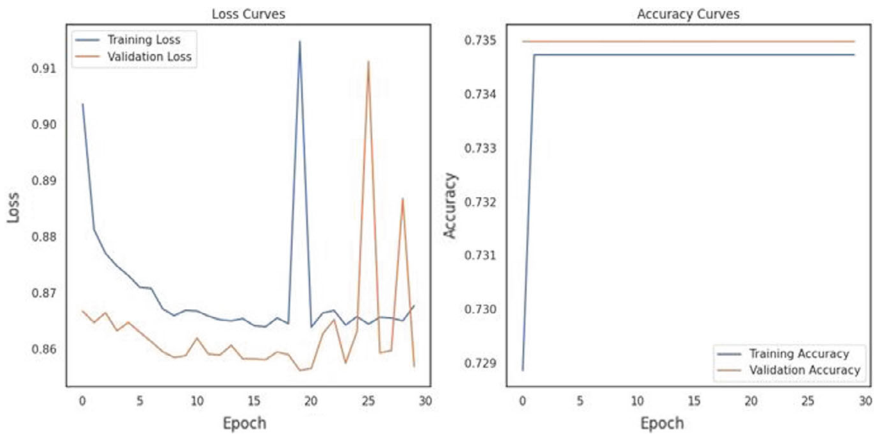


Fig. 4 Training and validation metrics for the CNN model: loss and accuracy trends

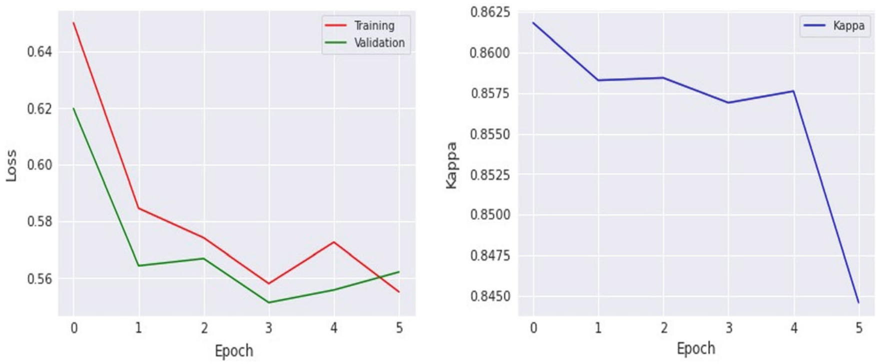


Fig. 5 EfficientB7 model's training and validation loss with Kappa coefficient analysis

well when faced with new, unseen samples. These numbers are precious tools for assessing the models' training dynamics and overall performance.

5 Conclusion

The number of people diagnosed with diabetes has increased yearly, making it a significant health issue. One of the complications that can arise from diabetes is damage to the eyes. Detecting diabetes early is critical to managing and potentially curing the condition. Studies have shown that approximately 30% of individuals with diabetes may develop DR. Timely diagnosis can significantly improve the chances of successful treatment, while delayed diagnosis can worsen the situation and result in irreversible blindness. Traditionally, diagnosing DR requires a manual examination by specialists, which can be time-consuming. Therefore, there is a growing interest in developing automatic diagnostic methods that can provide substantial assistance. This research utilized the EfficientNetb7 model and data augmentation techniques, resulting in notably improved outcomes. Although we determine the necessity to analyze the performance of other transfer learning or CNN-based models, our findings demonstrate that data augmentation with EfficientNetb7 outperformed other models, producing highly satisfactory results. Our research will have a positive and meaningful impact on advancing this field.

References

1. Qiao L, Zhu Y, Zhou H (2020) Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. IEEE Access 8:104292–104302
2. Alyoubi WL, Shalash WM, Abulkhair MF (2020) Diabetic retinopathy detection through deep learning techniques: a review. Inform Med Unlocked 20
3. Math L, Fatima R (2021) Adaptive machine learning classification for diabeticretinopathy. Multimedia Tools Appl 80(4):5173–5186
4. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama 316(22):2402–2410
5. Salongcay RP, Silva PS (2018) The role of teleophthalmology in the management of diabetic retinopathy. Asia-Pac J Ophthalmol 7(1):17–21
6. Gangwar AK, Ravi V (2021) Diabetic retinopathy detection using transfer learning and deep learning. In: Evolution in computational intelligence: frontiers in intelligent computing: theory and applications (FICTA 2020), vol 1. Springer, pp 679–689
7. Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using deeplearning. Ophthalmology 124(7):962–969
8. Scotland GS, McNamee P, Fleming AD, Goatman KA, Philip S, Prescott GJ, Sharp PF, Williams GJ, Wykes W, Leese GP et al (2010) Costs and consequences of automated algorithms versus manual grading for the detection of referable diabetic retinopathy. Br J Ophthalmol 94(6):712–719
9. Khan Z, Khan FG, Khan A, Rehman ZU, Shah S, Qummar S, Ali F, Pack S (2021) Diabetic retinopathy detection using VGG-NIN a deep learning architecture. IEEE Access 9:61408–61416
10. Zhang C, Lei T, Chen P (2022) Diabetic retinopathy grading by a source-free transfer learning approach. Biomed Signal Process Control 73:103423
11. Deepthi K, Josephine MS, Jeyabalaraja V (2022) Fundus image classification systems for accurate prediction of retinopathy. J Harbin Inst Technol 54(10):161–165. ISSN: 0367-6234

Learning and Inference of Graph Discrete Structures for Graph Neural Networks in ML

Talari Swapna, Rajendhar Reddy Gaddam, S. Dhanalakshmi, B. Pradeep, Balusupati Anil Kumar, and Pirangi Vijay Kumar

Abstract Graph-structured data is modeled and analyzed by graph neural networks (GNNs) in a variability of domains, such as recommendation systems, social networks, and bioinformatics. Nevertheless, their efficacy is largely dependent on their capacity to represent and deduce distinct structures that are intrinsic to graphs. This paper investigates current developments in inference and learning strategies designed to deal with discrete structures in graph data. We explore encoding strategies for graph structures into continuous representations, utilizing deep learning and probabilistic modeling approaches. We also cover methods for effective inference over discrete structures, including matching subgraphs and testing graph isomorphisms. We also study how inference and learning interact in GNNs and discuss future possibilities and difficulties in this quickly developing subject. To promote improvements in graph representation learning and graph-based reasoning problems.

Keywords Graph discrete structures · GNNs · Deep learning · Machine learning · Social networks

---

T. Swapna (✉) · R. R. Gaddam · S. Dhanalakshmi · B. Pradeep · B. A. Kumar · P. V. Kumar
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: swapnat@cmritonline.ac.in

R. R. Gaddam
e-mail: rajendarreddy.gaddam@cmritonline.ac.in

B. Pradeep
e-mail: pradeepb26@cmritonline.ac.in

B. A. Kumar
e-mail: anilkumarb@cmritonline.ac.in

P. V. Kumar
e-mail: vijaykumar@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_25

1 Introduction

Graph neural networks have emerged as powerful tools for learning representations of graph-structured data [1], enabling tasks such as node classification, graph classification, and link prediction. Central to the effectiveness of GNNs is their ability to capture complex relationships and dependencies within graphs. However, traditional GNNs are primarily designed for continuous node attributes and edge features, limiting their applicability to discrete graph structures. To address this limitation, recent research has focused on extending GNNs to handle discrete structures, such as molecular graphs in chemistry or social networks in sociology, where nodes represent discrete entities and edges denote pairwise relationships. This endeavor involves developing methods for learning and inference that are tailored to discrete graph structures. One of the main features of GNNs is their capacity to grasp complex graph interactions [2]. Nodes, edges, and the patterns of connection between them are separate structures in graphs, and a thorough comprehension of them is required for this. For GNNs to be successful in their many uses, it is essential to learn representations that encode these structures well. GNNs use reasoning about discrete structures throughout the inference process in order to classify or predict.

A. *Graph Neural Networks (GNNs)*

Graph neural networks (GNNs) are a class of artificial neural networks [3] designed to operate on graph-structured data. Unlike traditional neural networks that process data arranged in grids (like images) or sequences (like text), GNNs can handle non-Euclidean data represented as graphs, where nodes represent entities and edges represent relationships between them.

GNNs leverage graph convolutional layers to iteratively aggregate information from neighboring nodes, enabling them to learn representations that capture both local and global graph structures. This allows GNNs to perform tasks such as node classification, link prediction, and graph classification (Fig. 1).

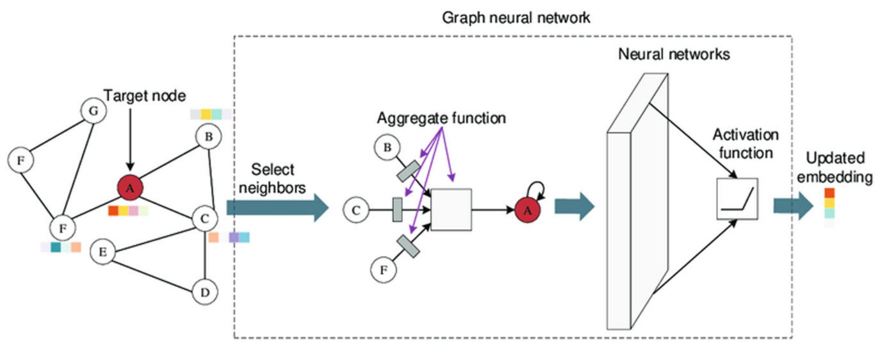


Fig. 1 Architecture of GNNs [4]

GNNs have found applications in various domains including social network analysis, recommendation systems, drug discovery, and cybersecurity. They offer advantages like the ability to model complex relationships, handle heterogeneous data, and generalize to unseen graph structures.

However, challenges remain in scaling GNNs to large graphs efficiently, handling noisy or incomplete data, and addressing over-smoothing issues where information from distant nodes becomes indistinguishable. Ongoing research focuses on developing more scalable architectures, improving representation learning, and enhancing the interpretability and robustness of GNNs.

2 Literature Review

A. *Machine Learning (ML) Methods on Graph Neural Networks*

In [5] stated that graph data, which included extensive connection information among pieces, was required for many learning tasks. Learning molecular fingerprints, illness classification, protein interface prediction, and physical system modeling necessitated models that could learn from graph inputs. Another major area that required graph reasoning models was learning from non-structural data, such as texts and photographs. This involved reasoning on extracted structures, such as phrase dependency trees and image scene graphs. Neural models known as graph neural networks (GNNs) were able to represent the interdependence of graphs by means of messages sent between their nodes.

One study by Wu et al. [6] mentioned that from picture categorization and video processing to voice recognition and NLP, deep learning had recently transformed several machine learning jobs. In most cases, the information gathered from such endeavors was shown in a geometrical sense. But there were more and more uses for graphs with complicated interactions and interdependencies between items, representing data from non-Euclidean domains. They offered a new taxonomy that classified them as either recurrent, convolutional, graph autoencoders, or spatial–temporal. They went on to summarize the GNN open-source scripts, benchmark datasets, and model assessment processes, and they talked about the many fields where GNNs had found use. They concluded by suggesting future avenues of inquiry into this dynamic and expanding area.

Waikhom and Patgiri [7] stated that the area of machine learning had been reenergized by deep learning throughout the past ten years or so. With cutting-edge efficiency, it had resolved several issues in computer vision, voice recognition, NLP, and other fields. The data was often shown in these domains using the coordinate system of Euclidean space. Graphs were ideal representations of non-Euclidean spaces, which many other domains corresponded to. When trying to depict the interconnections and interdependence between different things, graphs worked well. Historically, features for graphs that were hand-crafted had failed miserably in drawing the inferences needed for different tasks from this complicated data representation.

B. *Evolution of GNNs*

In research [8] proposed GNN-based meta-learning methods for a range of graph issues demonstrated the efficacy of combining the two techniques. Architectures, shared representations, and applications were the three main criteria by which we classified the literature. Lastly, we went over a number of intriguing open problems and potential avenues for further investigation.

In [9] noted that deep learning had revitalized the machine learning profession in the past decade. This article provided an overview of GNNs and their background information before doing a thorough assessment of GNNs in various learning models, including semi-supervised, self-supervised, supervised, unsupervised, and few-shot or meta-learning. A taxonomy of all the graph-based learning settings together with the methods that belonged to each one was seen. Both theoretical and empirical considerations were given to the methods used for each learning activity. Additionally, broad architectural design principles were offered for constructing GNN models. Along with the many outstanding problems that continued to affect GNNs' generalizability, a number of examples of their use and benchmark datasets were offered.

The extensive usage of deep learning on Euclidean data had sparked a renaissance in the study of graph representation learning and other non-Euclidean domain neural network architectures. Now that these graph neural networks (GNN) had proven successful in a static environment, researchers could move on to more realistic situations where the graph changed over time. They proposed EvolveGCN model for graph convolutional networks (GCNs) that evolved along the temporal axis without using node embeddings. By training the GCN parameters with an RNN, the suggested method was able to capture the graph sequence's dynamic nature. For the parameter evolution, two architectures were taken into consideration. The suggested method was tested on a variety of tasks, such as node classification, edge classification, and link prediction. The experimental results showed that when compared to comparable methodologies, EvolveGCN typically performed better [10].

In [11], this study presented the geographical-temporal graph social network (STGSN), a framework for graph neural networks that incorporated both geographical and temporal aspects into the modeling of social networks. A new method was used to capture the temporal properties of social networks by utilizing the temporal attention mechanism. The method's interpretability was enhanced by developing a technique that analyzed the distribution of attention across time. The study concluded by doing comprehensive tests on six publicly available datasets to demonstrate the efficacy of the approaches [11].

Graph structure learning (GSL) was a new approach that aimed to address this issue by teaching graph neural networks to learn a graph structure that was consistent with the ground truth, thereby improving their performance. But GSL's present approach involved optimizing the ideal network structure repeatedly using a single GNN, which led to training issues like overfitting and susceptibility. This paper introduced an evolutionary graph neural network (EGNN), a new GSL technique, to boost GNN performance and strengthen defenses against adversarial assaults [12].

C. *Applications and Successes of Graph Neural Networks*

In [3] recommend GNNs for graph representation learning. As said, GNNs calculated a node's representation vector by repeatedly merging and changing its nearby nodes' vectors. On both node and graph classification tasks, various GNN versions achieved state-of-the-art results. GNNs revolutionized graph representation learning, but their limitations and properties remained unclear. A theoretical approach for assessing GNNs' expressive graph topology understanding was provided. It was proven that popular GNN versions like graph convolutional networks and GraphSAGE could not learn to distinguish between fundamental graph topologies, which was used to define their discriminative capabilities. Then, a basic design was shown to be as effective as the Weisfeiler–Lehman graph isomorphism test and the most expressive GNN class. The model achieved state-of-the-art performance, and numerous graph classification benchmarks validated theoretical conclusions. Shen et al. [13] discussed current wireless communications developments and the efficacy of deep learning technologies in solving tough problems. Early attempts failed due to a lack of generalizability and scalability in large-scale networks and a reliance on neural network topologies from computer vision. Due to its ability to employ domain knowledge—specifically the graph topology—to solve wireless communications challenges, GNNs have become popular. GNN-based systems may generalize well across varied system contexts and perform near-optimal in large-scale networks, but the absence of strong theoretical underpinnings and design standards may hamper their implementation. Theory and practice were expected to be bridged in the paper. According to theoretical assumptions, GNNs performed near-optimal in wireless networks with less training samples than traditional neural architectures. GNNs were able to optimize a -node network (where nodes may be users, base stations, or antennas) with a generalization error and training sample needs times lower than unstructured multi-layer perceptrons. The suggested wireless network design recommendation system included graph modeling, neural architecture design, and theory-guided performance enhancement. Comprehensive simulations of a variety of key challenges and network topologies proved the theory and practicality of the recommended design framework.

Zhou et al. [14] noted that graph neural networks (GNNs) extended classical neural networks to graph-structured data and performed well. Despite these successes, Euclidean models were naturally confined by their ability to represent Euclidean geometry, especially for datasets with highly non-Euclidean latent anatomy. A fresh wave of research and novel applications has focused on hyperbolic spaces as an alternative to standard approaches for processing graph data with a tree-like structure or power-law distribution. Euclidean space grew polynomially, while hyperbolic space extended exponentially with radius, making it ideal for representing complex real-world data. Thus, abstracting hierarchical or power-law distributed tree-like networks was thought to help it organically. For anybody interested in hyperbolic graph neural networks (HGNNs), this course provides a complete review of the methodology, applications, and problems in this interesting and quickly evolving discipline. An introduction to graph neural networks, Riemannian manifolds, and hyperbolic geometry was planned as the initial phase. The created HGNNs' technical details would be combined into a generic framework and component versions described after rigorous examination. Additionally, other industries will launch applications. It was concluded that various impediments will be addressed and solutions presented to help the research community succeed. It featured their first attempts [14].

GNNs can handle non-Euclidean data such scanned point clouds and building information models in the construction business. Despite its potential, there was a lack of comprehensive scholarly research on GNNs' application in building. Based on an intensive review of 34 articles, the article reportedly met this demand by summarizing construction sector GNN research. The paper examined existing literature to identify the pros and downsides of employing GNNs in building. This paper shows GNNs' potential future in construction and discusses graph data generation from common data sources. The research was also lauded for boosting awareness of GNNs in construction and offering practical solutions to real-world issues [15].

According to [16], social media users often exchanged photographs, videos, and music. Social media has opened new study options for many scholars. Recently, several research have demonstrated that graph neural networks (GNN) outperform classical deep learning. For online social network analysis, graph neural networks were recommended. The paper examined how VKontakte users' profiles may be utilized to train graph convolutional neural networks with several convolution layers (GCNConv, SAGEConv, GraphConv, GATConv, TransformerConv, GINConv) to predict professional success on the site. Professional performance was measured by VKontakte measures including friends, subscribers, interesting pages, and user connection graphs. This paper claimed to have classified graphs using graph convolutional neural networks, which employ many convolutional layers. The graph isomorphism network (GIN) layer gave the graph convolutional neural network its best accuracy, 0.88. This study provided measurements for OSN user profiles and social graphs evaluated using neural network methods, which may guide future social success research.

D. *Challenges and Limitations of Current Graph Neural Networks*

Graphs in molecular chemistry, computer vision, pattern recognition, social networking, and other scientific and technical domains. The study detailed graph neural networks (GNN) as a model for analyzing this data. GNN processed graph-based data using neural networks. Due to its success, GNN is a popular architecture to work on and has many applications. This article explained graph neural networks, their uses, and how we utilize them daily. We all use social media everyday, and one of the most popular applications is the recommendation system, which leverages our past purchases to propose people and products. In addition, this study revealed the underlying GNN implementation challenges [17].

This paper extensively evaluated graph neural network-based recommender systems. They then discussed why recommender systems should employ graph neural networks, including high-order connectivity, data structural features, and a better supervision signal. We methodically examined network creation, model optimization, propagating and aggregating embeddings, and computer efficiency. A comprehensive overview of graph neural network-based recommender system research was provided using the classification above. Open concerns and future pathways in this area were examined. Summaries were created from representative articles and code sources [18].

The ability to develop with graph data and their extensive use in practical applications have garnered academic interest in graph neural networks (GNNs) in recent years, according to [19]. As data privacy concerns arose, global neural networks (GNNs) had to adapt. Federated learning (FL) clients may have links, and therefore, better tools were required to exploit implicit information to increase performance. Due of this, federated GNNs (FedGNNs) are a new field that has evolved significantly. Researchers interested in this diverse subject found it challenging to grasp. Lack of a thorough investigation made entry harder. This article met that demand by covering this new field thoroughly. A two-dimensional FedGNN literature taxonomy was proposed: (1) The major taxonomy showed how GNNs enhanced FL training and how FL assisted GNN training. (2) The auxiliary taxonomy illuminated FedGNN methods for client-specific heterogeneity. Future research that solves current limitations and investigates essential themes may build more robust, explainable, efficient, fair, inductive, and comprehensive FedGNNs [19].

In [20], tested graph neural networks in heterophily and low homophily semi-supervised node categorization tasks. Multilayer perceptrons, which ignored graph structure, outperformed numerous well-known GNNs, demonstrating they could not generalize. The restriction led to the discovery of ego-and neighbor-embedding separation, higher-order neighborhoods, and intermediate representations that enhanced graph structure learning under heterophily. The combined graph neural network H2GCN was used to test the designs. The empirical study found that the discovered designs improved GNN accuracy by 27 and 40% on synthetic and real networks with heterophily, respectively, and produced comparable results with homophily.

Deep reinforcement learning (DRL) has helped pattern recognition, robotics, recommendation systems, and gaming, according to [21]. GNNs also outperformed other supervised learning algorithms for graph-structured data. GNN with DRL in graph-structured contexts has garnered attention recently. This article thoroughly examined these hybrid works. Two groups were formed for these works: (2) application-specific contributions that employed GNN-DRL to tackle application-specific problems, and (1) algorithmic contributions that helped DRL and GNN overcome each other's constraints. This merger solved many technical and biological science problems. Based on the review, the benefits of integrating these sectors were examined. The goal was to improve generalizability and simplify computation. Anyone interested in machine learning should read this paper since it outlines the main barriers to integrating DRL and GNN and proposes further research.

3 Methodology

Creating a mathematical model for learning and inference of graph discrete structures for graph neural networks (GNNs) typically involves defining equations that describe the update rules for node and edge features in the network. GNNs operate on graph-structured data, where nodes represent entities and edges represent relationships between these entities. Here is a general framework for a GNN model:

Let us denote:

- $G = (V, E)$ as the input graph, where V is the set of nodes and E is the set of edges.
- $h_v^{(1)}$ as the feature vector associated with node v at layer l.
- $h_e^{(1)}$ as the feature vector associated with edge e at layer l.
- $N(v)$ as the set of neighbors of node v.

The update equations for node and edge features in a typical GNN can be formulated as follows:

A. *Node Update Equation*:

$$h_v^{(l+1)} = U\left(h_v^{(l)},\ \left\{h_u^{(l)} : u \in N(v)\right\}\right)$$

where U is a function that aggregates information from the node's neighbors and possibly its own current state to update the node features. This function can vary based on the specific GNN architecture (e.g., graph convolutional networks, GraphSAGE, etc.). Common aggregation functions include mean aggregation, max-pooling, attention mechanism, etc.

B. *Edge Update Equation*:

$$h_e^{(l+1)} = V\left(h_v^{(l)},\ h_w^{(l)}\right)$$

where V is a function that updates the edge features based on the features of its incident nodes. This function can also vary based on the GNN architecture and may involve concatenation, element-wise operations, or other transformations.

These equations represent the propagation of information through the graph in each layer of the GNN. Typically, multiple layers of updates are applied iteratively to capture information from distant nodes in the graph. Apart from these updated equations, the GNN model may include additional components such as readout functions for graph-level predictions, loss functions for training, and optimization algorithms for parameter updates (e.g., stochastic gradient descent).

C. *GNNs-Graph Discrete Structures*

Graph discrete structures, such as graphs, are fundamental mathematical representations used extensively in the realm of graph neural networks (GNNs). A graph G is typically denoted as $G = (V, E)$, where V represents the set of nodes (or vertices) and E represents the set of edges. Each edge e is essentially a pair of nodes (u, v) where u, v are elements of V.

Graph G contains the adjacency matrix A, which is a square matrix of size $|V| \times |V|$, where $|V|$ signifies the number of nodes in the graph. This matrix helps indicate the presence or absence of edges between nodes. For instance, in a binary scenario where the graph is unweighted, $A_{ij} = 1$ if there is an edge between nodes i and j, and $A_{ij} = 0$ otherwise.

Nodes and edges in the graph may have associated feature vectors to represent attributes. The node feature matrix X has dimensions $|V| \times d$, where d is the dimensionality of the node features, with X_i representing the feature vector of node i. Similarly, the edge feature matrix R has dimensions $|E| \times f$, where f is the dimensionality of the edge features, with Re representing the feature vector of edge e.

In GNNs, information propagation across the graph is often achieved through message passing. A common operation in this context is operation of the graph convolution, given by the equation:

$$h_i^{(l+1)} = \sigma\left(\Sigma_{j \in N(i)} \frac{1}{C_{ij}} W^{(l)} h_j^{(l)}\right)$$

Here, h_i^l represents the representation of node i at layer l, $N(i)$ denotes the neighbors of node i, $W^{(\mathrm{l})}$ is the weight matrix at layer l, σ is the activation function, and c_{ij} is a normalization constant (e.g., degree of node i).

Graph pooling operations are employed to aggregate information across nodes and edges at different hierarchical levels. For example, a common pooling operation is graph max-pooling, which aggregates information by selecting the maximum representation across nodes:

$$h_G = \max_{i \in V} h_i.$$

These equations collectively form a foundational mathematical model for graph discrete structures of GNNs framework.

D. *GNNs in ML*

Graph neural networks (GNNs) represent data as nodes and edges, fostering versatile representation learning. Nodes encapsulate features, while edges denote relationships. GNNs utilize message passing among nodes, aggregating information from neighbors to update node representations iteratively. Typical architectures include graph convolutional networks (GCNs), graph attention networks (GATs), and GraphSAGE. GCNs perform convolutions in the spectral domain, leveraging graph Laplacian eigenvalues. GATs employ attention mechanisms to weight neighbor contributions adaptively. GraphSAGE samples and aggregates information from node neighborhoods. These models facilitate tasks like node classification, link prediction, and graph classification, demonstrating efficacy across diverse domains like social networks, biology, and recommendation systems.

A subset of machine learning models known as graph neural networks (GNNs) are developed to process data that is structured in a graph. Non-Euclidean data, such as social networks, molecular structures, and recommendation systems, can be handled by GNNs, in contrast to traditional neural networks that process data organized in grid-like structures (e.g., images).

GNNs iteratively update node representations based on information from neighboring nodes, leveraging the inherent relationships between entities in a graph. In doing so, they are able to extract the data's structural information and intricate dependencies.

Graph neural networks (GNNs) are made up of layers that relay messages between nodes; each layer updates the node representations based on the information gathered from nearby nodes. To make GNNs adaptable to different graph structures and tasks, these layers usually have learnable parameters.

Graph neural networks (GNNs) have many potential uses, some of which include traffic prediction, social network analysis, drug discovery, and recommendation systems. Thanks to their remarkable relational data modeling capabilities, they have been instrumental in driving remarkable progress in various domains, solidifying their position as a potent instrument in the realm of contemporary machine learning.

4 Conclusion and Future Work

This paper has provided a comprehensive overview of the methods and techniques employed in the learning and inference of graph discrete structures for graph neural networks (GNNs). Through a systematic examination of existing literature, we have identified key challenges and trends in this domain, including the representation learning of graph structures, graph generation, and inference tasks such as node classification and link prediction. We have also highlighted the importance of incorporating discrete structures into GNNs to enhance their performance and applicability across various domains, including social networks, bioinformatics, and recommendation systems.

Several potential directions for further study have been recognized. Firstly, approaches for learning and inferring graph discrete structures that are more durable and scalable are needed. This is especially true when dealing with large-scale and dynamic graphs. It is also crucial to continue looking into new methods and designs that may make good use of discrete structures while keeping the underlying network topology intact. Adding domain-specific information and limitations to GNN models may also increase their performance in practical settings. Last but not least, improving the current state-of-the-art will need testing and comparing current approaches on various datasets and workloads. All things considered, this study lays the groundwork for further studies that will try to improve GNNs' learning and inference skills for graph discrete structures.

References

1. Berrone S, Della Santa F, Mastropietro A, Pieraccini S, Vaccarino F (2022) Graph-informed neural networks for regressions on graph-structured data. Mathematics 10(5):786
2. Alon U, Yahav E (2020) On the bottleneck of graph neural networks and its practical implications. arXiv preprint. arXiv:2006.05205
3. Xu K, Hu W, Leskovec J, Jegelka S (2018) How powerful are graph neural networks?. arXiv preprint. arXiv:1810.0082
4. Zeng Y, Tang J (2021) Rlc-gnn: an improved deep architecture for spatial-based graph neural network with application to fraud detection. Appl Sci 11(12):5656
5. Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, Sun M (2020) Graph neural networks: a review of methods and applications. AI open 1:57–81
6. Wu Z, Pan S, Chen F, Long G, Zhang C, Philip SY (2020) A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst 32(1):4–24
7. Waikhom L, Patgiri R (2021) Graph neural networks: methods, applications, and opportunities. arXiv preprint. arXiv:2108.10733
8. Mandal D, Medya S, Uzzi B, Aggarwal C (2022) Metalearning with graph neural networks: methods and applications. ACM SIGKDD Explor Newsl 23(2):13–22
9. Waikhom L, Patgiri R (2023) A survey of graph neural networks in various learning paradigms: methods, applications, and challenges. Artif Intell Rev 56(7):6295–6364
10. Pareja A, Domeniconi G, Chen J, Ma T, Suzumura T, Kanezashi H, Leiserson C (2020) Evolvegcn: evolving graph convolutional networks for dynamic graphs. In: Proceedings of the AAAI conference on artificial intelligence, vol. 34, no. 04, pp 5363–5370
11. Min S, Gao Z, Peng J, Wang L, Qin K, Fang B (2021) STGSN—a spatial-temporal graph neural network framework for time-evolving social networks. Knowl-Based Syst 214:106746
12. Liu Z, Yang D, Wang Y, Lu M, Li R (2023) EGNN: graph structure learning based on evolutionary computation helps more in graph neural networks. Appl Soft Comput 135:110040
13. Shen Y, Zhang J, Song SH, Letaief KB (2022) Graph neural networks for wireless communications: from theory to practice. IEEE Trans Wirel Commun
14. Zhou M, Yang M, Xiong B, Xiong H, King I (2023) Hyperbolic graph neural networks: a tutorial on methods and applications. In: Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining, pp 5843–5844
15. Jia Y, Wang J, Shou W, Hosseini MR, Bai Y (2023) Graph neural networks for construction applications. Autom Constr 154:104984
16. Gafarov F, Berdnikov A, Ustin P (2022) Predicting success of the online social network user by graph neural networks. Int J Adv Intell Inform 8(3)
17. Gupta A, Matta P, Pant B (2021) Graph neural network: current state of art, challenges and applications. Mater Today: Proc 46:10927–10932
18. Gao C, Zheng Y, Li N, Li Y, Qin Y, Piao J, Li Y (2023) A survey of graph neural networks for recommender systems: challenges, methods, and directions. ACM Trans Recomm Syst 1(1):1–51
19. Liu R, Xing P, Deng Z, Li A, Guan C, Yu H (2024) Federated graph neural networks: overview, techniques, and challenges. IEEE Trans Neural Netw Learn Syst
20. Zhu J, Yan Y, Zhao L, Heimann M, Akoglu L, Koutra D (2020) Beyond homophily in graph neural networks: current limitations and effective designs. Adv Neural Inf Process Syst 33:7793–7804
21. Munikoti S, Agarwal D, Das L, Halappanavar M, Natarajan B (2023) Challenges and opportunities in deep reinforcement learning with graph neural networks: a comprehensive review of algorithms and applications. IEEE Trans Neural Netw Learn Syst

Intruder Detection for Safe Homes Through Real-Time Security Monitoring Using CNN

T. Aparna, S. Nanitha Reddy, S. V. L. Santhoshi Pavani, and P. Renuka

Abstract In the contemporary world, with a growing number of individuals choosing to live alone, the security has emerged as a paramount concern. Although existing surveillance systems provide continuous monitoring capabilities, it is a significant challenge for humans to maintain a constant and vigilant watch over safety. To address this concern, the proposed system presents a swift, robust, resilient, and highly secure solution for safeguarding properties and valuable assets. It leverages advanced face recognition algorithms, specifically convolutional neural networks (CNN), in conjunction with motion detection and alert mechanisms. The system incorporates a passive infrared (PIR) sensor, which detects motion through infrared radiations. When human presence is detected within the monitored area, the system employs facial recognition technology. This technology cross-references the captured facial images with a secure database of registered faces. In the event that an unrecognized face is identified, an immediate alert is generated and sent to the user via email. By combining motion detection with advanced facial recognition technology, this system not only offers robust security but also minimizes false alarms, making it highly effective and reliable. It provides homeowners and property managers with unparalleled peace of mind. Consequently, this innovative security solution not only addresses the security needs of those living alone but also enhances the overall safety and protection of their premises and assets.

Keywords PIR sensor · Image processing · Facial recognition · Real-time security monitoring · Proactive threat detection

T. Aparna (✉) · S. Nanitha Reddy · S. V. L. Santhoshi Pavani · P. Renuka
Department of Information Technology, G. Narayanamma Institute of Technology and Sciences, Shaikpet, Hyderabad 500104, India
e-mail: aparna.tanam@gnits.ac.in

S. Nanitha Reddy
e-mail: snanitha624@icloud.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_26

1 Introduction

Security still remains as a major concern in the current technological era. Traditional security cameras, although effective in surveillance are not much advanced in providing alerts and often inundate users with false alarms. The security alert camera system mentioned integrates the camera surveillance with advanced facial recognition algorithms, this integration not only revolutionizes the field of security alert systems but also sets the stage for a more efficient and effective means of safeguarding individuals and their assets. The system leverages the use of passive infrared (PIR) sensors for motion detection and integrates advanced facial recognition algorithms to distinguish between authorized individuals and potential intruders. The salient feature of our security solution is that the camera remains vigilant while conserving energy by activating only when motion is detected. By focusing on motion-triggered alerts, we alleviate the common problem of nuisance alarms associated with traditional surveillance systems. Through this research, we aim to offer enhanced security by identifying potential threats and send alerts whenever a potential intruder is found. We provide comprehensive insights into real-world application of our solution, highlighting its potential to revolutionize security surveillance. In the following sections, we delve into the architecture, operation, and performance of this innovative security alert camera system.

2 Related Study

After referring various articles and research papers, we found that was a scope to improvise security features by integrating video surveillance with facial recognition algorithms.

In [1–4] they proposed a system which employs webcam-based face detection, which is activated when the user enters the correct password. The user is required to memorize a static password. The system does not incorporate a mechanism for users to change or update their passwords. Moreover, the idea of integrating face recognition into a system where users already possess knowledge of a password may appear somewhat redundant.

In [5–7] designed an automated attendance system that will take attendance of a student who is present in a classroom after periodic time. Using this system as a reference, we designed a model for ensuring security at remote places like home, office, etc. The enhanced system provides alerts for the user when an intruder is detected.

In. [8–11], designed system integrates home automation and video surveillance systems. The system might not function accurately as it overloads the raspberry pi with home automation and video surveillance features, it does not provide proactive alerts to the user when an intruder is detected.

In [12–14] the system employs a face recognition system without any database or pi sensor, which works continually and increases load on the system, which is not practical.

Our objective is to address the shortcomings of existing surveillance systems by introducing several innovative features. Key enhancements is the establishment of a dedicated database, meticulously designed for storing the facial data of known individuals. This database forms the cornerstone of our system's facial recognition capabilities, allowing for precise face detection and comparison.

In the event of a potential security breach, our system goes beyond the conventional by promptly transmitting alert messages enriched with attached images of the detected intruders. This real-time information empowers users with critical visual evidence for immediate response and decision-making.

Moreover, we have integrated a passive infrared (PIR) sensor, a vital component that activates the camera exclusively when motion is detected. This smart utilization of the PIR sensor not only conserves resources but also significantly elevates system efficiency and speed, ensuring that the system operates with optimal performance.

3 Proposed Approach

The core functionality of the system relies on a passive infrared (PIR) sensor, which serves as its motion detection component. When the PIR sensor detects motion, it triggers the system to capture a series of images using an integrated camera. Subsequently, an algorithm comes into play to determine whether the detected motion corresponds to the presence of a human. This algorithm is designed to distinguish between human and non-human movement patterns, ensuring that the system does not take much computation load.

If the algorithm confirms the presence of a human, the system proceeds to the next crucial step, facial recognition. This facial recognition algorithm is responsible for comparing the facial features of the detected individual with a database of known faces stored within the system. These known faces typically belong to authorized individuals who are allowed access to the secured area.

If a match is found between the detected face and a known face in the database, the system deems it safe for the guest or authorized person to enter. However, if no match is found, indicating the presence of an unrecognized individual, the system takes immediate action to alert the user or security personnel. This alert mechanism is typically executed by sending an email notification, providing real-time information about the potential intrusion.

In summary, the system's operation involves a multi-step process: motion detection, human presence confirmation, facial recognition, and access decision. It combines these elements to create a robust security solution that minimizes false alarms, enhances safety, and ensures efficient access control. This comprehensive approach ensures that only authorized individuals are granted access to the secured area while promptly notifying users of any potential security breaches (Fig. 1).

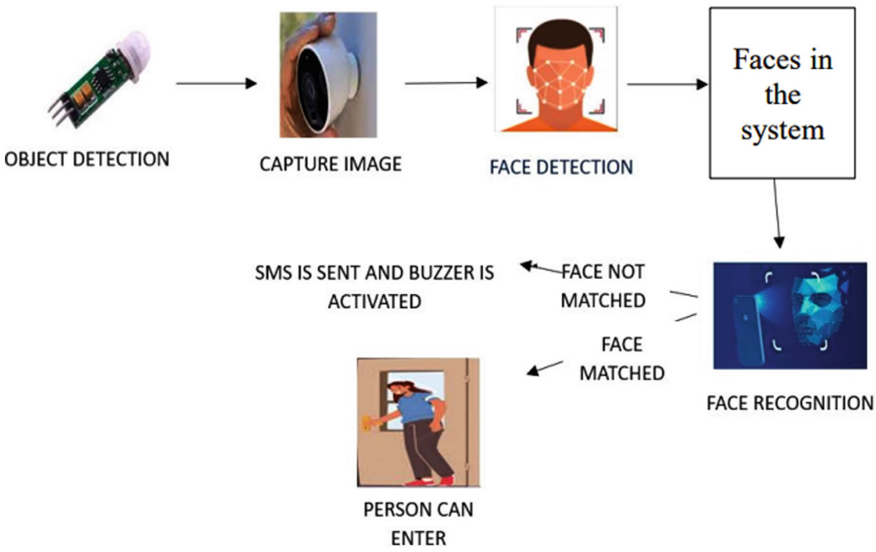


Fig. 1 System architecture

3.1 Implementation Details

Sensor Module: At the heart of the system's core functionality is the sensor module, anchored by the remarkable passive infrared (PIR) sensor. When an object, such as a person, moves within the monitored area, the PIR sensor consistently detects subtle changes in infrared radiation. This heightened sensitivity transforms the system into a responsive sentinel, instantly reacting to human presence and motion. By skillfully utilizing the interplay between thermal emissions and its own perceptive abilities, the PIR sensor seamlessly initiates subsequent system modules.

Face Detection Module: When activated by the PIR sensor's precise detection, the system seamlessly shifts its focus to the face detection module. This module, like a skilled investigator, employs OpenCV's deep learning face detection algorithm. Real-time images from the monitored area are collected, forming a digital canvas. The algorithm takes the lead, carefully analyzing the visual data. Each image holds the potential to reveal human faces, and the algorithm's cognitive abilities come into play as it identifies facial features within intricate patterns. This moment initiates the essential process of facial recognition, a vital aspect of the system's primary function.

Face Recognition Module: Having confirmed the presence of a face in the previous module, our attention now turns to the face recognition module, powered by the convolutional neural network (CNN) algorithm. This digital expert, meticulously trained like a seasoned virtuoso, is poised to decipher the intricate language of human facial features. The CNN algorithm embarks on a complex process of comparison, navigating through the captured images with precision. Informed by the nuances learned from a diverse array of authenticated facial images, the algorithm identifies the subtle distinctions that set individual faces apart. The subsequent comparison of these identified features with the respected records stored in the system's database creates a comprehensive verification or inquiry process. Each comparison represents a testament to the system's sophisticated capabilities.

Alert Module: The pinnacle of the system's operation is reached with the alert module, a vigilant entity ready to initiate rapid responses at any sign of a security breach. If the detected face does not match any of the authorized faces stored in the system's database, a sequence of strategic actions is instantly triggered. First and foremost, a critical SMS alert is sent to the designated system administrator. This SMS goes beyond mere text communication by including a visual image as evidence. This image serves as immediate confirmation or raises suspicion of a potential breach, demanding swift attention.

System Integrity and Enhancement: The robust integration of these modules culminates in a comprehensive and responsive smart security system. To maintain system integrity, continuous improvement strategies are implemented. Regular updates to the CNN algorithm, database optimization, and enhancements in the sensor module's capabilities ensure that the system remains at the forefront of security technology. In conclusion, the multifaceted architecture of the proposed smart security system is composed of distinct yet intricately interconnected modules. These modules seamlessly collaborate to ensure accurate motion detection, precise facial recognition, prompt alerts, and user-defined responses. The harmonious synergy of technology, algorithms, and user empowerment solidifies the system's position as a dependable and sophisticated solution for enhancing home security and providing peace of mind.

3.2 Design

The depicted use case diagram in Fig. 2. illustrates how various users and external systems interact with the system, highlighting the diverse functionalities it offers. In this diagram, the actors involved primarily include visitors and users, who actively interact with the system to achieve two primary objectives: detecting intruders and ensuring safety. Visitors and users are the key participants who actively engage with the system. Visitors represent individuals who interact with the system, possibly as guests or temporary users, while users are individuals with authorized access and usage rights.

The activity diagram described in Fig. 3 serves as a visual representation of the step-by-step journey that the system follows during its operation. The sequence is initiated when the system detects motion and ends by enhancing security through the generation of alerts, facilitated by the application of facial recognition technology to identify potential intruders. Motion detection acts as the trigger for subsequent

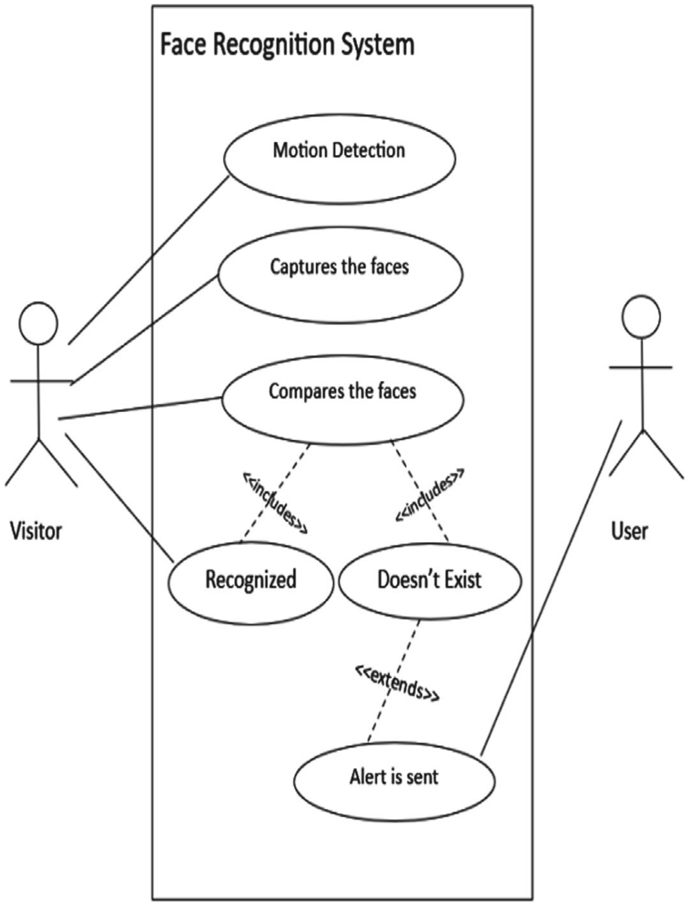


Fig. 2 Use case diagram

actions. As the process unfolds, the system seamlessly integrates facial recognition technology to identify any individuals within the monitored space.

4 Results

The image in Fig. 4 illustrates the system's capability to recognize faces that are already stored in its database. In instances where recognized faces match those in the database, the system operates without triggering any alerts, ensuring that there are no security breaches. This feature minimizes unnecessary notifications or alerts for known individuals, thus optimizing its performance and ensuring a more streamlined and efficient operation.

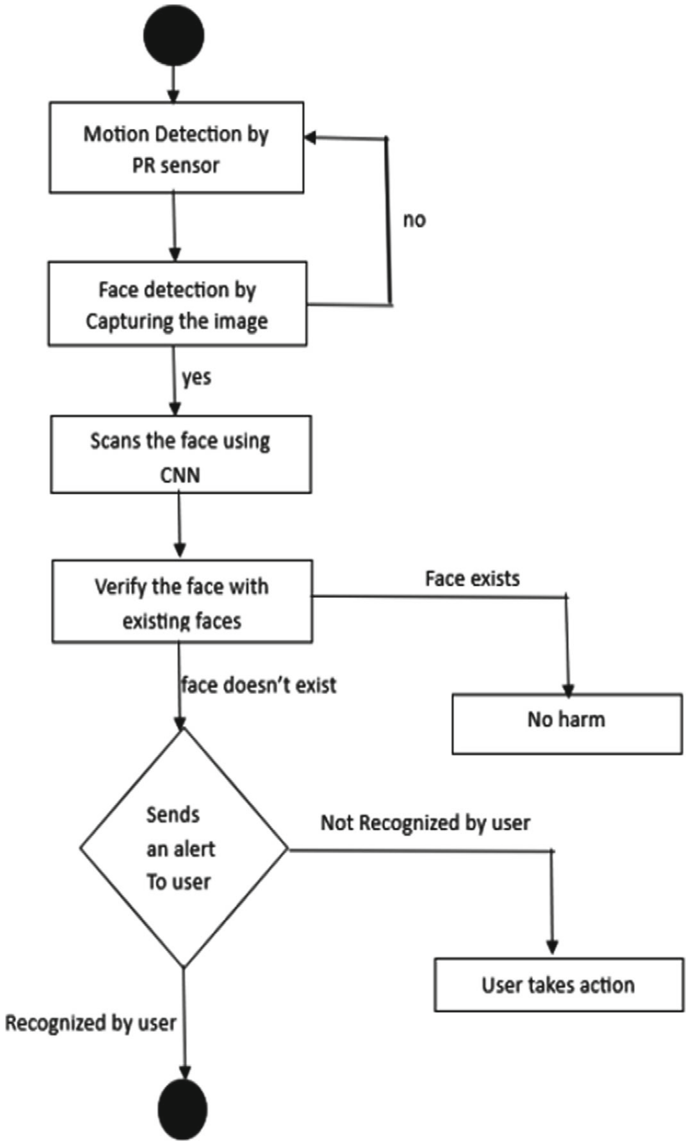


Fig. 3 Activity diagram

The image as shown in Fig. 5 depicts how the system responds when an unauthorized individual is detected. In a case where the intruder's face is not recognized by the system, it promptly sends an alert message to the user, with attached images providing visual evidence of the intruder's presence. This approach significantly enhances the security features of the system.

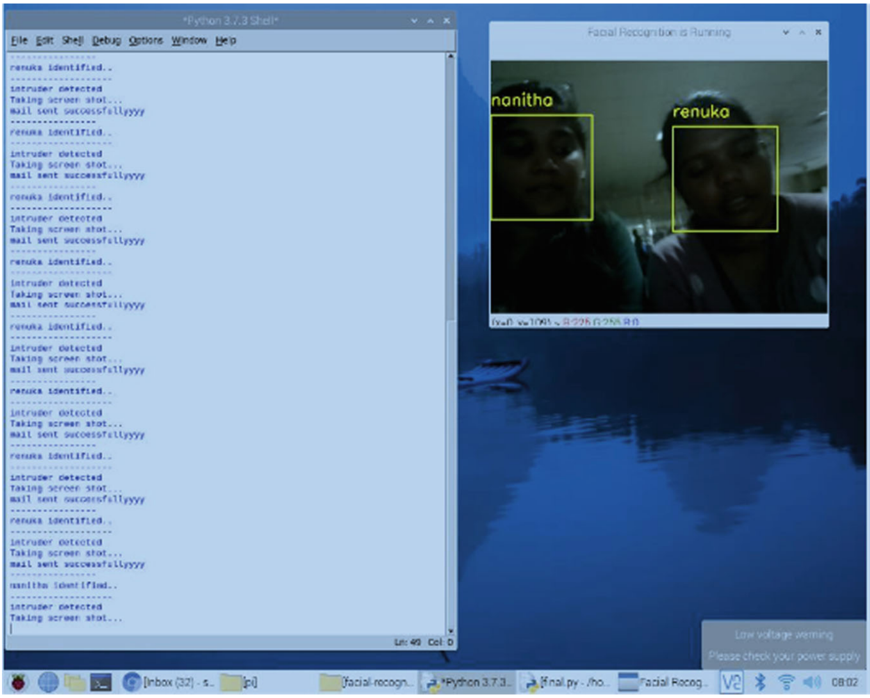


Fig. 4 Detecting the faces

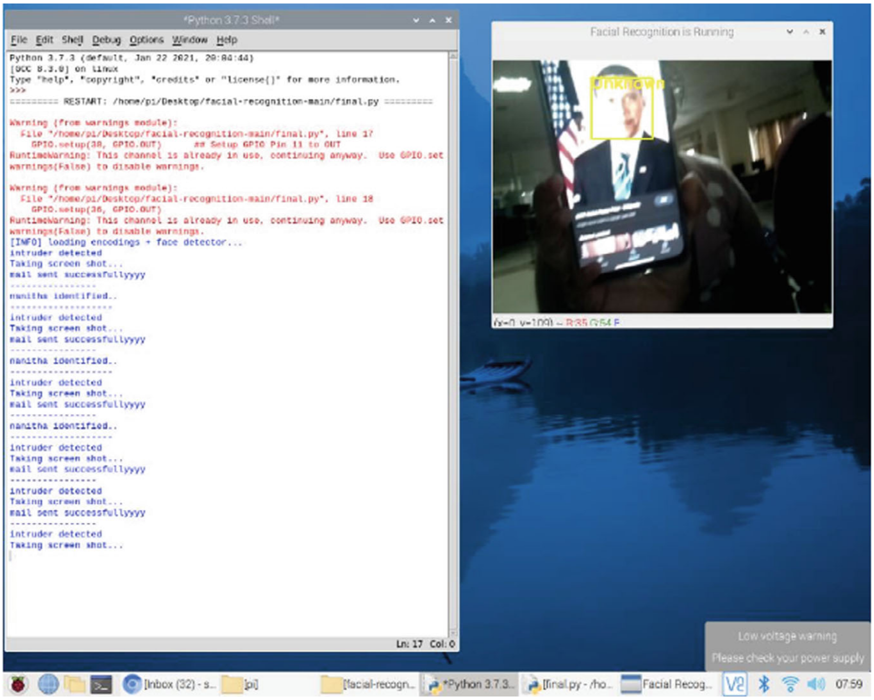


Fig. 5 Identifying intruders

5 Conclusion

In conclusion, the security warning camera system presented here represents a blend of advanced technology and security requirements. Its successful integration of motion detection and facial recognition offers a holistic approach to security that can redefine the way we protect our spaces. As we continue to face evolving security challenges, this project serves as a testament to the power of innovative solutions to ensure the safety and peace of mind of individuals and organizations. The future of security is here, and it is intelligent, efficient, and responsive. By combining motion detection using PIR sensors with advanced facial recognition algorithms, this system provides a proactive and comprehensive solution for the security of residential and commercial spaces.

The use of PIR sensors ensures that the system is triggered only when relevant motion is detected, reducing false alarms and conserving computational resources. Once motion is detected, the system uses a sophisticated algorithm to determine whether human presence is involved, further reducing the likelihood of false alarms due to non-human movement. The user-friendly interface and seamless email alert system make it accessible and convenient for users to monitor and secure their spaces remotely. Our security alert camera system addresses this need by providing real-time monitoring, proactive threat detection, and user-friendly controls. By harnessing the power of technology to protect our spaces, we take an important step toward creating a safer and more secure environment for individuals and businesses alike. The project underscores the potential for smart surveillance systems to revolutionize the way we approach security, ultimately contributing to peace of mind and security in an ever-evolving world.

References

1. Rizki RP, Hamidi EAZ, Kamelia L (2020) Image processing technique for smart home security based on the principal component analysis (PCA) methods. In: 2020 6th International conference on wireless and telematics (ICWT)
2. Tanwar S, Patel P, Patel K, Tyagi S, Kumar N, Obaidat MS An advanced internet of thing based security alert system for smart home. In: 2017 International conference on computer, information and telecommunication systems (CITS)
3. Ehsan MA, Mahfuj MMI, Hossain GMI (2016) Smart home automation and security system. Research pp 1–3. https://doi.org/10.13140/RG.2.2.19635.84000
4. Kumar M, Shimi SL (2015) Voice recognition based home automation system for paralyzed people. Int J Adv Res Electron Commun Eng (IJARECE) 4(10). Paul Barry (2006) Head first python. Shroff Publishers & Distributors Pvt. Ltd
5. Veer ND, Momin BF (2016) An automated attendance system using video surveillance camera. In: 2016 IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)
6. Kumar I, Kothari Y, Kapoor H, Bhatnagar A, Mishra P (2018) Securing smart homes using face recognition techniques. Int J Adv Res search Comput Sci 9(2):280–283
7. Patel MM, Jajal MA, Vataliya DB (2015) Home automation using raspberry pi. Int J Innov Emerg Res Eng 2(3)
8. Akhmetzhanov BK, Gazizuly OA, Nurlan Z, Zhakiy N (2022) Integration of a video surveillance system into a smart home using the home assistant. In: 2022 International conference on smart information systems and technologies (SIST)
9. Poongothai M, Subramanian PM, Rajeswari A (2018) Design and implementation of iot-based smart laboratory. In: 2018 5th International conference on industrial engineering and applications, ICIEA 2018, pp 169–173. https://doi.org/10.1109/IEA.2018.8387090
10. Yang Q (2019) Parameters optimization and application of SVM based on PCA-particle swarm algorithm. Int J Sci Res Sci Eng Technol 6(4):325–330. https://doi.org/10.32628/ijsrset196431
11. Aishu'eili H, Gupta GS, Mukhopadhyay S (2011) Voice recognition based wireless home automation system mechatronics (ICOM), 2011 4th International conference on 17–19 May 2011
12. Basavaprasad B, Ravi M (2014) A study on the importance of image processing and its applications. IJRET Int J Res Eng Technol 3(3):155–160
13. Sheeba B, Nair AS, Akhil SR, Abraham A, Abraham AA (2016) MIA-my intelligent assistant. Int J Adv Res Comput Commun Eng 5(4)
14. Divyarajsinh P, Brijesh M (2013) 3D Face recognition system based on 3D Eigenfaces. LAP Lambert Academic Publishing

Detection and Monitoring of Impurities in Water Using IoT

A. Vijaya Krishna, S. RamaCharan, Datta Manvita Krovvidi, Pranathi Gaddam, and Harika Satya

Abstract Water contaminants must be identified and monitored in order to provide clear, secure water for a variety of uses. Internet of Things (IoT), which has profoundly transformed this industry, now enables remote and real-time water quality monitoring. This research will examine how the Internet of Things (IoT) can be utilized to monitor and detect contaminants in water sources. Internet of Things (IoT) is a network of sensors and devices that capture and share real-time data on water quality indicators such as pH, turbidity, dissolved oxygen, and pollution levels. These sensors are strategically deployed in water bodies, distribution networks, and treatment facilities to provide comprehensive coverage and rapid response to any anomalies. The collected data is relayed to a central system that uses data analytics to access water trends, predict potential contaminations, and trigger alerts in case of deviations from preset norms. The benefits of IoT-enabled impurity detection are manifold. Real-time monitoring aids in the optimization of water treatment processes, reducing operational costs and resource wastage. Moreover, the seamless integration of IoT with mobile applications and online platforms empowers authorities to engage communities in water quality management, fostering awareness and collective responsibility. In conclusion, the integration of IoT technology in water impurity detection and monitoring holds great promise for enhancing water safety and sustainability.

Keywords Internet of things · Water quality · Sensors · pH value

A. Vijaya Krishna (✉) · D. M. Krovvidi · P. Gaddam · H. Satya
Department of Computer Science and Technology, GNITS, Hyderabad, India
e-mail: Vijayakrishnaakula2022@gnits.ac.in

S. RamaCharan
Department of Information Technology, GNITS, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_27

1 Introduction

Water is a resource that is essential for supporting life and preserving public health. However, as risks like industrialization, urbanization, and climate change become more prevalent, water contamination has increased, creating difficult management problems for water quality [1, 2]. Integration of Internet of Things (IoT) technology with environmental monitoring systems has come to light as a viable remedy for these problems in recent years [3]. A vital resource, groundwater is especially useful for drinking and irrigation [4]. The vast majority of people who live in rural areas rely on groundwater supplies for their everyday needs. Even though 60 to 85% of India's groundwater is used for drinking and farming, both natural and human activities have a large impact on how much of it is contaminated [5, 6]. Toxic substances in rocks and soils, such as fluoride-rich minerals, are the cause of natural contamination. Nitrate NO3 and fluoride F ions are the most pervasive pollutants in groundwater, and in recent years, the chemical composition of groundwater and the resulting health risks have been the main areas of research attention [7]. High fluoride 1.5 mg/L and nitrate45mg/L concentrations in drinking water cause heath impact on the human body [8].

2 Literature Survey

This study examines the concerns, issues, difficulties, and knowledge gaps in the water quality literature from 2018 to 2022 [9]. It also provides a taxonomy of IoT-based water monitoring systems based on their architecture, connection, data processing, and application domains. But his work mainly lacks on implementation in real-time scenarios.

A wireless, intelligent pH sensor that connects to a cloud platform for data collecting and calibration can be used to continually monitor the pH of a feeding solution [10]. Water contaminants, salts, and hydrogen ions in fresh water and wastewater are also discussed, and linear regression is utilized to interpret the results using machine learning. The main concern of this work is communication protocols.

This paper offers a system that can employ IoT devices to monitor water quality indicators such as pH, turbidity, conductivity, dissolved oxygen, and temperature in subsurface pipelines [11]. This work suffers from accurate values due to sensor readings.

Vaishnavi et al. [12], conducted a literature review on sensor-based and IoT-based smart water quality monitoring systems in this work. The benefits of several strategies, techniques, and technology for monitoring water quality measures are discussed. But his work mainly lacks on communication protocols which suits for IoT.

This research describes a system that employs the Internet of Things to detect water impurities such as temperature, pH, turbidity, and water flow [13]. If any of the parameters exceeds the threshold, the system can send the data to a cloud server and notify users by SMS or email. Preprocessing of data in the cloud is a hectic one.

3 Proposed System

This work focuses on interfacing an ESP32 microcontroller with a pH sensor, turbidity sensor, and an OLED display. The pH sensor is connected to analog input pin 32, while the turbidity sensor is connected to analog input pin 33. The OLED display is connected to pins 21 and 22. The collected pH sensor data is sent to the Blynk IoT app, which provides live text data for both sensors, as well as graphical historical data. The Blynk app also enables alerts and notifications for both turbidity and pH, including email alerts for low and high pH values. The entire hardware setup is powered by a 5 V USB power supply, and the programming is done using the Arduino IDE with the necessary libraries. This paper offers a comprehensive solution for monitoring pH and turbidity levels, providing real-time data visualization, and sending alerts to ensure water quality control (Fig. 1).

The water quality monitoring system comprises several essential components interconnected to ensure efficient data collection and remote monitoring. At its core is the ESP32 microcontroller, a versatile unit that not only controls the system but also facilitates Wi-Fi and Bluetooth connectivity, ideal for IoT applications. Two key sensors, the pH sensor and turbidity sensor, are connected to the ESP32, with each providing analog data on water acidity and suspended particle levels, respectively. This data is then processed and sent to the OLED display, which presents user-friendly visualizations of pH and turbidity readings. The Blynk IoT app serves

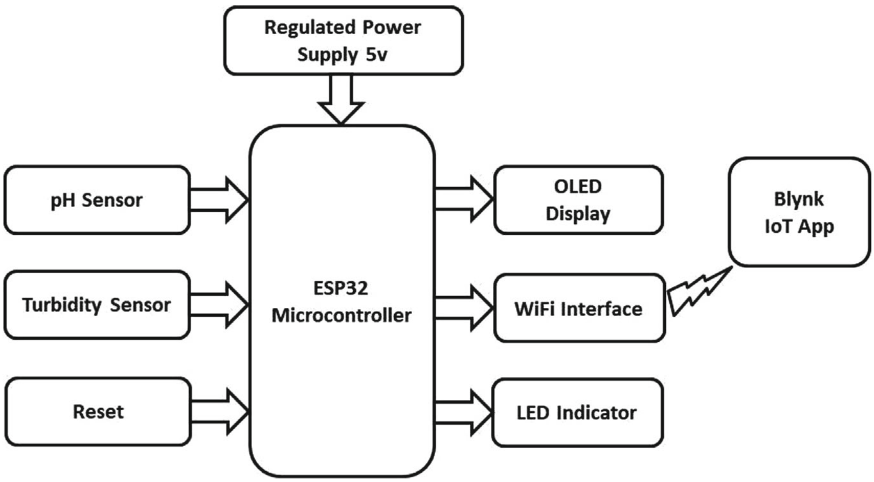


Fig. 1 System architecture

Table 1 Measurement values

Test cases	Turbidity value (trbt)	pH value (pH)	Alert
1	100	7.6	trbt: High turbidity, Impure water
2	20	7.7	–
3	40	9.6	pH: High pH value
4	20	3.2	pH: Low pH value
5	45	6.5	–

as the remote interface, allowing users to monitor and control the system. The ESP32 establishes a Wi-Fi connection to the Blynk server, enabling real-time data visualization and interaction. Lastly, the system is powered by a 5 V USB source, ensuring consistent operation. This block diagram showcases the seamless flow of data and communication between components, with the ESP32 at the center, orchestrating the entire water quality monitoring process.

4 Results

This paper effectively integrates an ESP32 microcontroller with pH and turbidity sensors, displaying data on an OLED screen and transmitting it to the Blynk IoT app for remote monitoring. The Blynk app offers a user-friendly interface for real-time and historical data visualization, including alerts for pH levels exceeding set thresholds. Overall, this system proves efficient for smart water purification and monitoring, ensuring prompt action to maintain water quality standards (Table 1).

Turbidity function = Converted an analog sensor value (presumably sensor value) from a range of 0 to 2200 to a new range of 100 to 0.

- **If** $0 < \text{Turbidity} < 50$—Pure Water

  else—Impure
- **If** $5 < \text{pH} < 8$—Pure Water
- **If** $\text{pH} > 7$—High pH
- **If** $\text{pH} < 5$—Low pH

5 Conclusion

The monitoring system presented in this paper signifies a successful fusion of IoT technology with water quality management. Through the integration of an ESP32 microcontroller, pH, and turbidity sensors, as well as an OLED display and the Blynk IoT application, a comprehensive, user-friendly solution has been realized. This system offers real-time monitoring of pH and turbidity levels, supplying precise, current data on water quality. The OLED display provides convenient, on-site access to sensor readings, while the Blynk app extends remote accessibility, enabling users to oversee water quality from virtually anywhere and at any time. The incorporation of alert functionality elevates the system's effectiveness by issuing notifications, including email alerts, when pH values surpass predefined thresholds. This ensures that users receive timely notifications about potential water quality issues, empowering them to take immediate corrective actions.

In essence, this paper serves as a testament to the immense potential of IoT technology in water purification and monitoring. The synergy of sensors, microcontrollers, and IoT connectivity offers an efficient means of upholding water quality standards. With further enhancements and refinements, this system stands poised to play a vital role in ensuring clean and safe water across diverse applications, encompassing drinking, agriculture, and industrial use.

References

1. Geetha S, Gouthami SJSW (2016) Internet of things enabled real time water quality monitoring system. Smart Water 2(1):1–19
2. Rao, AS, Marshall, S, Gubbi, J, Palaniswami, M, Sinnott, R, Pettigrovet, V (2013) Design of low-cost autonomous water quality monitoring system. In: 2013 International conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 14–19
3. Ward RC, Loftis JC, McBride GB (1991) Design of water quality monitoring systems. John Wiley & Sons
4. Prasad, AN, Mamun, KA, Islam, FR, Haqva, H (2015) Smart water quality monitoring system. In: 2015 2nd Asia-Pacific world congress on computer science and engineering (APWC on CSE). IEEE, pp 1–6
5. Chowdury MSU, Emran TB, Ghosh S, Pathak A, Alam MM, Absar N, Hossain MS (2019) IoT based real-time river water quality monitoring system. Procedia comput Sci 155:161–168
6. Dong J, Wang G, Yan H, Xu J, Zhang X (2015) A survey of smart water quality monitoring system. Environ Sci Pollut Res 22:4893–4906
7. Chen Y, Han D (2018) Water quality monitoring in smart city: a pilot project. Autom Constr 89:307–316
8. Martínez R, Vela N, El Aatik A, Murray E, Roche P, Navarro JM (2020) On the use of an IoT integrated system for water quality monitoring and management in wastewater treatment plants. Water 12(4):1096
9. Zulkifli CZ, Garfan S, Talal M, Alamoodi AH, Alamleh A, Ahmaro IYY et al (2022) IoT-based water monitoring systems: a systematic review. Water 14(22), p 3621. IoT Projects with ESP32: Build exciting and practical IoT projects using the powerful ESP32 by Agus Kurniawan
10. Pyingkodi M, Thenmozhi K, Nanthini K, Karthikeyan M, Kalpana T, Deepak, PV (2022) Examination of water impurities using IoT and machine learning techniques. In: IoT based control networks and intelligent systems: proceedings of 3rd ICICNIS 2022. Springer Nature Singapore, Singapore, pp 675–687. Internet of things with Arduino blueprints by Pradeeka Seneviratne
11. Nadu T (2020) An underground pipeline water quality monitoring using IoT devices. Eur J Mol Clin Med 7 2046–2054
12. Vaishnavi V, Varshitha RC, Tejaswini M, Biju NR, Kumar K (2018) Literature survey on smart water quality monitoring system. Int J Innov Eng Sci 3(3):20–24
13. Bhisekar S, Meghare C, Parate H, Prajapati S, Meshram, B, Bobade NP (2018) An IoT based water monitoring system for smart city

A Unique Algorithm About Global Enhancement Problems Present in African Vulture Optimization Algorithm

Suraya Mubeen, Ganapathi Antharam, B. Kavitha Rani, Abdul Subhani Shaik, P. Raveendra Babu, and Venkataiah

Abstract Developing an arrangement or structure as entirely excellent functional or effectively as feasible is called optimization. Strategies that optimize or reduce certain research variables are found using methods of optimization. A higher-level heuristic called a metaheuristic, for short, is used in optimizing to create a process that effectively solves issues like limited computing power and partial, incomplete data. Estimating techniques were offered as a novel method for resolving complicated and multidimensional situations. Heuristic and metaheuristic algorithms are two prominent subcategories of these. Metaheuristic approaches are utilized for a variety of challenging and intricate optimal problems, while heuristic techniques are employed for local traps and solving specific optimal issues. Exploring and profiting are both goals of metaheuristic. This approach initially uses exploration to find solutions before moving on to exploitation, which increases the precision of the exploration phase. The use of metaheuristics in continuous, discrete, and optimization has increased significantly. Intelligent beings and organisms in nature create these kinds of algorithms. Many of these metaheuristic algorithms have recently

S. Mubeen
Department of Electronics and Communication Engineering, CMR Technical Campus, Hyderabad, Telangana, India

G. Antharam
Department of Computer Science and Engineering, CMR Technical Campus, Hyderabad, Telangana, India

B. Kavitha Rani (✉)
Department of Information Technology, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: cmrtc.paper@gmail.com

A. S. Shaik · P. Raveendra Babu
Department of Electronics and Communication Engineering, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India
e-mail: praveendrababu@cmrcet.ac.in

Venkataiah
Department of Computer Science and Engineering, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_28

been presented, and one of them is known as “African Vulture Optimization Algorithm (AVOA).” This suggested a lifestyle modeled by African vultures. Based to the fundamental ideas regarding vultures, the AVOA method is going to be developed in a step-by-step manner, with every requirement met and points defined in each stage. There are four phases in the AVOA technique: phases 1, 2, 3, and 4. Phase 1 is used to identify the best vultures for each group, Phase 2 is used to determine the pace at which vultures starve, Phase 3 is used for exploration, and Phase 4 is used for exploitation. African vulture algorithm for optimization holds optimization employing these stages. Consequently, we anticipate that the “African vulture optimization algorithm” will produce greater outcomes than any other algorithm that has been presented so far.

Keywords Optimization · Approximation algorithms · Metaheuristic · Heuristic · Exploration · Exploitation · African vulture optimization algorithm · African vulture life style · Phases

1 Introduction

There exist numerous optimization issues in the current industrial sector. These optimization issues provide us with specific criteria to attain our objective in specific circumstances. Technology is advancing quickly today, and numerous industries are using strategies to solve optimization challenges in order to create high-quality solutions. Various algorithms have been developed for creating these solutions by numerous scientists, while these methods have also been developed by observing the behavior of various animals. As a result, such algorithms have been successful in providing excellent solutions to optimization issues.

Numerous algorithms had previously been in use, and they were categorized using various techniques like heuristic and metaheuristic techniques. Heuristic is defined as a procedure that uses the algorithm and depends on the time and space-intensive specifications. It is said that metaheuristic approach is an enhanced variation on the heuristic technique. This offers instructions for the heuristic procedure. The heuristic approach can provide a workable solution to the optimization problem, but metaheuristics cannot promise to do the same, but they can still tackle challenging optimization problems.

Additionally, solitary solution-based as well as population-based metaheuristic optimization algorithms are subcategories of metaheuristic optimization methods. Because single solution-dependent metaheuristic algorithms for optimization can only be useful for a single optimal solution rather than for all optimum problems, they have limitations when it comes to improving the best possible solutions. However, all solutions to optimal challenges are suitable for based on population’s metaheuristic optimization techniques.

In order to avoid arriving at the best possible outcome and to improve the solution's achievement of the desired/expected aim, metaheuristic optimization approaches typically undergo exploration as well as utilization.

Such algorithms are created by looking at nature. There was numerous methods in use earlier, especially the Cuckoo search technique (CAO), and as a result of minimizing the downsides, many more algorithms also came into existence. This process led to the development of the teaching learning based algorithm (TLBO). African vulture optimization algorithm (AVOA), which is modeled after the behavior of African vultures, is created by restricting the flaws in these current methods. The architecture of this algorithm was inspired by ways African vultures obtain food, either alone or through other vultures.

2 Literature Review

According to Abdollahzadeh et al. in 2021 [1], the success of machine learning approaches is crucial to the feature selection (FS) phase of data mining. Processing is accelerated, and categorization accuracy is enhanced. In this study, three distinct strategies to FS are proposed. In the subsequent solution, the Harris Hawks Optimization (HHO) technique was multiplied, in the first solution, the Fruitfly Optimization Algorithm (FOA) have been multiplied, and in the final solution, both of these solutions are hybrided and assigned the name MOHHOFOA. On 15 typical data sets, the performance of the MOPSO, NSGA-II, BGWOPSOFS, and B-MOABC strategies for FS was evaluated employing mean, best, the very least, and standard deviation (STD) parameters.

The Arithmetic Optimization Algorithm (AOA), that is presented in the current study by Abualigah et al. [2], is a novel metaheuristic method that uses the distributional conduct of each of the four basic arithmetic operators in mathematical concepts (Multiplication (), Division (), Subtraction (), and Addition ()). AOA is theoretically modeled and put into use to accomplish the optimization operations in a number of search domains. AOA's effectiveness is compared to 29 standard functions and a variety of real engineering design difficulties in order to show how applicable it is. The suggested AOA's evaluation of performance, convergent behaviors, and computation complexity was all put through the test using various scenarios. Experimental results show that the AOA beats 11 other popular optimization algorithms in tackling challenging optimization challenges.

The Aquila Optimizer (AO), a unique population-based optimization method suggested in this article, takes its cues from the typical actions of aquilas as they chase their prey, according to Abualigah et al. [3]. The four techniques are thus: high soar about vertical stoop to choose the area of search space; contouring flight about short glide attack to discover throughout a divergent looking space; low flight about slow descent attack to take advantage throughout a converge in search space; while walk as well as grab prey to swoop. The newly developed optimizer's ability to find the optimal solution for diverse optimization problems is tested through a series of experiments.

The clustering of data constitutes a subfield of unsupervised learning that entails classifying samples as groups, wherein participants are comparable to one another, according to Rahnema et al. [4]. While the method known as K-means is a simple and quick clustering technique, it comes with a few initial problems. For successful clustering, for example, a lot hinges on the beginning value. It is also susceptible to imbalanced clusters and outliers. Numerous artificial bee colony (ABC) technique is one of the metaheuristic techniques utilized today to address many optimization problems, including clustering. This method has two main problems: investigation and late convergence. In this paper, the exploration and delayed convergence problems in ABC are addressed using an algorithm called ABCWOA which integrates Random Memory (RM) with Elite Memory (EM). In feeding ABCWOA technique, RM has taken advantage of the whale optimization algorithm's (WOA) search phase for the bait, and EM uses it as well to increase convergence. Furthermore, we continuously control how EM is used.

Fan et al. [5] one of the best techniques for resolving challenging engineering issues is the use of metaheuristic optimization algorithms. However, a metaheuristic algorithm's performance is influenced by its capacity for exploration and exploitation. Therefore, an enhanced African vulture optimization algorithm based on tent chaotic mapping and time-varying mechanism (TAVOA) is presented in order to further enhance the African vulture optimization algorithm (AVOA). A tent chaotic map is first created to initialize the population. Second, the person's previous ideal position is noted and used to update their location.

Shaddeli [6] The feeding and orienting habits of African vultures serve as the basis for the African vulture optimization algorithm (AVOA). It has strong operators and balances exploration and effectiveness when addressing optimization tasks. This algorithm must be discretized in order to be applied in discrete applications. In this study, two new variants are presented, based on the S- and V-shaped transfer functions of AVOA and BAOVAH, respectively. Also avoided is a rise in computational complexity. To enhance the performance of this model, disruption operator and bitwise approach have also been employed. The BAVOA-v1 is a multi-strategy AVOA version that is provided.

Abdollahzadeh [7], Skandha [8], Venkataiah et al. [9] the majority of these algorithms take their cues from the natural world's animal life and collective intelligence. The lifestyle of African vultures is used as inspiration for a new metaheuristic in this paper. The system, known as the African vulture optimization system (AVOA), models the foraging and navigational habits of African vultures. The performance of AVOA is then assessed using 36 common benchmark functions. The proposed algorithm's superiority over a number of other algorithms is then shown through a comparative analysis.

3 African Vulture Optimization Algorithm (AVOA)

Abdollahzadeh et al. presented the AVOA method in 2021 [1], and the approach has since been used to solve a variety of real-world engineering challenges. A novel metaheuristic technique called AVOA was developed as a result of study on where to locate and how to feed different vultures in Africa [7]. This population-based AVOA optimization solving method is capable of resolving any optimal issues. The four steps of this AVOA technique's operation are selecting the most suitable vulture from among a variety of vultures, starving vultures, investigation, and exploiting. The AVOA techniques address the numerous difficult optimization issues depicted in Fig. 1 according to these four stages.

At first, vultures gather in groups. Following arrangement, it moves into additional stages.

3.1 Phase 1-Detection of Best Vulture Among Various Vultures

Following arrangement, the algorithm's first phase occurs when the best vulture is discovered. The groups known as the first best vulture group and second best vulture group are chosen from across several groupings, and the remaining vultures are then formed into a third group. The following equations are used to choose the most suitable groups

$$R(i) = \begin{cases} \text{Best Vulture } 1 & \textit{if } \; p(i) = \text{L1} \\ \text{Best Vulture } 2 & \textit{if } \; p(i) = \text{L2} \end{cases} \tag{1}$$

while L1 and L2 are specific random numbers within a particular set of borders, and $R(i)$ is considered to represent a best vulture group across numerous best vulture

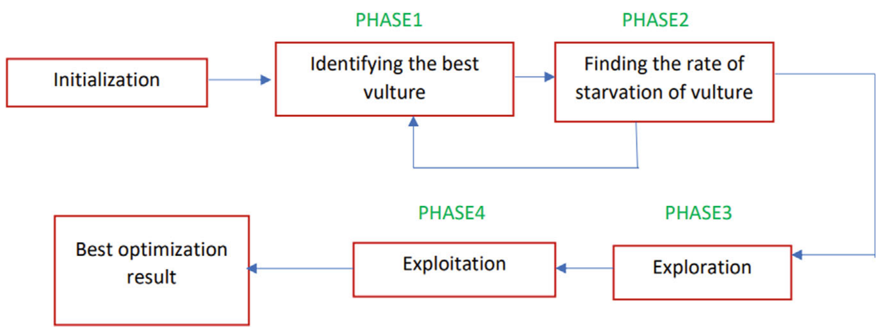


Fig. 1 AVOA block diagram

groups, and $p(i)$ is considered to represent the probability about best vultures group. Either L1 or L2 has to add up to one. The aforementioned roulette wheel equation yields the result indicated below, which is $p(i)$.

$$p(i) = \frac{F}{\sum_{i=1}^{n} F} \tag{2}$$

while n is referred to as the total amount of vultures and F is referred to as the vulture's satisfying factor

3.2 Phase 2-Starvation of Vultures

The second phase, whereby vulture famine occurs, begins after finding the best vulture. The vulture technique is used for locating food during this phase. Certain vultures with the energy to fly can go great distances in search of food, while some vultures with insufficient energy to accomplish so will be reliant on the other powerful vultures with sufficient energy. Because they are dependent on powerful vultures for food, weak vulture are going to attack strong vultures and behave aggressively in an effort to steal food from them. The aforementioned equation, which is shown here, is therefore applicable for estimating this starvation.

$$F = (2 * K + 1) * z * (1 - (\text{iteration } i1/\text{max iterations})) + t \tag{3}$$

F is referred to as the adequate factor for vulture starvation, K is referred to as a number that is selected at random, t is referred to as the vulture's time in search of food, and z is referred to as a random value that ranges from [0, 1].

When vultures are found to be starving, the research stage begins. Exploration and profit-seeking are used to conduct this study stage. According to the vulture malnutrition factor, one of them is chosen out of the group. Vultures go through the exploration phase if the amount of F is higher than 1; otherwise, they move on to the extraction stage.

3.3 Phase 3-Exploitation

Although this is still the inquiry stage, the vultures would primarily search for food in various locations and take some time to do so. In this instance, two approaches are taken into account: first, they will look for food locally, and then that they will travel farther. The parameter P1 is responsible for this misuse. According to the calculations above, this variable determines where the plan for the vulture should progress.

$$P(i+1) = \begin{cases} \text{Equation} & \textit{if } \text{P1} \geq \text{KP1} \\ \text{Equation} & \textit{if } \text{P1} < \text{KP1} \end{cases} \tag{4}$$

Hence, $P(i+1)$ is referred to as the $i+1$th vulture's placement. One claims that K is a randomized number.

The formula beneath calculates one of both strategies.

$$P(i+1) = R(i) - D(i) \times F \tag{5}$$

While F is the starving percentage of vultures, $R(i)$ is considered to indicate a picking of greatest vultures across multiple vultures groups, and $D(i)$ is considered to be distances between vulture and current ideal rating of the vultures. $D(i)$ is determined using the formula shown below.

$$D(i) = |X \times R(i) - P(i)| \tag{6}$$

When one approach fails, a different strategy is used, as demonstrated in the calculation below.

$$P(i+1) = R(i) - F + \text{K2} \times ((u - l) \times \text{K3} + l) \tag{7}$$

wherein the upper and lower borders are denoted by u and l.

Excavation phase follows if the famine factor fails to fulfill this phase of investigation.

3.4 Phase 4-Exploitation

The exploitative phase begins whenever the famine index is less than 1. Various approaches will be used throughout this phase, and any of them may be used. According to the P2 and P3 variables as indicated in the following equations, there are going to various approaches in those two methods, and thus, these might be chosen.

$$P(i^{\cdot} + 1) = \begin{cases} \text{Equation} & \textit{if } \text{P2} \geq \text{KP1} \\ \text{Equation} & \textit{if } \text{P2} < \text{KP1} \end{cases} \tag{8}$$

Formula ensues whenever P2 is bigger and equivalent to the component P1's randomized number.

$$P(i+1) = R(i) - D(i) \times F \tag{9}$$

$$D(i) = |X \times R(i) - P(i)| \tag{10}$$

Since this example fails, it moves on to a different scenario, as stated in the calculation below.

$$P(i + 1) = D(i) \times (F + K) - d(t) \tag{11}$$

$$d(t) = R(i) - P(i) \tag{12}$$

In this section $P(i)$ is referred to as the ith vulture's their position, $D(i)$ is determined at random, and the position of the $i + 1$th vulture is specified.

Consequently, competition and rotating flying strategies are present in exploitation.

4 Cuckoo Search Algorithm (CAO)

Xin-She Yang and Suash Deb created an optimized algorithm known as Cuckoo Search in 2009. Cuckoo Search is just one among several nature-inspired algorithms widely utilized to address optimization issues in various engineering disciplines as shown in Fig. 2.

4.1 *Host Discovering of Cuckoo Eggs*

The home bird has a chance of finding cuckoo eggs having a probability of $p(a)(0,1)$, while $p(a)$ is the probability of finding foreign eggs. If an intruding bird finds the cuckoo bird eggs, it might discard the egg, depart the nest as see in Fig. 3, and create a brand-new nest.

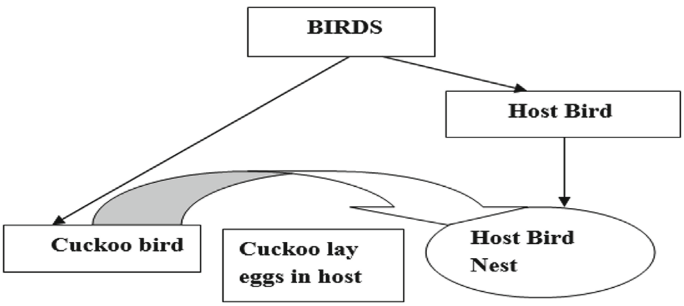


Fig. 2 Cuckoo search algorithm

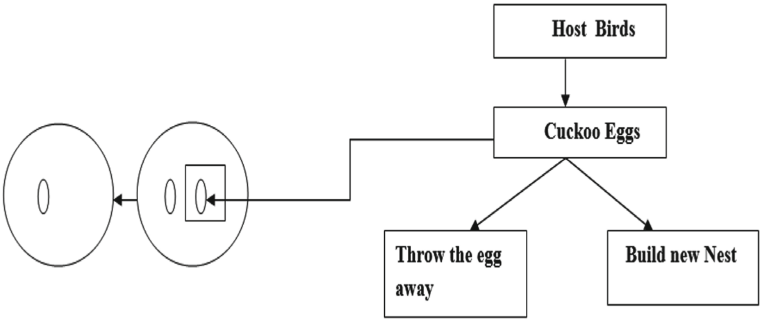


Fig. 3 Discovering of cuckoo eggs

In this section, the cuckoo bird uses the Levy Flights to produce a Fresh Solution (Egg) in the host bird nest.

$$[[x]]\_i^\wedge(t+1) = x\_i^\wedge t + \alpha \otimes \mathrm{Levy}(\lambda) \tag{13}$$

where x_i^{t+1} is a new solution, x_i^t is the current location, α is step size, Levy (λ) Levy exponent, where Levy $(\lambda) = t^{-\lambda}$.

5 Teaching Learning Based Algorithm (TLBO)

It depends on how a teacher's influence affects students' performance in a class. The algorithm's two most important elements are the instructor and the students.

5.1 Teacher Phase

This stage models how students (i.e., learners) interact with teachers. In this phase, an instructor imparts information to the students and works to raise the class mean score. There are several steps in the instructor's phase. Creating a novel solution, determining fitness, and using the avaricious choice

$$X\_\mathrm{new} = X + r[[(X]]_{\mathrm{best}} - T_f X_{\mathrm{mean}}) \tag{14}$$

While X_{new} is referred to as the most recent solution, X is the current answer, X_{best} is the most qualified instructor, and X_{mean} is the average student score. T_f stands for instructor factor, and it might be 1 or 2. Calculations under compute T_f and X.

$$T\_f = \mathrm{round}(1 + K) \tag{15}$$

$$X = l + K. * (u - l) \tag{16}$$

5.2 Learning Phase

This stage represents the learning that occurs when students engage with one another. Additionally, the students can learn through conversing with and relating to their fellow classmates. There are two steps in this educational period. Preliminary new solutions are produced by utilizing partner solutions (X_p).

For maximum

$$X_{\text{new}} = X + r\left(X - X_p\right) \quad \text{if} f > f_p \tag{17}$$

$$X_{\text{new}} = X - r\left(X - X_p\right) \quad \text{if} f < f_p \tag{18}$$

For minimum

$$X_{\text{new}} = X + r\left(X - X_p\right) \quad \text{if} f < f_p \tag{19}$$

$$X_{\text{new}} = X - r\left(X - X_p\right) \quad \text{if} f > f_p \tag{20}$$

While f is considered to be the fitness of the current student, X_p is considered to be the other person's solution, and f_p is considered to be the fitness of the partnership solution. Following the discovery of the novel solution, aggressive selection based on the highest or lowest condition occurs, as illustrated here.

$$\text{For maximum} : f_{\text{new}} > f_{\text{old}} \tag{21}$$

$$\text{For minimum} : f_{\text{new}} < f_{\text{old}} \tag{22}$$

6 Simulation Results

Since AVOA is a recently approved technique while there has been many other algorithms, these prior techniques are used in the standard equations (Unimodal Functions). Presently, the variance in erasing the ideal errors is seen, as indicated in the tables above.

Unimodal Benchmark Equations:

$$\text{F1} - \text{De Jong's} - f(x) = \sum_{i=1}^{d} x_i^2$$

$$\text{F2} - \text{Rastrings} - o = 10.0 * \text{size}(x, 2) + \text{sum}(x.\hat{}2 - 10.0 * \cos(2 * \text{pi}. * x), 2$$

$$\text{F3} - \text{Step} - o = [(\text{sum}(x)) + 0.5].\hat{}2$$

6.1 Observation of African Vulture Optimization Algorithm

The mistake has been optimized according to De Jong's function. In De Jong's function, the initial mistakes was too large at 133346 over the given function, yet during the subsequent iteration, mainly the mistake was corrected, bringing the overall amount of iterations down to zero and clearing the optimal issues. The technique's error minimization is shown in Fig. 4.

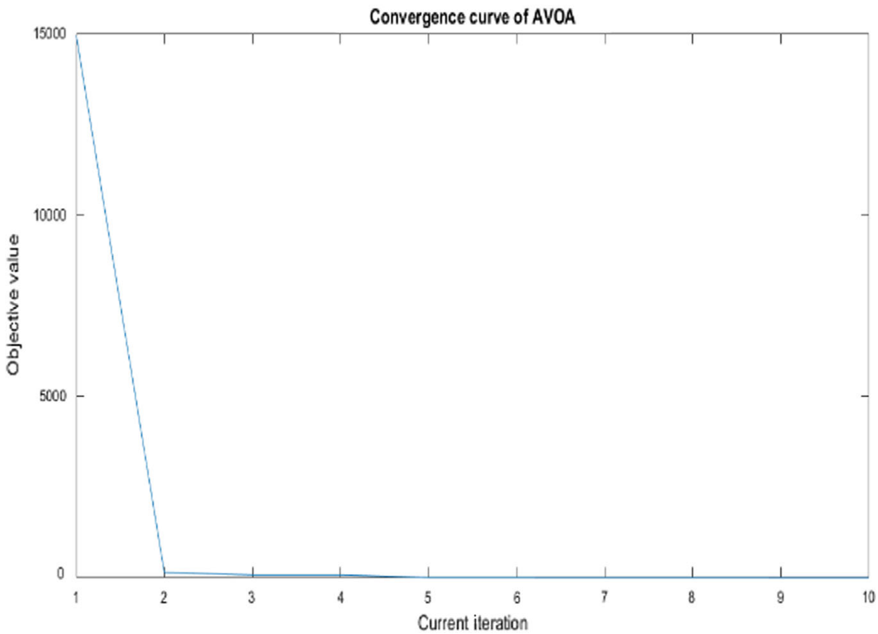


Fig. 4 De Jong's equation explanation for AVOA

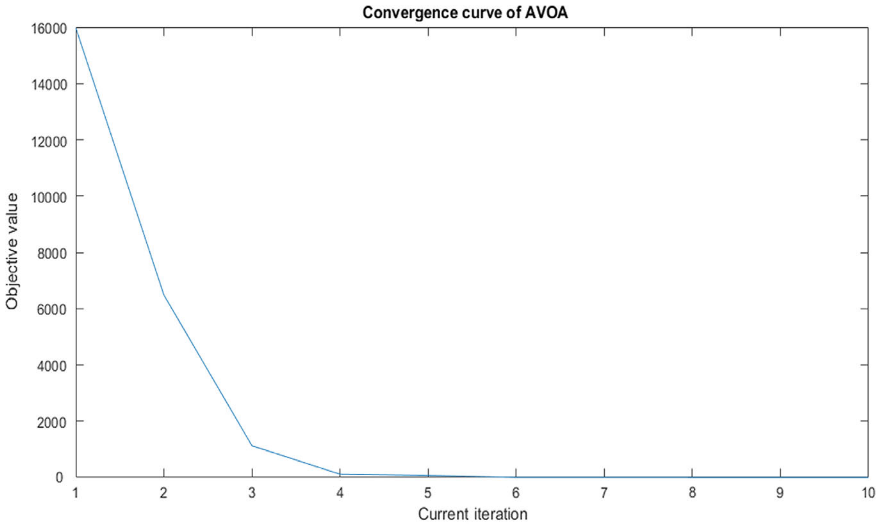


Fig. 5 Rastring's equation solution for AVOA

The error is reduced with the Rastring's function. In this section in the Rastring's function, the error was lowered after less rounds. In the beginning, the error was excessively large for the given function, at 14386.32, but after 14 iterations, it was only corrected to 0.006 and cleared the optimal issues. The algorithm's error minimization is shown in Fig. 5.

Error has been optimized for a particular function step. In this phase of the function, the initial mistake is too large at 16.975 compared to the other functions, yet during the second iteration, the error was only reduced to 0.5 and the algorithm eventually solved the ideal issues. The system's error optimization is shown in Fig. 6.

6.2 *African Vulture Optimization Algorithm Was Contrasted Against Two Other Algorithms (Cuckoo Search and Instructional Learning-Based Techniques)*

As shown in Fig. 7, De Jong's functional is compared to the three techniques:

$$\text{AVOA} - \text{optimal solution} - \text{2nd iterations}$$

$$\text{CAO} - \text{optimal solution} - \text{no rectification}$$

$$\text{TLBO} - \text{optimal solution} - \text{9th iteration}$$

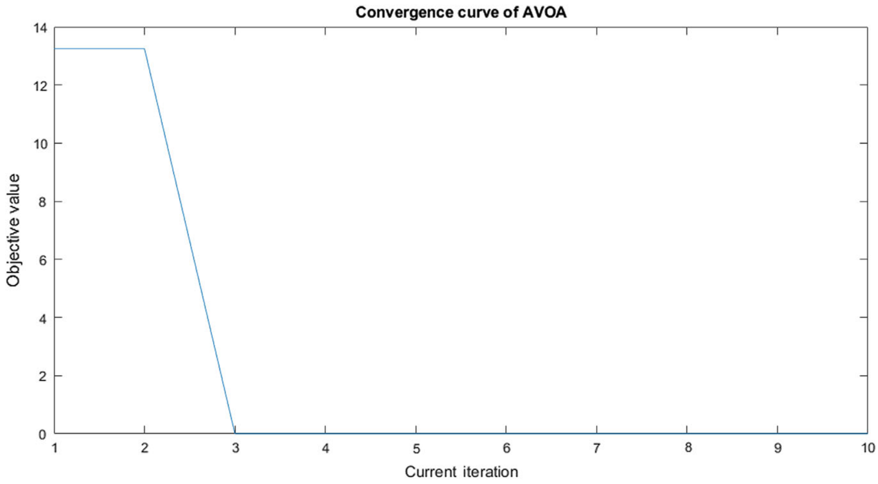


Fig. 6 Step equation solution for AVOA

As show in Fig. 8, the blue graph denotes AVOA, the yellow graph Cuckoo, and the red graph TLBO. It is evident from the foregoing that AVOA corrects errors more quickly when compared to either of the two methods.

Comparing of Rastring's function with all three algorithms.

$$\text{AVOA} - \text{optimal solution} - 14\text{th iterations}$$

$$\text{CAO} - \text{optimal solution} - 15\text{th iteration}$$

$$\text{TLBO} - \text{optimal solution} - 15\text{th iteration}$$

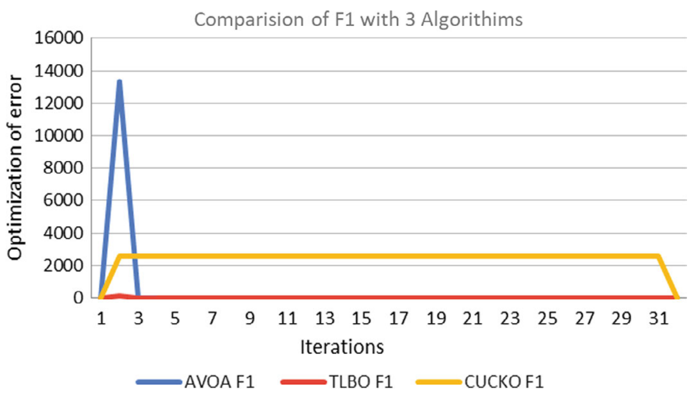


Fig. 7 Comparison of De Jong's equation

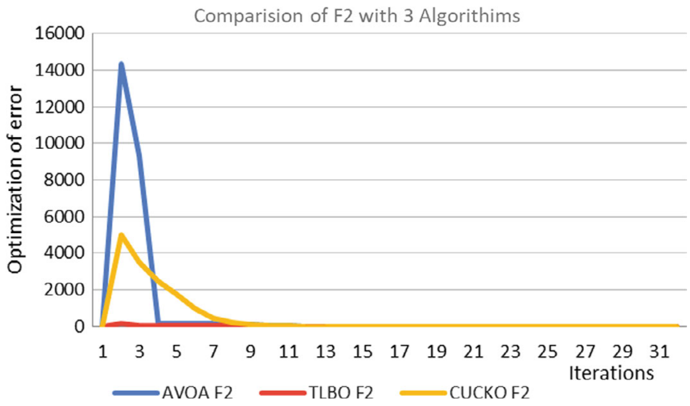


Fig. 8 Comparison of Rastring's equation

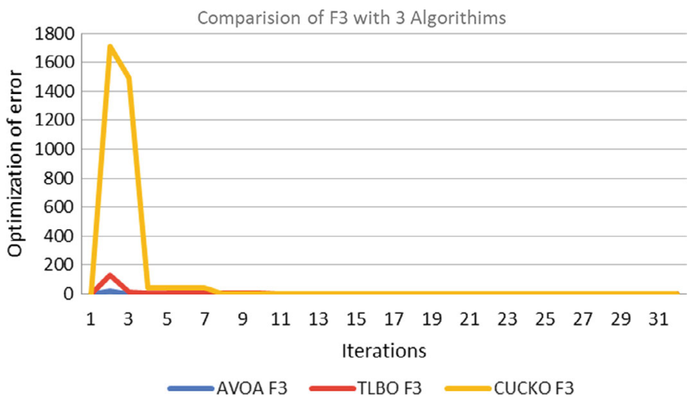


Fig. 9 Comparison of step equation

Comparing of step function with all three algorithms:

As shown in Fig. 9, the blue graph denotes AVOA, the yellow graph Cuckoo, and the red graph TLBO. It is evident from the foregoing that AVOA corrects errors more quickly when compared to either of the two methods.

$$\text{AVOA} - \text{optimal solution} - \text{2nd iterations}$$

$$\text{CAO} - \text{optimal solution} - \text{8th iteration}$$

$$\text{TLBO} - \text{optimal solution} - \text{10th iteration}$$

7 Conclusion

According to the aforementioned analysis of the three techniques, AVOA is the best one since it corrects the optimum errors within a minimum number of repetitions and, while containing a large amount of error, and it does so more quickly than the other two. AVOA additionally solves a few benchmark functions that are not addressed by other methods. Therefore, it can be said that AVOA is the most effective algorithm when compared to previously implemented algorithms.

References

1. Abdollahzadeh B, Gharehchopogh FS (2021) A multi- objective optimization algorithm for feature selection problems. Eng Comput 1–19
2. Abualigah L et al (2020) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376. Article 113609
3. Abualigah L et al (2021) Aquila optimizer: a novel meta-heuristic optimization algorithm. Comput Ind Eng. Article 107250. Anderson DJ, Horwitz RJ (1979) Competitive interactions amo
4. Rahnema N, Gharehchopogh FS (2020) An improved artificial bee colony algorithm based on whale optimization algorithm for data clustering. Multimedia Tools Appl 79(43):32169–32194
5. Fan J, Li Y, Wang T (2021) An improved African vultures optimization algorithm based on tent chaotic mapping and time-varying mechanism
6. Shaddeli A (2022) An improved African vulture optimization algorithm for feature selection problems and its application of sentiment analysis on movie reviews. Big Data Cogn Comput 6(4)
7. Abdollahzadeh B (2021) African vultures optimization algorithm: a new nature-inspired metaheuristic algorithm for global optimization problems. Comput Ind Eng 158
8. Skandha SS, Gupta SK, Saba L, Koppula VK, Johri AM, Khanna NN, Mavrogeni S, Laird JR, Pareek G, Miner M, Sfikakis PP, Protogerou A, Misra DP, Agarwal V, Sharma AM, Viswanathan V, Rathore VS, Turk M, Kolluri R, Viskovic K, Cuadrado-Godia E, Kitas GD, Nicolaides A, Suri JS (2020) 3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: Atheromaticâ„¢ 2.0. Comput Biol Med 125
9. Venkataiah V, Nagaratna M, Mohanty R (2022) Application of chaotic increasing linear inertia weight and diversity improved particle swarm optimization to predict accurate software cost estimation. Int J Electr Electron Res. https://doi.org/10.37391/IJEER.100218

Object Detection and Alerting System in Car Using IoT

A. Vijaya Krishna, V. Sesha Bhargavi, Bakkuru Hima, Mamilllapalli Vyshnavi, and Mareddy Pravalika

Abstract The "object detection and alerting system in car using IoT" for enhanced safety project introduces a novel approach to bolster safety by seamlessly integrating Internet of Things (IoT) technology into object detection and alerting systems. This project addresses the critical need for proactive detection of objects and obstacles in various environments, enhancing user awareness and minimizing potential hazards. Nowadays, there are many safety and security systems in the car such as airbags, anti-lock systems, etc. Whenever a hand is placed on the car door panel or frame and the car door is closed unnotified or unchecked or accidentally or in another case like whenever the children are playing and place their hand on the car door panel accidentally while the car door is closing, this causes injuries. So, we came up with a system "object detection on the car door panel and alerting system" in a car which detects the obstacle while closing the car door and prevents these types of injuries or accidents from occurring. The project's design embraces the principles of IoT scalability and adaptability. The modular architecture allows for seamless expansion of the network to cover larger areas or new environments. Additionally, it can integrate with existing IoT ecosystems, smart homes, or industrial setups, amplifying its usability across diverse applications.

Keywords IoT · Ecosystem · Alerting systems

1 Introduction

The car door detection system is an innovative project designed to enhance the safety and convenience of vehicle passengers by detecting hands or objects in the path of a closing car door. Once an object or hand is detected in the proximity of the closing car

A. Vijaya Krishna (✉) · B. Hima · M. Vyshnavi · M. Pravalika
Department of Computer Science and Technology, GNITS, Hyderabad, India
e-mail: Vijayakrishnaakula2022@gnits.ac.in

V. Sesha Bhargavi
Department of Information Technology, GNITS, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_29

door, the system triggers a warning mechanism [1–3]. This can involve audio alerts, buzzer sounds, or both, to notify passengers and prompt them to remove their hands or objects from the danger zone [4]. This system uses an Arduino microcontroller and an ultrasonic sensor to detect the presence of hands or objects near a closing car door.

These sensors are strategically placed around the vehicle, creating a protective sensor perimeter. When an object or obstacle enters this zone, the ultrasonic sensors detect it, and the Arduino microcontroller processes this data in real time [5]. If the detected object is too close to the vehicle, the Arduino triggers a buzzer to emit a distinct audible warning for the driver.

Sensors are strategically placed in the car door frame or surrounding areas. These sensors can include two ultrasonic sensors. The sensors continuously monitor the door frame area for any obstructions. The sensor data is processed to detect the presence of an object or obstruction in the door frame [6]. If the measured distance is below a certain threshold, it indicates the presence of an object. The system activates a buzzer to alert the driver or passengers when an object is detected. The buzzer can emit a loud and distinct sound, drawing attention to the potential obstruction. The volume and pattern of the buzzer can be designed to effectively communicate the urgency of the situation. Upon hearing the buzzer, the driver or passengers can take necessary action to prevent accidents [7, 8]. They can check the door frame area, identify the object or obstruction, and remove it before attempting to close the door again. The buzzer serves as a reminder and prompts immediate attention to ensure safety.

2 Related Study

In [9], Liu et al. proposed a novel car door control system that enhances safety by automatically locking car doors when objects approach from the front or rear. The system utilizes ultrasonic sensors, servo motors, and microcontrollers to detect nearby objects and prevent accidents and injuries. Previous research in this area has explored a similar concept involving child locks and ultrasonic sensors. However, this earlier model was limited in its functionality and outdated in terms of technology.

In [10], Yan et al. proposed a design of car reversing anti-collision warning system vehicle door opening warning safety system by preventing accidents caused by opening doors into the path of oncoming vehicles. In this work, the main limitations are communication protocols.

In [11], Lai et al. proposed a car door safety system with wireless alert transmission. The system utilizes a single digital camera and a field programmable gate array chip to implement motion detection, speed estimation, and collision warnings. But implementation of this wireless protocols in the car detection system is highly complicated.

In [12], Liu et al. design of intelligent car door control system. By employing a comprehensive arrangement of sensors, microcontrollers, actuators, and power modules, the system aims to enhance safety on highways and busy roads, particularly in India, where road accidents are a major concern. The system's core objective is to address the issue of negligent door opening, a significant contributor to accidents, injuries, and fatalities. But the main drawback of this work is implementation of communication protocols.

3 Proposed Approach

The primary motivation is to prevent accidents or injuries caused by car doors closing on hands, fingers, or other objects. Such incidents can result in severe injuries or damage. By implementing a reliable detection system, the risk of these accidents can be significantly reduced, ensuring safety. This system also minimizes the potential for injuries and prevents unnecessary harm to passengers. Overall, the motivation behind the proposed car door detection system lies in prioritizing safety, preventing injuries and damages, enhancing user convenience, and utilizing technological advancements to create a reliable and efficient solution for detecting hands or objects while closing car doors.

The development of a car door detection system is driven by an unwavering commitment to safety, with the overarching goal of averting accidents and injuries that result from car doors closing on hands, fingers, or other objects. The potential consequences of such incidents are severe and encompass a spectrum of injuries and damages, making the implementation of a dependable detection system not just a luxury but a necessity. By harnessing the power of cutting-edge technology, this system strives to drastically curtail the hazards associated with inadvertent car door closures (Figs. 1 and 2).

In our proposed work, we are using the ultrasonic sensors on car doors, and with the help of sensor, we can calculate the hand distance. Then, the sensed information is passed onto the Arduino microcontroller. From this, the information is passed onto the car dashboard and alerts the user.

Fig. 1 Hand accidentally placed in the car door

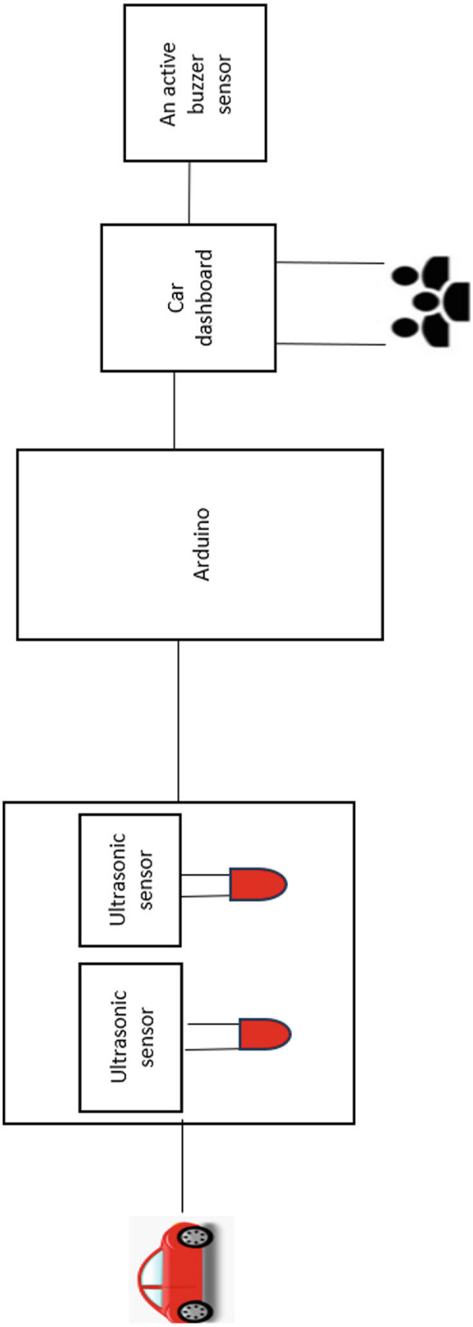


Fig. 2 System architecture

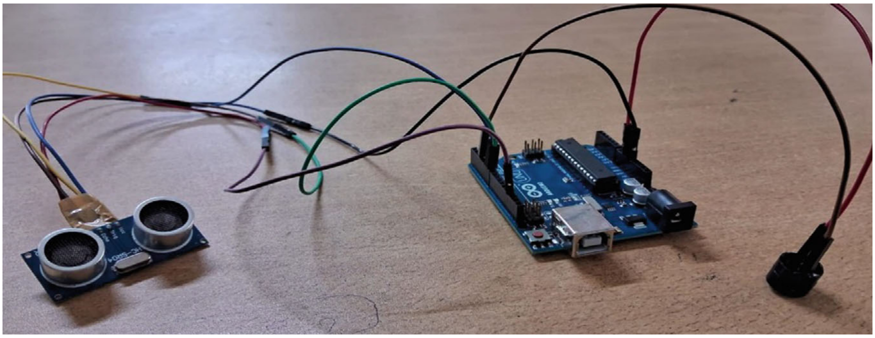


Fig. 3 Arduino connected with an ultrasonic sensor and a buzzer

3.1 Implementation and Results

The hardware connection of our proposed work is represented in Fig. 3 The same prototype is deployed in the car environment it is going to give good results.

In Fig. 4, we can calculate the distance of the object by using the ultrasonic sensor and sends this sensed information to alert the user.

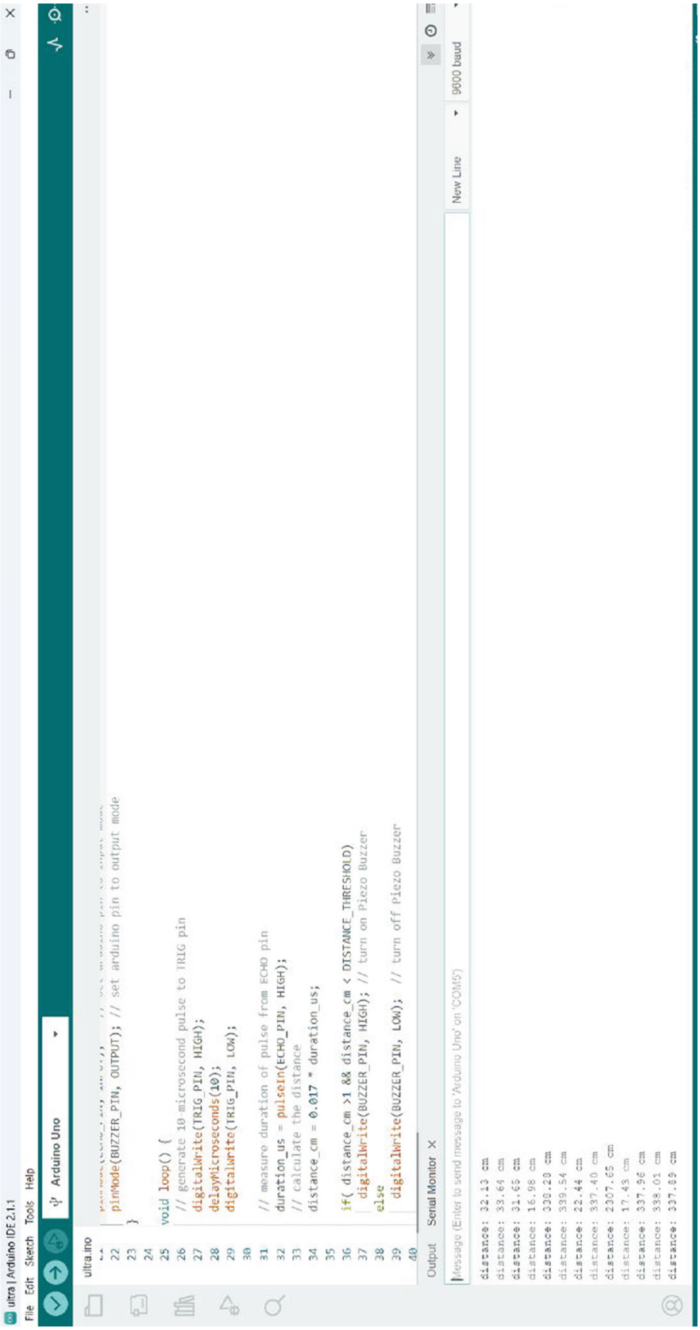


Fig. 4 Calculating the distance of the object

4 Conclusion

In conclusion, the object detection and alerting system designed for the car door closing using an ultrasonic sensor, buzzer, and an Arduino in IoT provides an effective solution to enhance safety and prevent potential accidents. By incorporating the ultrasonic sensor, the system can accurately detect the presence of objects, such as obstacles or hands, near the car door during the closing process. When an object or hand is detected within a specified distance, the Arduino triggers the buzzer, generating an audible alert to the driver or passengers. This warning system serves as a crucial safety measure, as it notifies users about potential hazards and prevents unintended harm to individuals or damage to the vehicle. The IoT aspect of the system further enhances its functionality and connectivity. It allows real-time monitoring and data exchange between the device and other connected components, enabling seamless integration into the broader ecosystem of smart vehicles and smart transportation infrastructure.

References

1. Lalitha VL, Raju SH, Sonti VK, Mohan VM (2021) Customized smart object detection: statistics of detected objects using IoT. In: 2021 International conference on artificial intelligence and smart systems (ICAIS), Coimbatore, India
2. Nyamati V, Chaudhuri T, Jayavel K (2017) Intelligent collision avoidance and safety warning system for car driving. In: 2017 International conference on intelligent computing and control systems (ICICCS), Madurai, India
3. Mukhopadhyay D, Gupta M, Attar T, Chavan P, Patel V (2018) An attempt to develop an iot based vehicle security system. In: 2018 IEEE international symposium on smart electronic systems (iSES) (Formerly iNiS), pp 195-198, Hyderabad, India. https://doi.org/10.1109/iSES.2018.00050
4. Hemaanand M, Rakesh Chowdary P, Darshan S, Jagadeeswaran S, Karthika R, Parameswaran L (2020) Advanced driver assistance system using computer vision and IOT. In: Smys S, Tavares J, Balas V, Iliyasu, A (eds) Computational vision and bio-inspired computing. ICCVBIC 2019. Advances in intelligent systems and computing, vol. 1108. Springer, Cham
5. Al-Youif S, Ali MAM, Mohammed MN (2018) Alcohol detection for car locking system. In: 2018 IEEE symposium on computer applications and industrial electronics (ISCAIE), Penang, Malaysia
6. Zhang Z, Xu H, Chao Z, Li X, Wang C (2015) A novel vehicle reversing speed control based on obstacle detection and sparse representation. IEEE Trans Intell Transp Syst 16(3):1321–1334. https://doi.org/10.1109/TITS.2014.2360337
7. Kulkarni AU, Potdar AM, Hegde S, Baligar VP (2019) RADAR based object detector using ultrasonic sensor. In: 2019 1st International conference on advances in information technology (ICAIT), Chikmagalur, India, pp 204–209. https://doi.org/10.1109/ICAIT47043.2019.8987259
8. Romadhon AS, Widyaningrum VT (2022) Application of sensors in Arduino as a control in smart home. In: 2022 IEEE 8th Information Technology International Seminar (ITIS), Surabaya, Indonesia, pp 130–133. https://doi.org/10.1109/ITIS57155.2022.10010217
9. Liu G-R, Zhou M-Z, Wang L-L, Wang H (2016) A radar-based door open warning technology for vehicle active safety. In: 2016 International conference on information system and artificial intelligence (ISAI), Hong Kong, China, pp 479–484. https://doi.org/10.1109/ISAI.2016.0108

10. Yan X, Gao W (2012) The design of car reversing anti-collision warning system. In: 2012 Fourth international conference on computational intelligence and communication networks, pp 866–869. https://doi.org/10.1109/CICN.2012.197
11. Lai C-H, Lin C-C, Tsai G-L, Hwang Y-S (2018) The car door safety system with wireless alert transmission. In: 2018 7th International symposium on next generation electronics (ISNE), Taipei, Taiwan, 2018
12. Liu Q, Qu S (2020) Design of intelligent car door control system. In: 2020 IEEE 3rd International conference on automation, electronics and electrical engineering (AUTEEE)

Blood Donation Management Application

Sanjivani Adsul, Mrunal Dhulap, Advay Dhule, Adityaraje Dhumal, Ajinkya Dhumal, Gayatri Dhumal, and Girish Dhurve

Abstract Availability of blood is a crucial requirement in many cases such as accidents and surgeries. Medical disorders such as cancer, hemophilia, sickle cell disease, and anemia frequently require blood transfusions for their regular treatments. Thus, many times people need to suffer a lot due to the unavailability of blood during emergence. Also, surveys found that 6 lakh units of blood get wasted every year due to improper storage systems. This paper presents the development of an Android application for blood donation management, named 'Jeevanam' that is markedly designed as a connecting link between donors and people in need of blood. The application will help the user to search for the blood donor of their desired blood group near their location. The application maintains a database of blood donors belonging to different blood groups. The application is built using Java, and the database is managed using Firebase. It has a geolocation feature that enables the user to look for the proximity of the donor from their location. Additionally, the application has a provision to locate the nearest blood banks in the vicinity. The voluntary donors will get notified about the blood donation drives being carried out. Moreover, the application maintains the donation history of the donor and monitors their eligibility for upcoming donations

S. Adsul · M. Dhulap · A. Dhule · A. Dhumal · A. Dhumal · G. Dhumal (✉) · G. Dhurve
Vishwakarma Institute of Technology, Pune, Maharashtra, India
e-mail: gayatri.dhumal221@vit.edu

S. Adsul
e-mail: sanjivani.adsul@vit.edu

M. Dhulap
e-mail: mrunal.dhulap22@vit.edu

A. Dhule
e-mail: advay.dhule221@vit.edu

A. Dhumal
e-mail: adityaraje.dhumal221@vit.edu

A. Dhumal
e-mail: ajinkya.dhumal221@vit.edu

G. Dhurve
e-mail: girish.dhurve22@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_30

by carrying out certain virtual checkups. The checkups include a questionnaire that will examine the health of the donor before accepting the donation requests. Thus, the application helps in minimizing the wastage of blood, enhances the awareness of blood donation, and users can obtain their required blood group; hence, the efficiency of the blood donation process is increased.

Keywords Blood donation · Android application · Firebase · Geolocation · Notifications · Donation history

1 Introduction

To meet a nation's fundamental blood needs, the World Health Organization (WHO) considers that a minimum of 1% of the population should donate blood. Surveys stated that broadly, 14.6 million units of blood are needed every year, but there is a scarcity of about 1.5 million. The exact optimal number of blood banks remains uncertain, but it is evident that the current count falls short. To address this issue and cater to the growing demand, India must enhance its number of blood banks and implement measures to facilitate blood donation. Henceforth, there is a severe need to induce awareness of blood donation among the people. The implementation of blood-based applications would facilitate absolute connections between blood donors and individuals in need, making it easier for both parties to find each other. Traditional blood banks lack observability inventory resulting in shortage and wastage of blood. Blood banks serve as an extensive utility to store and manage information related to volunteer blood donors. Patients in need of blood can search for the blood donors of their desired blood group. The integration of geolocation capabilities within the application will empower users to effortlessly locate the nearest available donors and can send requests through email and messages. The person who is requested will get notifications, and alerts through the application and thus can donate blood. Individuals who receive requests will be promptly notified and alerted via the application, facilitating seamless blood donation. The application is intricately connected with both hospitals and blood banks, thereby enabling these institutions to assist users in dire emergencies, ensuring swift access to donors. The application maintains a comprehensive donor history, guaranteeing adherence to a recommended three-month interval before individuals who have previously donated blood are eligible for their next donation.

2 Literature Review

The paper proposes the development of an Android app for blood donation. The app allows users to locate blood donors, arrange appointments for blood donation, and monitor their blood donation history. The application utilizes geolocation to identify blood donors in the vicinity. It utilizes the Google Maps API to allow users to locate blood donors in their vicinity. Additionally, it uses the Firebase real-time database to store user data and blood donation records [1]. This study outlines a methodological approach to streamline the integration of blood donors with nearby blood banks, emphasizing its importance in the donation process. The registration process involves comprehensive donor information collection, extending to the registration of blood banks and medical institutions. With a focus on managing the blood inventory database, a mobile application, and cloud technology integration establish a centralized repository for data on blood inventories, categorized by blood groups. This system, initially overseen by hospitals or blood banks, ensures consistent tracking of blood unit quantities. Leveraging geolocation capabilities, it identifies nearby donors and aids in the optimal repositioning of blood units to align with storage demands [2].

This paper delineates the application's system architecture, optimizing location accuracy through geofencing and the fused location provider. Secure logins are ensured with Firebase authentication. The donor database relies on Cloud Firestore, offering scalability, performance, and real-time synchronization. The search module aids recipients in identifying recent blood donors meeting the 90-day eligibility criteria for blood donation [3].

The states to design an app that alleviated the public's concerns include educational awareness about donation, blood centre's location, donor eligibility, reminders, scheduling appointments, and being able to ask donors when needed. We have incorporated these features in our application for the convenience of the user. This paper insists on allowing requests for blood via email and SMS using JavaMail API libraries and messaging permissions in the manifest of the application [4]. The mobile application was developed using Android Studio, Firebase, and Jira-like tools to manage user stories and track project progress. This paper focuses on the scrum methodology used to develop the application [5] . The application uses Firebase Cloud Messaging to send timely and efficient notifications about blood donation requests to both blood donors and blood recipients. Blood donors are alerted to blood donation requests as soon as they are created, so they can respond quickly. Blood recipients are also notified when a blood donor has agreed to donate blood [6].

The application's donor history tracking system allows users to keep track of their own blood donation history and the blood donation history of others. It also ensures that donors follow the recommended three-month gap between donations [7]. The paper proposes the machine-learning-based implementation of a health assessment questionnaire to be completed by potential blood donors prior to donation. The questionnaire would be reviewed by medical experts to assess the donor's eligibility. The results of the assessment would be provided to the donor and to the person who requested the donation [8]. A mobile application called B-Door allows NGOs and hospitals to find blood donors in their local area. The application has a database of blood donor information that is verified using the J48 decision tree algorithm and WEKA. Users can register as donors, hospitals, blood banks, or NGOs. Authorized users can create requests for blood donations, and the application will notify potential donors in the nearby area [9].

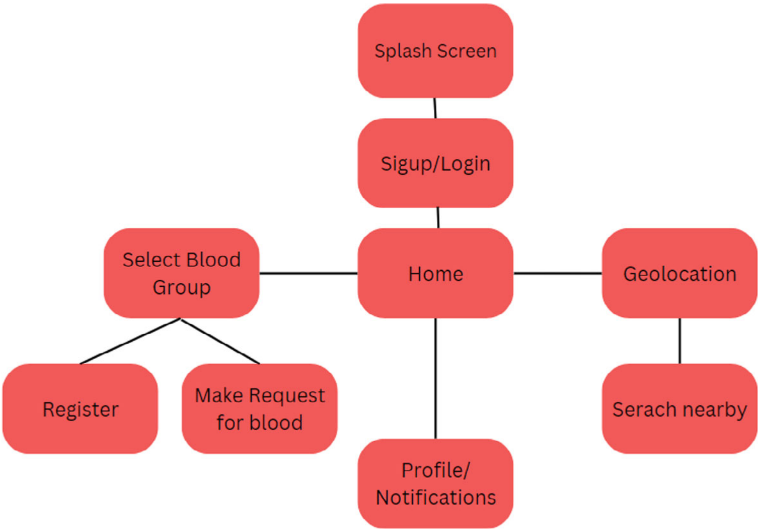


Fig. 1 Flow of the application

A questionnaire was developed to collect information about the donor's health, including whether they had a temperature, any infections or diseases, and their hemoglobin level. This information was used to predict whether the donor was eligible to donate blood. The algorithms used for the prediction were decision trees, logistic regression, and support vector machines [10].

3 Methodology

3.1 Flowchart

Figure 1 maps out the step-by-step workflow of the application, providing a comprehensive overview of how the various components and processes contribute to the overall functionality of the system.

3.2 Tools and Technology

Android Studio Android Studio is a robust IDE tailored for Android app development, offering tools for UI design, coding, testing, and debugging. Its built-in emulators, real device testing, and integration with Android SDK and Google Play services enhance app creation efficiency.

Firebase Google presents Firebase, a potent backend-as-a-service (BaaS) platform, comprising an array of services and tools like Authentication and Cloud Firestore databases. These offerings can be effectively leveraged for the development of applications.

3.3 Implementation

User Authentication This involves designing Signup and Login for registration of new users and password recovery. Firebase authentication enables easy integration of secure user authentication into the applications. Email/password authentication is applied in the application**.**

Registration Module The user has to create a profile and register their blood group. Cloud Firestore serves as an optimal real-time cloud database solution. It effectively structures donor data within collections and documents, facilitating instant updates on donor availability. Utilizing queries, donors can be filtered based on attributes such as blood type and location.

Notification Module The user can find the donor of their desired blood group, nearby blood banks, and hospitals and can send the requests to them. Firebase Cloud Messaging (FCM) is used to streamline communication in a blood donor app. FCM can send notifications to blood donors in real time about urgent blood needs, ensuring that they are aware of the situation and can respond quickly.

Geolocation Google Maps service is integrated into the application using the Google Maps API which enables navigation and efficient donor location. This feature allows users to effectively search for nearby donors and pinpoint the locations of donation centers, enhancing user experience and convenience**.**

Donor History Tracking a donor's historical contributions, the donor history module showcases details such as donation dates and blood types. This valuable data is sourced from the donor database and thoughtfully displayed within the donor's profile.

4 Results and Discussions

The integration of geolocation functionality within the blood donation app has yielded noteworthy outcomes, significantly enhancing the user experience and optimizing the blood donation process. Through real-time updates and notifications, the app can facilitate rapid communication between donors, blood banks, and recipients. This efficient communication led to quicker blood matching and allocation, particularly in emergencies. By harnessing technology to connect donors with specific

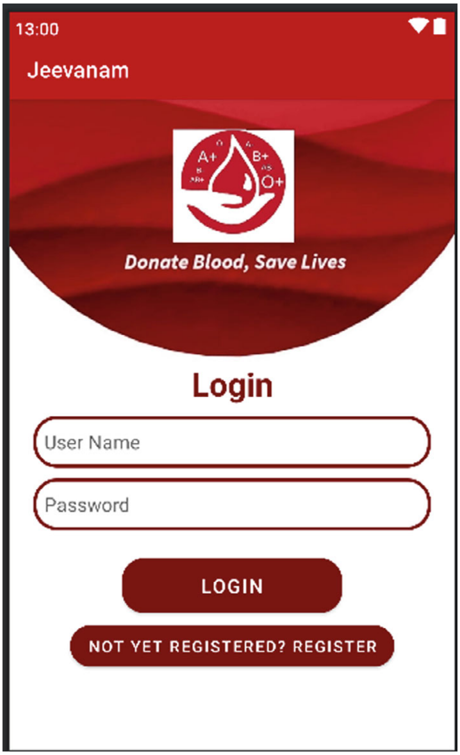


Fig. 2 Login page

blood types to patients requiring the same, the app contributes to more targeted and effective blood supply management. Figures 2, 3, and 4 show the interfaces of the application.

5 Future Scope

The current application primarily functions as a communication conduit connecting people in need of blood. However, ensuring the eligibility of blood donors is of paramount importance. According to WHO guidelines, individuals suffering from contagious illnesses like colds, flu, sore throat, cold sores, or stomach bugs are not suitable blood donors. Guaranteeing the donor's well-being is crucial. To address this, an innovative approach involves integrating a machine-learning-driven questionnaire within the app. This questionnaire will employ user-friendly questions to assess the donor's health status. The data collected will be analyzed using machine learning algorithms, providing doctors with comprehensive insights into the donor's health. This informed feedback will enable medical professionals to make well-informed

Fig. 3 Home page

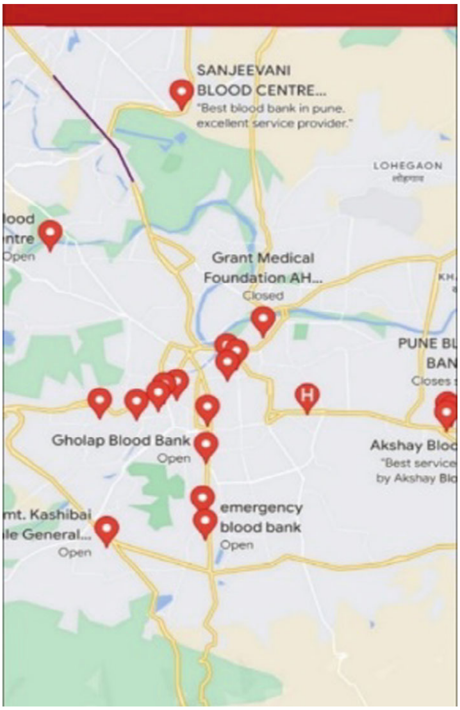


Fig. 4 Searching blood banks

decisions regarding the donor's eligibility. This addition aligns with the app's aim of enhancing donor safety and contributing to a healthier blood donation process.

6 Conclusion

By enhancing accessibility, efficiency, and user convenience, this innovative blood donation application effectively mitigates obstacles that often deter potential donors. This holistic approach broadens the donor base, addressing scarcity. Incorporated features such as appointment scheduling, personalized reminders, and incentives foster a culture of consistent contributions. Moreover, leveraging Firebase Cloud Messaging (FCM) facilitates real-time alerts, engaging donors promptly. Google Maps integration streamlines location-based services, simplifying blood donation center searches. This amalgamation optimizes logistics by unifying communication between blood banks, donors, and recipients, ensuring effective coordination. Consequently, the app not only transforms blood availability dynamics but also significantly elevates healthcare outcomes while alleviating crises. Its cross-platform adaptability ensures ease of blood searches, catering to diverse device preferences.

References

1. Abhijith A, Abinaya V (2023) Android app for blood donation EPRA. Int J Res Dev (IJRD) 8(5):35–39
2. Ismael RD, Hussein HA, Salih MM (2022) Finding the blood donor and inventory monitoring via mobile application. J Algebraic Statistics 13(1):810–816
3. Sangeetha M et al Blood donor searching android application. Int J Health Sci I 13492–13503
4. Law Y, Mon CS, Keoy KH (2022) Mobile blood bank application-blood. In: 2022 1st International conference on computational science and technology (ICCST). IEEE
5. Hummady MM (2022) Design mobile application for blood donation system. In: 2022 International conference on computer and applications (ICCA). IEEE
6. Bandyopadhyay A et al (2021) Blood donation management system using android application. In: Proceedings of international conference on communication, circuits, and systems: IC3S 2020. Springer, Singapore
7. Dombe A, Bhutkar G, Dongre A (2021) Prototype design of android app for blood donation. Digital literacy and socio-cultural acceptance of ICT in developing countries, pp 261–272
8. Piraviendran R, Othman M, Mohamed SA (2021) Blood donation management application on android platform. Appl Inf Technol Comput Sci 2(2):1827–1845
9. Periyanayagi S et al (2021) BDoor app-blood donation application using android studio. J Phys Conf Ser 1917(1)
10. Joshi AJ, Siddhesh SA Blood bank app management using android application

Navigating Future Coronary Heart Disease Scenarios with Ensemble Learning

Navya Jettoji, L. Arokia Jesu Prabhu, and Vijender Solanki

Abstract The primary driver of mortality in both industrialized and non-industrial countries is coronary illness, or CHD. It is very unsafe for your wellbeing. The objective of the review is to coordinate patient consideration and focus on early location. To increment execution, an ideal democratic classifier expectation model is utilized, along with a hyperparameter. The AdaBoost, Decision Tree, Gradient Boosting, XG Boost, CatBoost, and Light GBM with Focal Loss utilizing complex hyperparameter improvement utilizing OPTUNA. The democratic classifier using random forest and AdaBoost got an accuracy of 99.3%, precision of 98.2%, recall of 98.1%, and F_1-score of 98.2% in our trial research. As an outcome, when estimated against different classifiers, the Democratic Classifier accomplishes the best score.

Keywords Coronary heart disease · Machine learning · Classification · Voting classifier · Hyperparameter

1 Introduction

As one of the primary drivers of death universally and a serious general medical condition, coronary illness presents a considerable obstacle [1–3]. This disorder is joined by various side effects, including angina or chest torment, dyspnea, and cardiovascular breakdown. Extreme coronary illness can bring about a cardiovascular

N. Jettoji · L. Arokia Jesu Prabhu (✉)
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, Telangana, India
e-mail: arokiajeruprabhu@gmail.com

N. Jettoji
e-mail: navyajettoji@gmail.com

V. Solanki
Computer Science and Engineering Department, Stanley College of Engineering and Technology for Women (Autonomous), Hyderabad, India
e-mail: spesinfo@yahoo.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_31

failure that definitely diminishes an individual's personal satisfaction and may for all time hurt the heart muscle. Along these lines, it's critical to recognize and treat coronary illness with the fitting clinical consideration and way of life alterations. [4, 5] The chance of a fix and treatment costs can be diminished by early recognition of coronary illness. In the clinical region, various information mining methods and ML calculations have been broadly involved as of late because of progressions in ML strategies and a huge lessening in the expense of information stockpiling. Information mining innovation has been fundamental in the field of medical services information mining, which incorporates biomedicine, drug mining, assistant diagnostics, and ailment analysis. Information mining apparatuses empower the production of ailment expectation models, investigation of results, and revelation of concealed illness data from a lot of unstructured clinical information [6–8].

Coronary illness is as yet a significant overall medical condition, in this way precise forecast models are fundamental. Nonetheless, existing models frequently neglect to accomplish ideal execution with regards to exactness and effectiveness. This work proposes a democratic classifier for coronary illness expectation to close this hole. The objective of the examination is to fortify the ongoing models' shortcomings and increment the precision of judgments. Through a total execution study and examination, the exploration looks to give a valuable methodology to advance the early conclusion of coronary illness, working on quiet results simultaneously.

The model's exactness is additionally affected by the misfortune capability [9, 10]. This work proposed an engaged misfortune capability that incorporated the class weight α and the example trouble weight changing variable γ to the cross-entropy misfortune. The [11] reason for this study was to resolve the issue of lopsided conveyances of positive and negative examples. Besides, the center misfortune capability can upgrade the model's general usefulness. This study utilized the engaged misfortune capability to refresh the LightGBM [12] model's default misfortune capability, which was then used to gauge CHD.

2 Literature Review

S. Nematzadeh, F. Kiani, M. Torkamanian-Afshar, and N. Aydin et al. [13] **Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: A bioinformatics study on biomedical and biological cases:**

The dataset and preparing methods impact a model's presentation in ML challenges. The narrative of a model might be modified by choosing the fitting preparation system. While specific calculations in all actuality do outstandingly well in certain datasets, they could battle in other datasets. Moreover, execution might be upgraded by changing a calculation's hyperparameters, which control the preparation strategies. The proposed approach just takes a dataset as info and returns the most completely researched calculation along with significant contentions.

A M. Liang, B. An, K. Li, L. Du, T. Deng, S. Cao, Y. Du, L. Xu, X. Gao, L. Zhang, J. Li, and H. Gao, et al. [14] Improving Genomic Prediction with Machine Learning Incorporating TPE for Hyperparameters Optimization:

ML has been viewed as the most intense apparatus for examining genomic information from high-throughput sequencing as a result of its prevalent expectation capacity. The discoveries showed that KRR-TPE was the most down to earth and had the most grounded expectation capacity across all gatherings. In correlation with GBLUP, the expectation exactness of KRR-TPE shown a normal increment of 8.73 and 6.08% for the Chinese Simmental hamburger dairy cattle and Loblolly pine populaces, separately.

M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, et al. [15] Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation:

A critical advancement in clinical picture examination is the utilization of programmed division procedures. We show that our recommended misfortune capability is strong to class awkwardness and reliably beats the other misfortune capabilities by looking at its exhibition against six Dice or cross entropy-based misfortune capabilities, across 2D parallel, 3D double, and 3D multi-class division assignments.

C. A. U. Hassan, J. Iqbal, R. Irfan, S. Hussain, A. D. Algarni, S. S. H. Bukhari, N. Alturki, and S. S. Ullah, et al. Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifiers:

One of the main sources of mortality overall is coronary illness. Perhaps of the hardest thing to do in clinical information examination is to analyze a heart condition. With regards to deciding and making expectations in light of information created by the worldwide medical care industry, machine learning (ML) is useful for symptomatic guide. A few component mixes and notable grouping techniques were utilized to present the forecast model. With angle supported trees and multi-facet perceptrons, we had the option to achieve 95% exactness in the expectation of coronary illness. With a precision pace of 96%, the Random Forest model performs better in the prediction of heart disease.

L. Lin, J. Zhang, N. Zhang, J. Shi, and C. Chen, et al. Optimized LightGBM Power Fingerprint Identification Based on Entropy Features:

The huge volume of force finger impression information habitually experiences imbalanced classes and is trying to transfer because of the obliged information move rate for Internet of Things interchanges. In conclusion, the LightGBM classifier's hyperparameters were improved utilizing the Optuna strategy. By advancing the LightGBM model's misfortune capability, the influence finger impression distinguishing proof model's arrangement execution on uneven datasets was essentially improved. The results of the examinations exhibit that the strategy may effectively bring down the computational intricacy of component extraction and lower the transmission force of finger impression information.

3 Methodology

The methodology predicts coronary illness utilizing ML procedures and a recommended casting a ballot classifier. There are preparing and test sets inside the coronary illness dataset. Various AI techniques, including irregular woodland, strategic relapse, *k*-closest neighbor, and XGBoost, were utilized and stood out from the Framingham Risk Score (FRS) (Fig. 1).

Data exploration: We will stack information into the framework with the assistance of this module.

Processing: We will peruse information for handling by utilizing the module.

Utilizing this module, information will be parted into two classes: train and test.

Model generation: Construct a few ML classifiers by using the qualities that were removed from the dataset. This involves:

AdaBoost: Versatile Helping, or AdaBoost, is a popular gathering learning procedure for ML errands including relapse and characterization.

The figure of an AdaBoost for order might be communicated numerically as:

$$\text{AdaBoost}(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t X h_t(x)\right) \tag{1}$$

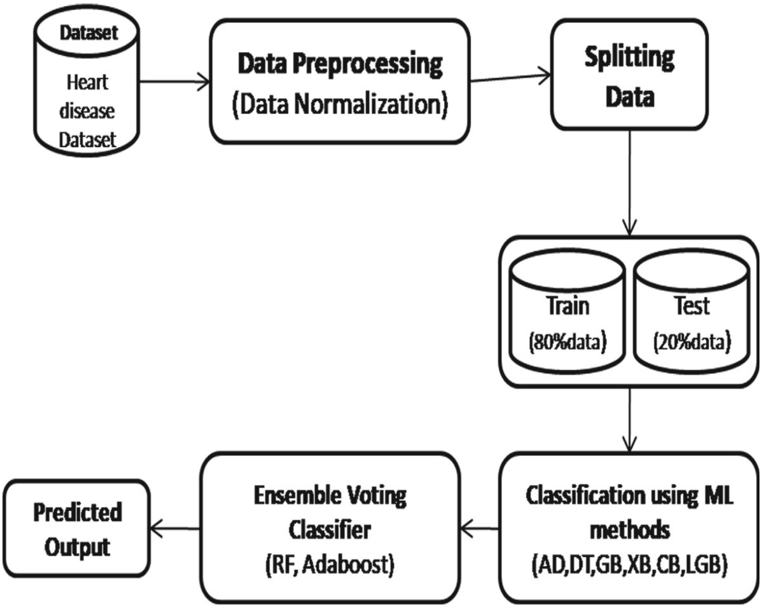


Fig. 1 System architecture

For instance, $h\_(t)$ (x) is the expectation of feeble student $h\_(t)$, and T is the all out number of frail students. α_t is the weight designated to powerless student $h\_(t)$ x.

Decision Tree: A popular directed ML strategy for grouping and relapse applications is the choice tree. It partitions the information space recursively into more modest areas and makes decisions as indicated by the attributes of the information.

With the decision tree, we might utilize a case x to expect its result y in the accompanying manner:

$$Y = \sum_{m=1}^{M} (I(x \text{ follows path to leaf } L_m) \text{value of leaf } L_m) \tag{2}$$

A marker capability called where $I(x$ follows the way to leaf $L\_m)$ returns 1 if the occasion x follows the way to leaf L_m and 0 without a trace of such a case.

Gradient Boosting: An intense ML strategy for both relapse and grouping applications is inclination helping. It develops the possibility of gathering realizing, which joins many models to make a more intense expectation model. In particular, slope helping centers around fostering a progression of feeble students (normally choice trees) consistently in a bit by bit style, with each new model fixing the errors of the earlier ones.

As far as math, it is communicated as

$$F_m(x) = F_{m-1}(x) + \lambda.h_m(x) \tag{3}$$

where the ensemble model's forecast following the n-th iteration is denoted by $F_m(x)$. Following the $(m - 1)$-th iteration, the ensemble model's forecast is expressed as F_{m-1} (x).The learning rate, denoted as λ and sometimes referred to as shrinkage or step size, regulates how much each weak learner contributes to the ensemble.

XG Boost: A proficient and very much upgraded form of the Inclination Helping method, XG Boost is otherwise called Outrageous Slope Supporting. It was made by Tianqi Chen and is notable for its speed and viability in both useful applications and ML challenges.

Another superior exhibition inclination helping library made by Yandex is called CatBoost. "Categorical Boosting" is a big motivator for it, and it is made particularly to manage all out qualities well.

For paired and multi-class characterization issues, LightGBM with Central Misfortune is an execution of the central misfortune. The connected Medium article is accessible here. Simpler to code: deffocal_loss_lgb(y_pred, dtrain, alpha, gamma): a, g = alpha, gamma y_true = dtrain is the Central Misfortune for LightGBM.

This is the meaning of the center misfortune capability:

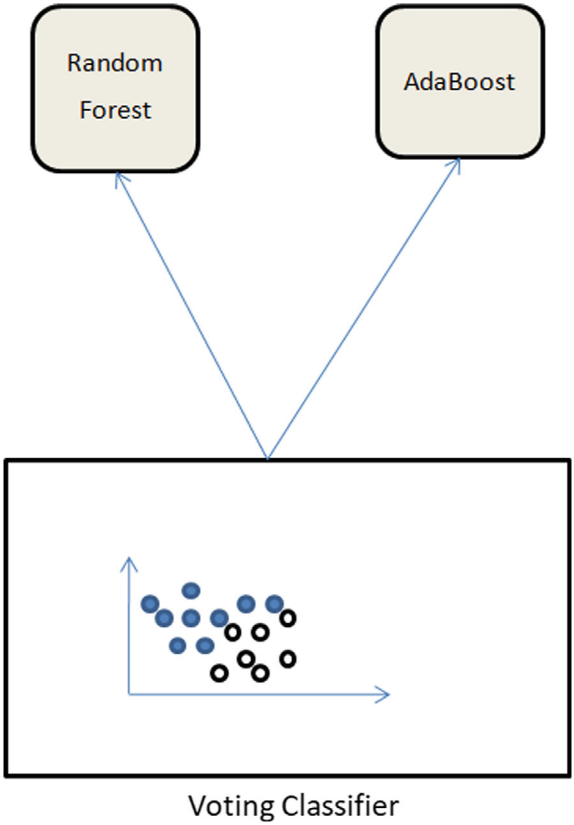
$$\mathrm{FL}(p_t) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t) \tag{4}$$

where p_t the normal likelihood of the genuine class is p_tis. The class adjusting weight, signified as α_tis, is habitually utilized to level the commitments of each and every class.

Voting Classifier (RF + AdaBoost): An ML assessor known as a democratic classifier prepares different base models or assessors and makes expectations by adding the consequences of each base assessor. Voting classifier choices for every assessor result can be converged to frame the collecting rules (Fig. 2).

User signup and login: Using this module will result in login and enlistment.
User input: This module's utilization will give contribution to estimating.
Prediction: The last conjecture is shown.

Fig. 2 Voting classifier [14]



4 Implementation

It is recommended to utilize Random Forest and Adaboost to make a democratic classifier anticipating model for coronary illness. The model's exhibition is surveyed utilizing different rules.

Data Preparation: Burden the dataset with information on the different coronary illness risk factors. The "training" section ought to be eliminated as it is irrelevant to the analysis. Look for any missing qualities, then erase any lines that incorporate them. This utilized the expected libraries, which included sci-pack learn, Pandas, NumPy, Statsmodels, Scipy, and Matplotlib. For exercises including information handling, factual investigation, representation, and ML, these libraries are much of the time used. Loads the "framingham.csv" dataset into the heart_df Pandas DataFrame. The "training" section is eliminated from the DataFrame as it is considered unnecessary changes the name of the "male" segment to "Sex_male" to make the dataset's missing qualities more justifiable and to eliminate any lines that have them.

Exploratory Data Analysis (EDA): Use count plots and histograms to perceive how different qualities are conveyed visually. Using a couple plot, delineate the pairwise joins between different characteristics. For the dataset, figure elucidating measurements.

Model Evaluation: Utilizing the train_test_split() strategy, partition the dataset into preparing and testing sets. Using the preparation set of information, make a calculated relapse model. Forecast the consequences of coronary illness in light of the test data. Accuracy, disarray lattice, responsiveness, particularity, positive prescient worth, negative prescient worth, positive probability proportion, and negative probability proportion ought to be in every way used to survey the model's performance.

Threshold Adjustment: Alter the likelihood forecast characterization limit and evaluate the model's presentation at different levels. For every edge, show the awareness, explicitness, and disarray grid.

Visualization of Model Accuracy: In conclusion, a plot is made to show how exact different models are.

5 Experimental Results

Accuracy: A test's not set in stone by how well it can recognize debilitated and sound examples. We ought to register the level of genuine up-sides and genuine negatives in each examined occurrence to survey the Accuracy of a test. As far as math, this is communicated as:

$$\text{Accuracy} = \text{True Positive} + \text{True Negative}/\text{True Positive} \\ + \text{True Negative} + \text{False Positive} + \text{False Negative} \quad (5)$$

Precision: Precision estimates the level of accurately classified examples or cases among the positive examples. Thus, coming up next is the equation to decide the Precision:

$$\text{Precision} = \text{True Positives}/(\text{True Positives} + \text{False Positives}) \quad (6)$$

Recall: In ML, review is a measurement that evaluates a model's ability to find all relevant occurrences of a given class. It gives data about how well a model catches instances of a specific class. It is determined as the proportion of appropriately anticipated positive perceptions to the all out number of genuine up-sides.

$$\text{Recall} = \text{True Positive}/\text{True Positive} + \text{False Negative} \quad (7)$$

***F_1*-Score**: An evaluation metric for ML that measures a model's accuracy is known as the F_1-score. It coordinates a model's accuracy and review evaluations. The exactness measure works out the times a model accurately anticipated the entire dataset.

$$F_1\text{-Score} = 2/(1/\text{Precision} + 1/\text{Recall}) \quad (8)$$

Table 1 records the exactness execution of ML algorithms. Through relationship investigation, we discovered that the democratic classifier outflanked other ML classifiers, accomplishing accuracy of 99.3%, accuracy of 98.2%, review of 98.1%, and F_1-score of 98.2% (Fig. 3).

The voting classifier with the best exactness, precision, review, and F_1-score is shown in the examination chart for the result table, which is seen previously.

Table 1 Comparison table

ML mode	Accuracy	Precision	Recall	F_1-score
AdaBoost	75.6	76.3	75.2	76.6
Decision tree	81.4	84.6	79.6	82.6
Gradient boosting	76.5	76.5	77.9	77.6
XGBoost	90.5	87.9	91.6	90.3
CatBoost	67.3	72.6	60.5	68.6
LightGBM	89.5	86.6	92.6	88.6
Voting classifier	99.3	98.2	98.1	98.2

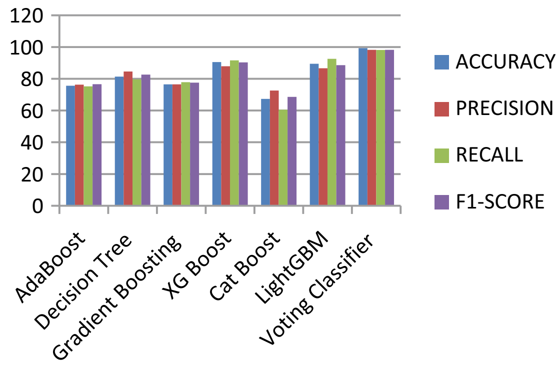


Fig. 3 Performance graph

6 Conclusion

The paper proposed an expectation model for early recognizable proof of coronary illness utilizing the ideal democratic classifier. Data on coronary illness from the Framingham Heart Organization were used to survey the viability of the recommended model utilizing different measures. The proposed model beat other looking at models as far as AUC esteem. The discoveries demonstrate the way that the proposed methodology can lessen the costs connected with treating patients with coronary illness and increment the pace of early recognition of the condition in everybody.

References

1. Katta N, Loethen T, Lavie CJ, Alpert MA (2021) Obesity and coronary heart disease: epidemiology, pathology, and coronary artery imaging. Curr Probl Cardiol 46(3), Article no. 100655. https://doi.org/10.1016/j.cpcardiol.2020.100655
2. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H (2018) EDoctor: machine learning and the future of medicine. J Intern Med 284(6):603–619. https://doi.org/10.1111/joim.12822
3. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T (2017) Artificial intelligence in precision cardiovascular medicine. J Amer Coll Cardiol 69(21):2657–2664. https://doi.org/10.1016/j.jacc.2017.03.571
4. Rauschert S, Raubenheimer K, Melton PE, Huang RC (2020) Machine learning and clinical epigenetics: a review of challenges for diagnosis and classification. Clin Epigenetics 12(1):51. https://doi.org/10.1186/s13148-020-00842-4
5. Kim JK, Kang S (2017) Neural network-based coronary heart disease risk prediction using feature correlation analysis. J Healthc Eng 2017, Article no. 2780501. https://doi.org/10.1155/2017/2780501
6. Romm EL, Tsigelny IF (2020) Artificial intelligence in drug treatment. Annu Rev Pharmacol Toxicol 60(1):353–369. https://doi.org/10.1146/annurev-pharmtox-010919-023746
7. Vercio LL, Amador K, Bannister JJ, Crites S, Gutierrez A, MacDonald ME, Moore J, Mouches P, Rajashekar D, Schimert S, Subbanna N, Tuladhar A, Wang N, Wilms M, Winder A, Forkert

ND (2020) Supervised machine learning tools: a tutorial for clinicians. J Neural Eng 17(6), Article no. 062001. https://doi.org/10.1088/1741-2552/abbff2
8. Arfat Y, Mittone G, Esposito R, Cantalupo B, De Ferrari GM, Aldinucci M (2022) Machine learning for cardiology. Minerva Cardiol Angiol 70(1):75–91. https://doi.org/10.23736/s2724-5683.21.05709-4
9. Goldman O, Raphaeli O, Goldman E, Leshno M (2021) Improvement in the prediction of coronary heart disease risk by using artificial neural networks. Qual Manage Health Care 30(4):244–250. https://doi.org/10.1097/qmh.0000000000000309.
10. Du Z, Yang Y, Zheng J, Li Q, Lin D, Li Y, Fan J, Cheng W, Chen X-H, Cai Y (2020) Accurate prediction of coronary heart disease for patients with hypertension from electronic health records with big data and machine-learning methods: model development and performance evaluation. JMIR Med Inf 8(7), Article no. e17257. https://doi.org/10.2196/17257
11. Ke G et al (2017) LightGBM: a highly efficient gradient boosting decision tree. In: Proceedings on 31st international conference neural information processing system, Long Beach, CA, USA, pp 3149–3157
12. Akiba T, Sano S, Yanase T, Ohta T, Koyama M (2019) OPTUNA: a next-generation hyper-parameter optimization framework. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, Anchorage, AK, USA, pp 2623–2631
13. Nematzadeh S, Kiani F, Torkamanian-Afshar M, Aydin N (2022) Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: a bioinformatics study on biomedical and biological cases. Comput Biol Chem 97, Article no. 107619. https://doi.org/10.1016/j.compbiolchem.2021.107619.
14. Liang M, An B, Li K, Du L, Deng T, Cao S, Du Y, Xu L, Gao X, Zhang L, Li J, Gao H (2022) Improving genomic prediction with machine learning incorporating TPE for hyperparameters optimization. Biology 11(11):1647. https://doi.org/10.3390/biology11111647
15. Yeung M, Sala E, Schönlieb C-B, Rundo L (2022) Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. Comput Med Imag Graph 95, Article no. 102026. https://doi.org/10.1016/j.compmedimag.2021.102026

Exploring the Impact of Artificial Intelligence on Business Value in the Era of Digital Transmission

Y. V. Naga Kumari, B. Ramesh Babu, G. Kamal, T. Hemanth Kumar, M. Vinod, and K. S. V. Akhil

Abstract Artificial intelligence, also referred to as machine learning, is a rapidly expanding commercial technology. Today's businesses and daily lives use artificial intelligence. In business, artificial intelligence can enhance precision, speed, and cost. Modern times are the most enthralling in human history due to the rapid development of new discoveries and technologies. Artificial intelligence has enabled industrial automation, self-driving vehicles, fitness activity monitors, and online education. There are implications for society, the economy, and individuals resulting from the pervasive adoption of artificial intelligence. In order to integrate artificial intelligence into their operations, the majority of businesses are revising their strategies and business models. Numerous organizations are oblivious of the consequences of utilizing AI; therefore, its effects must be considered. It's tough. Researchers in Andhra Pradesh, India, will conduct interviews with company decision-makers and regular employees to better comprehend how artificial intelligence (AI) impacts businesses. An online questionnaire captures sample data. Using frequency tables, graphs, and one-way analysis of variance, we analyzed the data. Analyzing the impact of AI on organizations will be aided by a key business factor. This study demonstrates that artificial intelligence (AI) offers numerous opportunities and has the potential to transform employment in a variety of contexts. Now everyone is aware of this. As technology advances swiftly in all areas of life and business, artificial intelligence will help corporations and other organizations prepare for new challenges. All aspects of an organization benefit from artificial intelligence, particularly market leadership and sustainability. This is one of the strongest AI arguments.

Keywords Artificial intelligence · Business · Automation · ANOVA

Y. V. Naga Kumari (✉) · G. Kamal · T. Hemanth Kumar · M. Vinod · K. S. V. Akhil
Department of Business and Management Studies, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, India
e-mail: nagakumari73@gmail.com

B. Ramesh Babu
Department of Information Technology, Sir C R Reddy College of Engineering, Eluru, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_32

1 Introduction

Artificial intelligence (AI) lacks a clear, unambiguous definition despite its expanding popularity. Intelligence is the capacity to function appropriately and anticipate its environment. Artificial intelligence is the science of creating intelligent automata. The process of integrating cloud computing, network devices, robotics, computers, digital content production, and various corporate processes, systems, and daily activities constitutes artificial intelligence. AI-based computation has occurred in the past, continues to occur currently, and will continue to occur in the future. According to John McCarthy, "Artificial Intelligence" refers to the science and engineering underlying the production of intelligent devices, specifically intelligent computer programs. McCarthy is credited with developing artificial intelligence (AI), which he defines as "the science and engineering of creating intelligent devices and extremely intelligent computer programs."

AI refers to the imitation of human intelligence by machines. AI is the study of "intelligent agents," which is a subset of software engineering. When most people refer to "artificial intelligence," they are referring to a machine's capacity to learn and make decisions in a manner similar to that of humans.

In the recent years, a great deal of artificial intelligence-based software has been released. In the recent years, tech giants have prioritized subfields of artificial intelligence such as machine learning, natural language processing, image processing, and data mining. Google's predictive search bar, Gmail's spam filter, and Netflix's program recommendations all use machine learning. Google Voice and Siri both utilize Natural Language Processing. Facebook's software for labeling photographs with people's identities and Google's self-driving cars require image processing. Information mining is a colloquial term for the field of programming. This is owing to the daily generation of enormous amounts of data. Businesses such as Facebook and Google receive a large quantity of data from their consumers on a regular basis, and they need a way to interpret this data. In today's tech-driven world, artificial intelligence has proven to be a valuable new tool [1].

2 History of Artificial Intelligence

Here's an overview of how artificial intelligence has evolved over the previous six decades.

- In 1956, John McCarthy came up with the phrase "artificial intelligence" and organized the first conference on the subject of artificial intelligence.
- In 1969, Shakey, the world's first mobile robot designed for a variety of tasks, was constructed. It is no longer limited to merely carrying out a set of instructions but may now perform actions with a specific goal in mind.

- In a match, the 1997 supercomputer known as Deep Blue was victorious over the reigning world chess champion. The creation of this enormous computer by IBM was a significant step in the right direction.
- It wasn't until 2002 that the world saw its first robotic vacuum cleaner that was a commercial success.
- 2005 marked the beginning of the era in which voice recognition, robotic process automation (RPA), a dancing robot, and smart housing were first introduced.
- In 2020, Baidu provides the Linear Fold AI algorithm to medical and scientific teams during the early stages of the SARS-CoV-2 (Covid-19) epidemic. These organizations are developing a vaccine to combat the Covid-19 epidemic. The program can determine the RNA sequence of the virus in less than 27 s, which is 120 times faster than the methods previously used.

3 Importance of AI to the Economic World

According to the PwC Global Artificial Intelligence Report, AI could add $15.8 trillion to the global economy by 2030. By 2030, 45% of all economic benefits will be the result of product modifications [2]. All of these alterations will be prompted by an increase in customer demand. This is due to the fact that AI will result in a greater variety of products that can be tailored to each individual's preferences and that are more attractive and less expensive than other products. True, this contribution will benefit everyone, but more than 70% of AI's global economic impact will be felt in China (26% GDP growth by 2030) and North America (14.5%).

4 What Contributions Does AI Make to Different Industries?

AI is a cutting-edge technology with numerous applications in a variety of fields. All organizations are becoming more aware of the advantages of utilizing technology. AI applications are already widely used in a variety of industries. Technology is changing the world in a digital way. AI will enable many companies and sectors to streamline and optimize their processes. By incorporating this technology into your organization, you will be able to take benefit of artificial intelligence. It is currently used by a lot of enterprises and corporations. Discover how artificial intelligence can impact the following businesses and industries.

4.1 Education

Artificial intelligence can help the education industry in a lot of ways. Technology makes it easier to make and use a wide range of educational programs [3]. It can also be used for making games and software. Using artificial intelligence, the whole school system and ways of teaching can be changed and improved. It starts with schools and colleges that offer degrees and certificates.

Apps that use AI can help both institutions and students. Putting them to use in education has the potential to change the way people teach and learn. This helps to make the whole process better. It changes and improves learning activities so that all students can improve.

4.2 Automotive

You may already know about cars that can drive themselves. It shows how things that were once unthinkable have become real. Today, nothing is just science fiction. A new study says that there will be about 33 million self-driving cars on the road by 2040.

This improvement is due to artificial intelligence. There are many companies around the world that make self-driving cars. These self-driving cars use data science and artificial intelligence to figure out where to go. These companies' vehicles collect more than a petabyte of data every day to figure out the best ways to drive, routes, and safety precautions [4].

4.3 Navigation and Transportation

AI has become a huge trend in the travel and transportation industry as a whole. It not only tells the drivers which company has the shortest route, but it also makes it easier to plan trips to remote places. Businesses utilize AI to navigate. Some travel companies have also added AI to their systems to make money off of the use of smartphones [5]. 82% of people, according to research, use their smartphones to find out about local sights and restaurants.

You have probably heard of Google Maps, but you may not have heard of AI in many fields. It can use a sophisticated algorithm to scan roads and find better routes. It can also tell you in real time what the traffic is like on the way, whether you are taking a bus, train, or walking.

4.4 Manufacturing

Artificial intelligence helps manufacturers solve internal problems and improve their work. With the help of AI, manufacturing facilities can easily automate everything, make decisions, integrate, and channel [6]. Technology can have a big effect on how things are made and help them get better. Experts think that AI will help boost production by 40% by the end of the year 2035. AI's benefits include the ability to make decisions based on data, which speeds up the process, increases productivity, and lowers operational costs. It also helps with product development and scaling up.

4.5 Agriculture

Agriculture is one of the hardest places to do things right. AI makes it easier to predict regional customer demand and the behavior of the supply chain, as well as to choose the right raw materials for a given yield. With this method, you can figure out the large-scale weather patterns that can affect the harvest. For example, too much rain or heat can hurt plants.

4.6 Health Care

CB Insights says that in the near future, 86% of health businesses will use technologies that use artificial intelligence. Using technology in healthcare will help in a lot of different ways, that's for sure [7]. It can, for example, be used to store medical records, evaluate tests and X-rays, create virtual nurses, and enter data, among many other things.

Artificial intelligence is a great technology when used in hospitals and clinics. Experts say that putting artificial intelligence and the Internet of Medical Things together holds a lot of promise. The technology will have a big effect on both diagnosing diseases and finding out more about them. Applications driven by AO are assisting individuals in living healthier lifestyles.

4.7 Consumer Support

One of the best things about AI is how it helps the customer service industry. AI has been used a lot in many different fields, such as retail, finance, and insurance. The AI-powered technology solution will let companies help customers in real time and in a way that fits their needs. The technology can also help staff be more productive and happy [8].

4.8 Cyber Security

When artificial intelligence is used in the field of cybersecurity, big changes can happen. This will make it possible to figure out who the hackers are before they attack. AI can learn and adapt to the dangers of the current business environment. So, better security is the only way to make sure that AI technology can be used. Cyber security experts are using the benefits of Artificial Intelligence to reduce the number of frauds and attacks.

4.9 Software Engineering

With new tools and libraries that make it possible to write code that uses natural language, AI is sure to shake up the software development industry in a big way. Also, technology makes it possible to automate anything. Using insights from AI-based technologies can open up new ways to think about how software is made and how it can help an organization grow.

4.10 Finance

Financial institutions employ artificial intelligence to improve customer ideas and make better decisions. Researchers think that AI and the business of money are a great match. In the financial sector, the most important factors are accuracy, reporting data in real time, and processing large amounts of data [9]. The technology is perfect for these tasks, which is why the financial industry is recognizing its accuracy and using machine learning, adaptive intelligence, algorithm trading, automation, and chatbots in their work.

Robo advisors are now available from several online financial platforms. These advisors use AI to create a unique profile for each investor based on what their financial goals are [10].

4.11 E-Commerce

Have you ever looked for the same product on two different e-commerce sites and found the same image? If you answered yes, you have AI to thank.

Companies use algorithms for machine learning to improve how they interact with their customers. The algorithms can help tailor the experience for each client, and the technology can also help businesses make more sales.

Amazon is one of the most powerful e-commerce platforms and is known for using AI in every part of the customer life cycle. In fact, the company has built its whole business on AI, and many of its projects are related to AI.

5 Common Real-World Applications of AI

The following are some popular uses of AI that you may encounter.

5.1 Navigation and Maps

The application of artificial intelligence has tremendously facilitated travel. You may now enter your position directly into the Google or Apple Maps app on your phone, removing the need for paper maps or directions.

To what end, then, does the program seek its objectives? What about the best route, barricades, and traffic jams? Once upon a time, satellite-based GPS was the only choice, but AI is now being used to provide consumers with a far improved experience.

Machine learning techniques that recall the edges of the structures it has learned provide better map visualizations, as well as detection and understanding of home and block numbers. The software has also been trained to detect and account for changes in traffic patterns, allowing it to recommend other routes that avoid bottlenecks.

5.2 Recognizing Faces

AI is now embedded in many aspects of our life, such as the use of virtual filters on our faces while shooting images and the use of facial ID to unlock our phones. The former has a feature called "facial detection," which allows it to recognize any human face. The latter makes use of technologies that enable the identification of individual human faces. Government buildings and airports also utilize facial recognition for security and surveillance purposes.

5.3 Autocorrect and Other Text Editors

Maybe you used Grammarly to double check your term paper before turning it in, or maybe you use it today to make sure your email to the boss is error-free. There is artificial intelligence at work here as well. In word processors, texting applications, and wherever else you could type something, AI algorithms are using machine learning, deep learning, and natural language processing to spot errors in your writing and provide suggestions for improvement. In order to teach computers grammar, linguists and computer scientists collaborate. Incorrect usage of commas will be flagged by the editor thanks to the training algorithms received from high-quality linguistic data.

5.4 Algorithms for Searching and Suggesting

Have you noticed that the recommendations made to you when you wish to watch a movie or purchase online are often related to your hobbies or recent searches? These sophisticated recommendation algorithms have observed your online behavior and habits in order to understand more about you and what you're interested in. Machine learning and deep learning are used to store and evaluate the data that is gathered at the front end (from the user). The system can then make educated guesses about your tastes and suggest products or music you would enjoy.

5.5 Chatbots

Dealing with customer service may be a hassle and a waste of time. It's a pricey and difficult-to-manage division for businesses. The employment of AI chatbots is one prominent AI-based answer to this problem. The methods used to construct these functions allow for automated responses to commonly asked queries, order taking and tracking, and call routing.

Using NLP, chatbots are trained to mimic the speech patterns of human customer service agents (NLP). Modern chatbots can process a wide variety of inputs, not only yes/no questions. They are able to provide in-depth replies to difficult queries. In fact, the bot learns from its mistakes and improves with each review it receives, so you can be assured that you'll always get the best possible service.

5.6 Electronic Helpers

We frequently ask our digital helpers to do things for us when we're too busy to do them ourselves. Don't text and drive, kids; instead, have the assistant give your mother a call while you're behind the wheel. AIs can do things like search your phone's contacts, look for the term "Mom," and then dial your mother's number for you, as demonstrated by virtual assistants like Siri. These helpers make decisions about your requests using natural language processing (NLP), machine learning (ML), statistical analysis (SAS), and algorithmic execution (AlgoExec). The processes behind voice search and picture search are quite similar.

5.7 Online Networks

To keep you interested and "plugged in," social media apps are increasingly relying on AI to do everything from content monitoring and connection suggestions to ad targeting.

Using the use of keyword identification and visual picture recognition, AI algorithms can quickly identify problematic posts that violate terms of service and remove them. While deep learning's neural network design is a key part of the process, it is not the only factor [11].

As social media platforms view their users as the commodity they sell, they employ AI to link those users with the advertisers and marketers who have found their profiles to be particularly appealing. Artificial intelligence for social media may also learn what a person like and recommend more of the same.

5.8 Electronic Transactions

You haven't visited a bank branch in five years since it would be a huge time sink to dash there for every transaction. This is made possible in part by artificial intelligence. Today's banks use AI to streamline the payment process for its consumers.

With the use of AI for things like security, identity management, and privacy restrictions, it is now feasible to make deposits, transfer funds, and even start accounts remotely.

Credit card spending trends can also reveal probable fraud. Likewise, this is an application of AI. The algorithms may deduce the user's usual product preferences, purchasing habits, and average purchase price.

The system can provide a warning or request user input to confirm any financial activities that seem suspicious when compared to the user's profile.

6 AI Comes in Many Different Forms

Here are the different kinds of AI.

6.1 Only the Reactive

These machines have no memory or data to work with and specialize in a single subject. For example, in a chess game, the machine monitors the movements and makes the best decision possible to win.

6.2 Insufficient Memory

These gadgets collect historical data and store it in their memory. They have enough memory or experience to make good decisions, but their memory is limited. This system, e.g., can propose a restaurant based on the geographic data obtained.

6.3 The Mind–Body Connection

This sort of AI can understand concepts and emotions, as well as interact socially. This type of machine, however, has yet to be developed.

6.4 Machines that Are Self-aware

The next wave of revolutionary technology will be self-aware machines. They will be sentient, intelligent, and conscious.

7 Commonest Obstacles in AI

There are a number of issues with artificial intelligence, and we will discuss these obstacles and their solutions.

7.1 Availability of Data

Most organizations have isolated, segregated, inconsistent, and low-quality data. For firms wanting to get value from AI, such data presents a huge problem.

Companies must have a clear plan from the start in order for AI required data to be obtained from data sources in an orderly and consistent manner.

7.2 A Scarcity of Talent

Companies that want to reap AI's benefits may struggle to find workers who have the knowledge and expertise to properly fulfill the technology's potential. It is abundantly obvious that due to a dearth of artificial intelligence (AI) competence and a scarcity of appropriately trained technical personnel.

To successfully deploy and manage AI systems, businesses should consider allocating resources to educating their staff.

7.3 Implementation and Cost Time

The time and money required to create an AI solution is a major consideration when considering whether or not to pursue such initiatives. Businesses that don't have the resources or expertise to employ AI on their own.

Experts in AI can solve the integration problems that prevent systems and platforms from functioning smoothly together.

7.4 Safety and Regulations

The volume and security of acquired data is a significant problem for AI in commercial settings.

Strong authentication, data protection, and security governance in accordance with applicable regulations are all upheld.

7.5 Organizational Support

Organizational support is critical to the effective deployment and adoption of new technologies. To eliminate any resistance to this transition, business leaders must drive organizational transformation and industry leaders must provide the necessary infrastructure, training, education, and other types of employee support.

7.6 Job Reduction

AI helps businesses reduce costs and strain, and can also render certain duties unnecessary. Humans and AI must collaborate in order to maximize productivity. AI is a tremendous asset for repetitive, low-level duties. This entails employing AI in a supporting capacity as opposed to substituting human capabilities.

8 The Rapid Development of AI in the Corporate World

Businesses today have access to a greater volume of data than ever before. Forbes reports that the quantity of data generated and consumed will increase by a factor of five thousand between 2010 and 2020. Companies can now accumulate user information, allowing them to make more informed business decisions. This capability is enabled by the development of novel technologies.

A significant number of businesses have avidly embraced AI in recent years as part of their efforts to make better use of the data they are currently collecting. Take into consideration the following:

- According to Grand View Research, the value of the worldwide artificial intelligence market in 2020 was $62 billion, and it is anticipated to expand at an annual growth rate of 40.2% from 2021 to 2028.
- McKinsey's poll on "The state of AI in 2021" found that 56% of respondents report using AI, which is an increase from the 50% of respondents who used it in 2020.
- According to a survey by Accenture on artificial intelligence (AI), C-suite executives believe that harnessing AI will help them accomplish their development targets 84% of the time.
- According to a TechJury report, 35% of businesses are now employing AI and 42% are looking into it for potential future use.
- Leading companies spend consistently in AI (91.5%).
- IT analyst company Info-Tech Research Group claims that 44% of private sector enterprises want to invest in AI systems by 2023. Gartner also predicts that by 2023, consumer satisfaction would increase by 25% in organizations that employ AI.

These numbers demonstrate that artificial intelligence is no longer a cutting-edge technology that is reserved for a select few companies. It has developed into an indispensable component of the operations of a significant number of businesses all around the world.

9 How Does AI Fit into the Corporate World?

AI can be utilized in a variety of ways by businesses, with the majority of applications focusing on growth. By adopting AI and machine learning, businesses are discovering novel methods to improve business performance. Among the commercial advantages of AI are:

- Increasing productivity through process automation.
- Improving service's efficiency or consistency.
- Utilizing customer insights to guide decision making.
- identifying opportunities for novel product and service development.

AI is compatible with the vast majority of business strategies. Understanding the function of data acquisition and analysis in artificial intelligence is essential for getting started with AI. By analyzing the AI's underlying methodology, you can determine how AI might benefit your industry. An introductory AI course, such as Wharton Online's Artificial Intelligence for Business, is an excellent starting point for anyone interested in learning more about how AI is reshaping business.

10 Review of Literature

According to Demis Hassabis, the founder of Deepmind-Google's AI Business, "artificial intelligence is the science of making computers intelligent" [12]. Because AI is a foreign term used for a variety of meanings, this is the most common and most appropriate explanation. Several subcategories of artificial intelligence, such as machine learning and deep learning, provide real-world AI applications such as search suggestions, voice recognition, virtual assistants, and image recognition. The term artificial intelligence (AI) refers to the computer-assisted analytic course that aims to construct intelligent automated systems. It is the automated system that inputs data to perform the responsibilities of intelligent beings in a manner that increases their success rate.

Today, artificial intelligence can comprehend, learn, solve problems, and make decisions [13]. Robots, computers, and other AI-programmed systems can reason. According to IBM's AI research supervisor Guruduth Banavar, artificial intelligence can be viewed as a compilation of technologies.

These technologies perform varied duties and have varying costs, but their goal is nearly identical: to emulate human intelligence in technology in order to make the functions intelligent. AI advancements and implementations in many areas have led in the growth of AI technologies that are clearly effective and advantageous for marketing professionals. Because there is greater focus on digital marketing than traditional marketing tactics, there is abundant data for the use of AI technology. AI technologies have the potential to assist marketing managers in a variety of roles, including lead generation, market research, social media management, and user experience personalization [14, 15] say that artificial intelligence might be able to copy human traits or habits. Also, at the moment, some artificial intelligence is made without cutting-edge ways to learn on its own. Instead, it is told what to do. AI will end up like this [10]. In the end, AI machines will be able to understand how people act and feel and will change their roles accordingly.

AI is transforming nearly every aspect of human existence, including work, the economy, communication, wars, protection, security, ethics, and social insurance [12]. We still do not know how it will evolve over time or whether it will make the world a better or more dangerous place to live. Each technology has its own advantages and disadvantages, but the advantages must always outweigh the disadvantages for the technology to be commercially successful. We do not yet know whether the positive effects of artificial intelligence will always outweigh the negative effects in the long term. If this is not the case, we are in a difficult position. When we glance around, it appears that we comprehend how technology has changed things, whether it is “smart homes,” “smart health care,” “Industry 4.0,” or “self-driving cars.” Nevertheless, we frequently disagreed with the government regarding unemployment, fees, security, etc. As AI improves, more robots and self-operating systems are being developed to perform tasks previously performed by humans. This is the current situation, but the long-term outlook is more promising [16].

11 Objectives

1. To study how these AI advancements will impact many sectors of businesses and our daily lives.
2. To examine the role of artificial intelligence (AI) and its impact on business structures and strategies.

12 Research Methodology

The author conducted interviews with one hundred employees as well as those in authority positions at an organizations located in Andhra Pradesh. With the objective of obtaining information on people’s beliefs and knowledge regarding artificial intelligence (AI), as well as how it impacts society both now and in the future, an online survey is being used. This information will be utilized to inform the decisions that will be made on policy. A poll was carried out to ascertain people’s and company owners’ perspectives on the significance of artificial intelligence (AI) as well as the anticipated implications of AI across many sectors of the economy and in our day-to-day lives. The responses to this poll will be analyzed in order to ascertain the degree to which artificial intelligence currently has an influence on both our personal and professional life. Through the use of an online poll, we also reached out to Indian businesses, both decision-makers and regular employees, in industries such as financial services, technology, and manufacturing, to inquire about their perspectives on AI and how it affects society both now and in the future [13]. These industries include manufacturing, technology, and financial services. The replies to the poll were from people who were either self-employed, worked for one of the organizations being considered either full-time or part time, or were employed elsewhere. Business decision-makers and influencers were characterized as executives and managers in organizations who affected crucial decisions about the development of technology and services, as well as other crucial aspects of the business. The poll included responses from professionals in a variety of fields, including those in the manufacturing, technology, and financial services industries. When conducting research, both primary and secondary sources are utilized to collect data. In order to study the aforementioned issue, the researcher started by conducting a survey in order to collect primary data for the first time. In addition, papers, journals, books, websites, and blogs are examples of secondary sources of knowledge. The researcher employed a method known as convenience sampling, in which subjects were selected for participation for one of many predetermined reasons. In order to accomplish the goals of this inquiry, the author has zeroed in on a particular facet of business that is being changed as a direct result of the influence of AI. Investigating the ways in which artificial intelligence is influencing these aspects will offer a more complete view of the ways in which AI is influencing business models and strategy.

13 Analysis and Result

Hypothesis

H01: AI has no big effect on the growth of a business's economy.

Cronbach's Alpha was calculated on 100 responses and found to be 0.869; as a result, the Cronbach alpha value for the 100 employees of various organizations evaluated has been determined. A fantastic figure is 869, which is a superb indication of the data's quality and verifies the discovered data's almost 86% dependability. As a result, the data is sufficiently reliable for further investigation (Table 1).

Gender was the first key demographic question asked, as seen in Table 2 and Fig. 1. Based on the frequency of gender, the table revealed that 74 (74%) of the 50 people who responded were men and 26 (26%) were women. It revealed that the number of males working exceeded the number of women working.

According to Table 3 and Fig. 2, 46 (46%) of the 100 respondents who responded to the study believe AI has a very high impact on economic growth. 28 individuals (28%) believe AI has a large impact on economic growth, while 16 people (16%) believe AI has a medium impact. Six (6% of respondents) believe the effect is minor,

Table 1 Reliability statistics

Reliability statistics	
Cronbach's alpha	Number of respondents
0.869	100

Source Author's compilation

Table 2 Frequency table of gender

Gender					
		Frequency	Percent	Valid percent	Cumulative percent
Valid	Male	74	74	74	74
	Female	26	26	26	100
	Total	50	100	100	

Source Author's compilation

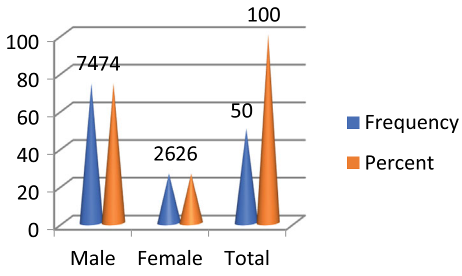


Fig. 1 Participants are classified based on their gender

while the remaining four (4% of respondents) did not react. As a result, the majority of those who responded to the study believe that AI has a significant impact on business and economic growth.

Table 4 gives the variation (Sum of Squares), degrees of freedom (*df*), and variance (Mean Square) within and across groups, as well as the *F* value (*F*) and its significance (Sig.). The sign indicates whether or not the null hypothesis, which states that the population means the same, should be rejected.

The preceding table demonstrates that there is a substantial difference between the mean squares (2.011), although it is not statistically significant ($F = 2.202$).

Because the sig. value for H01 is less than 0.05, these hypotheses must be rejected. Instead, they must accept their alternatives, which claim that AI has a substantial impact on corporate growth.

Table 3 Frequency table of the effects of artificial intelligence on the growth of business

Impact of AI on the economic growth of business					
		Frequency	Percent	Valid percent	Cumulative percent
Valid	Very high	46	46	46	46
	High	28	28	28	74
	Medium	16	16	16	90
	Low	6	6	6	96
	NA	4	4	4	100
	Total	100	100	100	

Source Author's compilation

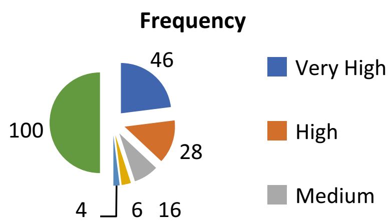


Fig. 2 Effects of AI on the growth of business

Table 4 ANOVA

ANOVA						
		Sum of squares	df	Mean square	*F*	Sig.
Impact of AI on the economic growth of business	Between groups	2.011	1	2.011	2.202	0.02
	Within groups	47.686	48	0.894		
	Total	100	49			
Impact of AI	Between groups	2.351	1	2.351	2.512	0.042

Source Author's compilation

14 Conclusion

If you look around, you'll see that our culture is evolving as a result of the widespread adoption of AI's many benefits. A machine that uses artificial intelligence to monitor your heart rate can be found in almost any hospital, and online shoppers can expect to see a list of products tailored to their own preferences provided by an AI recommendation engine. These are just a few of the ways in which AI has improved our lives. In the future, a lot more will occur. And as time goes on, more and more sectors will implement this technology to better their operations.

Rapid progress has been made in the last several decades in the area of artificial intelligence. New applications for artificial intelligence (AI) are constantly being discovered, and the field is expected to grow and develop fast in the years ahead. The use of AI in business is rapidly becoming one of the field's most promising frontiers. It paves the path for more self-driving and remote-controlled production equipment, meaning fewer humans will be needed. The benefits of this include better accuracy, flexibility, and safety. Autonomous vehicles are another hot topic in the field of industrial AI.

References

1. Bakkalbasi N, Bauer K, Glover J, Wang L (2006) Three options for citation tracking: google scholar, scopus and web of science. BMC Biomed Digit Libr 3(7):1–8. https://doi.org/10.1186/1742-5581-3-7
2. Burkhalter B (1963) Applying artificial-intelligence to pattern-cutters problem. Oper Res 11:39
3. Borkar VR, Carey MJ, Li C (2012) Big data platforms: what's next? ACM Crossroads 19(1):44–49
4. Callan JP, Croft WB, Harding SM (1992) The INQUERY retrieval system. In: Database and expert systems applications. Springer, Vienna, pp 78–83. https://doi.org/10.1007/978-3-7091-7557-6_14
5. Cavalcante R, Brasileiro RC, Souza VLF, Nobrega JP, Oliveira ALI (2016) Computational intelligence and financial markets: a survey and future directions. Expert Syst Appl 55:194–211. https://doi.org/10.1016/j.eswa.2016.02.006

6. Chadegani AA, Salehi H, Yunus M, Farhadi H, Fooladi M, Farhadi M, Ale Ebrahim N (2013) A comparison between two main academic literature collections: web of science and scopus databases. Asian Soc Sci 9(5):18–26. https://doi.org/10.5539/ass.v9n5p18
7. Chan FTS, Jiang B, Tang NKH (2000) Development of intelligent decision support tools to aid the design of flexible manufacturing systems. Int J Prod Econ 65(1):73–84. https://doi.org/10.1016/S0925-5273(99)00091-2
8. Cho V (2003) A comparison of three different approaches to tourist arrival forecasting. Tour Manage 24(3):323–330. https://doi.org/10.1016/S0261-5177(02)00068-7
9. Choi JJ, Ozkan B (2019) Innovation and disruption: industry practices and conceptual bases. In: Choi JJ, Ozkan B (eds) Disruptive innovation in business and finance in the digital world, vol 20. Emerald Publishing Limited, pp 3–13. https://doi.org/10.1108/S1569-376720190000020003
10. Martínez DM, Fernandez-Rodríguez JC (2015) Artificial intelligence applied to project success: a literature review. IJIMAI 3(5):77–84
11. Altman EI, Marco G, Varetto F (1994) Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience). J Bank Finan 18(3):505–529. https://doi.org/10.1016/0378-4266(94)90007-8
12. Ahmed I, Singh S (2015) AIML based voice enabled artificial intelligent chatterbot. Int J u-and e-Service Sci Technol 8(2):375–384
13. Zhang X, Chan FT, Yan C, Bose I (2022) Towards risk-aware artificial intelligence and machine learning systems: an overview. Decis Support Syst 159:113800
14. Sterne J (2017) Artificial intelligence for marketing: practical applications. John Wiley & Sons
15. Özgül E, Turan N, Tınar MY (2017) Obstacles of the development of innovative business ideas: a survey on students of entrepreneurship education. Girişimcilik İnovasyon ve Pazarlama Araştırmaları Dergisi 1(2):120–135
16. Tyagi A (2016) Artificial intelligence: boon or bane? Available at SSRN 2836438
17. Kapoor N, Bahl N (2016) Comparative study of forward and backward chaining in artificial intelligence. Int J Eng Comput Sci 5(4):16239–16242

Exploring Water Quality Prediction Using Machine Learning—An Efficient and Emerging Method of Evaluation

Rajeshwarrao Arabelli, Arasam Pooja, Sripathi Sai Nanditha, Kalluri Deepa, Mondithoka Ajaybabu, and Syed Musthak Ahmed

Abstract Water contamination poses a severe risk to India's water resources' sustainability and can result in insufficient water supply for all residents, despite the country's abundance of water resource. A significant portion of the population is impacted by different infections brought on by consuming contaminated water. Aquatic life forms, from fish to bacteria, depend on water for their home in addition to being essential for human nutrition. Water quality analysis is vital for maintaining ecological balance and biodiversity since aquatic ecosystems' health is closely associated with the quality of the water they live in. In this project, machine learning is used to analyze and forecast the water quality using Python tools and the dataset that was fed into the model. This project aids in determining the potability of the water. Numerous factors, including pH, conductivity, hardness, particulates, turbidity, chloramines, sulfates, trihalomethanes, and organic carbons, are included in the dataset that affect the water's cleanliness.

Keywords Water quality · Classification · Decision tree

1 Introduction

Water is an essential element for life, playing a crucial role in various bodily functions, including digestion, circulation, temperature regulation, and waste removal. Therefore, ensuring the quality of water is vital for maintaining human health and well-being. Water quality monitoring is crucial for human health and the environment, as it helps to ensure the safety of water for consumption and other uses [1]. The importance of clean water for human life cannot be overstated, as it is essential

R. Arabelli (✉) · A. Pooja · S. S. Nanditha · K. Deepa · M. Ajaybabu
Department of ECE, School of Engineering, SR University, Warangal, India
e-mail: rajeshwarrao432@gmail.com

S. M. Ahmed
Department of ECE, Vagdevi Engineering College, Warangal, India
e-mail: musthak.ahmed@vecw.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_33

for bodily functions and overall health [2]. However, water pollution is a significant global issue, with industrialization and urbanization contributing to its severity [3]. Therefore, the development and implementation of cost-efficient water quality monitoring systems are essential to safeguard human health and the environment.

Water must be important because Earth is the only planet where life exists. The oceans contain around 96% of the planet's water, with only 2.5% of the planet's surface being covered in water. The average person needs more than three liters of water every day. Despite the presence of rivers and lakes all around the city, we still have to deal with the issue of water contamination. Human activities such industrial effluents, agricultural practices, problems with urban waste management, and increased urbanization have posed major challenges to the freshwater ecosystem in recent years. Additionally, the consequences of climate change have an influence on the typical functions of aquatic ecosystems, such as eating and reproducing.

These pollution levels have also affected the habitats of aquatic plants and wildlife. Determining the primary pollution activities, their sources, and their fate in the aquatic environment, including their spatial distribution, is essential to maintaining the regular functioning of ecosystems and protecting freshwater aquatic species globally. Therefore, water quality surveillance is crucial since improper monitoring can have detrimental effects on living things. Although there are a number of manual techniques for predicting the quality of water, they are costly and time-consuming. There are restrictions on this approach as well. For instance, a farmer must wait a week or longer for all the laborious labor involved in measuring the water quality before he can check it for use in his crops.

However, sensors can be used to manually assess the water quality by measuring a variety of attributes. Since this will take a long time, one way to deliver the forecasts could be through the use of ML technology. Because technology is advancing so quickly these days, advancement and change can occur more quickly, leading to a rise in acceleration. A number of methods are proposed for the modeling and water quality prediction. The few technologies that are developing the fastest are artificial intelligence, machine learning, deep learning, etc. As a result, techniques such as statistical methods, visual modeling, analysis, and predictive algorithms, among others, can be applied to the forecast.

Mechanism models and nonmechanism models are the two categories of models that are available for simulating and forecasting water quality. This study examined various machine learning techniques for water quality prediction.

2 Related Work

Al-Sulttani et al. [4], a thorough analysis using a feature extraction technique was carried out to compare the effectiveness of integrated models with solo models. The study provides a comprehensive analysis of how different model architectures, when combined with feature extraction techniques, influence predictive accuracy in water quality prediction tasks. By assessing the efficacy of both integrated and independent models, the research offers valuable insights into optimizing predictive performance in complex environmental systems. The relative benefits and drawbacks of different modeling techniques are clarified by this comparative analysis, which will help guide future developments in water quality forecast techniques.

Salisu Yusuf Muhammad et al. developed a trustworthy classification model for evaluating the quality of water in [5] using machine learning approaches. Their research highlights the usefulness of machine learning in solving problems related to water quality evaluation.

By leveraging sophisticated classification algorithms, the research introduces a systematic approach for categorizing water quality parameters, facilitating more accurate and efficient quality evaluation. This endeavor highlights the pivotal role of machine learning in enabling data-driven methodologies for managing and safeguarding water resources.

Mohammed et al. [6], lake water underwent treatment procedures, followed by comprehensive quality testing. The study delved into various aspects of water treatment processes, aiming to ascertain the efficacy of the treatment methods employed. Through meticulous evaluation of the treated lake water quality, the study provides important information on how well the treatment procedures work to fulfill the necessary requirements for water quality. This endeavor underscores the importance of ensuring water safety through systematic testing and treatment measures, crucial for maintaining ecological balance and public health.

In [7], the evaluation of harvest performance, growth, and yield in the prawn farm is contingent upon the analysis of five key water quality parameters. These parameters serve as critical indicators, offering insights into the ecological conditions essential for optimal prawn cultivation. By considering factors such as dissolved oxygen levels, pH, temperature, ammonia, and nitrate concentrations, the study aims to comprehensively assess the environmental suitability for prawn farming. This approach underscores the significance of water quality management in maximizing production efficiency and ensuring sustainable aquaculture practices.

Random Forest, as described in reference [8], represents a supervised machine learning technique characterized by an ensemble of decision trees. By combining the forecasts from several decision trees, this approach enhances predictive accuracy and robustness. It likely elaborates on the methodology and application of Random Forest in various domains. Its efficacy stems from the diversity of constituent trees and the randomization employed during training, contributing to reliable and versatile predictive modeling capabilities.

Liao and Sun [9], the decision tree is a popular supervised learning method that is frequently applied to classification tasks. This approach involves recursively partitioning the input space into regions, guided by features that optimize classification accuracy. It likely provides insights into the principles, methodologies, and applications of decision trees in classification scenarios. Its widespread use stems from its interpretability, scalability, and effectiveness in handling complex classification problems across various domains.

Two hybrid decision tree-based machine learning models complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN)-RF and

CEEMDAN-XGBoost are used to estimate the quality of water is done in [10]. Osim predicting the quality of drinkable water is essential for efficient water distribution and environmental preservation programs, claim Pal et al. [11].

By focusing on forecasting water quality fit for consumption, the study highlights the potential to mitigate health risks and optimize resource allocation in water management systems. This approach underscores the proactive measures needed to ensure safe and sustainable access to potable water, promoting both public health and environmental stewardship. This study used deep learning and the Long Short-Term Memory (LSTM)–Convolutional Neural Network (CNN) [12].

Hayder et al. [13], the dataset contains attributes including turbidity, potability, pH value, hardness, solids content, chloramines, sulfate concentration, conductivity, levels of organic carbon, and trihalomethane status that are important for evaluating the quality of water. These varied characteristics enable precise predictive modeling for potability assessment and environmental preservation by offering thorough insights into the many aspects influencing water quality.

Machine learning algorithms are able to identify patterns and correlations with previous data and pertinent information in order to generate precise projections regarding the surroundings of water quality indicators [14]. The purpose of this work is to investigate supervised machine learning techniques for determining the water quality index (WQI) and water quality class (WQC) [15]. Pathogens and potentially contaminated substances need to be continuously monitored due to their high concentration in water A_1. Presented a predictive neural network model that may be used to determine whether water is safe for human consumption. The model employed neural networks to identify water that was 70% accurate for human consumption [16].

Kurra et al. [17], evaluated the suitability of KNN and DT as data mining approaches in a different study. This study looked at the potential applications of machine learning algorithms in the prediction of water quality by utilizing projected values for future capacities. Ahmed et al. [18], the study made use of a broad variety of machine learning algorithms, including support vector machine methods, multiple linear regression, gradient boosting, Random Forest, and polynomial regression. This extensive methodology made it possible to compare and thoroughly explore several modeling methodologies or the water quality forecast. Shaik et al. [19], study aimed to identify the most effective method for accurately forecasting water quality characteristics by utilizing a variety of algorithms. This thorough review provides insightful information about the effectiveness and applicability of several machine learning techniques for handling challenging environmental prediction problems. Such diverse methodologies offer robust solutions for optimizing water resource management and environmental protection efforts.

3 Technologies Used

Python Libraries

A numerous Python libraries are accessible through the Python Notebook. The Pandas library is used to obtain the input file from the user and performs data analysis. To perform operations on arrays and matrices, utilize the NumPy library. Two data visualization libraries that offer a variety of graphs and plots for data analysis are Matplotlib and Seaborn. The SKLearn library is utilized for pre-processing data and separating it into training and testing sets.

Anaconda and Python Notebook

An open-source Python distribution is called Anaconda. It is used in several domains, including deep learning, machine learning, and data science. With over 300 data science libraries at their disposal, working with Anaconda for data science turns into a really beneficial option for every coder. The Python notebook required to do the assignments is available from Anaconda.

Data

The dataset, which is in the form of a CSV file, includes information on several variables, including pH, hardness, solids, chloramines, and potability. The project's outcome is the potability.

4 Implementation

Random Forest classification is applied to this project. To boost the dataset's projected accuracy, a Random Forest classifier employs numerous decision trees on various dataset subsets and averages them, uses the majority vote of each tree's projections to predict the outcome in advance. Accuracy is increased and overfitting is prevented when there are more trees in the forest. The proposed system implementation flowchart is shown in (Fig. 1).

5 Results

The dataset we deemed to have more non-potable water data than potable water data in the form of a bar graph can be consumed, as shown in Table 1, which counts the values of the potability attribute. A zero indicates unfit for drinking, while a one indicates water.

The histogram in Fig. 2 provides a graphical representation of the attribute analysis.

All the bar graphs above represent the attribute analysis of all the samples in the dataset. Sample values are taken on the *x*-axis of the bar graph, and attributes values are taken on the *y*-axis of the bar graph. For all the samples in the data we calculated

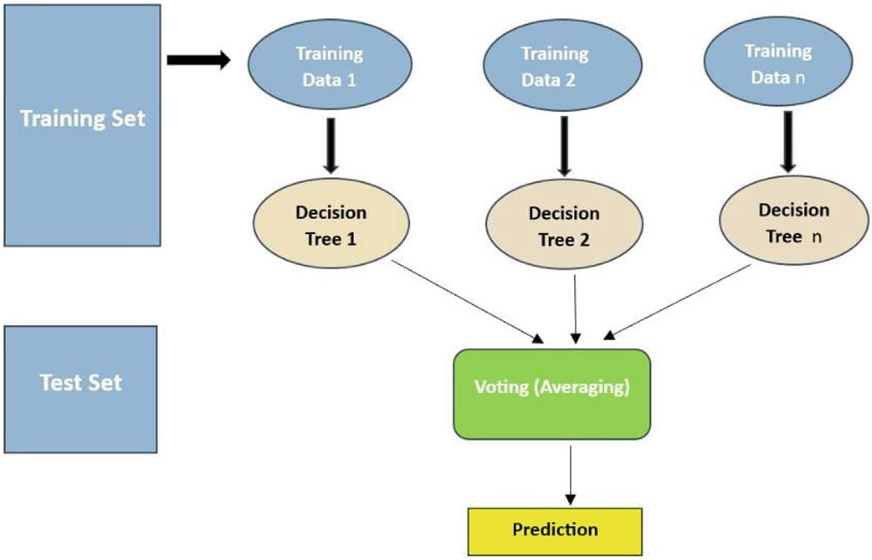


Fig. 1 Flowchart of the proposed system

pH (Fig. 3), Hardness (Fig. 4), solids (Fig. 5), chloromines (Fig. 6), sulfates (Fig. 7), conductivity (Fig. 8), organic carbon (Fig. 9), trihalomethanes (Fig. 10), turbidity (Fig. 11) from these values we got a result potability graph which indicates the quality of the water. If the potability of water is 1 it is suitable for drinking and if 0 it is not suitable for drinking.

A heatmap graphical depiction of the properties using a color-coded method is shown in Fig. 12.

Compared to previous classification methods, the Random Forest approach improves the model's accuracy as shown in Table 2.

Table 1 Summary of characteristics

	pH	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_ carbon	Trihalomethanes	Turbidity	Potability
Count	2785.00	3276.00	3276.00	3276.00	2495.00	3276.00	3276.00	3114.00	3276.00	3276.00
Mean	7.08	196.37	22,014.09	7.12	333.78	426.21	14.28	66.40	3.97	0.39
Std	1.59	32.88	8768.57	1.58	41.42	80.82	3.31	16.18	0.78	0.49
Min	0.00	47.43	320.94	0.35	129.00	181.48	2.20	0.74	1.45	0.00
25%	6.09	176.85	15,666.69	6.13	307.70	365.73	12.07	55.84	3.44	0.00
50%	7.04	196.97	20,927.83	7.13	333.07	421.88	14.22	66.62	3.96	0.00
75%	8.06	216.67	27,332.76	8.11	359.95	481.79	16.56	77.34	4.50	1.00
Max	14.00	323.12	61,227.20	13.13	481.03	753.34	28.30	124.00	6.74	1.00

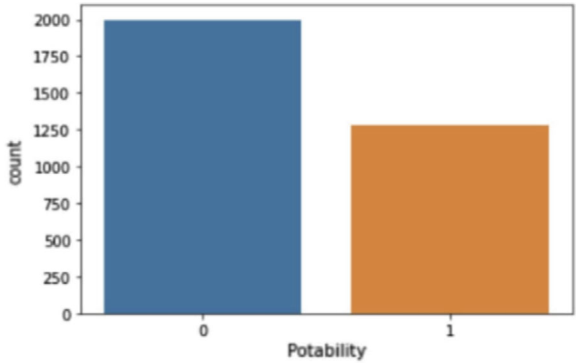


Fig. 2 Potability graph

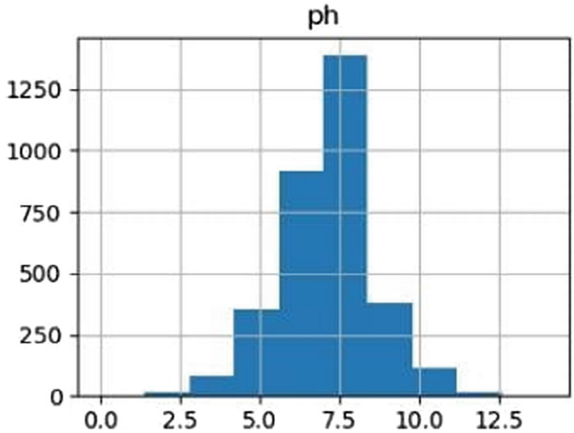


Fig. 3 pH graph

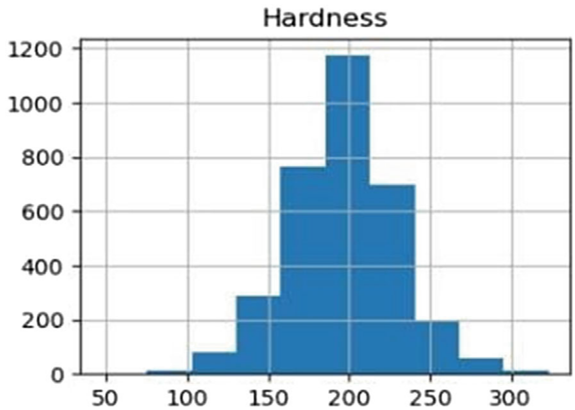


Fig. 4 Hardness graph

Fig. 5 Solids graph

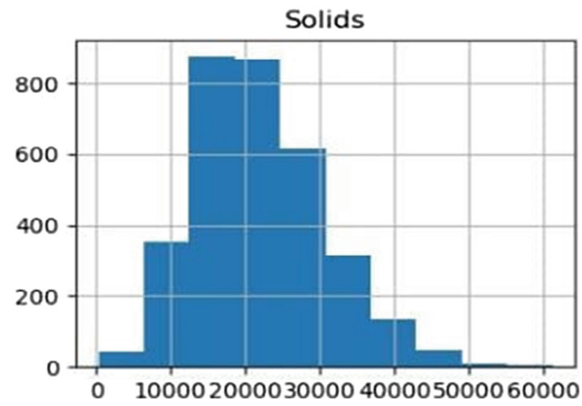


Fig. 6 Chloramines graph

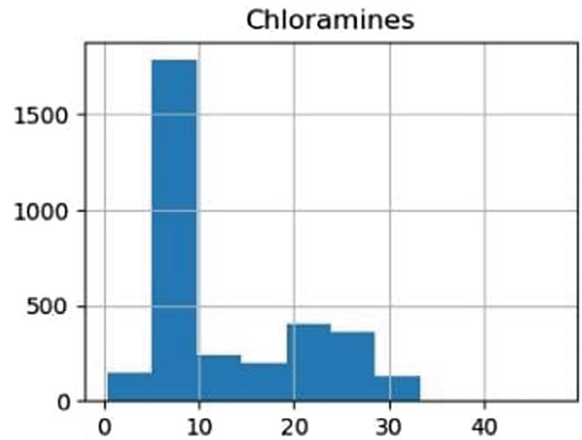
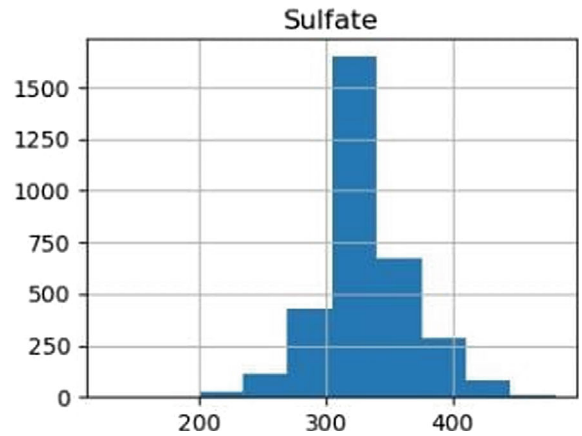


Fig. 7 Sulfate graph



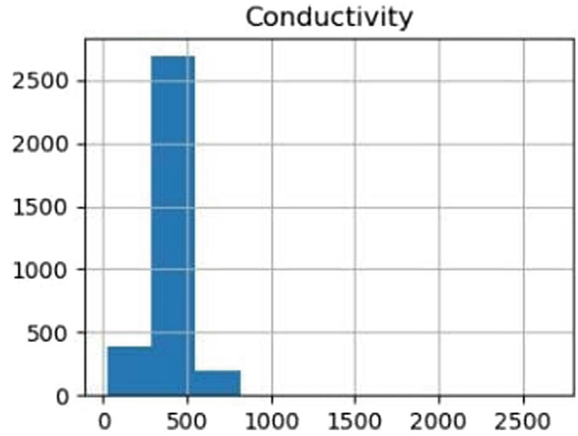


Fig. 8 Conductivity graph

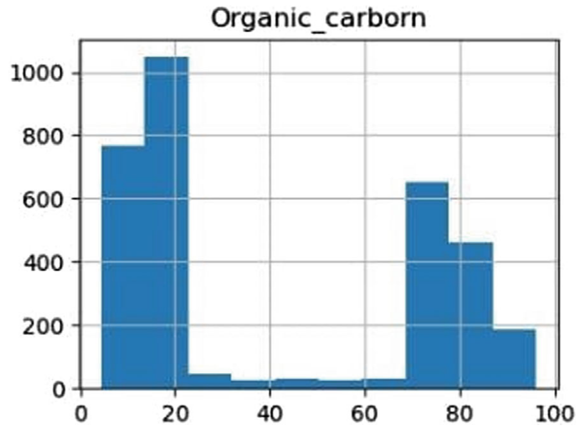


Fig. 9 Organic carbon graph

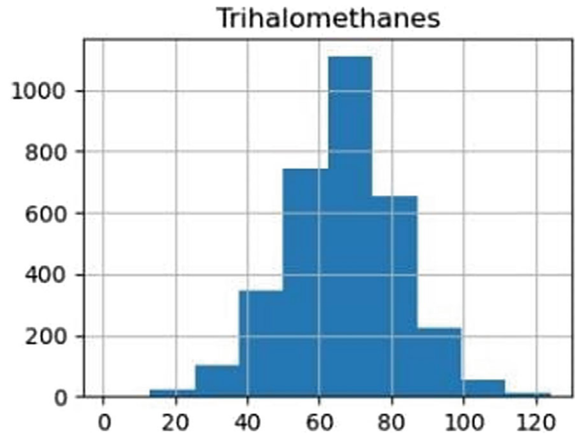
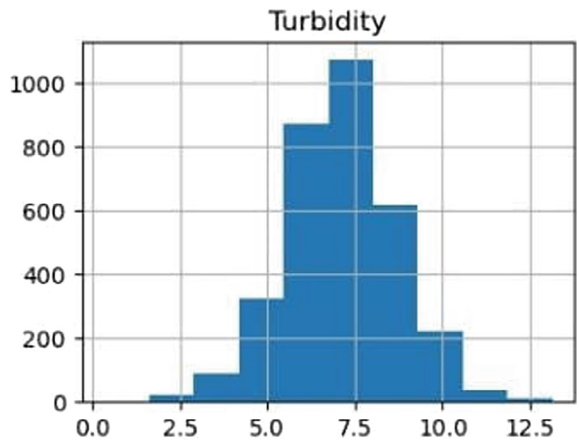


Fig. 10 Trihalomethanes graph

Fig. 11 Turbidity graph



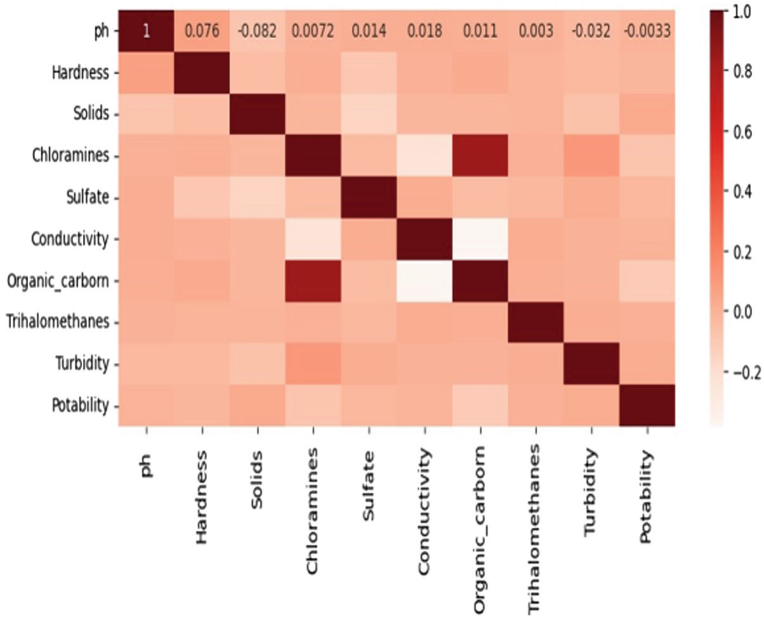


Fig. 12 Heatmap representation

Table 2 Accuracy comparison of various models

Model name	Accuracy
Decision tree classifier	92.98780487
Random forest classifier	98.17073170
Logistic regression	62.43292682

6 Conclusion

The paper emphasizes the importance of water quality monitoring for human health, environmental sustainability, and aquatic life, highlighting the challenges posed by contamination and the role of machine learning in predicting water quality. Water contamination poses a significant risk to water resources in India, impacting human health and biodiversity. Machine learning techniques can effectively predict water quality indicators and classify water quality, aiding in ensuring clean water for human health and well-being.

The developed models show promising output concerning water quality index prediction and classification. As a result, we can learn how the characteristics affect the water's purity. This is applicable to all regions of the nation. It can also be improved much more. The approach assists people in determining whether the water they are drinking is safe or not. Through collaborative efforts and continued innovation, we can continue to refine this vital process, securing clean and safe water for all.

Acknowledgements We would like to express our sincere appreciation to SR University for their invaluable support and guidance through the research process. Their contributions have significantly enriched the quality of this paper.

References

1. Ptiček Siročič A, Ojdanić K, Dogančić D, Plantak L (2023) Water quality for human consumption from the public water supply system. Environ Sci Proc 25(1):21
2. Ranjith J, Rajeev H (2023) Aqua check-water quality monitoring system. Indian J Image Process Recogn (IJIPR) 3(3):1–5
3. Wang X, Yang W (2019) Water quality monitoring and evaluation using remote sensing techniques in China: a systematic review. Ecosyst Health Sustain 5(1):47–56
4. Al-Sulttani AO, Al-Mukhtar M, Roomi AB, Farooque AA, Khedher KM, Yaseen ZM (2021) Proposition of new ensemble data-intelligence models for surface water quality prediction. IEEE Access 9:108527–108541
5. Salisu Muhammad Y, Makhtar M, Rozaimee A, Aziz AA, Jamal AA (2015) Classification model for water quality using machine learning techniques. Int J Softw Eng Appl 9(6):45–52
6. Mohammed H, Tornyeviadzi HM, Seidu R. Emulating process-based water quality modeling in water source reservoirs using machine learning. J Hydrol 609
7. Rana M, Rahman A, Dabrowski J, Arnold S, McCulloch J, Pais B (2021) Machine learning approach to investigate the influence of water quality on aquatic livestock in freshwater ponds. Biosyst Eng 208
8. Ahmed M, Mumtaz R, Anwar H, Shah A, Irfan R, Garcia-Nieto J (2019) Effective water quality prediction using supervised machine learning. Water 11
9. Liao H, Sun W (2010) Forecasting and evaluating water quality of Chao lake based on an improved decision tree method. Procedia Environ Sci 2:970–979
10. Lu H, Ma X (2020) Hybrid decision tree-based machine learning models for short-term water quality prediction. Chemosphere 249:126169
11. Pal OK (2022) The quality of drinkable water using machine learning techniques. Int J Adv Eng Res Sci

12. Baek SS, Pyo J, Chun JA (2020) Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. Water 12(12):3399
13. Hayder G, Kurniawan I, Mustafa H (2020) Implementation of machine learning methods for monitoring and predicting water quality parameters. Appl Chem Biointerface Res 11:9285–9295
14. Ighalo JO, Adeniyi AG, Marques G (2020) Internet of things for water quality monitoring and assessment: a comprehensive review. In: Artificial intelligence for sustainable development: theory practice and future applications, pp 245–259
15. Ahmed U, Mumtaz R, Anwar H, Shah AA, Irfan R, García-Nieto J (2019) Efficient water quality prediction using supervised machine learning. Water 11(11):2210
16. Chafloque R (2021) Predictive neural networks model for detection of water quality for human consumption. IEEE Xplore
17. Kurra SS, Naidu SG, Chowdala S, Yellanki SC, Sunanda DBE (2022) Water quality prediction using machine learning. Int Res J Modernization Eng Technol Sci
18. Ahmed U, Mumtaz R, Anwar H, Shah AA, Irfan R, García-Nieto J (2019) Efficient water quality prediction using supervised machine learning. Water 11(11)
19. Shaik MA, Manoharan G, Prashanth B, Akhil N, Akash A, Reddy TRS (2022) Prediction of crop yield using machine learning. AIP Conf Proc 2418(1)

Electric Vehicle Battery Management System with Charge Monitoring and Its Protection

Rajababu Durgam, Sreeja Adilapuram, Vinod Kumar Akula, Rohith Vemula, Shiva Prasad Gokarapu, and Rejeshwarrao Arrabelli

Abstract Worldwide, the population is increasing, and they are using different types of vehicles like petroleum and diesel consumption vehicles for their transportation. Therefore, the vehicles can release various types of environmentally harmful gases like carbon dioxide, monoxide gases, etc. The gases can pollute the environment, leading to pollution and increasing global warming. To decrease pollution in the environment and reduce global warming effects, most of the countries are looking toward electric vehicles (EVs). In EVs, energy storage batteries play a vital role. But the batteries are facing some problems like overheating, fast discharging, and slow charging. The issues can be minimized by using an effective monitoring and control system. This paper presents the effective charging and discharging of the proposed EV battery management system. The proposed system is designed and modeled using MATLAB and Simulink, and results are presented.

Keywords Battery management system · Charging of battery · Electric vehicle · Pulse generator · State of charge

1 Introduction

In these modern days, technology is the one-time solution to many problems. In the same way, in order to protect the environment, electric vehicles (EVs) are introduced into the market. The environment is mostly polluted by various means of transportation, which are driven using internal combustion engines (ICE), like cars, bikes, buses, etc. These vehicles consume fuels like petrol and diesel and release harmful gases like carbon dioxide, carbon monoxide, etc., which lead to air pollution. To overcome this problem, electric vehicles play a crucial role by replacing IC engines

R. Durgam (✉) · S. Adilapuram · V. K. Akula · R. Vemula · S. P. Gokarapu
EEE Department, SR University, Warangal, India
e-mail: durgamrajababu@gmail.com

R. Arrabelli
ECE Department, SR University, Warangal, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_34

with electric vehicles. Electric vehicles are more widely adopted in many parts of the world because of their great performance and advantages.

The EVs are lighter in weight because of their composite components. Electric vehicles help reduce greenhouse gas emissions and carbon footprints by replacing conventional internal combustion engines. As compared to IC-engine vehicles, electric vehicles have a faster pickup rate. EVs have a lower maintenance cost because they require one-time battery charging, whereas IC engines have a higher maintenance cost because fuel costs are higher. EVs are more environmentally friendly in spite of air and noise pollution, in particular. The primary requirement for battery charging is renewable energies like solar and wind. Electric vehicles have some specific requirements to operate efficiently. Such electric vehicles require rechargeable batteries for continuous energy supply with a better battery management system, and rechargeable batteries are of different types, like lithium-ion batteries, nickel-metal hydride batteries, lead acid batteries, etc.

A battery management system is an important component in electric vehicles; its major role is to control, monitor, and protect the battery. The battery requires continuous cell monitoring, like detecting voltage and temperature levels, for balanced charging and discharging. The battery requires the estimation of the state of charge (SOC), which gives accurate information about the level of charge present in the battery. It also requires continuous monitoring of the state of health (SOH), such as the capacity and internal resistance of the battery, which is helpful to know the battery performance and battery life. Also, in batteries, it is important to manage the thermal stability of the battery. The battery mainly requires protection from overcharging and over discharging problems.

In this article, the state of charge (SOC), state of health (SOH), and state of function (SOF) of an electric vehicle battery management system are proposed and explained. It is observed that SOC is estimated by the present maximum capacity to avoid errors in SOC estimation [1–4]. Lithium-ion batteries are widely used in electric vehicles. They are available in packs that are configured with a number of cells connected in series and parallel. These are used to deliver the power and force required to move the electric vehicle [5–8]. Any battery ages and loses its performance at some point. To avoid this problem, it is suggested to use twin-based technology, which is so helpful for the efficient working of the batteries in electric vehicles. This technology is proposed in two types of models: the Extreme Gradient Boost model and the Extended Kalman Filter, which involves the prediction of SOC in the batteries [9, 10]. By integrating technology and upgraded versions of processors, the SOC of batteries can be estimated, like Digital Signal Processors (DSP) with updated features that do passive cell balancing in electric vehicles. This type of technology is widely used because of the simple and efficient functioning of batteries [11, 12]. Arduino is the most commonly used device in many other appliances for sensing and generating an action output, like a smartphone. In the same way, it is also used to monitor the battery life in electric vehicles by sensing voltage, current, and temperature and displaying them on the user interface [13]. As everyone knows, there is a saying that "prevention is better than cure." In the same way, it is important to note that to prevent battery fire accidents, we use battery management systems for optimal use and safety of the battery in EVs [14]. The protection of the battery is very important in EVs. When the batteries are overheated, they can release smoke. Batteries also release smoke at fault conditions, which are sensed by various sensors like temperature sensors, smoke sensors, and current sensors [15]. The Internet of Things (IoT) was the popular technology used for smart grid developments, and it is also useful for battery management and SOC estimation of batteries, so this technology is mainly used for a better and smarter way of charging the battery in an optimal and safer way [16]. Specifically, when we talk about SOH (state of health), it focuses on the charging amount of the battery while it is working and gives clear information regarding how much distance an EV should travel [17]. In order to protect electric vehicles from fire accidents, there is another way to control the vehicle system using vanadium air flow batteries. They help reduce fire flames coming out of the battery by using nitrogen gas diffusing through a battery box fixed to it during overcharging or overheating conditions. Batteries can also be monitored using analog and digital sensors with microcontrollers [18].

The span of analysis for this study was set from 2007 to 2023 based on various criteria such as the number and types of publications, the keywords used, and so on. A total number of 50 sources, including journals and books were identified in this period, which is a significant number. From these sources, 57 documents were selected, comprising articles, review papers, books, book chapters, and excluding other publications. The bibliometric data shown that the annual publication rate in this domain increased by 10.58%, indicating that the topic or the keywords were widely researched or explored. The number of authors who published in this field each year reflected the academic interest in this topic. The publication to author ratio of this study was 15.79%, suggesting that there were many active authors, and electric vehicle trend was growing. Figure 1 summarizes the research projects done using the keywords 179. The average age of the documents was 7.28, which implies that the research was relatively recent. The average international coauthorship rate of 15.7% implies that this field required extensive research collaboration, as it was dominated by smaller teams. The analysis of the yearly publication output showed that 57 articles were published in 2023, which served as the input for this study. Worldwide the researchers and scientist were searched about various keywords shown in Fig. 2.

This paper describes the battery management system described in Sect. 2. For system analysis and simulation, mathematical modeling is required. The mathematical modeling of the system is explained in Sect. 3. The simulation of the system and results are explained in Sect. 4. The conclusion of the system is described in the next session. Here in this article, the observations are proposed as a battery management system that plays a vital role in protecting and improving battery efficiency. It provides continuous control and monitoring over charging and discharging to improve the lifespan and efficiency of the batteries in electric vehicles. It focuses on the battery's temperature, voltage, and current monitoring, which prevents overcharging and deep discharging in the battery. This technology also gives complete balancing and monitoring over individual cells in the battery pack, and it ensures uniform performance, which increases the battery capacity. Also, BMS in electric vehicles provides deep protection and safety from overheating and fire diffusion. It

Fig. 1 Research information about electrical vehicles from 2007 to 2023

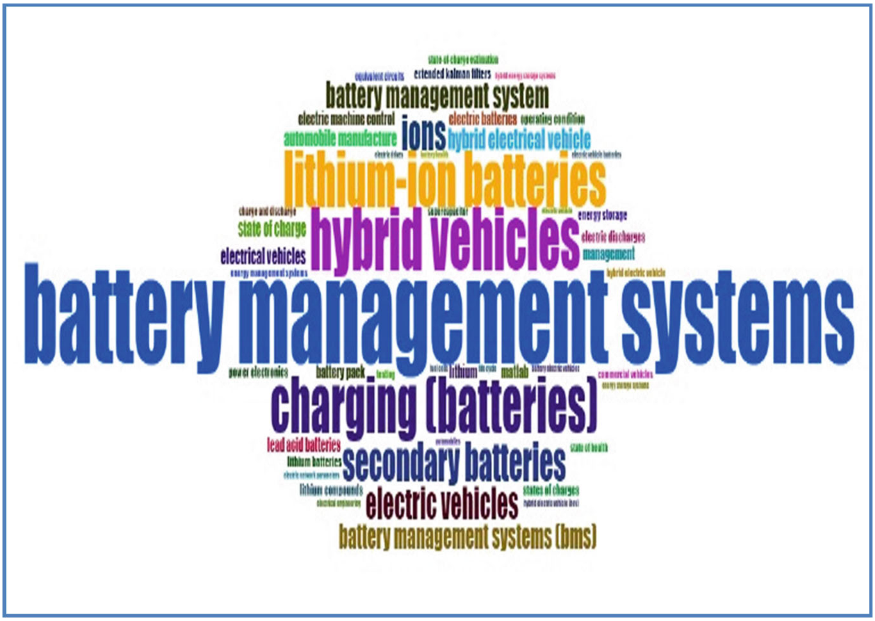


Fig. 2 World cloud data from bibliometric analysis

also monitors the battery's health by detecting faults that occur in the battery like short circuits and overheating. This type of battery management system is achieved and observed using MATLAB/Simulink software. So, in this way, the battery management system (BMS) provides complete protection for electric vehicles in order to reduce the battery blasting rate in EVs. So, using these most efficient and safer electric vehicles with a better battery management system helps the environment by reducing pollution and the release of toxic gases into nature from the ICE's only by replacing them with electric vehicles. This system also provides a better solution to control global warming. The detailed simulation process and modeling of a battery management system are explained in Sect. 2.

2 System Description

In the EV BMS, the study of battery at different stages is observed and controlled; battery energy management is very important for optimal utilization of storage energy and protection of the battery from overheating. They can be achieved by systematic control and monitoring of the system. In this paper, for the effective monitoring, a system is proposed which is shown in Fig. 3 using MATLAB/Simulink.

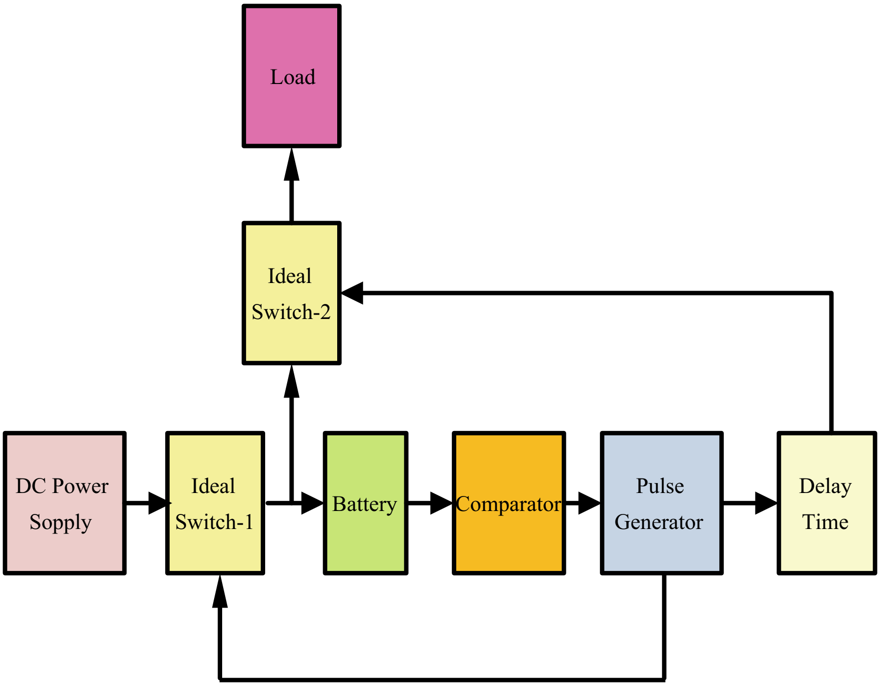


Fig. 3 Block diagram of effective monitoring of proposed system

2.1 Constructions

For the simulation of the system, initially a DC voltage supply is considered. Two ideal switches are connected between the source and the battery. Battery output is connected to the comparator (chart). Here in the chart, conditions are given according to the battery application. If the given condition is satisfied, then it generates a pulse, which is again fed to the two ideal switches, which are connected in series with battery input terminals; otherwise, no pulse is generated and the switches are kept at their initial condition. After the pulse generator, a time delay is provided. Based on the initially specified conditions, i.e., the storage charge is less than 20% or greater than 80% of the battery rating, the switches can be opened, and the battery is isolated from the DC supply. If the storage charge of the battery is within the limits (20–80% of the battery storage capacity), the switches are closed, and the battery will be charged. Based on the storage charge, the pulses will be generated, and they are given to the switches (switch 3 and 4), which are connected in series with the load. When switches 3 and 4 are closed, it means that the load is connected to the battery, and it will discharge the storage energy.

2.2 Working Procedure

The step-by-step procedure of the proposed system flow chart is shown in Fig. 4. And its algorithm is explained as:

Step 1: Read the initial conditions of the battery. The nominal voltage, battery capacity, region of operation, load condition, and time.
Step 2: If the battery is fully charged, then ideal switches 1 are open; otherwise, the battery is connected to supply through the switches.
Step 3: The battery condition is given to the input of the comparator.
Step 4: If the battery charge is greater than 80% of the rated capacity or overheating, the ideal switches 1 are open; otherwise, if the charge of the battery is less than 20%, the battery is connected to the supply to get charge.
Step 5: If the battery condition is within the limits, the operating pulse is generated.
Step 6: The generated pulse is given to switch 2, and the load is connected or disconnected from the battery.

3 Mathematical Modeling

In the mathematical modeling of the battery management system, the state of charge of the battery (SOC) is calculated by using Eq. (1).

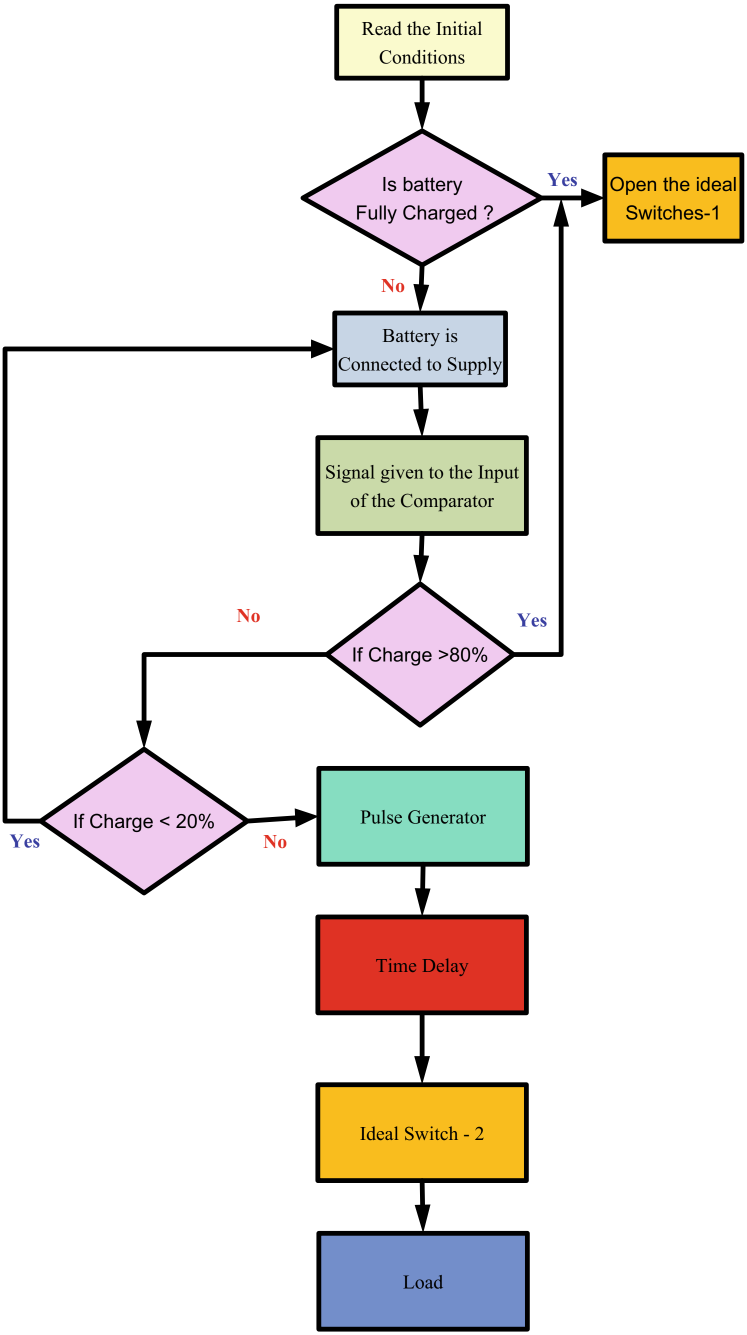


Fig. 4 Flowchart of the battery charge monitoring system

$$\frac{q(t)_{\text{residual capacity}}}{q_{\text{present maximum capacity}}} \tag{1}$$

where $q(t)_{\text{residual capacity}}$ is the residual capacity of the battery, $q_{\text{present maximum capacity}}$ is the present maximum capacity of the battery. $q(t)$ residual capacity is measured by OCV and coulomb integral as Eq. (2).

$$q(t_n)_{\text{residual capacity}} = q(t_0)_{\text{residual capacity}} - \int_0^{t_n} i(t)\mathrm{d}t \tag{2}$$

Then, $q_{\text{present maximum}}$ capacity can be written as Eq. (3).

$$q_{\text{present maximum capacity}} = q_{\text{nominal capacity}} * \text{ SOC} \tag{3}$$

According to the charge/discharge cycle of the battery specification (ICR18650-26H) and our experimental data, a lookup table is developed for SOH.

SOF is a digital signal stating whether the battery has sufficient power capability to reach a pre-defined task. The output of the SOF based on the SOC, SOH, and, e.g., forecast distance is described in Eqs. (4–6).

$$\text{Remaining Power train} = \text{SOC} * \text{ nominal capacity } * \text{ SOH} \tag{4}$$

$$\text{Remaining Power train} = \text{SOC} * \text{ nominal capacity } * \text{ SOH} \tag{5}$$

$$\text{SoF} = \begin{cases} 1 \text{ if storage power is with in the limits} \\ 0 \text{ if storage power is not with in the limits} \end{cases} \tag{6}$$

4 Simulation Results

The proposed battery management system for EVs is simulated using MATLAB/Simulink, and the results of the system is shown in Fig. 5.

From the results of the system, it can be observed that at time $t = 0$ s, switches 1 are closed. Therefore, the battery will charge up to $t = 20$ s. From 20 to 50 s, the charge on the battery is constant, and the storage charge is within the limits. In this condition, the power is supplied by the battery to the load through switches II. If the storage charge on the battery is exceeding its limit, i.e., 80%, the battery will overheat. Because of the overheating, the battery may get damaged, which leads to a fire accident in the vehicle. In order to protect the battery, the switches are open, and the battery is isolated from the supply.

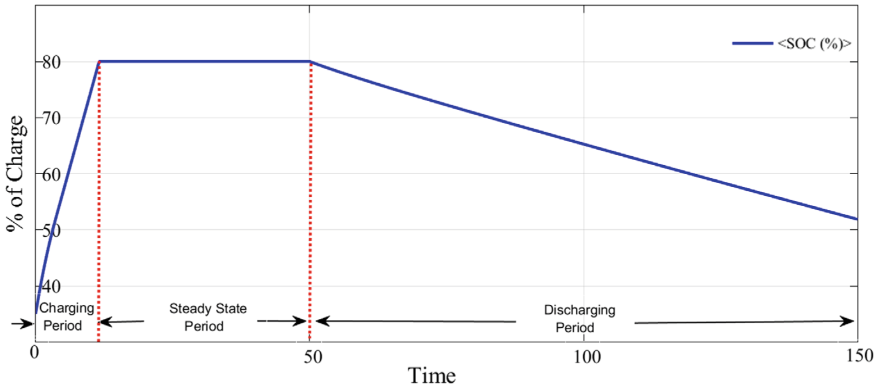


Fig. 5 Simulation result of the battery management system

5 Conclusion

In EVs, the battery is the major part, and its cost is also higher. For the effective utilization of storage charges and battery monitoring, a suitable system is required. In this paper, a battery management system is proposed and simulated by MATLAB/ Simulink. From the results, it can be observed that when the charge in the battery is greater than 80% of the full charging capacity, even though the charging switch is closed, the battery cannot be charged because the electronic switch will be open. Therefore, the battery will not get overheated. Similarly, if the charge in the battery is less than 20% of the rated charge, to avoid battery failure, it is isolated from the battery. In this procedure, the proposed method protects the battery from battery death due to low charge and damage to the battery from overheating.

References

1. Hou Z-Y, Lou P-Y, Wang C-C (2017) State of charge, state of health, and state of function monitoring for EV BMS. In: 2017 IEEE International conference on consumer electronics (ICCE), Las Vegas, NV, USA, pp 310–311. https://doi.org/10.1109/ICCE.2017.7889332
2. Mi CC, Buja G, Choi SY, Rim CT (2016) Modern advances in wireless power transfer systems for roadway powered electric vehicles. IEEE Trans Ind Electron 63(10):6533–6545
3. Priyanka, Dhuliya P, Kathait SS, Semwal S, Joshi T, Pathela M (2022) Electric vehicles (EV's): a brief review. In: 2022 International conference on advances in computing, communication and materials (ICACCM), Dehradun, India, pp 1–3, https://doi.org/10.1109/ICACCM56405.2022.10009125
4. Cao Z, Mahmoudi A, Kahourzade S, Soong WL (2021) An overview of electric motors for electric vehicles. In: 2021 31st Australasian universities power engineering conference (AUPEC), Perth, Australia, pp 1–6. https://doi.org/10.1109/AUPEC52110.2021.9597739
5. Naguib M, Kollmeyer P, Emadi A (2021) Lithium-ion battery pack robust state of charge estimation, cell inconsistency, and balancing: review. IEEE Access 9:50570–50582. https://doi.org/10.1109/ACCESS.2021.3068776

6. Guo Y, Wang L, Zhu Q, Liao C, Li F (2016) Applying LCC compensation network to dynamic wireless EV charging system. IEEE Trans Ind Electron 63(10):6557–6567
7. Guo Y, Wang L, Zhu Q, Liao C, Li F (2016) Switch-on modeling and analysis of dynamic wireless charging system used for electric vehicles. IEEE Trans Ind Electron 63(10):6568–6579
8. Satheesh S, Mathavan L (2020) Transient analysis of an interconnected power system incorporating smes along with electric vehicle utilizing optimized controller tuning
9. Jafari S, Byun Y-C (2022) Prediction of the battery state using the digital twin framework based on the battery management system. IEEE Access 10:124685–124696. https://doi.org/10.1109/ACCESS.2022.3225093
10. Fernández G et al (2018) EV charging infrastructure in a petrol station, lessons learned. In: 2018 International symposium on industrial electronics (INDEL). IEEE
11. Shigeta N, Hosseini SE (2020) Sustainable development of the automobile industry in the United States, Europe, and Japan with special focus on the vehicles' power sources. Energies 14:78
12. Dalvi S, Thale S (2020) Design of DSP controlled passive cell balancing network based battery management system for EV application. In: 2020 IEEE India council international subsections conference (INDISCON), Visakhapatnam, India, pp 84–89, https://doi.org/10.1109/INDISCON50162.2020.00029
13. Manikandan BV, Kiruthickroshan P, Kumar MV, Chandrasekeran L (2023) Battery management system with charge monitor and fire protection for electrical drive. E3S Web Conf (ICSERET-2023) 387, Article no. 01002. https://doi.org/10.1051/e3sconf/202338701002
14. Sai Durga Prasad K, Kushal GC, Harijan BD, Basavaraj MS, Sowmya G (2023) EV BMS with charge monitor and fire protection. Int J Creative Res Thoughts (IJCRT) 11(5). ISSN: 2320-2882
15. Prakash A, More UK, Kushwaha S, Gholap AB, Muley K (2022) EV battery protection system. Int Res J Eng Technol (IRJET) 09(06). e-ISSN: 2395-0056
16. Hariprasad A, Priyanka I, Sandeep R, Ravi V, Shekar O (2020) Battery management system in electric vehicle. Int J En Res Technol (IJERT) 9(05). ISSN: 2278-0181. https://doi.org/10.17577/ijertv9is050458
17. Barelli L, Bidini G, Ottaviano PA, Pelosi D, Perla M, Trombetti L, Gallorini F, Serangeli M (2021) Electric vehicles fire protection during charge operation through Vanadium-air flow battery technology. Heliyon 7(9):e08064. https://doi.org/10.1016/j.heliyon.2021.e08064
18. Shinde R, Aglave D, Kamble P (2023) EV BMS with charger monitor and fire protection. Int J Res Publ Rev 4(5):4354–4356

UART Implementation for RFID Modules with Various FPGA Technologies

S. Fayaz Begum, Shaik Karimullah, Fahimuddin Shaik, Syed Javeed Basha, and P. Syamala Devi

Abstract The serial communication protocols include the Universal Asynchronous Receiver Transmitter (UART) standard**.** It is an integrated circuit that is used to asynchronously transmit and receive data over the computer's serial port. Bytes of data are taken and transmitted by the UART as sequential bits. The second UART resembles the bit into whole bytes at the destination. UART is primarily used to transmit data cheaply, quickly, and over small distances among two embedded devices. This article's primary focus is on the UART implementation for RFID modules with various FPGA technologies. Transmitter, Receiver, and Baud Rate Generator are the three main components of the design. UART protocol is now frequently used by RF-Microcontroller modules to connect to other electronics. On many FPGAs, a simplified Finite State Machine (FSM) UART design has been put into practice. The design is made more stable and dependable using FSM. With its RF components, UART offers extremely reliable long-distance transmission. Verilog HDL is utilized to implement the suggested design, and Xilinx ISE 12.1, 14.7, Xilinx Vivado 2018 and Intel Quartus Prime 19.1 are used for simulation. These FPGAs are used to determine the highest speed that may be achieved and the associated power consumption.

Keywords Speed · Power consumption · FPGA · FSM · And UART

1 Introduction

A serial communication system called Universal Asynchronous Receiver Transmitter (UART) requires few transmission lines and offers excellent stability for long distances. Asynchronous transmission was employed by the UART. Asynchronous transmission refers to the transmission of groups of bits without data connection synchronization as independent units with a start/stop flag [1]. The main element of a computer's serial communication subsystem is the UART controller. Bytes of data

S. F. Begum · S. Karimullah (✉) · F. Shaik · S. J. Basha · P. S. Devi
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: munnu483@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_35

are taken and transmitted by the UART as sequential bits. A second UART reassembles the bit into full bytes at the destination. 16 bytes can be saved in both receive and transmit modes when internal FIFOs are activated in this mode [2]. Additionally, it is now widely employed by embedded systems with RF microcontrollers and computer peripherals. Full-duplex mode serial communication is available with UART. A single UART chip just needs a few aspects of the UART protocol to transmit data. The main characteristics of a hypothetical UART device are described in this paper, along with how they would be implemented on an FPGA [3]. The Finite State Machine (FSM) is utilized to increase the stability and dependability of this system. The three main components of this design are the transmitter, receiver and clock signal for generating the baud rate. Field programmable gate array is a type of semiconductor device that can be configured to operate in accordance with a predetermined HDL design using the Hardware Description Language (HDL). The "logic blocks" that make up FPGAs' programmable logic components are organized hierarchically and contain adjustable interconnects to make it easier to connect the blocks [4]. The UART data frame is seen in Fig. 1. The fundamental UART design transmits data in full duplex over just two wires (TX, RX). The receiver's input is called RX, and the transmitter's output is called TX. These lines might be in either a HIGH or a low state. Transmitter, or TX line, remains at HIGH state in an idle state of 2. A "Start Bit" is an additional bit that is inserted to the beginning of a word during transmission [5]. The receiver is informed by this "Start Bit" that the transmitter has begun data transmission. When all the data bits have been received, when the data transmission is over, a "Stop Bit" is sent to the receiver at the HIGH logic level.

Start bit—It informs the receiver that a new character is about to appear.

Data bit—The character is represented by the following five to nine bits, depending on the code set used.

Stop bit—The next one or two bits are referred to as the stop bit since they are always in the mark (logic high, or "1" condition) state (s). They let the receiver know when the character is finished. There are always at least two guaranteed signal changes between characters because the start bit is logic low (0) and the stop bit is logic high (1).

The UART can identify a break condition if the line is maintained in the logic low condition for longer than a character time.

The receiver has received all bits, it discards the "start bit" and "stop bit" and changes the serial data's format to that of parallel data. The receiver and transmitter

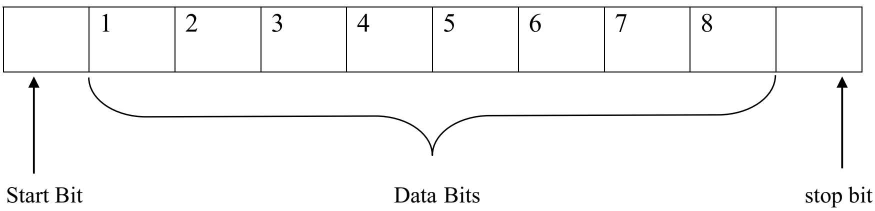


Fig. 1 UART data frame

are synchronized using a “baud rate generator.“ The hardware that creates the clock frequency for the transmitter and receiver at a particular baud rate is a straightforward frequency divider. Standard 9600bps baud rate is applied in our design [6].

Need and Significance

The major application for UART is the low speed, short distance and inexpensive data transmission among two embedded devices. Remote meter reading, access control systems, wireless data communications, industrial data acquisition, wireless remote control telemetry, security systems, robot control and other applications all regularly use UART RF modules [7].

Baud Rate

The number of data bits—0 s and 1 s—transmitted in a second is known as the bit rate. By definition, baud rate refers to how frequently a signal in a communications channel changes state. Baud stands for line state changes per second. The number of signal or symbol changes per second is referred to as the baud rate. One of numerous voltage, frequency, or phase shifts is a symbol. Two symbols in NRZ binary denote the voltage level. Baud rate displays how many bits are actually being sent over the media, as opposed to monitoring how much data is being transported from one UART device to another. The overhead Start, Stop, and Parity bits that are created by the sending UART and eliminated by the receiving UART are included in the baud count. As a result, the transmission of seven-bit words of data actually requires 10 bits.

$$\text{Baud Rate} = \text{No of Bits Transmitted/Receiver Per Second}$$

The transmitter and receiver’s baud rates are generated using a baud rate generator. Not necessary for reads or writes or for any other purpose. Crystal or an external clock DLM/DLL registers with a 16-bit divisor are programmed. To write data to the transmit holding register in transmitter FIFO mode, the TXFIFO transmit data queue is used. When the transmit shift register (TSR) is empty, data from the TXFIFO is moved to the TSR. On the TX output pin, TSR shifts the data. Incoming data from the receiver is queued in the RX FIFO and is received in the receiver shift register (RSR) [8].

2 Literature Review

In order to solve testability and data integrity, this study presents a novel strategy using a status register and an 8-bit UART. The entire design is implemented in VHDL, completely simulated in Modalism, and synthesized using Xilinx ISE 14.5 software [9]. The previous study uses the VHDL description language to demonstrate the UART using approach for implementation of UART using FIFO. The UART protocol, or serial communication method, is used to communicate information or data in serial form. A special computer chip called a universal asynchronous receiver transmitter serves as the interface between the serial port transmission and the parallel computer bus transmission [10].

3 Methodology

3.1 Existing Method

Asynchronous transmission was used by the UART. Asynchronous transmission refers to the transmission of groups of bits without data connection synchronization as independent units with a start/stop flag. The main element of a computer's serial communication subsystem is the UART controller. Bytes of data are taken and transmitted by the UART as sequential bits. A second UART reassembles the bit into full bytes at the destination. To reduce system overhead and increase system efficiency, all the logic is housed on a single chip. The functionalities of two pins have been modified to provide DMA transfer signaling (Fig. 2).

The UART converts data characters transferred from serial to parallel from a peripheral device or a MODEM, The characters which are received from CPU is then converted into serial. Throughout functional operation, the CPU has access to the UART's whole status at any time. The kind and state of the transfer operations being carried out by the UART, as well as any error situations, are included in the status information given (parity, overrun, framing, or break interrupt). The internal transmitter logic is driven by a 16 c clock produced by the programmable baud rate generator in the UART. Additionally, provisions are made so that the receiver logic can be driven by this 16 c clock. The UART has full MODEM control

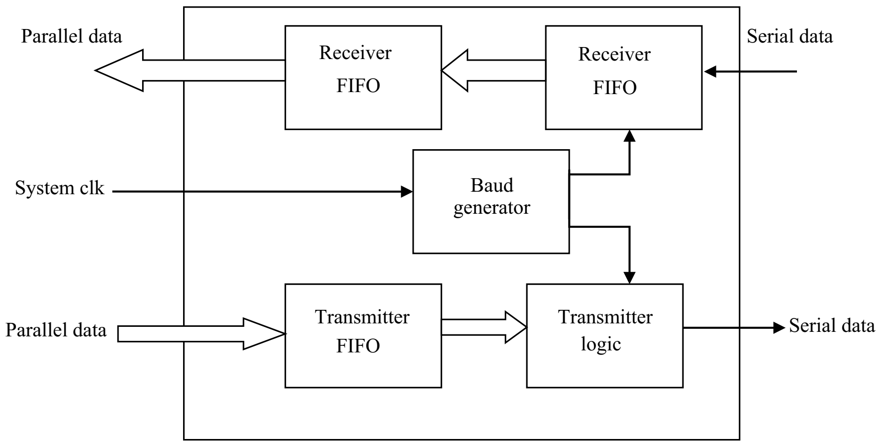


Fig. 2 Configuration of UART with FIFO

capabilities in addition to a processor-interrupt system. Programming interrupts to meet the demands of the user can reduce the computation required to manage the communications channel.

3.2 *Proposed Method*

UART typically just has two wires, does not require a clock signal, and has a parity bit for error checking in the UART data frame. The size of data frame is restricted above 9 bits, and the buad rate of each UART must be within 10% of the other UART's buad rates. However, UART does not support multiple slave or multiple master systems. We are employing FIFO and RFID to solve the current issue.

The suggested approach is highly adaptable to various chips and communication protocols thanks to its customizable structure.

These modules function as follows.

- The RFID reader module reads the data from the RFID tag when it is close to it.
- Uses the UART communication protocol to send data in a serial format.
- A UART receiver sends bytes in byte format to the data management module after reading the data in serial form.
- The write module receives data from the Data Management module and verifies it.
- FIFO means First In First Out.
- Data will be read using the read controller module after being written into the FIFO. The data is transmitted in serial form using a UART transmitter module (Fig. 3).

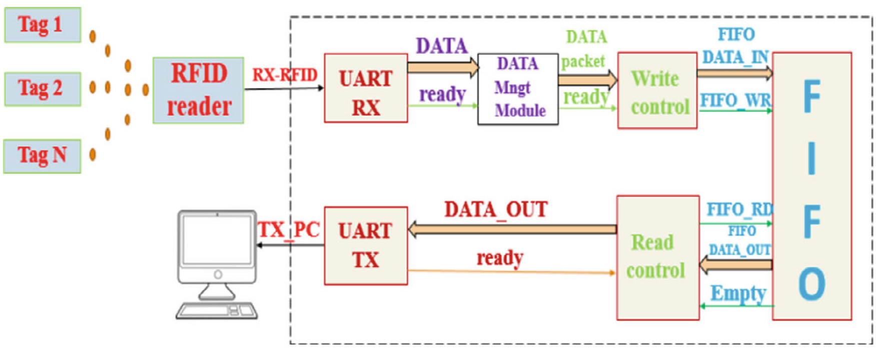


Fig. 3 Block diagram for RFID

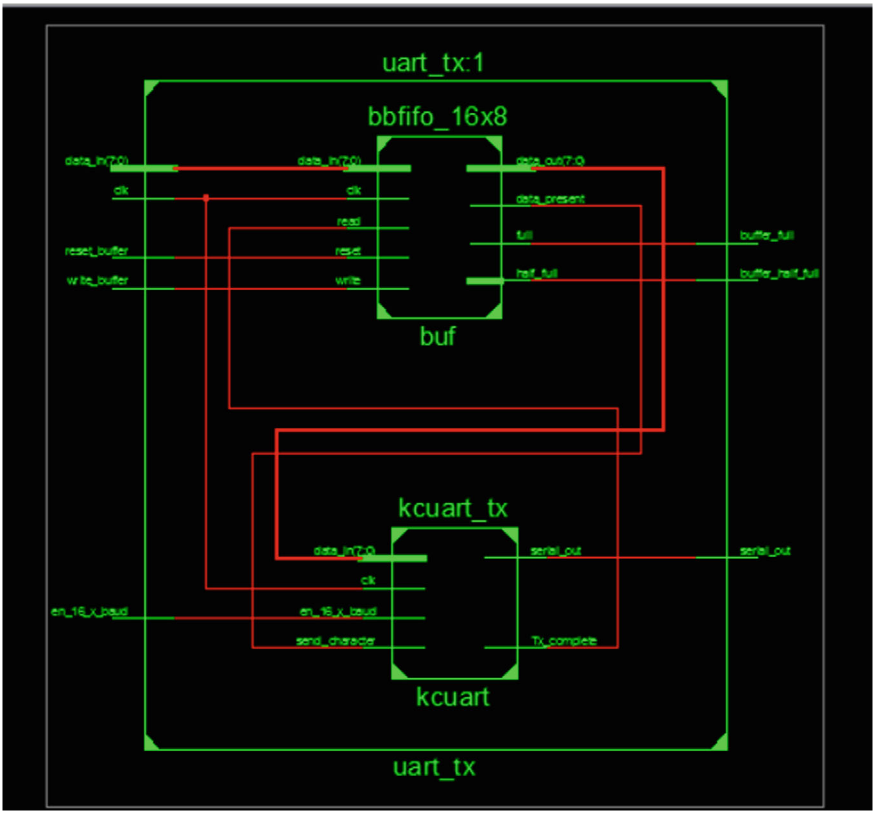


Fig. 4 Block of UART_TX

3.2.1 Uart_tx

UART TX is a mix of the constant (k) compact UART transmitter and the FIFO buffer "bbfifo_16*8" modules. Byte data for transmission is accepted in the FIFO buffer. The "kcuart_tx" transmitter is in charge of the FIFO's output side. To send the data to the serial line, the "kcuart_tx" circuit automatically reads the buffer. According to Fig. 4, there are numerous input ports and output ports in the UART_TX module. Ports on the Spartan 3E will be coupled with ports on UART_TX.

3.2.2 Uart_rx

UART RX is a mix of the constant (*k*) compact UART receiver and the FIFO buffer "bbfifo_16*8" modules. The device will store the data in its FIFO buffer. The "kcuart_rx" receiver is in charge of the FIFO's input side. The UART_RX module, as seen

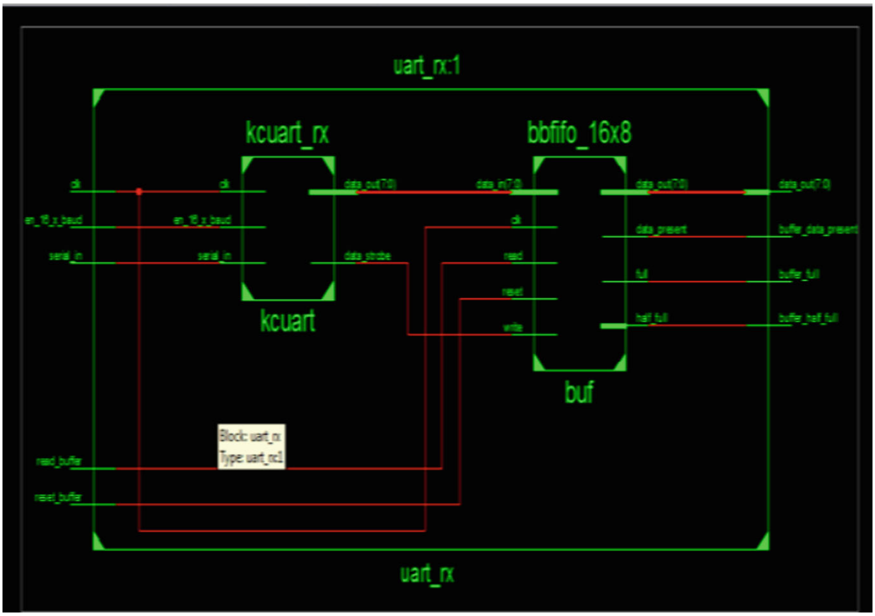


Fig. 5 Block of UART_RX

in Fig. 5, includes a number of input ports and output ports. Ports on the Spartan 3E will be coupled with ports on UART_RX.

4 Results and Analysis

The simulation software is XILINXA 14.7 and selected device is Vivado 2018.

4.1 Transmitter Module Simulation:

According to the simulation report, the timing diagram shown has an 8-bit start bit, an 8-bit input data bit, and a stop bit, for a total of 10 bits of data in the transmitter, which converts parallel data into serial data. The data input here is 1001011. When the start bit is logic low, the transmission begins. The transmission is stopped by passing an active high signal on the stop bit. The below Fig. 6 is the simulation of UART transmitter.

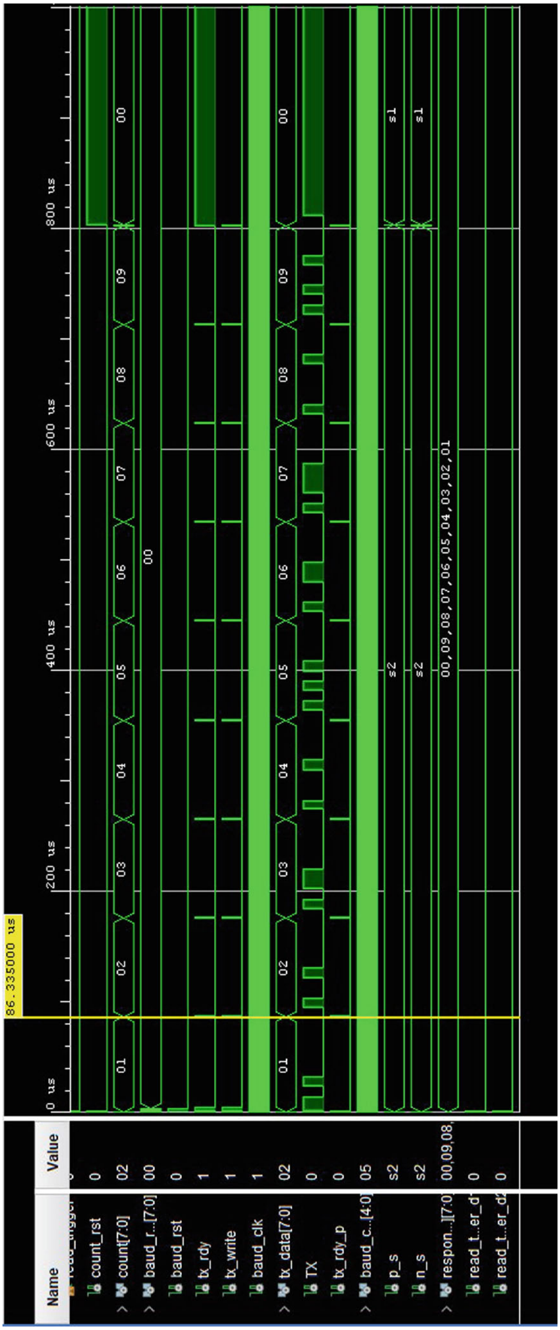


Fig. 6 Simulation of UART transmitter

4.2 Receiver Module Simulation:

In order to determine if the data from the transmitter is accurate or not, the receiver converts parallel data into serial data. The Fig. 7 is the simulation of UART receiver.

5 Conclusion

We proposed a UART implementation for RF modules in this work. Its core components include transmitters, receivers, baud rate generators, and an asynchronous FIFO. These FPGAs are used to determine the highest speed that may be achieved and the associated power consumption. Using ISE 14.7 from Xilinx, the design is successfully synthesized and simulated. The results are consistent and dependable, demonstrating proper functionality.

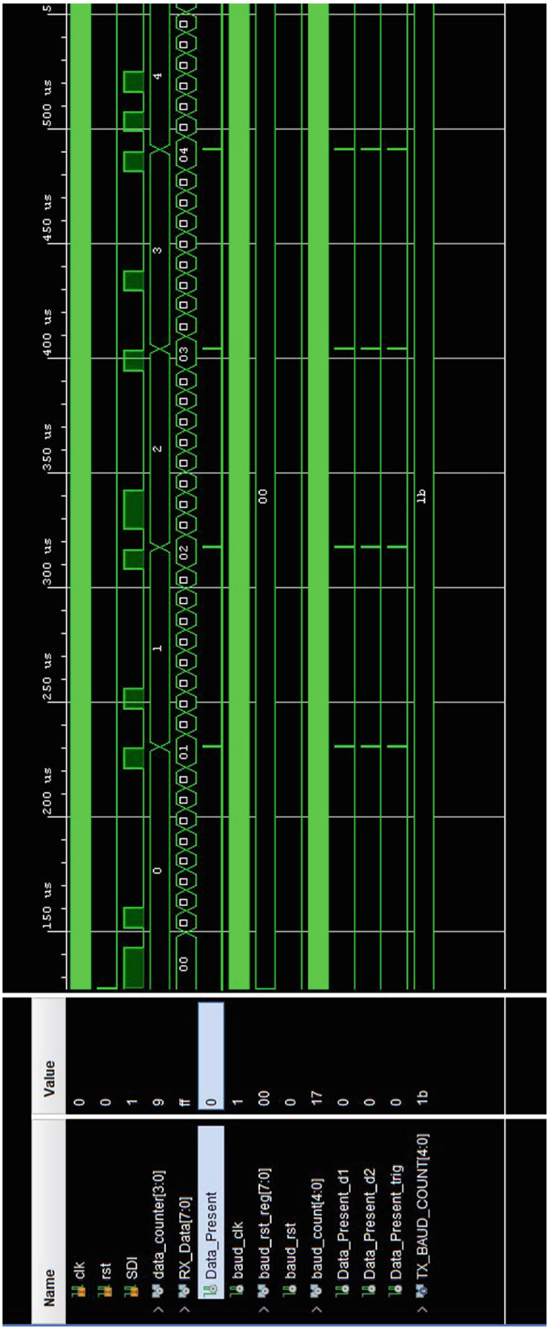


Fig. 7 Simulation of receiver

References

1. Anu Priya D, Vishnuvardhan D, Mamatha G, Shaik F, Karimullah S (2023) VIP development of SPI controller for open-power processor-based fabless SOC. In: Kumar A, Ghinea G, Merugu S (eds) Proceedings of the 2nd international conference on cognitive and intelligent computing. ICCIC 2022. Cognitive Science and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_2
2. Karimullah S, Vishnuvardhan D, Bhaskar V (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Pers Commun 127:3041–3059. https://doi.org/10.1007/s11277-022-09909-2
3. Krishna SLV, Abdul Rahim B, ShaikF, Soundara Rajan K (2010) Lossless embedding using pixel differences and histogram shifting technique. In: Recent advances in space technology services and climate change 2010 (RSTS & CC-2010). Chennai, India, pp 213-216. https://doi.org/10.1109/RSTSCC.2010.5712850
4. Karimullah S, Sai Sumanth Goud E, Lava Kumar Reddy K (2023) Spectral efficiency for multi-bit and blind medium estimation of DCO-OFDM used vehicular visible light communication. In: Kumar A, Senatore S, Gunjan VK (eds) ICDSMLA 2021. Lecture notes in electrical engineering, vol 947. Springer, Singapore. https://doi.org/10.1007/978-981-19-5936-3_83
5. Malviya A, Sharma VK (2016) An improved approach of UART implementation in VHDL using Status Register. Int J Digit Appl Contemp Res Indore, India
6. Kumar A, Shaik F, Rahim BA, Kumar DS Signal and image processing in medical applications. Springer, Heidelberg
7. Longadge R, Mahindra M, Dahake V (2015) Design and implementation of UART Int J Adv Res Comput Commun Eng Nagpur, India
8. Shaik F, Giri Prasad MN, Rao J, Abdul Rahim B, SomaSekhar (2010) A Medical image analysis of electron micrographs in diabetic patients using contrast enhancement. In: 2010 International conference on mechanical and electrical technology. Singapore, pp 482–485. https://doi.org/10.1109/ICMET.2010.5598408
9. Karimullah S, Lalitha M, Mounika, Mahesh Reddy S, Mahammad Waseem S, Irfan SM (2023). Performance enhancement of wireless systems using hybrid RIS technique. In: Proceedings of the 4th international conference on data science, machine learning and applications. ICDSMLA 2022. Lecture notes in electrical engineering, vol 1038. Springer, Singapore. https://doi.org/10.1007/978-981-99-2058-7_17
10. Shaik F, Sharma AK, Ahmed SM (2016) Detection and analysis of diabetic myonecrosis using an improved hybrid image processing model. In: 2016 2nd international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB). Chennai, India, pp 314–317. https://doi.org/10.1109/AEEICB.2016.7538298

Automatic Segmentation of Retinal Blood Vessel Using the Homomorphic Filter and Multilevel Algorithm

G. Obulesu, M. Ravi Kishore, Fahimuddin Shaik, Shaik Karimullah, and CH. Naga Raju

Abstract In the automatic diagnosis of numerous ophthalmic and cardiovascular illnesses, segmentation of retinal blood vessels plays a significant role. Medical analysis and the detection of associated disorders depend on the ability to segment thin and thick retinal vessels. The fundus image of the eye exhibits variable vascular thickness, the latter of which is not segregated by the majority of the previously suggested algorithms. This technique aims to distinguish thick and thin vessels in order to produce two images individually, which can then be combined to provide the final segmentation result. The process includes a number of steps. In addition to applying filters for input retinal image smoothing, primary processing entails many stages. The first one uses an updated top-hat, homomorphic filter to segment large vessels, while the other one presents an optimal top-hat, homomorphic filtering, gray-level conversion, matching filter, and segmentation using multilevel algorithm to segment narrow vessels. Morphological image operations are lastly performed. With this suggestion, a minimal number of false positives are obtained. This technique effectively divides the vessels and raises the performance standards. Implementation part was done in MATLAB R2022b. DRIVE datasets were used to evaluate the suggested method.

Keywords Blood vessels · Fundus image · Optimized top-hat · Multilevel algorithm

1 Introduction

One of the most delicate human body organ is eye, which allows us to see everything around us. It is said to be the most sophisticated and complicated organ in the human body despite its small size. The three primary components of the human eye are the

G. Obulesu · M. Ravi Kishore · F. Shaik · S. Karimullah (✉) · CH. Naga Raju
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: Munnu483@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_36

retinal blood vessels, optic nerve, and macula [1]. Eye doctors as well as other health-care providers frequently be using the retinal fundus inspection as conventional diagnostic procedure for the detection, identification, and providing systematic of ocular illnesses such diabetes mellitus, Macula development, and others. [2]. Retina is only organ in human body where blood microcirculation can be observed non-invasively for the detection of systemic diseases. Subdivision of the neovascularization tree is thus extremely pertinent to the area of medical cinematography.

The scientific community continues to be very interested in this task due to the increasingly accurate results that have been obtained over the past few decades using a variety of automated techniques for segmenting the retinal vessels using the fundus examination [3]. Due to the time-consuming and expensive nature of manual blood vessel extraction, several automated segmentation approaches are used [4]. Both supervised and unsupervised automated techniques are categorized. Unsupervised methods have the advantage of requiring fewer computations and not needing manually annotated images to train the model. These include threshold-based approach, model-based methods, active contour methods, mathematical morphology, region growth, multiscale approach, vessel tracing, and matching filter. A feature detection and selection stage is followed by a classification stage in supervised approaches other than deep learning models. The classifier divides the fundus picture into vessel and non-vessel sections using binary classification [5]. For the majority of the works mentioned in the literature, the difficulty of segmenting thin vessels still poses a barrier to their full potential. In order to detect neovascular disorders, this thin vasculature provides information that is crucial [6]. Therefore, it is essential to be able to segment these veins more effectively in order to detect and diagnose eye problems.

This article describes a precise approach for segmenting the retinal vascular in eye fundus image. The suggested method is broken up into several stages. In the beginning, Gaussian filter was used to generate a standardized gray scale photograph of the fundus. Then, it uses a variety of filters, including the homomorphic and the optimized top-hat filter, to segment thicker vessels, while another step uses multilayer segmentation to segment thinner vessels. The proposed technique also makes use of morphological image processing.

The following is summary of this article's key contributions:

- An original approach is proposed for the identification of both thin and large retinal capillaries.
- The optimal top-hat is offered as a fresh variation on the traditional top-hat procedure.
- This idea yields the fewest possible false positives.
- This methodology competes with classification techniques while exceeding the most current unregulated algorithms in terms of specificity and accuracy, and it comes with a limited processing cost.

2 Literature Survey

In the recent years, a lot of scholars have published their work in this field. Over the years, research into and examination of retinal blood vessels has become one of fundamental methods for identifying various eye illnesses. Different methods for extracting retinal blood from a fundus picture have been proposed, helping doctors identify the condition and identify various eye disorders early on.

A vessel segmentation using a Multi-Threshold and Morphological Operation offered by [1], this approach uses AHE, CLAHE, and average filtering also employs morphological operations to eliminate any flaws. This method is quite expensive. MATLAB 2005a was used to evaluate the effectiveness of the DRIVE database.

Another method for automatically segmenting the blood arteries from retinal images has been suggested by [3], two deep neural networks have been proposed: Unet, which uses EfficientNet as its backbone, and EfficientNet encoder and LinkNet decoder. The pre-processing phases used are gamma adjustment and contrast limiting histogram equalization. Results are assessed using benchmark fundus image datasets as DRIVE, STARE, HRF, and CHASE DB1. Performance evaluation criteria produce poor results.

In the paper [5], discussed an innovative technique for segmenting retinal pictures. In this machine assessment, retinal fundus images splitting is done in order to retrieve pictures of retinal blood vessels. FCM and the morphological opening procedure are coupled to segment individual vessels. CLAHE strengthens the blood vessels, which causes the retinal image to become more streamlined. After employing FCM to segment the retinal picture, clusters are created. Segmentation has been done using photos from DRIVE databases. Low accuracy has been found in the results on many factors related to the segmentation of retinal pictures.

In their study, [6] proposed a novel deep learning approach for segmenting and centerline extraction of retinal blood vessels that integrated the capsule network with the CNN architectures. They performed a quantitative analysis for centerline extraction and vessel segmentation using the DRIVE dataset. Capsule networks require less training data and have fewer parameters.

There are supervised and unsupervised versions of the previously suggested algorithms [7]. These strategies employ a number of multi threshold techniques as well as machine learning and deep learning algorithms [8]. The majority of studies described in literature are still constrained their ability to operate at their best by the difficulty of segmenting narrow vessels [9]. The identification and treatment of vision problems therefore rely on achieving a better segmentation of these arteries. This article offers a precise approach for segmenting the retinal vasculature in order to deal with the bulk of the problems [3].

3 Methodology

This article presents the key ideas behind the proposed technique which is depicted in Fig. 1. This method's objective is to distinguish thick and thin vessels, which will result in the generation of two independent images, which will then be combined to get the final segmentation result.

3.1 *Feature Extraction (Green Channel)*

Initially, the image's green channel is extracted during pre-processing. Usually, there is not much contrast in retinal images. Micro-aneurysms can easily be seen due to the strong contrast in the green channel. It is a technique for separating an RGB color image into each of its individual pixel channels or color space components. A gray scale image is ultimately extracted.

3.2 *Image Pre-processing*

Pre-processing must be applied to the macular retinal fundus image shown in Fig. 2a before the thick and thin vessels processing stages are carried since the results are immediately and significantly impacted. In this stage, filtering are employed in order create an ocular smooth and efficient picture.

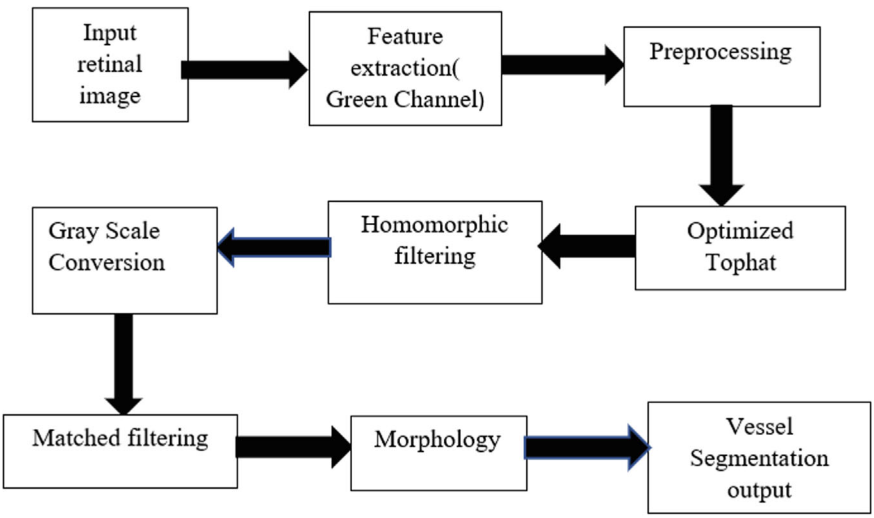
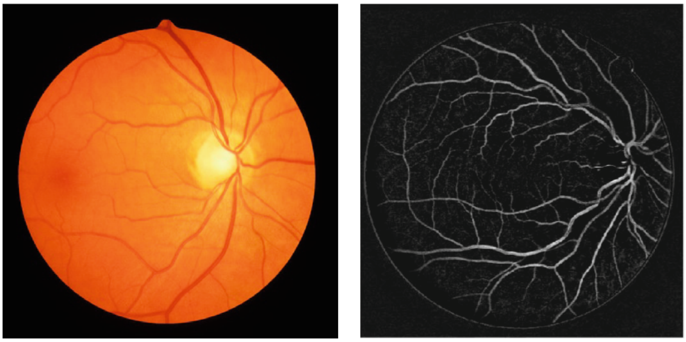


Fig. 1 Flowchart for the proposed vessel segmentation technique



(a) Eye fundus image (b) Optimised top-hat

Fig. 2 Variation of top-hat algorithm

3.3 Optimized Top-Hat

The primary processing stage's usage of top-hat transformation improves the algorithm's ability to extract blood vessels in the presence of features like fovea, exudates, etc. This approach suggested a new, improved top-hat methodology that resolves several issues with the traditional top-hat operation. After being implemented to the input fundus picture in Fig. 2a, the output of the recommended optimal top-hat filtration is shown in Fig. 2b. Even before improved top-hat and the morphological image procedures are reversed. On complementary image, the top-hat operator has been used. Here is how the suggested optimal top-hat is defined:

$$\mathrm{TO} = (\mathrm{Ic\ So})\ \mathrm{Ic} \bullet \mathrm{Sc}$$

where So and Sc are the structure aspects for the openness () and closure (•) drivers, respectively, and Ic is the picture perfect compliment of I.

$$\mathrm{Ic} = U \backslash I$$

where U denotes each of the possibilities for each pixel.

3.4 Homomorphic Filtering

It is possible to alter brightness and reflection strengths of picture over a spectrum domain by utilizing frequency domain filtering technology known homomorphic filtration, enabling the use of various frequency domain filters such as high pass filters for a variety of analyses. By reducing the light module's involvement to the segmentation, it is often essential to enhance the luminance from the source images.

In order to solve the problem, an effective terms filter is used, which weakens the low-frequency and strengthens the held fully.

3.5 Gray Scale Conversion

Despite the fact that contrary to the RGB source image, the green channel of such retinal fundus picture has a strong vessel backdrop brightness. Applying a noise removal step before proceeding to the next processes is advantageous and beneficial. Improving visual components, like formerly contrasting parts, is possible with the help of the Gaussian smoothing filtering technique. Finally, the outcome is gray scale image.

3.6 Matched Filtering

Matching filters are intended to enhance picture portions that suit a certain allocation. As a consequence, this sorting analysis plans for a particular level of correlation between both the local image area and the specific dispersion. When utilizing this method, the target is to spot in gray scale measurements of such pass of capillaries in retinal images simple linear blood vessel segments whose characteristics are anticipated by the Edge model. Usage of method reduces the risk of faulty vascular detection, hence minimizing the false positive rate.

3.7 Multilevel Algorithm

The coincident filtering output is next subjected to the multilevel algorithm approach, which is multilayer result in an additional used to streamline the pixel integrity in the image's components and make it simple to spot related sections. This process includes using meta-heuristic procedure to reduce defined criteria of cross-entropy as an optimal solution in order to find the optimal threshold values. For the multilayer technique, the minimum thresholding is determined in order to get an ultimate algorithm relies on the vector, which comprises the various threshold being used divide the image into sections.

This method is used till the end condition is satisfied and the function is evaluated till the minimal value is generated by a certain set of thresholds. For best separation of the thin vessels, the HHO settings employed are 250 reiterations as cutoff criteria, 30 hawks for the output generation. The picture from the four thresholds is quantized in order to segment thin capillaries.

3.8 Morphology

A group of nonlinear processes focused on the morphology or shape of features in an image make up morphological image processing. Morphological procedures are particularly well adapted to the processing of binary pictures since they only depend on the relative ordering of pixel values rather than their numerical values. Measurements of the retinal vasculature's size, length, breadth, branching, and angles can be made by segmenting the retinal arteries.

3.9 Vessel Segmentation Output

Finally, post-processing is performed to eliminate minute undesired components and noise produced by the earlier stages. To build an ultimate binary image which can then be evaluated, this stage involves feature extraction procedures like a connection test where binary elements less than a given pixel resolution are removed. After combining the large and small vascular classification in binary sequence using a binary OR, a performance of the proposed method image is produced that exhibits the results for parts of flowchart in Fig. 1.

4 Results

An 8 GB RAM device is used to do the complete analysis. An Intel(R) Core i5 @ 2.40 GHz CPU is used, while Windows 10 is the operating system. It is a 64-bit operating system. The entire study is carried out using the most recent, more advanced version of MATLAB R2022b. In this, MATLAB technology like communication tools are used.

The results obtained of the suggested approach used to segment retinal vessels over two eye fundus datasets are shown in this section using the DRIVE database. The image below displays the finalized images after the application of the novel suggested method for dividing up retinal image (Figs. 3 and 4).

Phrases that are most often used to describe a binary classification test are sensitivity, specificity, and accuracy. These terms statistically evaluate the test's effectiveness. The suggested method's average accuracy, sensitivity, specificity and code compilation time with are 0.9871, 0.7477, and 0.9789, 1.65 s respectively. The previously suggested supervised methods generated parameter results with specificity = 0.8823, sensitivity = 0.6234, accuracy = 0.8322 and compilation time of 2.75 s. The outcomes of the unsupervised algorithms are specificity = 0.9231, sensitivity = 0.7123, accuracy = 0.8963, and the software generation time is 2.23 s. The outcomes of various approaches are displayed in the table.

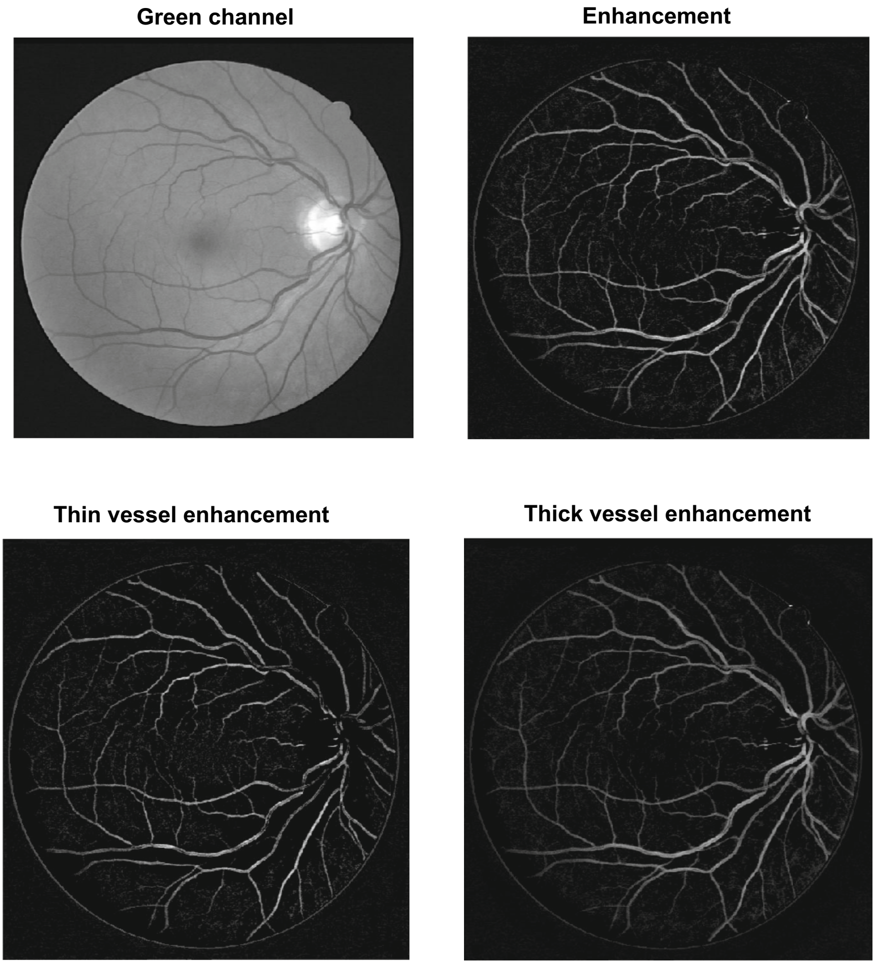


Fig. 3 Resultant pictures of proposed methodology

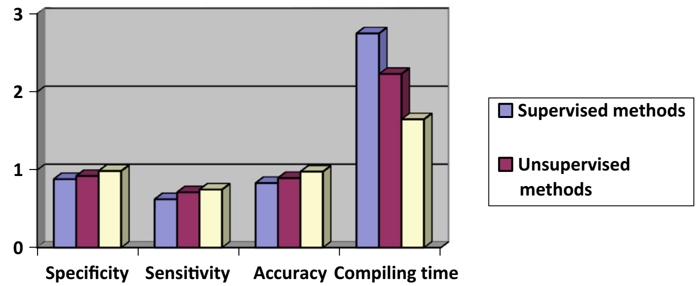


Fig. 4 Comparison of several characteristics for each technique

Table 1 Outcomes of implementing the proposed approach on the DRIVE dataset

	DRIVE dataset			
Methods	Specificity	Sensitivity	Accuracy	Compiling time
Supervised methods	0.8823	0.6234	0.8322	2.75
Unsupervised methods	0.9231	0.7123	0.8963	2.23
Proposed method	0.9871	0.7477	0.9789	1.65

The specificity of supervised techniques is lower than that of unsupervised methods, and the specificity of the suggested approach is higher than that of those earlier methods. Both supervised and unsupervised approaches produced results with lower sensitivity when compared to the novel recommended methodology. The suggested approach is more accurate than others. In comparison with earlier techniques, the suggested solution requires a lot less time to execute the code (Table 1).

5 Conclusion

The segmentation approach for both thin and thick artery detection serves as the fundamental basis for an unique methodology that improves the identification of endoscopic picture of the retinal arterial system of human eye. The recommended technique demonstrates high-specificity performance for having to learn or conditioning operations that demand outstanding performance without the necessity for a segmentation approach. The key component of the proposed framework is the parameter adjustment of the homomorphic and optimized top-hat filtering stages, which is dependent on the segmentation outcome, in this instance thin and thick vessels. To improve segmentation accuracy and specificity, one must have this functionality. The major weaknesses of the idea is the low sensitivity attained in contrast to the levels provided in the condition.

Finally, by separating the region of interest, it is anticipated to implement this technology as a post-phase in a reliable system in both healthy and diseased ocular fundi for illness diagnosis. This technique was thought about when this concept was being developed. A high categorization specificity value or even a small false positive value is required for achieving this.

References

1. Preity, Jayanthi N (2020) A segmentation technique of retinal blood vessels using multi-threshold and morphological operations. In: 2020 International conference on computational performance evaluation (ComPE), pp 447–452 https://doi.org/10.1109/ComPE49325.2020.9200042

2. Shaik F, Sharma AK, Ahmed SM (2016) Detection and analysis of Diabetic Myonecrosis using an improved hybrid image processing model. In: 2016 2nd International conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB). Chennai, India, pp 314–317 https://doi.org/10.1109/AEEICB.2016.7538298
3. Karimullah S, Vishnuvardhan D, Bhaskar V (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Pers Commun 127:3041–3059. https://doi.org/10.1007/s11277-022-09909-2
4. Ramos-Soto O, Rodríguez-Esparza E, Balderas-Mata SE, Oliva D, Hassanien AE, Meleppat RK, Zawadzki RJ (2021) An efficient retinal blood vessel segmentation in eye fundus images by using optimized top-hat and homomorphic filtering. Comput Methods Programs Biomed 201:105949. https://doi.org/10.1016/j.cmpb.2021.105949. Epub 2021 Jan 27. PMID: 33567382
5. Mathews MR, Anzar SM, Kalesh Krishnan R, Panthakkan A (2020) EfficientNet for retinal blood vessel segmentation. In: 2020 3rd International conference on signal processing and information security (ICSPIS), pp 1–4. https://doi.org/10.1109/ICSPIS51252.2020.9340135
6. Mou L, Chen L, Cheng J, Gu Z, Zhao Y, Liu J (2019) Dense dilated network with probability regularized walk for vessel detection. IEEE Trans. Med. Imaging
7. Kaur M, Kamra A(2020) Segmentation of retinal blood vessels via novel mechanism. In: 2020 2nd International conference on advances in computing, communication control and networking (ICACCCN), pp 997–1001 https://doi.org/10.1109/ICACCCN51052.2020.9362765
8. Kromm C, Rohr K (2020) Inception capsule network for retinal blood vessel segmentation and centerline extraction. In: 2020 IEEE 17th International symposium on biomedical imaging (ISBI), pp 1223–1226 https://doi.org/10.1109/ISBI45749.2020.9098538
9. Vaisnav M, Rishi S, Gopichand M, Ananthi G (2021) Identification of retinal diseases using blood vessel extraction. In: 2021 Sixth international conference on wireless communications, signal processing and networking (WiSPNET), pp 30–32 https://doi.org/10.1109/WiSPNET51692.2021.9419460

Multi-modal Medical Novel Image Fusion by Using DTNP Systems

G. Obulesu, M. Ravi Kishore, Shaik Karimullah, Fahimuddin Shaik, and CH. Naga Raju

Abstract This study will mainly concentrate on the novel mechanism used in the parallel and distributed computing model dynamic threshold neural *P* systems which involves the cooperative peaking of neurons inside a logical region. DTNP systems are relatively new and have a few intriguing properties. We intended to evaluate if these properties might be merged with existing NSCT in order to progress a special picture fusion technique for MMMI-multiple modality medical imaging which provides many advantages of NSCT-based multi-modal medical picture systems. The low-frequency NSCT coefficients will be extracted from the NSCT output for the DTNP process input used in the fusion rules. Additionally, WLE-INSML advantages are used for the extraction of high-frequency NSCT coefficients, and these coefficients are then used to create the fusion rules. A 12 pairs of multi-modality medical image pairs from an open dataset are used to evaluate the proposed fusion approach. The results of the quantitative as well as qualitative trials show the advantages of the suggested fusion process in terms of visual quality and fusion performance.

Keywords Distributed computing model · Multi-modal medical picture systems · Fusion technique

1 Introduction

1.1 Image Fusion

The process of combining many photographs into one enhanced image is to boost the information contained in the separate images [1]. An new field called "image fusion" combines the images taken by several sensors to create one informative image that may be used to make decisions.

G. Obulesu · M. R. Kishore · S. Karimullah (✉) · F. Shaik · CH. N. Raju
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: munnu483@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_37

High-resolution tomography highlights bony structures and other anatomical components (CT). Metabolic functional information is provided by computed tomography employing single-photon emission and positron emission tomography (PET and SPECT). But many of these photographs have poor resolution, and they are typically displayed in pseudo-color [2]. In order to improve diagnostic accuracy for the same patient using diverse approaches, studies from many different fields have combined their analyses of images from many sources. To assimilate two or additional images from various modalities keen on one fusion image, a variety of fusion techniques have been developed.

1.2 Motivation and Contribution

These studies combine the DTNP-dynamic threshold neural *P* schemes, a newly created DPNCM-computing model, which combine the increase and varied threshold mechanisms. We have demonstrated the universality of DTNP systems as turing machines [3]. This essay focuses on the fusion of pictures from various medical modalities utilizing DTNP system techniques, and it makes a unique fusion methodology for the NSCT domain based on the DTNP system.

The following is a description of the motivation for this study.

Old-fashioned image fusion approaches like NSCT-based methods perform badly when compared to contemporary, cutting-edge methods like SR-based and also DL-based techniques. NSCT is nevertheless advantageous in a number of ways that make image fusion possible. The NSCT, for example, is able to extract more data from multi-modality medical images. An exciting aspect of the recently developed model known as DTNP systems is the coordinated firing of neurons in a constrained region. We looked into the idea of combining this attribute with the NSCT in order to create a brand-new fusion technique for multi-modality medical imaging. The benefits of existing DTNP system could significantly increase fusion efficiency of the NSCT-dependent procedures for medical images with many modalities, even though the proposed fusion procedure can only slightly exceed the most modern fusion techniques in terms of performance.

2 Literature Survey

2.1 MST-Based Fusion Methods

This methods included wavelet transform, discrete wavelet transformations, gradient pyramid, Laplacian pyramid (LP), dual-tree, and wavelet transform. Utilize a complicated wavelet transform [4]. Even though, these methods have shortcomings from issues such inadequate spatiality, lack of shift invariance, and lack of time invariance. To solve these issues, a number of MST-based fusion approaches, like the curvelet, have been developed [5]. Surface let transform, non-subsampled contourlet transform, and non-subsampled shear let transform (NSCT).

2.2 SR-Dependent Fusion Methods

This method provided dictionary is first identified by SR-based dependent methods [6] before the depiction coefficients are integrated in accordance with a fusion rule and the fused image is produced. The provided dictionary and the fused sparse coefficients are then combined to form the fused image. Recently, many have been investigated. Picture fusion component examination is also known as convolutional sparsity-based morphology [7].

2.3 ID-Dependent Fusion Methods

This methods is dependent on the knowledge of real images may be successfully reduced to acceptable structural elements. To discover the most imperative information and preserve the textural information, a new color-gray picture fusion technique based on morphological component analysis (MCA) has been proposed [8] to completely take advantage of morphological diversity. Based on the characteristics of the photographs and the advantages of MCA, a special approach of multi-component fusion was developed to create enhanced fused images [9].

3 Methodology

3.1 NSCT

It is used to generate multi-resolution and multi-directional scenarios and has advantages in smoothness processing [10]. It is not the best solution for picture fusion, though, as it lacks translation invariance and is prone to producing pseudo-Gibbs phenomenon (artifact) close to the singular point of the reconstructed image, which distorts the picture. Additional comprehensive investigations have been carried out toward this goal by other academics. After contourlet transformation, Cunha et al. suggested a multi-scale disintegration technique that is superior to contourlet change, known as non-subsampled contourlet change [11].

NSCT has the properties of interpretation invariance and avoidance of ghostly associating obtained through the NSCT opposite [12]. The following stage addresses the dislodging difference issue created in the first stage, resulting in a merged image with solid relevance and high impact [13]. Bhateja et al. also found a combination of fixed wavelet change and non-subsampled shape change space. This calculation reduces the number of merged images and improves the differentiation of indicative highlights [14].

3.2 Proposed Methodology

Models for distributed parallel computing are a subclass which is known as DTNP systems that combine the SNP schemes. DTNP schemes are modeled as a matrix form of neurons along with local links that are capable of performing multi-modality picture fusion.

Consider a medicinal image of size mn and INSML mn as a feature matrix that includes both the images of the NSCT coefficients. As a result, as illustrated in Fig. 1, it is developed as a matrix of size mn neurons, where the feature value of each neuron in the network has an external input named INSML *ij* and an internal connection with the other neurons in its r-neighborhood. Neuron *ij* receives spikes from neighboring neurons and, applying the firing rule, sends the spikes it has produced to the other neurons.

Each neuron *ij*, which functions as both an input and an output neuron, receives its external input from the NSCT constant at location, or its INSML value. The neuron *ij*'s firing rate for the spiking rule $(E)/((\mathrm{au,a}))$ is defined by the formula $E(u_{ij}(t)\ ij(t))(u_{ij}(t)\ u)(ij(t))$. When the maximum condition is met, neuron *ij* excites to indicate the consumption of the high state with values *u* and in data segment and in dynamic threshold segment. The formulas for the equation

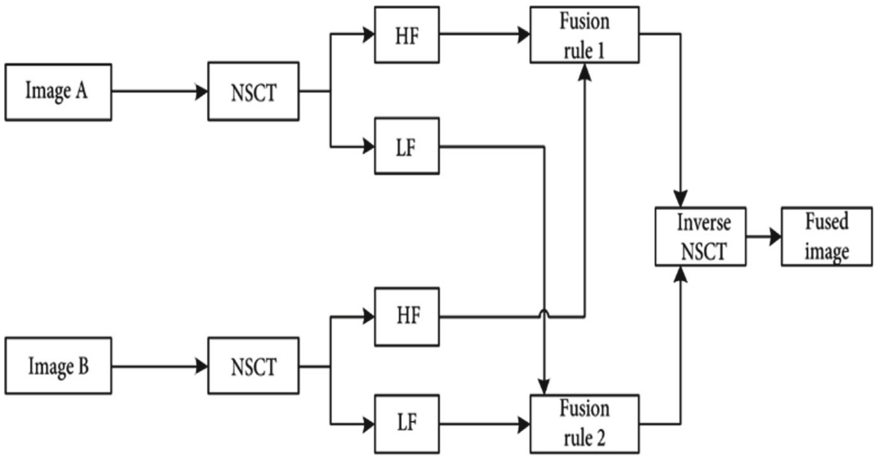


Fig. 1 NSCT fusion algorithm schematic diagram

$$u_{ij}(t+1) = \begin{Bmatrix} u_{ij}(t) - u + V_{ij} + \sum \sigma_{\mathrm{kl}} \in \delta r W_{\mathrm{kl}} P_{\mathrm{kl}}(t), & \text{if } \sigma_{ij} \text{ fires} \\ u_{ij}(t) + V_{ij} + \sum \sigma_{\mathrm{kl}} \in \delta r W_{\mathrm{kl}} P_{\mathrm{kl}}(\mathrm{t}), & \text{otherwise} \end{Bmatrix} \tag{1}$$

$$\tau_{ij}(t+1) = \begin{Bmatrix} \tau_{ij}(t) - \tau + p, & \text{if } \sigma_{ij} \text{ fires} \\ \tau_{ij}(t), & \text{otherwise} \end{Bmatrix} \tag{2}$$

V_{ij} = INSML, where A neuron's response to an external impulse is represented by the letters ij, p, w_{kl} (t), the conforming to local weight w, and p, which represents the value of the spikes that neuron ij receives from neighboring neuron kl. The system functions as follows: We initially set u_{ij} $(0) = v_{ij}$ $(0) = 0$, ij $(0) = 0$, and W_{rr} for each neuron ij. An external input matrix serves as the beginning state, INSML m n, and continues iteratively until it reaches the maximum number of iterations, t. After that, the system shuts down. Each neuron's total quantity of spikes delivered during calculation is taken into consideration, and this output is recycled through control signal aimed at the fusion of MMMI images. Easiness, the firing rules are applied using the maximum consumption method (i.e., set $u = u_{ij}$ (t) and $= ij$ (t)).

This method operation is explained in the section below. For each and every neuron ij, we first set u_{ij} $(0) = v_{ij}$ $(0) = 0$, ij $(0) = 0$, and W_{rr} as the assignment. The starting point and an outside influence matrix INSML mn serve as the starting point for the DTNP system, which iterates until it reaches the maximum number of iterations ($t_{\max}$). As a result, the system turns off. The output of each neuron ij during computing is the sum of the spikes it sends, and this value serves as a medical picture fusion control signal from multiple modalities.

3.3 Fusion Framework for MMM Images

For multi-modality medical images in the exciting domain, as shown in Fig. 3, we recommend a framework for DTNP-based image fusion. The suggested fusion framework is composed of three elements: non sub-sampled transform, the fusion rules, and the inverse NSCT transform. Medical multi-modality images make up the two source images as shown in Fig. 2.

This pictures are divided into non-sub-sampled constant expending NSCT transformation scheme. The generated NSCT constants are then created by fusing the NSCT coefficients. In contrast, distinct fusion rules are used to combine the two different coefficients individually. An INSML-WLE is utilized to combine the low-frequency NSCT coefficients, while a DTNP system-based fusion rule is employed to combine the high-frequency NSCT coefficients. The DTNP schemes, which are an essential element of the projected framework for fusion, are employed for the fusion of the high-frequency constants. This coefficients of two MMMI pictures, A and B, are to be merged, respectively, using two DTNP systems. The transformed NSCT coefficients from the fused NSCT are utilized to employing the inverse NSCT transform to reassemble a fusion picture. The stages involved are as follows, according to the proposed fusion.

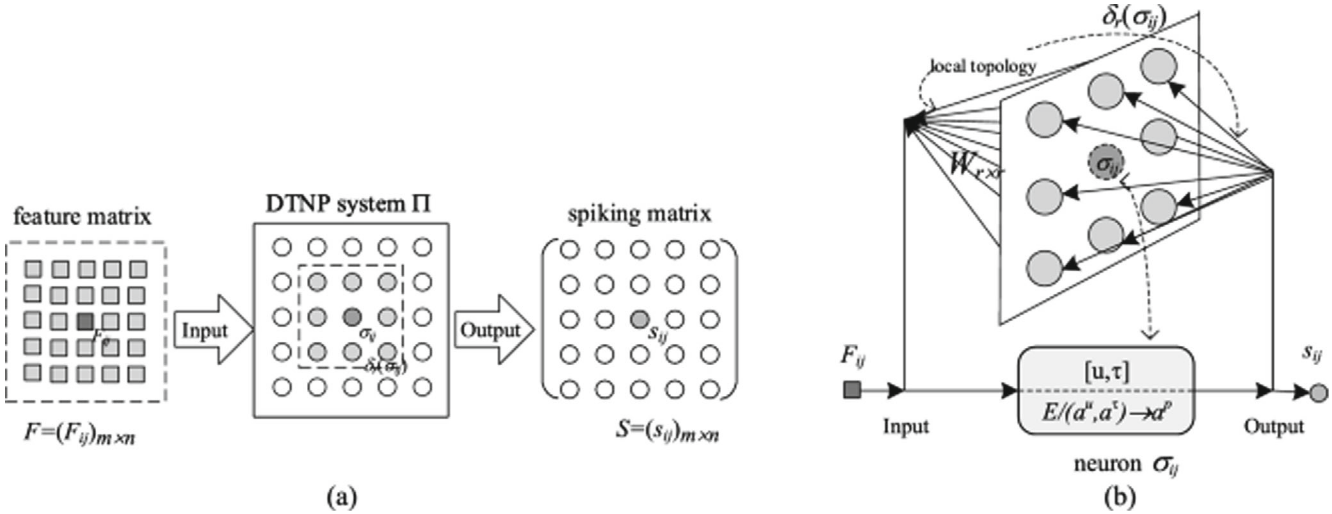


Fig. 2 DTNP system and its associated NSCT coefficients feature matrix, Vm x n

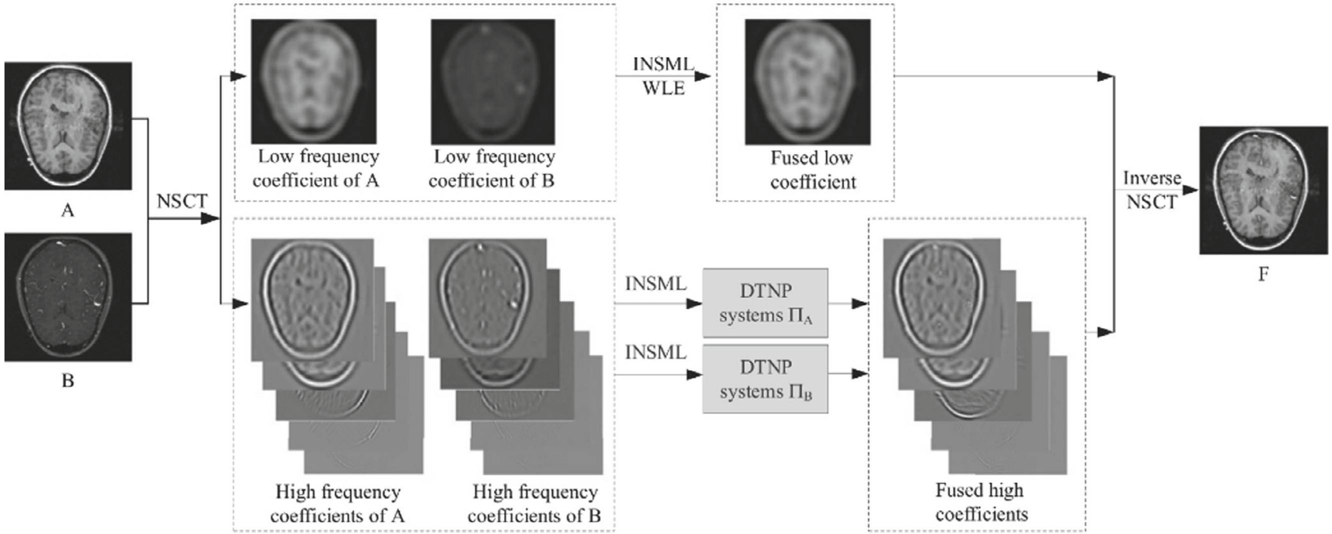


Fig. 3 Framework for proposed fusion using DTNP methods in the NSCT field with two multi-modal medical images as sources

I. By using NSCT transform, we can decompose the multi-modality medical pictures into non-subsampled counterlet transform. As a result, we are going to obtain two types of NSCT coefficients; they are high and low for each medical image that are used.
II. Low-frequency NSCT coefficients from many medical imaging modalities are combined using INSML-WLE, and the maximum frequency NSCT constants of multi-modality medical pictures' INSML feature matrices are taken into account as the supplementary inputs for dynamic threshold systems A and B.
III. Till the extreme number of iterations, t, is achieved, the two DTNP systems remain in the original state. At some point, they stop. In order to regulate the high-frequency fusion rules. Then, the two DTNP outputs are used.
IV. Utilizing the inversion of the NSCT to renovate the fusion image: The NSCT coefficients are transformed back to their original form after being mixed. The following part provides a detailed explanation of two fusion conditions for the coefficients.

3.4 Low-Frequency NSCT Coefficient Fusion Rules

The final fusion image is significantly influenced by the low-frequency coefficient fusion technique that was applied. The proposed fusion process takes into account energy conservation and detail extraction. Since the majority of that image energy is found in its low-frequency coefficients, it is widely known that the distinct identical position of the source pictures may have varied energy strengths due to different medical imaging techniques. It results to energy loss in merged images. As a result, the intensity can drastically decrease.

When it comes to WLE,

$$\mathrm{WLE}_{l0}(i, j) = \sum_{a} \sum_{b} w''^{(a,b)} C_{l0}(i + a,\, j + b)^2 \tag{3}$$

$W =$ matrix weight with the following definition, and $C_{l0}(i, j)$ is the low-frequency NSCT constant at location (I, j).

$$W = \frac{1}{16} \begin{matrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{matrix}$$

When combining MMM images, the low-frequency constants' capacity to occasionally hold a modest amount of distinctive information is typically useful. An INSML is utilized to obtain the precise information. This is how the INSML functionality is described:

$$\text{INSMLi}_{l0}(i, j) = \sum_a \sum_b w''^{(a,b)} \text{IML}_{l0}(i + a, j + b) \quad (4)$$

W is an equation for the matrix weight.

$W = \frac{1}{15} \begin{matrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{matrix}$ and IML is given by

$$\begin{aligned} \text{IML}_{l0}(i, j) = {} & |2C_{l0}(\text{i, j}) - C_{l0}(i - 1, j) - \text{C}_{l0}(i + 1, j)| \\ & + |2C_{l0}(i, j) - C_{l0}(i, j - 1) - C_{l0}(i, j + 1)| \\ & + |2C_{l0}(i, j) - C_{l0}(i - 1, j - 1) - C_{l0}(i + 1, j + 1)| \\ & + |2C_{l0}(i, j) - C_{l0}(i - 1, j + 1) - C_{l0}(i + 1, j - 1)| \end{aligned} \quad (5)$$

where C_{l0}(i, j) is a low-frequency coefficient (i, j). It should be underlined that $\text{IML}_{l0}(i, j)$ considers the diagonal coefficients and nearby information, and the square root coefficients will be a factor of 1/2.

This characteristics represent primary energy and partial descriptions of the corresponding images. As a result, the fusion rule, a combination of these two, is developed for the WLE-INSML functionality. Low-frequency coefficients of mixed modality medical pictures imagine A $(I\ j) =$ INSMLA.

WLEA for zero (i, j). A and B source pictures are linked by $B_{l0}(i, j) = l0(i, j)$ and INSMLB $\text{WLEB}_{l0}(i, j)$ $l0(i, j)$. Based on the WLE-INSML characteristics, the following definitions of the fusion conditions for low-frequency constants can be made (Table 1):

$$C^{\text{F}}_{l0}(i, j) = \begin{cases} c^A_{l0}(i, j), \text{if} \alpha^A_{l0}(i, \text{j}) \geq \alpha^B_{l0}(i, j) \\ c^B_{l0}(i, j), \text{Other wise} \end{cases} \quad (6)$$

$$\text{c}^F_{lr}(i, j) = \begin{cases} c^A_{lr}(i, j), \text{if} p^A_{ij} \geq p^B_{ij} \\ c^B_{lr}(i, j), \text{if} p^A_{ij} < p^B_{ij} \end{cases} \quad (7)$$

4 Results

These four other fusion methods were contrasted with the proposed method. Two sets of medical picture pairs from the dataset, “Group 1” and “Group 3” contain a variety of medical image pairs, including CT/MRI and $\text{MRI}_{\text{T1}}/\text{MRI}_{\text{T2}}$, which have been utilized to illustrate the experimental results. These approaches’ combined images are shown in Figs. 4, 5 and 6. Figure 5 demonstrates: (1) the wavelet-fused image has a generally low brightness and poorly preserves the unique details of the original images.

Table 1 Parameters of the fusion methods shown in the aforementioned photographs and comparisons of the existing and suggested ways

Image	Methods	Qabf	MI	FMI	SCD	MS_ SSIM	SD
Group (a)	Wavelet	0.4197	8.6867	0.8803	0.4904	0.8364	69.6428
	NSCT	**0.5904**	9.9105	0.8902	0.5404	**0.8877**	68.5398
	NSST-PAPCNN	0.5016	**11.5344**	0.8819	1.1824	0.8271	98.0668
	Proposed	0.5801	10.4887	**0.9039**	**1.3197**	0.8716	**100.2606**
Group (b)	Wavelet	0.5233	8.3320	0.8602	1.0607	0.9363	54.8943
	NSCT	**0.6185**	7.8842	0.8584	1.3508	**0.9456**	63.1007
	NSST-PAPCNN	0.5530	**11.5741**	**0.8726**	1.4069	0.9224	**75.0621**
	Proposed	0.6176	9.7471	0.8621	**1.5261**	0.9297	72.6617
Group (c)	Wavelet	0.4038	11.8110	0.9019	1.3350	0.8638	33.5205
	NSCT	0.7087	12.2556	**0.9050**	1.6033	0.9465	45.0023
	NSST-PAPCNN	0.6873	**13.9104**	0.9016	1.7223	0.9249	**56.0239**
	Proposed	**0.7371**	13.6377	0.9047	**1.7269**	**0.9478**	55.9198

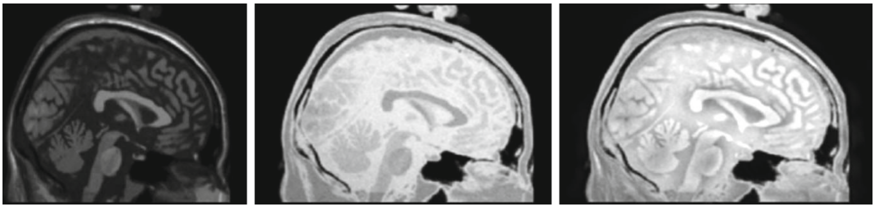


Fig. 4 Group (a): Sagittal MR scanner

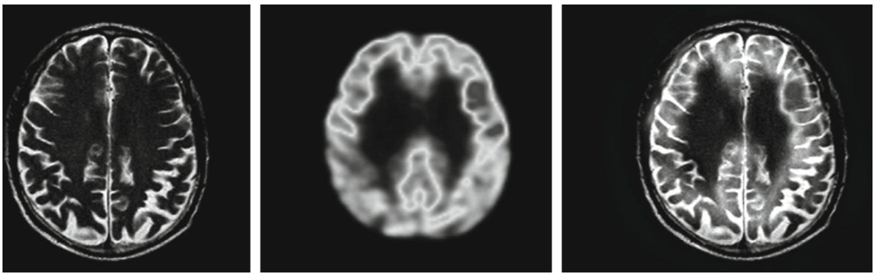


Fig. 5 Group (b): Axial plane

According to various metrics, the efficiency of the above DL-based fusion approaches can be approached or even partially surpassed by the suggested fusion method. The results of the individual and independent comparison show that the DTNP systems are superior in terms of enhancing the functionality of NSCT-centered fusion approaches.

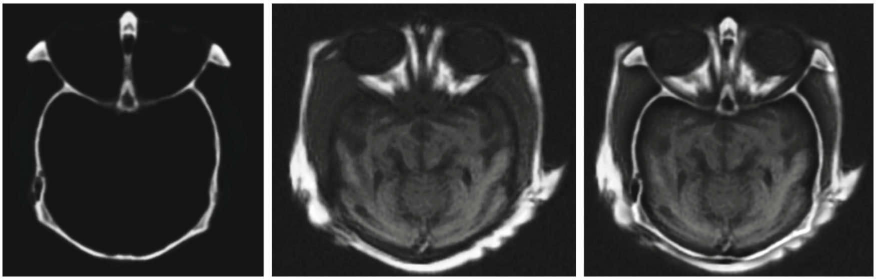


Fig. 6 Group (c): Coronal plane

5 Conclusion

DTNPS are universal models of distributed parallel computation that may simulate any calculation in the turing language. This study examined the use of a DTNP system with a fusion architecture to analyze multi-modality medical images in the NSCT field. The fusion framework that was proposed combined the features from the exciting methodologies. Medical images are expressed by an amalgamation of INSML along with WLE properties. The primary energy and certain minor characteristics associated with low-frequency NSCT coefficients, which are necessary to construct a fusion rule, are expressed. In order to extract specific information from medical images, like edges and contours, the INSML feature is employed with high-frequency NSCT coefficients. The proposed systems recognize this feature as an external input. The IN-SML features can produce more energy because of the neurons' supportive spiking mechanism. The cooperative spiking mechanism of the neurons enables the IN-SML characteristics to more successfully generate neural spiking in this systems.

For the combining of multiple medical imaging modalities, the NSCT + INSML + WLE + DTNP are combined in the suggested fusion architecture. This framework's drawback is this. The complementary information of other picture formats, such as multi-focus images, may not be properly recovered through the INSML characteristics because of the various imaging principles. The processing of other picture kinds could not be possible with this combination, as well. DTNP systems might therefore have to add additional characteristics to the proposed combination. Work in the future will focus on fusing various image forms.

6 Future Scope

Recent years have seen a lot of interest in color and 3D medical imagery. However, the grayscale medical images are where the proposed fusion architecture is most useful. The fusion outline using the INSML, NSCT, DTNP, and WLE and therefore can be used to instantly analyze color and three-dimensional medical images. Upcoming research will primarily concentrate on evolving a framework for the merging of colored and three-dimensional medical pictures using DTNP and other elements.

References

1. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. Springer Plus 5:507. https://doi.org/10.1186/s40064-016-2152-2
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. Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR in MIMO-OFDM using adaptive SLM and PTS technique. Int J Pure Appl Math Special issue 118(17):355–373 ISSN: 1311-8080 (printed version) ISSN: 1314–3395
4. 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, pp 599–610. Publisher Springer
5. Peng H, Wang J, Pérenz-Jiménez MJ, Riscos-Núñez A (2019) Dynamic thresholdneural p systems. Knowle Based Syst 163:875–884
6. Karimullah S, Vishnuvardhan D, Arif M, Gunjan VK, Shaik F, alamSiddiquee KN (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Communications and Mobile Computing 2022:10 Article ID 3016709. https://doi.org/10.1155/2022/3016709
7. Lewis JJ, OCallaghan RJ, Nikolov SG, Bull DR, Canagarajah N (2007) Pixel-andregion-based image fusion with complex wavelets. Inf Fusion 8(2):119–130
8. 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
9. Li H, Wang Y, Yang Z et al (2020) Discriminative dictionary learning-based multiplecomponent decomposition for detail-preserving noisy image fusion. IEEE Trans Instrum Meas 69(4):1082–1102
10. Karimullah S, Vishnu Vardhan D, Javeed Basha S (2020) Floorplanning for placement of modulesin VLSI physical design using harmonysearch technique. In: ICDSMLA 2019. Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
11. Jiang Y, Wang M (2014) Image fusion with morphological component analysis. Inf. Fusion 18(7):107–118
12. Tang H, Xiao B, Li W, Wang G (2017) Pixel convolutional neural network for multifocus image fusion. Inf Sci 433:125–141
13. Karimullah S, Vishnu Vardhan D (2020) Iterative analysis of optimization algorithms for placement and routing in asicdesign. In: ICDSMLA 2019. Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
14. Zhang Q, Levine MD (2016) Robust multifocus image fusion using multitask sparse representation and spatial context. IEEE Trans. Image Process 25:2045–2058 [CrossRef] [PubMed]

Image Segmentation Using a Robust Fuzzy Subspace Clustering Approach

S. Fayaz Begum, Fahimuddin Shaik, Shaik Karimullah, and Syed Javeed Basha

Abstract FSC LNML is an algorithm that integrates mean membership linkage, adaptive local variance, and non-local information for image segmentation in fuzzy subspace clustering noisy images. Although the Fuzzy C-means (FCM) clustering algorithm demonstrates exceptional performance in image segmentation, its noise resistance is enhanced by incorporating additional redundant visual data when non-local spatial information is incorporated. In contrast, under-segmentation may increase the quantity of noise throughout the non-local geographic domain, making the quantity of iteration steps critical in FCM. To compensate for underclassification in non-local data, local variability patterns are used. The durability of the FCM goal function is then improved by including non-local and local variance. To address the issues associated with voracious convergence, the objective function's denominator is changed to mean participation linkage, resulting in a decrease in the amount of iterations necessary. Furthermore, adaptive constraints are imposed on non-local knowledge about the original picture, and local variation based on the inverse relationship between the absolute magnitude difference. The setup of subspace improves the quality of color picture segmentation, which permits the assignment of adjustable weights to each dimension. The FSC LNML technique, as demonstrated by the simulation outcomes on chaotic color and grayscale images, exhibits superior performance compared to established fuzzy-based clustering methods.

Keywords Fuzzy subspace · Noise image · Membership linking · Robustness · Non-local data

S. F. Begum · F. Shaik · S. Karimullah (✉)
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: munnu483@gmail.com

S. J. Basha
Smartsoc Solutions Pvt. Ltd, Banglore, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_38

1 Introduction

Picture segmentation is a crucial area of study for future research in computer vision and picture interpretation. In recent decades, there has been significant progress in the field of picture segmentation algorithms. These algorithms have demonstrated their usefulness in a wide range of areas including resource surveying, artificial intelligence, remote sensing data analysis, and disease diagnosis [1]. The categories of supervised and unsupervised photo segmentation can be broadly divided into two. In contrast to unsupervised methods, certain supervised systems like as convolutional neural networks (CNN) and fully convolutional networks (FCN) allow for image segmentation through feature learning, but require a substantial quantity of training data. The clustering algorithm's exceptional resilience, minimal input parameters, and straightforwardness render it a typical unsupervised strategy for segmenting pictures. FCM is a highly used algorithm for fuzzy clustering. The fuzzy clustering technique [2] considers the uncertainty of image pixel assignment and is frequently employed in clustering methods. Noise interference can affect pictures during their acquisition and conversion in the virtual world due to the unpredictable nature of noise.

The modified version of FCM disregards the membership information related to the neighborhood, but utilizes histograms to accelerate the clustering process, while still considering the image neighborhood information. The relationship between pixels can be elucidated through the utilization of the Hidden Markov Random Field (HMRF) [3].

Therefore, it is recommended to consider techniques such as the Liu methodology, HMRFFCM, and FLDNICM as extensions of HMRF. HMRF outperforms FCM in picture segmentation due to its ability to collect spatial context data from local pixel blocks and consider the prior state of the present membership degree. The presence of a large number of noisy pixels within the local pixel block may diminish the sensitivity of noisy picture segmentation. However, noise can also obliterate contextual information inside an image. In order to mitigate the susceptibility of classical Euclidean distance to nonlinear data, the data are transformed into a high-dimensional space by the utilization of the kernel technique. The KWFLICM, ARKFCM, GKWFLICM, and KBFWCM algorithms were formulated using the kernel technique. The presence of significant uncertainty in the segmentation of images with heavy noise is a difficult problem to tackle. Despite the effective performance of the kernel approach in the task of nonlinear classification, it remains problematic. Merely including local information is inadequate, as the noise density rises, the neighboring pixels of the noisy central pixel may also be affected by noise. Moreover, it demonstrates that the data is uniformly processed across all dimensions by assigning equal importance to each pixel's dimensions. The influence of pixels on clustering may vary depending on the dimensions.

Optimizing the allocation of weights to each dimension has the potential to improve the segmentation performance of color images. Consequently, the concept of the subspace becomes more widely known. Three algorithms based on subspace are proposed in [4]: ESSC, FCS, and FSC. However, the following methods do not take into account the overall number of iterations that occur during the iteration process, leading to a higher number of iterations when minimizing the objective function. In order to address the difficulty highlighted before, a new FCM clustering approach known as FCM SICM was recently introduced [5]. The system relies on the connection between membership and local data. However, in cases where the image is heavily affected by noise, the method exhibits anomalies in the membership degrees, resulting in unusual values for the membership connections. This diminishes the precision of the segmentation and compels the algorithm to prematurely converge prior to achieving the desired clustering outcome.

Although the algorithms described above are noise resistant, the following issues have arisen in the past. "[6]" is the user's text. The method demonstrates an under-segmentation impact due to the substantial amount of noise in the picture and the limited effectiveness of utilizing local or non-local knowledge to minimize noise. In the case of color photographs, however, the contribution of distinct channels to clustering discoveries may not be equal, rendering the average technique inappropriate for segmenting pictures with color. To overcome these issues, the FCM SICM technique is used to minimize the number of iterations while improving the algorithm's segmentation accuracy. Due to erroneous membership values, this technique demonstrates early convergence. The clustering procedure is carried out using the aforementioned approaches, with equal weight given to each dimension of the fuzzy subdomain clustering technique (FSC LNML). Responsive local variance, non-local geographic data, and mean membership linkage are all part of the system. Here are the strategies and objectives. To solve the problem of under-segmentation induced by high noise density, we create a local variance template. Our method takes non-local spatial information into account and seeks to remedy the image's under-segmentation. The efficacy of the clustering process is strongly dependent on its repetition rate. We use mean membership connecting into the goal function to decrease the overall number of iterations needed and avoid premature convergence. The inclusion of their individual contributions is enabled by simultaneously exploiting the reciprocal nature of the original picture, local variation, along with non-local spatial information while adaptively restricting them. It is vital to notice that when pixels are clustered together, they can fail to have the same influence on all dimensions. The allocation of weights to specific dimensions may be optimized to improve the efficacy of color picture segmentation. To enhance the accuracy of segmentation of color photographs, this study applies appropriate weights to each aspect of the data and includes subspace grouping into the objective function. To conclude, our proposed technique outperforms existing clustering methods by a factor of three. The use of a local variance template helps to prevent under-segmentation in regard to non-local spatial information. The method should handle the convergence problem by lowering the total amount of iterations and changing the goal function to optimize the answer. In particular, the objective function's denominator should be the mean membership linkage. Improving the weighting of specific picture dimensions to improve the effectiveness of color segmentation of images. The subsequent sections of the essay are structured in the following manner. Section 2 encompasses the fuzzy subspace clustering algorithm, the conventional FCM clustering technique, and non-local spatial data. Section 3 provides an elaborate explanation of the suggested approach and how it is put into action. Section 4 presents the experimental findings for both authentic color and grayscale photographs, as well as the process of selecting parameters. The suggested approach, FSC LNML, is compared to previous similar work in Sect. 5. The final section, Sect. 6, presents several conclusions.

2 Literature Review

Using fuzzy logic, we describe a novel technique for the fuzzy subdivision of magnetic resonance imaging (MRI) information as well as the evaluation of intensity inhomogeneities. Inhomogeneities in MRI intensity can be attributable to faults in radio-frequency coils or issues with acquisition processes. As a consequence, the image displays gradual and continual shading chromatic aberration which can cause errors when using standard intensity-based categorization techniques. To account for changes in homogeneity, we modified the objective function of the traditional fuzzy c-means (FCM) algorithm. This modification allows a pixel's (voxel's) labeling to be influenced by the labels that are placed of its neighbors. The proximity effect is a regularization method that biases the final result towards piecewise-homogeneous labelings. This regularization method is excellent for segmenting scans that have been affected by the combination of pepper and salt noise. The suggested technique's applicability and efficiency are demonstrated by studies conducted on both synthetic pictures and actual MR data. The user's text is "[7]".

The goal of this research is to define various techniques for how technology might help with medical decision-making. This duty is both critical and difficult, since it involves several medical and technical issues. The most important feature, however, is the engagement of humans. First, there is the patient, who is suffering from a medical problem and need the most effective treatment. Second, regardless of the conditions, the physician is expected to make informed conclusions or forecasts about the patient's current and future medical status. Technology, including artificial intelligence (AI) tools, can help both parties, and applicable modeling tool theory can help as well. Artificial neural networks (ANNs) are a powerful tool in this domain [8].

This study investigates two basic computer vision problems: contour identification and picture segmentation. We use cutting-edge algorithms to achieve both of these goals. The contour detector we developed uses spectral clustering to include various local inputs within a globalization framework. In our segmentation technique, we apply a flexible mechanism to convert the output from every contour detector through a hierarchical region tree. We simplify the challenge of picture segmentation to the procedure of identifying contours by employing this method [9]. Our contour detecting and segmentation methods have undergone extensive testing and frequently outperform competitor algorithms. The automatically generated hierarchy segmentations can be improved further in an interactive way using user-specified annotations. We may use computation at different picture resolutions in order to link our system to recognized applications [10] [11].

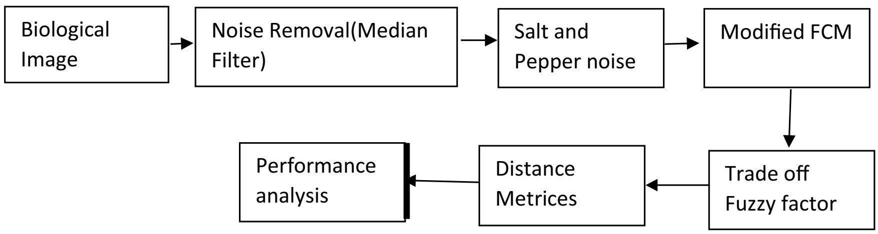
3 Methodology

4 Existing Method

The fuzzy c-means (FCM) clustering of data technique divides a data set into *N* clusters, and every point of information in the data set belongs to every cluster to some extent. A data point at the center of a cluster, for example, has a high degree of involvement in that cluster, whereas a data point far removed from the cluster center has a low degree of affiliation in that cluster.

The method allocates membership for every data point relating to every cluster center based on the basics of separation between the core of the cluster and the data point. The more membership the data has in that given cluster, the closer it is to the clusters center. Each data point's membership should, of course, equal one. Membership and cluster centers are adjusted based on the formula following each iteration. FCM has the benefit of being unsupervised and convergent, but it also has some drawbacks, including a lengthy calculation time, sensitivity to the initial assumption (speed, local minima), sensitivity to noise, and low (or even no) membership degree for outliers (noisy points).

5 Proposed Methodology



Choosing a portrayal space is the first step in doing image clustering. Next, we must apply a suitable separation measure (closeness measure), to coordinate between the images and group focuses in the selected portrayal space. Depending on how closely the images resemble the various bunch focuses, the picture grouping is then carried out either in a supervised approach with human intervention or in an unsupervised procedure. The analysis discussed above can be summed up as follows. (1) Despite being straightforward, FCM, FSC, and ESSC are not robust against noise because they do not take into account the spatial information of the image. (2) To increase robustness, the FCM NLS and FCM SNLS objective functions utilize non-local spatial information. Nevertheless, as the level of noise escalates, the algorithm fails to accurately segment the image, resulting in the loss of image features and the blurring of its boundaries as it attempts to obtain non-local spatial information. (3) Due to the high value and critical nature of the objective function during the clustering procedure, a considerable number of iterations are necessary. In order to preserve a portion of the image's edge features while preventing the undersegmentation of non-local spatial information, this article presents an innovative local variance template. Then, using fuzzy subspace aggregation, local variance when non-local location data are integrated, and a suitable weight is applied to each dimension to improve picture segmentation performance. Finally, the mean memberships linkage is used as the target function's denominator, resulting in a decrease of four repetition steps.

In contrast to clustering, which may not account for the equal influence of pixels on various dimensions, the traditional FCM method processes data uniformly across all dimensions. By allocating suitable weights to each component, color image segmentation can be enhanced in both quality and dimension. In consequence, introduced the fuzzy subspace clustering (FSC) algorithm, which effectively aggregates high-dimensional datasets through the appropriate allocation of weights to data points spanning various dimensions.

The objective function is

$$\mathrm{J_{FSC}} = \sum_{i=1}^{k}\sum_{j=1}^{N} u_{\mathrm{ij}} \sum_{z=1}^{D} w_{iz}^{\tau}(\mathrm{y}_{jz} - \mathrm{c}_{iz})^2 + \gamma \sum_{i=1}^{k}\sum_{z=1}^{D} w_{iz}^{\tau}, \sum_{i=1}^{k} u_{ij} = 1, \sum_{z=1}^{D} w_{iz} = 1 \quad (1)$$

where D is the data dimension, w_{iz}^{τ} is the fuzzy weighting factor, τ is the fuzzy index, and is always set $\tau > 1$.

Using Lagrangian multiplier method, u_{ij}, w_{iz} and c_{iz} are updated by

$$c_{iz} = \frac{\sum_{j=1}^{N} u_{ij} y_i}{\sum_{j=1}^{N} u_{ij}} \quad (2)$$

$$w_{iz} = \frac{\left(\sum_{j=1}^{N} u_{ij}(y_{jz} - c_{iz})^2 + \gamma\right)^{\frac{-1}{\tau-1}}}{\sum_{z=1}^{N}\left(\sum_{j=1}^{N} u_{ij}(y_{jz} - c_{iz})^2 + \gamma\right)^{\frac{-1}{\tau-1}}} \quad (3)$$

$$u_{ij} = \begin{cases} 1, & \sum_{z=1}^{D} w_{iz}^{\tau}(y_{jz} - c_{iz})^2 \le \sum_{z=1}^{D} w_{iz}^{\tau}(y_{jz} - c_{pz})^2, p = 1, 2, \ldots, k \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

It is clear from Eq. (1) that while FSC gives each dimension a different weight, it only takes into account intra-class information while ignoring inter-class information. Therefore, it is necessary to enhance the objective function. Deng et al. presented the extended soft subspace clustering (ESSC), which incorporates intra-cluster and inter-cluster information. The objective function of ESSC can be formulated as

$$J_{\text{ESSC}} = \sum_{i=1}^{k}\sum_{j=1}^{N} u_{ij}^{m} \sum_{z=1}^{D} w_{iz}\left(y_{jz} - c_{iz}\right)^{2} + \gamma \sum_{i=1}^{k}\sum_{z=1}^{D} w_{iz} \ln w_{iz}$$
$$- n \sum_{i=1}^{k}\left(\sum_{j=1}^{N} u_{ij}^{m}\right) \sum_{z=1}^{D} w_{iz}(c_{iz} - c_{lz})^{2} \tag{5}$$

where c_{lz} is the global center of the sample, γ and n are used to control the influence of entropy and weight on cluster separation, respectively.

6 Result and Analysis

The following are the figures for image segmentation using the robustness of the fuzzy subspace clustering approach (Figs. 1, 2 and 3).

These are the three different types of figures. The first figure is the air craft image. The second figure is the knee image. And the third figure is the symbol image. Each figure has three images. The three images are original image, noise image, result image. Original image is the source of the image. Noise image is the second image. In this, some noise is added. And the third image is the result image. That means it is the output of the image.

Tabular Column.

Air craft image	Knee image	Logo image
VPC = 99.17%	VPC = 91.71%	VPC = 92.71%
VPC = 1.51%	VPC = 12.82%	VPC = 11.52%
Sa = 99.45%	Sa = 81.15%	Sa = 88.93%

VPC and VPE are two metrics used to assess the ambiguity of membership. As noise increases, the VPC and VPE of each algorithm diminish and then increase, respectively, suggesting that noise has a general effect on membership uncertainty. When noise density is low, the FCM NLS algorithm operates more effectively at segmenting owing to non-local spatial information's improved suppression of simple noise.

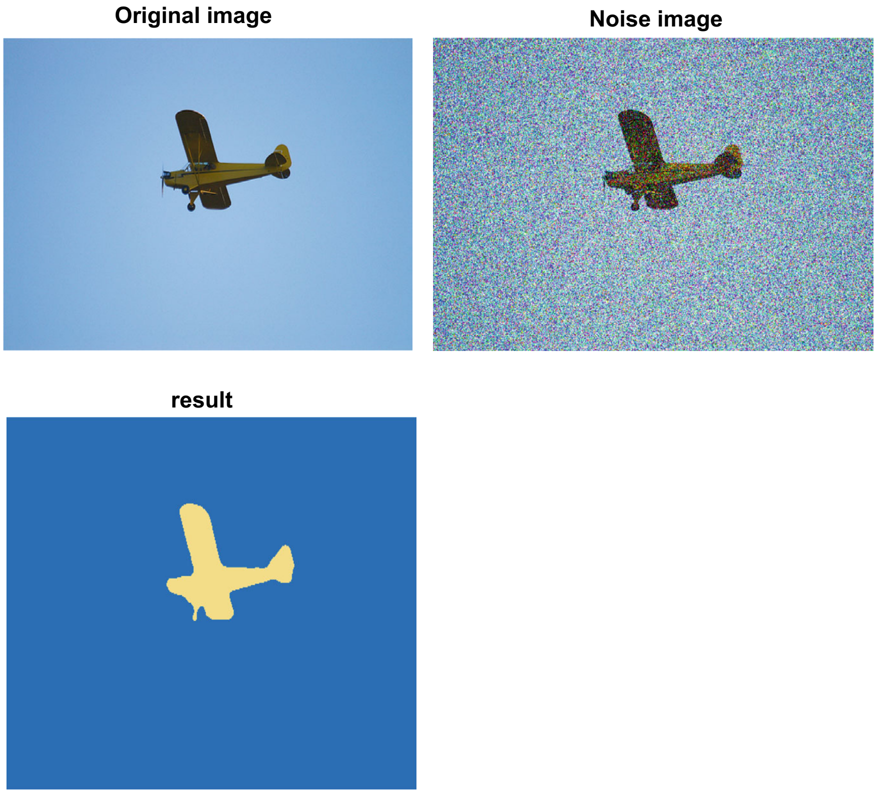


Fig. 1 Air craft image

7 Conclusion

This article provides a thorough analysis of picture segmentation using FCM (fuzzy C-means) algorithm. In order to tackle the unpredictability of chaotic picture segmentation, a wide range of FCM variables are offered. Non-local geographical information and membership linkage are two ways that effectively improve the resilience of FCM while minimizing the number of iterative cycles. This paper's fuzzy subspace clustering improvement incorporates mean membership linkage, adaptive local variance, and non-local spatial information to address the three aforementioned difficulties. Optimize the segmentation performance of color photos by assigning appropriate weights to each dimension in line with the subspace. The experiments are conducted using the following algorithms in the presence of mixed noise: FSC, FCM, KWFLICM, FCM NLS, ADFLICM, ARKFCM, FCM FWCW, FRFCM, SFFCM, CGFFCM, FCM SICM, and the suggested FSC LNML. The findings demonstrate that the suggested FCM LNML algorithm attains superior performance with minimal iterations, particularly in cases when the image is heavily affected by noise.

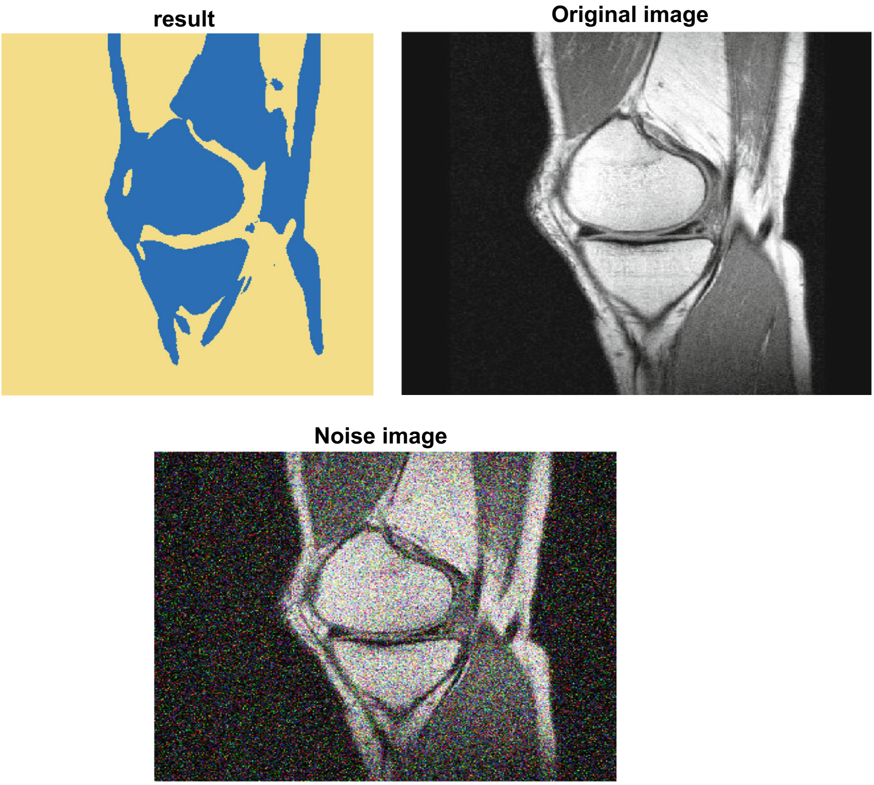


Fig. 2 Knee image

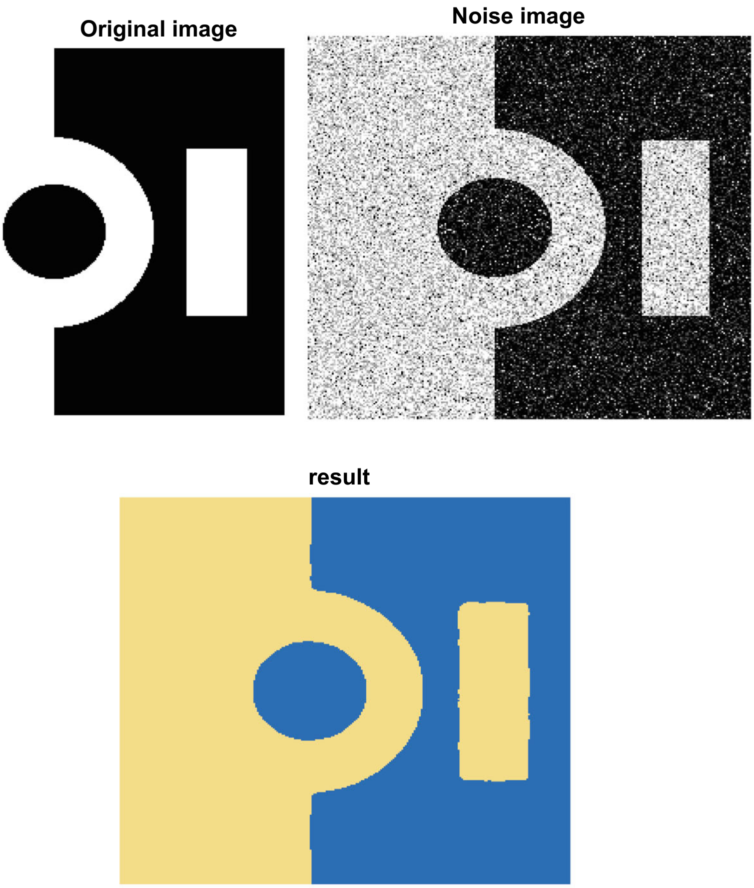


Fig. 3 Logo image

References

1. Ahmed MN, Yamany SM, Mohamed N, Farag AA, Moriarty T (2002) A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. IEEE Trans Med Imaging 21(3):193–199
2. Albu A, Precup R-E, Teban T-A (2019) Results and challenges of artificial neural networks used for decision-making and control in medical applications. Facta Univ. Ser. Mech. Eng 17(3):285–308
3. Anu Priya D, Vishnuvardhan D, Mamatha G, Shaik F, Karimullah S (2023) VIP development of SPI Controller for open-power processor-based fabless SOC. In: Kumar A, Ghinea G, Merugu S (eds) Proceedings of the 2nd international conference on cognitive and intelligent computing. ICCIC 2022. Cognitive Science and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_2
4. Karimullah S, Vishnuvardhan D, Bhaskar V (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Pers Commun 127:3041–3059. https://doi.org/10.1007/s11277-022-09909-2

5. Arbelaez P, Maire M, Fowlkes C, Malik J (2010) Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell 33(5):898–916
6. Shaik F, Sharma AK, Ahmed SM (2016) Detection and analysis of Diabetic Myonecrosis using an improved hybrid image processing model. In: 2016 2nd international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB). Chennai, India, pp 314–317. https://doi.org/10.1109/AEEICB.2016.7538298
7. Karimullah S, Vishnu Vardhan D, Basha SJ (2020) Floorplanning for placement of modules in vlsi physical design using harmony search technique. In: Kumar A, Paprzycki M, Gunjan V (eds) ICDSMLA 2019. Lecture notes in electrical engineering, vol 601. Springer, Singapore. https://doi.org/10.1007/978-981-15-1420-3_197
8. Shaik F, Giri Prasad MN, Rao J, Abdul Rahim B, SomaSekhar A (2010) Medical image analysis of electron micrographs in diabetic patients using contrast enhancement. In: 2010 International conference on mechanical and electrical technology. Singapore, pp 482–485. https://doi.org/10.1109/ICMET.2010.5598408
9. Karimullah S, Sai Sumanth Goud E, Lava Kumar Reddy K (2023) Spectral efficiency for multi-bit and blind medium estimation of DCO-OFDM used vehicular visible light communication. In: Kumar A, Senatore S, Gunjan VK (eds) ICDSMLA 2021. Lecture notes in electrical engineering, vol 947. Springer, Singapore. https://doi.org/10.1007/978-981-19-5936-3_83
10. Krishna SLV, Abdul Rahim B, Shaik F, Soundara Rajan K (2010) Lossless embedding using pixel differences and histogram shifting technique. In: Recent advances in space technology services and climate change 2010 (RSTS & CC-2010). Chennai, India, pp 213-216. https://doi.org/10.1109/RSTSCC.2010.5712850
11. Bai X, Chen Z, Zhang Y, Liu Z, Lu Y (2015) Infrared ship target segmentation based on spatial information improved FCM. IEEE Trans. Cybern. 46(12):3259–3271

Privacy Preservation on Social Networks Using Kangaroo Method

P. Deepthi and Nagaratna P. Hegde

Abstract In the dataset collected from various social networks, each user is matched with the closest target user to establish associations primarily within similar social circles, encompassing aspects such as education, family, and hobbies. It's worth noting that the utilization of social data, which may contain consumer information, can potentially expose privacy concerns. To address this, the paper presents a methodology based on the discrete logarithm, employing the Kangaroo method to compute within an arbitrary cyclic group. This method assumes the value lies within a predefined interval. Additionally, Wiener and Van Oorschot introduced a method called "linear speed-up," which minimizes storage requirements. It can also be effectively monitored and parallelized in its corresponding design, making the Kangaroo technique a valuable tool for addressing discrete logarithm issues. The Kangaroo method, offers a robust solution for discrete logarithm problems. Parameters can be fine-tuned to optimize its performance, preventing "useless collisions." This paper also introduces an analysis of the potential acceleration achieved by utilizing parallel resources beyond the algorithm's initial configuration. Through this approach, the Kangaroo method is applied to cryptosystems that safeguard user privacy when dealing with social data. To maintain privacy, the propagation direction is controlled by a limited set of indicators, ensuring that each user can access consistent functions based on homomorphic attributes, while remaining unable to access another user's data due to the random number concealment. This paper demonstrates that the proposed method is designed with consideration for communication cost and computation overhead.

Keywords Kangaroo method · Pollard rho method · Distinguished point method · Communication cost and · Computation overhead

---

P. Deepthi
Department of Computer Science and Engineering, Bhoj Reddy Engineering College for Women, Hyderabad, Telangana, India

N. P. Hegde (✉)
Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, Telangana, India
e-mail: nagaratnaph@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_39

1 Introduction

Over the past few years, the popularity of social networks has experienced a significant and remarkable surge. This surge has paved the way for exciting opportunities in the field of social network research and the development of novel applications [1], particularly in the domain of social network analysis [2, 3]. Social networks function as extensive reservoirs of information, encompassing a variety of attributes. This has consequently given rise to a broad spectrum of social services, including proximity-based queries aimed at measuring social distance. Earlier iterations of social software aimed to gauge the closeness between connected users based on their day-to-day interactions, which included actions like comments, likes, and shares. In this context, proximity typically refers to the weight assigned by subtracting the correspondence of users' characteristics from their mathematical social connection.

Solidarity, as a measure of common features, indicates that closer proximity suggests a stronger relationship with more shared features. The specific weight applied to proximity can vary depending on the case.

Notably, recent influence factors into every concurrently promoted message, and the smallest value represents the most favourable proximity. This distributed process is highly efficient, requiring minimal data exchange between nodes. However, while leveraging abundant social data yields valuable advantages, it also raises serious privacy concerns. Efficiently determining proximity requires the responsible handling of sensitive user information, which may encompass personal details such as birthdates, gender, place of residence, and interests, as well as social characteristics like sexual orientation, income level, geographical location, and social affiliations. Third parties, like social media platforms or marketers, have the potential to misuse this data, potentially resulting in privacy violations [4, 5]. As a result, individuals are growing more wary of divulging personal information, emphasizing the need for the creation of distributed privacy-preserving solutions. Homomorphic encryption is gaining popularity in this regard [6, 7], enabling users to compute on ciphertexts while maintaining privacy. However, secure procedures for calculating proximity, such as secure evaluation, can still incur substantial costs, especially for large graphs [8].

The primary goal of public-key cryptography is to facilitate secure communication between two parties, even in the absence of prior acquaintance and under the scrutiny of potential eavesdroppers. Commonly employed cryptographic methods, like RSA and Elliptic Curve public encryption, guarantee confidentiality, data integrity, authentication, and non-repudiation. The security of these systems depends on the computational complexity associated with solving the factorization problem for RSA and the challenge of computing discrete logarithms for elliptic curves. Differing from RSA, which demands substantially longer keys to achieve comparable security because of the sub-exponential factorization problem, elliptic curve encryption is the preferred choice.

Public-key methods are underpinned by the challenge of solving discrete logarithm problems. Digital signature systems, like ECDSA (Elliptic Curve Digital Signature Algorithm), are built upon the discrete logarithm problem within finite groups or their elliptic curve counterparts. The challenge of computing discrete logarithms involves finding an integer x such that $g^x = h$, where h represents a group element where G is a member of the cyclic finite group and G is generated by a specific generator g. This problem is intractable due to the multiplicative nature of the operation, making it suitable for cryptographic purposes. Within signature methods, x denotes the secret key utilized for generating signatures, while h functions as the public key employed for verification of signature.

Various methods attempt to tackle the challenge of computing discrete logarithms in signature schemes. In the two algorithms ECDSA and DSA (Digital Signature Algorithm), leading to the use of specialized techniques such as Pollard rho and kangaroo algorithms. These approaches are robust and do not necessitate precise knowledge of the underlying group's structure.

The rho algorithm is utilized when x can assume any non-negative integer value up to the order of g, for which $g^x = 1$. Conversely, the kangaroo method stores a compact, fixed set of group elements and provides a linear speed boost. While it does entail additional offline storage demands, it remains an efficient approach.

This paper, focus on efficiency and privacy in proposing a parallelized kangaroo-based encryption method for social networks. Specifically, many users aim to keep their information confidential while conducting proximity queries in social networks without revealing their data to others.

2 Literature Survey

Brickell and Shmatikov [8] explore scenarios where two parties jointly compute a task on interconnected graphs. Zhang et al. [10] introduce a distributed approach called knapsack where two parties collaboratively calculate benefit values without revealing their sensitive data, enabling decentralized distance calculations. This concept can be extended to secure multiparty computation methods [9], allowing each node to perform partial distance calculations. Consequently, distributed distance calculation has become a feasible option.

Recent developments in routing protocols have incorporated distance calculations. Henecka et al. [11] devise a routing protocol designed to protect privacy for distance calculations, wherein participants jointly calculate influence without directly examining specific routes or distances [12]. However, employing this technique to encrypt data using the public key of the intended recipient may not guarantee adequate security, as it can be susceptible to potential vulnerabilities.

Dynamic programming techniques have been utilized in the development of combinatorial auctions, as exemplified by Sakurai et al. [2] and Sandholm [11]. Nevertheless, as the quantity of goods and bidders grows, dynamic programming can result in non-polynomial computational expenses. Kikuchi and Thorpe [10] introduce a privacy-preserving combinatorial auction leveraging Shamir's secret sharing method. Subsequent advancements by Hu et al. [12] strive to diminish communication overhead and bolster resistance against passive and collusion-based attacks.

Several combinatorial auction protocols rely on homomorphic encryption within ciphertext domains [3]. Nonetheless, these protocols often incur higher computational costs. Palmer et al. [8] explore secure three-party computations for privacy-preserving combinatorial auctions, although the system may not be suitable when the inputs are not prearranged. Zhang et al. [7] employ the central idea of matrices and their inversion to perform scalar asymmetric product operations on encrypted data, an alternative to homomorphic encryption. Li et al. [6] adopt a random noise approach to conceal bid values.

Blockchain, as an emerging decentralized data management system, has garnered significant attention and found applications in electronic auctions. In conventional auction-based trading, participants might engage in collusion or act in a self-interested manner. Hence, utilizes the Ethereum framework to conduct trustless, secure, and decentralized auctions. Additionally, proposes a decentralized electricity transaction model for microgrids based on blockchain technology and continuous double auction (CDA) mechanisms, aiming to address issues in traditional management systems, including high operational costs and limited transparency.

3 Related Works

3.1 Distinguishing Between Kangaroo and Rho Methods

The Lambda Method, also known as the Kangaroo Method, can sometimes be mistakenly referred to as the Rho Method, especially with the growing trend of parallelizing the Rho Method.

3.2 Pollard's Rho Method

In Pollard's Rho Method for computing discrete logarithms, it generates a sequence (y_k) within the group G. It starts with the term $y_0 \in G$ and applies the following rule iteratively: y_{k+1} is determined by applying the pseudo-random mapping $F: G \to G$ to y_k, where k belongs to the set of natural numbers $(k \in N)$. Gradually, y_k is a sequence enters a repetitive pattern. When the terms are plotted the sequence on paper, commencing from the bottom and forming a cycle, the resulting shape closely resembles the Greek letter "rho" (as illustrated in Fig. 1).

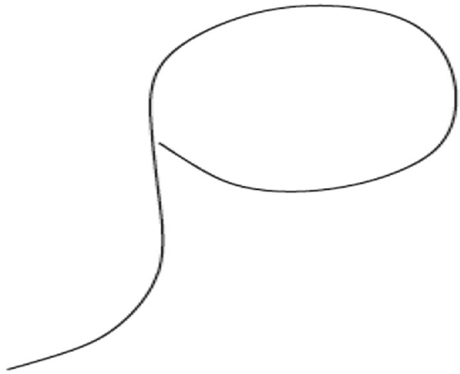


Fig. 1 Rho

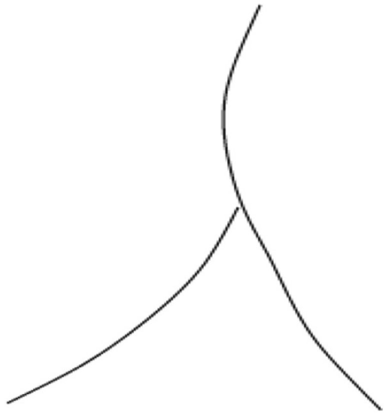


Fig. 2 Kangaroo lambda

3.3 Kangaroo Method

In the Kangaroo Method, the trajectories of two "kangaroos," one originating from the bottom left (ta_k) and the other from the bottom right (wi_k), they converge when they collide, creating a shape reminiscent of the Greek letter "lambda" (as depicted in Fig. 2).

3.4 Parallelized Rho Method

The Parallelized Rho Method functions with an array of sequences, each designated for a distinct processor, with the goal of having two of these sequences intersect or collide. When all these sequences' terms are used to create a picture, it exhibits a bundled "rho," and upon zooming in on the collision area, it appears as a "lambda" (Fig. 3).

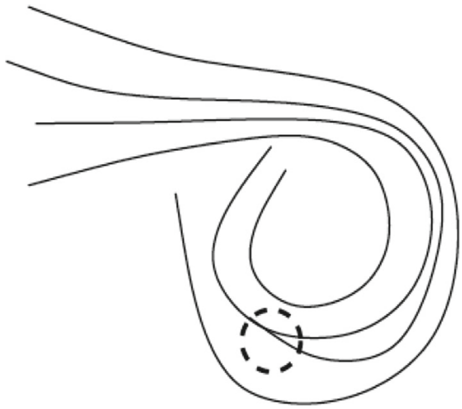


Fig. 3 Parallelized rho

3.5 Pollard's Kangaroo Algorithm

Originally known as the "Lambda Method for Capturing Kangaroos," the algorithm's name can be puzzling because when the parallel version of the Rho Method is employed, it produces a "lambda" shape, rather than a "rho."

3.6 The Distinguished Point Method

The Distinguished Point Method doesn't seek a match $(\mathrm{ta}_k = \mathrm{wi}_k)$ among all the terms of both the "tame" and "wild" kangaroo sequences but rather within a limited subset of terms that fulfill particular distinguishing criteria (Figs. 4 and 5).

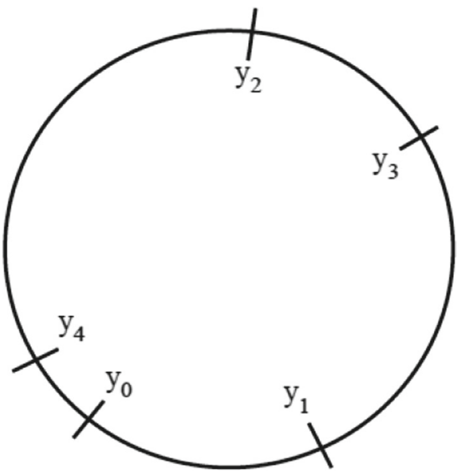


Fig. 4 The pseudo-random walk

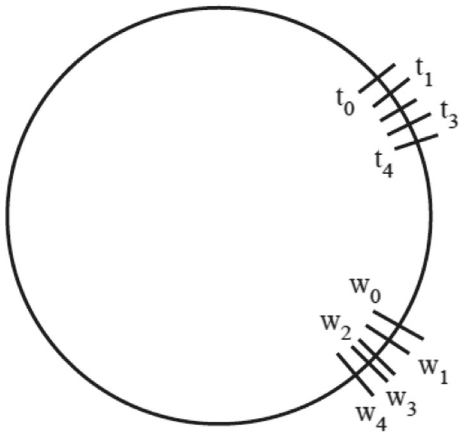


Fig. 5 The walks with short jump distances

Assume subset of G is D, and let a z is the group element be considered a special point if it belongs to D. To establish this, introduce f, which is an integer and define D as D_f, where D_f is the set of elements z in G that meet the criterion of having the least significant bit f in their binary representation set to zero.

z is as a group element denoted as a special point of degree f if it belongs to D_f. This definition provides the advantage of quickly discerning whether a group element holds distinguished status and permits the monitoring of D_f*'s* size by adjusting the parameter f.

Now, let's explore the process of employing distinguished points to identify collisions within the Kangaroo Method. It involves computing sequences (ta_k) and (wi_k) for the "tame" and "wild" kangaroos, with each ta_k expressed as $\mathrm{ta}_k = g^{(a+b)/2} + \delta_{k,\mathrm{tame}}$ and each wi_k as $\mathrm{wi}_k = h * g\ \delta^{k,\mathrm{wild}}$. Following each kangaroo jump, verify whether the current term is a member of D_f. Suppose a term qualifies as a distinguished point, to store the pair (ta_k, $\delta_{k,\mathrm{tame}}$) or (wi_k, $\delta_{k,\mathrm{wild}}$), and the process continues. Collision detection takes place when a new distinguished point is recorded, such as using a hash table, or by searching a dedicated file containing distinguished points for a pair $\mathrm{ta}_k = \mathrm{wi}_k$ through a specific routine.

The Distinguished Point Method is examined as: First denote Θ as the fraction of elements called as group elements that meet the criteria for being distinguished points. When D is equal to D_f, it follows that $\Theta = 2^{-f}$. The typical duration required for each kangaroo to identify a distinguished point amounts to $1/\Theta = 2^f$. Consequently, on average, $2/\Theta$ jumps are executed following the convergence of the two kangaroos' paths, leading to the detection of their collision.

4 Proposed Method

4.1 Steps of the Kangaroo Algorithm

Take into account G as an n-dimensional finite cyclic group produced by the element α.The objective is to determine the element's discrete logarithm x in regard to the base α. In simpler terms, we seek an $x \in Z_n$ where $\alpha x = \beta$. The lambda algorithm provides a method for searching for x inside the range$[a, \ldots, b] \subset Z_n$. By setting $\alpha = 0$ and $n-1$, search through all possible logarithms.

Here's the algorithm broken down into steps:

1. Begin by selecting integers of set S and defining a pseudorandom mapping function f: $G \to S$.
2. Next, N *is* chosen as an integer and calculate group elements $\{x_0, x_1, \ldots, x_N\}$ *in* a sequence of according to:

$$x_0 = \alpha^b$$
$$x_{i+1} = x_{i\alpha}^{f(x_i)} \text{for } i = 0,\ 1\ldots,\ N-1$$

3. Compute

$$d = \sum_{i=0}^{N-1} f(X_i)$$

Then:

$$x_N = x_0\alpha^d = \alpha^{b+d}$$

4. Initiate the computation of a second sequence of group elements $\{y_0, y_1, \ldots\}$ as follows:

$$y_0 = \beta$$
$$y_{i+1} = y_i\alpha^{f(y_i)}, \text{for } i = 0,\ 1, \ldots,\ N-1$$

And generate a corresponding sequence of integers $\{d_0, d_1, \ldots\}$ in accordance with:

$$\text{d} = \sum_{i=0}^{N-1} \text{f(yi)}$$

Observe that:

$$y_{i+1} = y_i\alpha^{f(y_i)}, fori = 0, 1, \ldots, N-1$$

5. When either of the following circumstances occurs, stop computing terms of $\{y_i\}$ and $\{d_i\}$:

(A). For some j, $y_i = x_N$. The following results are obtained if the sequences x_i and y_i "collide" in this way:

$$x_N = y_j \Rightarrow \alpha^{b+d} = \beta\,\alpha^d j \Rightarrow \beta \Rightarrow \alpha^{b+d-dj} \Rightarrow x \equiv b + d - d_j(\bmod n)$$

(B). In the event that $d_i > b-a + d$, this indicates that the algorithm has been unsuccessful in locating x. Additional attempts can be made by altering the selection of S and/or f.

The kangaroo algorithm, unlike the rho method, can solve the "interval discrete logarithm problem" where $y{=}x^k$

4.2 Parallelization of the Kangaroo Method

In comparison to Pollard's Rho algorithm, one of the main advantages of the Kangaroo method is its ability to achieve linear acceleration through parallelization. This feat entails splitting up a single pair of "tame" and "wild" kangaroos into many pairs, creating kangaroo "herds," and then analyzing kangaroo crashes among various herds. Although collisions can still happen within the same herd, these "unproductive" collisions waste computer resources because they do not contribute to the final solution. Many efforts have been made to reduce or get rid of these "unproductive collisions."

However, the results of these efforts have either retained some "useless collisions," leading to performance degradation, or have required a fixed and predetermined number of parallel threads. The latter approach may not align with the reality of consumer hardware, which often faces unpredictable and inconsistent workloads. Assuming that only a few cores will be available under normal loads is reasonable, but this assumption cannot be extended to all cores. Nevertheless, it's plausible that one or more cores may become temporarily available during the program's runtime.

Consider a finite cyclic group denoted as G, with its generation attributed the group element g, and suppose that h is an element within G. In this context, there exists a positive integer x, known as the discrete logarithm of h with respect to the base g, denoted as $h = g^x$. It's worth noting that a unique case arises when $h = 1$, and in such instances, x is referred to as the order of g. The main objective is to determine the value of x.

The methodology involves introducing two kangaroos into the scenario: a tame kangaroo Ta, with an initial position at $t_0 = g^b$, and a wild kangaroo W_i, with an initial position at $w_0 = h$. With regard to the exponents pertaining to g, Ta commences its journey from the interval $[a, b]$ as the upper limit and Wi embarks from an undisclosed position represented by x. The initial distances from the origin are characterized by $\delta_0(\text{Ta}) = b$ for Ta and $\delta_0(W_i) = 0$ for W_i.

Let's examine a collection of jumps, denoted as $S = \{g^{s1},\ldots,g^{sr}\}$ $(s_i > 0)$, and a hash function $v: G \rightarrow \{1, \ldots, r\}$. These s_i exponents signify the distances travelled, and the intention is for these distances to be notably smaller than the interval's length, $b-a$. It is proposed, in accordance with Pollard's suggestion, that the powers of two, which starts from 2^0 and extending to a size that will be determined later, can be suitable choices for the s_i values. Now, let the trained kangaroo follow a predetermined path through the herd.

$$t_{j+1} = t_j * g^{sv(t_j)}, j \in N_0$$

During the computation of the path (t_j), monitor the distances $\delta j(Ta)$,

$$\delta_{j+1}(T_a) = \delta_j(T_a) + s_v\left(t_j\right), \mathrm{j} \in \mathrm{N}_0$$

Hence, $t_j = g^{\delta j(\mathrm{Ta})}$, for each $j \in N_0$. The tame kangaroo eventually stops and sets up a trap at its destination after making a particular amount of jumps, according to tM.

Then, the kangaroo in the wild departs down the trail.

$$w_{j+1} = w_j * g^{sv(w_j)}, j \in N_0$$

with the distances δ_j (W_i) given by

$$\delta_{j+1}(W_i) = \delta_j(W_i) + s_v\left(w_j\right), j \in N_0$$

Then

$$w_j = h * g^{\delta j(Wi)}(j \in N_0)$$

Regarding exponents of 'g,' with each step, the exact knowledge of 'Ta's location, have lack of information about the initial position 'x' of '$\mathrm{W_i}$'. This is the reason why 'W_i' is denoted as "wild". Following each jump, assess whether 'W_i' has encountered the capture point. If such an event transpires, for instance at 'w_N,' the equation '$t_M = w_N$' immediately provides a solution for '$g^x = h$,' specifically.

$$x = \delta_M(T_a) - \delta_N(W_i)$$

If 'W_i' evades capture by 'Ta' after a specific number of jumps, suspend the wild kangaroo, and subsequently, initiate a new wild kangaroo to resume the pursuit, commencing from a starting point with position '$w_0 = h * g^z$' with an initial distance '$\delta_0(W_i) = z$ (z small)'.

When 'W_i' eventually becomes ensnared by 'Ta,' it implies that the trajectories of the two kangaroos must have crossed paths at some earlier point in their journeys. From that juncture onward, their paths are indistinguishable. In such a scenario, envisioning two paths traced on a sheet of paper, one originating from the lower-left corner and the other from the lower-right corner, forms the shape of the Greek letter lambda (‘λ’). This elucidates why this method is also referred to as the lambda method.

5 Performance Analysis

In this paper, performance evaluation and an analysis and the scheme was specified, with a specific emphasis on communication overhead and computational expenses. The paper present benchmarks using the Cryptosystem with a $|p|$ value of 192 bits, resulting in a cipher text size of $2|p|$-*bits*.

5.1 Communication Overhead

This paper delves into user communication, wherein each user is responsible for transmitting messages containing cipher texts to others.

The communication overhead (Table 1) can be divided into three principal components:

Initialization phase, Construction phase, and Query phase.

Each of the user transmits their public key to neighbouring users, in the initialization phase. This operation necessitates $q{\cdot}2|p|$ bits, an average of neighbours is assumed as q per user. Every user sends vertex information which is encrypted, during the construction phase, using $q\ l{\cdot}2|p|$ bits. Additionally, the signature operation adds an extra $q{\cdot}l{\cdot}2|p|$ bits to the overhead.

In the query phase, each of the user disseminates messages to their neighbours, requiring q·p·5| | bits, primarily composed of cipher text and signature data. This propagation process is repeated K times, with K set to 8.

Table 1 Communication overhead

Protocols	Communication overhead (bits)
Initialization	483qn
Construction	6745qn
Query	5462qn

Table 2 Computational costs

Terms	Computational cost (ms)
C_{enc}	0.1320
C_{dec}	0.1412
C_{mul}	0.0012
C_{exp}	0.0353
C_{sig}	0.1236
C_{ver}	0.1538

5.2 Computational Costs:

The computational costs(Table 2) associated with various operations such as decryption (dec), encryption (C_{enc}), multiplication (C_{mul}), exponentiation (C_{exp}), signature generation (C_{sig}), and requalification validation (C_{val}) are as follows.

For decryption, utilize a giant-step Baby-step approach, limiting the plaintext space to 20 bits.

During the production phase, every user signs and encrypts the feature vector using l bits, resulting in associated costs of $l(C_{\text{enc}} + C_{\text{sig}})$ per user. Additionally, each user is required to calculate neighbouring influence, involving a cost of $q(2lC_{\text{mul}} + 2lC_{\text{exp}} + C_{\text{ver}})$.

Each user generates messages in query phase. In forward message generation phase, every user necessitates: $q(21C_{\text{enc}} + 21C_{\text{exp}} + C_{\text{ver}})$ for computing the original ciphertext and $q(C_{\text{sig}} + C_{\text{ver}})$ for the signature.

With an increasing number of iterations, the linear cost escalates due to heightened data transmission. However, this increase remains manageable until a significant number of users progress to the fourth iteration. In subsequent iterations of the backward phase, messages will begin to accumulate.

6 Conclusion and Future Work

The privacy preservation system for optimizing nearness queries in social networks employs a decentralized approach. By leveraging the homomorphic attributes, parameter unpredictability, and signature validity, the propagation of messages can securely and effectively convey influence data. Over multiple iterations, we can achieve the optimal nearness. In future research, our model takes into account the ambiguity of user uniqueness, providing a more rational outlook on social networks. Our research will also delve into other privacy challenges and additional social frameworks that consider specific user characteristics and proximity, while developing efficient countermeasures against various types of attacks.

References

1. Abdaoui A, Erbady A, Al-Ali A, Mohamed A, Guizani M (2021) Fuzzy elliptic curve cryptography for authentication in internet of things. IEEE Internet Things J (Early Access) 1–1
2. Meneghetti A, Sala M, Taufer D (2020) A new ECDLP-based PoW model. Mathematics 8(8)
3. Pollard JM (1978) Monte carlo methods for index computation $({\rm mod}\p)$. Math Comput 32:918–924. https://doi.org/10.1090/S0025-5718-1978-0491431-9
4. Hankerson D, Vanstone S, Menezes AJ (2004) Guide to elliptic curve cryptography 1st edn. Springer-Verlag, New York, pp 311 ISBN-10: 038795273X
5. Teske E (2014) On random walks for pollard's rho method. Math Comput 70(234):809–825. 2000 of pollard's rho attack on elliptic curve cryptosystems J Comput Sci 10(8):1575–1581
6. Liu S, Yao H, Wang XA (2018) Fast elliptic curve scalar multiplication for resisting against spa. Int J Comput Sci Eng 17(3):343–352
7. Boyd C, Hale B, Mjolsnes SF, Stebila D (2016) From stateless to stateful: generic authentication and authenticated encryption constructions with application to TLS, In: Proceedings cryptographers' track RSA conference. Springer, Cham, Switzerland, pp 55–71
8. Tyagi M, Manoria M, Mishra B (2019) A framework for data storage security with efficient computing in cloud. In: Proceedings international conference on advanced computing networking and informatics. Springer, pp 109–116.
9. Louw J, Niezen G, Ramotsoela TD, Abu-Mahfouz AM (2016) A key distribution scheme using elliptic curve cryptography in wireless sensor networks. In: Proceedings IEEE 14th international conference on industrial informatics (INDIN), pp 1166–1170
10. Almajed HN, Almogren AS, Altameem A (2019) A resilient smart body sensor network through pyramid interconnection. IEEE Access 7:51039–51046
11. Zuccherato R (2000) Elliptic curve cryptography support in entrust Entrust. Datacard, Ottawa, ON, Canada, Tech. Re. 1.0
12. Kanda G, Antwi AO, Ryoo K (2018) Hardware architecture design of AES cryptosystem with 163-bit elliptic curve. In: Advanced multimedia and ubiquitous engineering. Springer, Singapore, pp 423–429

Deep Learning Techniques for Scene Classification and Labeling Images

Polavarapu V. N Rishitha Chowdary, Mareedu Geethika, and G. Geetha

Abstract In the image processing domain applications, which includes image retrieval, autonomous navigation, and content recommendation, scene categorization is crucial. Deep learning procedures have demonstrated the tremendous skill in overcoming the obstacles associated with visual scene categorization. This work gives a thorough investigation of deep learning-based visual scene categorization. Convolutional neural networks (CNNs) are used in the suggested method to extract features and recognize hierarchical patterns. To leverage the discriminative characteristics found in diverse scene categories, we investigate a range of architectures, including cutting-edge pre-trained models. The usefulness of data augmentation approaches like geometric transformations, color, and intensity transformations to improve the model's generalization capacity is also examined in this work, particularly in situations when the availability of labeled training data is constrained.

Keywords CNN · ResNet · DenseNet · Data augmentation · Neural networks · Algorithms · Pooling · Fully connected layers

1 Introduction

Automated land use analysis, urban planning, and environmental monitoring all depend on image scene classification, a critical task in the fields of remote sensing and computer vision. It has important ramifications for a variety of applications, such as crisis management, resource allocation, and urban planning, to be able to classify photos into predetermined scene classifications. The goal of picture scene classification is to allow computers to determine whether an image portrays a beach, a forest, a cityscape, a kitchen, or any other identifiable scene type. Fig. 1 approach relies on

P. V. N. R. Chowdary (✉) · M. Geethika · G. Geetha
Department of IT, VR Siddhartha Engineering College, JNTU-Kakinada, Vijayawada, India
e-mail: pvn.rishitha.c@gmail.com

G. Geetha
e-mail: geetha.g@vrsiddhartha.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_40

the extraction of numerous visual features, patterns, and contextual information from photos to reach educated judgements about picture classification. The availability of high-resolution satellite photography has made it possible to classify scenes more precisely and in-depth as remote sensing technology develops.

The WHU RS-19 dataset, a well-known benchmark dataset that captures the subtle subtleties of several land use categories, is extensively employed during research to organize photographic settings, with a focus on its applicability.

1.1 Problem Statement

Identifying predefined labels for images in accordance with the events they depict is a critical task in remote sensing and computer vision. The classification of remote sensing data is essential for plethora of applications. The main goal is to apply several algorithms from the CNN domain on a well-used dataset to discover which method predicts the labels and provides the best accuracy.

1.2 Objectives

- The aim of this study is to assess the functionality of the ResNet-50 and DenseNet architectures within the framework of picture scene categorization.
- The project is to evaluate alternative CNN architectures using the WHU RS-19 dataset in order to understand their capacity to gather information, recognize variations between categories, and reliably predict labels.

1.3 Applications

- Accurate picture scene categorization has important ramifications for urban planning and development.
- The categorization of visual scenes based on remote sensing is a key component of environmental monitoring and conservation activities.
- In the event of a natural disaster, such as fire, prompt and accurate information is essential for effective response to emergency and recovery operations.
- The categorization of image scenes using cutting-edge architectural techniques offers significant advantages in agricultural planning and monitoring.

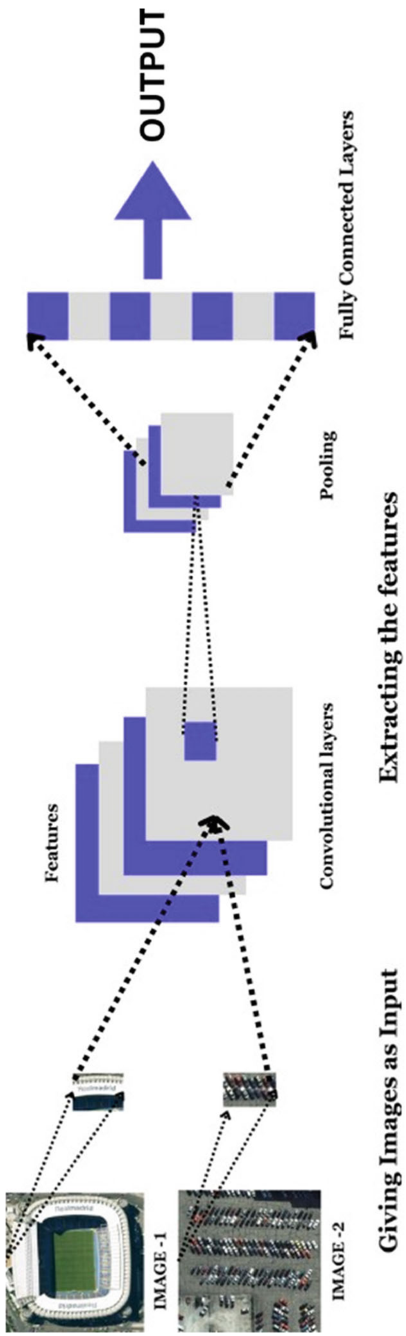


Fig. 1 Architecture of convolutional neural networks

2 Literature Survey

Sensing Image Scene Classification Based on Fusion Method (2021) [1]: Due of its numerous uses, remote sensing picture scene categorization is a popular study topic. Fusion-based techniques have gained a lot of interest more lately since they are seen to be an effective strategy for representing scene features. This study takes a different look at the fusion-based approach to remote sensing picture scene categorization.

Remote Sensing Image Scene Classification Based on Object Relationship Reasoning CNN (2017) [2]: Traditional CNN designs have performed admirably in image classification applications, it is possible that they fall short in capturing the nuanced connections between items in a scene. This is particularly pertinent for aerial imaging given that the sceneries may have complicated contextual relationships.

Remote Sensing Image Scene Classification: Benchmark and State of the Art (2020) [3]: It seems to be concentrated on giving an overview of the state of photographic landscape categorization using sensing techniques research and supplying a benchmark dataset for assessing the effectiveness of various classification techniques. Establishing a variety of scene categories and reviewing the most cutting-edge techniques are the main goals of the study.

Remote Sensing Image Classification: A Comprehensive Review and Applications (2022) [4]: The paper gives a thorough introduction and also an examination for the many methods, approaches, and developments in remote sensing image fusion. The goal of this study might be to present a comprehensive overview of the most recent approaches for fusing data in order to boost the accuracy and informative value of imagery.

2.1 Dataset

The ‘WHU RS-19’ dataset is developed by The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS) at Wuhan University, a high-resolution remote sensing dataset. The dataset includes numerous urban and suburban settings and offers 19 separate classes that reflect various land use types. It was created for surface area and land usage classification applications. The dataset acts as thorough baseline for assessing remote sensing picture scene classification algorithms, with tagged photos comprising residential, industrial, commercial, and natural environments. [5] The recently released public dataset WHU RS-19, which was gathered from Google Earth, comprises 950 photos with a resolution of 600 by 600 pixels and is uniformly dispersed over 19 different scene classifications.

2.2 Algorithms Used

A. ResNet

ResNet-50, or Residual Network 50, is the abbreviation for a potent and well-known deep neural networks topology generated for image categorization and other computer vision applications. ResNet-50 is an architecture that is made up of numerous convolutional layers connected by skip paths and global average pooling, making it deeper than typical networks at 50 layers. A residual block is a crucial thing in building component of residual neural networks (ResNet), sometimes known as skip connection. Fig. 2, design contributes to the solution of the vanishing gradient problem and allows for the training of very deep neural network. The previous layer's characteristic map or activation is utilized as an input to the residual block. Residual blocks are frequently made up of one or more convolutional layers that execute spatial modifications on the input feature map.

Activation Function: After each batch normalization process, ReLU (Rectified Linear Unit) activation functions are widely utilized. ReLU brings nonlinearity into the network, allowing it to understand complicated data correlations.

The skip connection (identity mapping) is the essential component of a residual block. This link lets the block's original input to bypass the convolutional layers and be added straight to their output.

[6] The feature vector from the CNN models is extracted using BaseNet. Next, we use two dense layers to generate two parallel attention branches. The attention probabilities are computed by the first attention branch and then applied to the feature vector by the second attention branch.

Using an 'ImageDataGenerator', which preprocesses pictures and employs data augmentation strategies to improve model generalization, this customized architecture is then trained on the dataset. Accuracy and loss measures are produced.

B. DenseNet

Following data preprocessing to guarantee homogeneity, [7] through autonomous learning, the network determines the weight of each feature channel. Based on this weight, it boosts the characteristics that are relevant to the job at hand and suppresses the features that are not. The channel-wise feature recalibration of the convolutional layer is explicitly modeled in the design.

In Fig. 3, Global Pooling Average structure and a layer of Dense with activation of softmax used to reduce dimension for creating class predictions are two significant changes. The specification of the optimizer (such as Adam), loss function (usually categorical cross-entropy), and evaluation measure (accuracy) is required after the prototypes were developed. During model training with an 'ImageDataGenerator', strategies for data augmenting are implemented to improve model generalization and resilience. The testing dataset is employed in the evaluation of the trained model, giving information on model's accuracy, precision, recall, and F1-score for various scene classes.

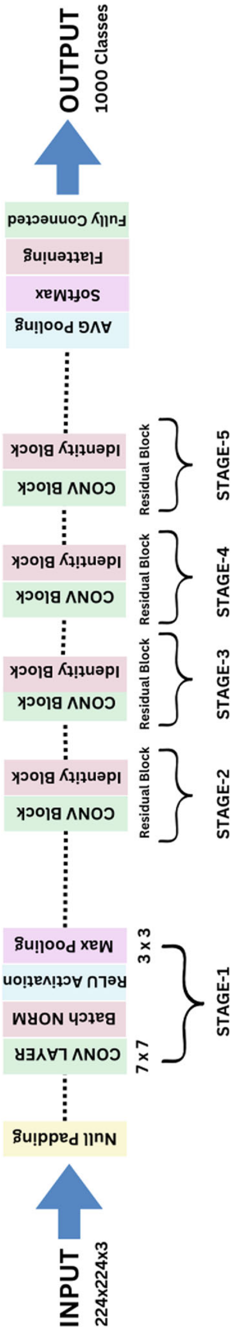


Fig. 2 Architecture of ResNet-50

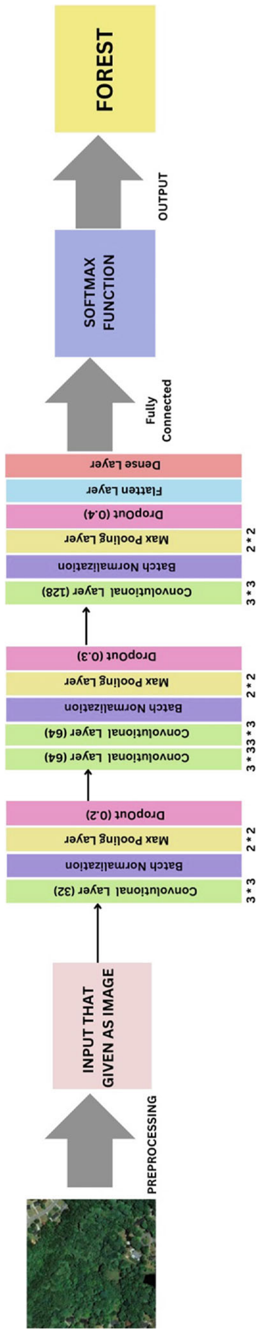


Fig. 3 Architecture of DenseNet

It includes bottleneck layers that, by employing 1 x 1 convolutions rather than 3 x 3 convolutions, lessen the computational load. To be able to regulate the network's complexity and manage the expansion of feature maps, DenseNet also adds transition layers.

3 System Architecture

Fig. 4 shows the pre-trained CNN (Convolutional Neural Network) model is the topology that trained on a large and diverse dataset, generally for tasks, such as image classification, prior to become accessible for other purposes. Transfer learning allows you to apply what a model has learnt on a big dataset (pre-training) to different but related task. CNN architectures, DenseNet and ResNet-50, may both be utilized as pre-trained models for transfer learning in plethora of computer vision applications. In the domain machine learning or deep learning setting, label prediction entails using trained model to make predictions or estimates about the target variable points.

Testing pictures with a trained CNN model entails a sequence of processes to guarantee that the model can make predictions on new, unseen data. The procedure guarantees that the machine is capable of generalize its learnt patterns to generate predictions on fresh, previously unknown data. Accurate label predictions are critical for the model's usefulness and are measured using measures like accuracy, precision, and recall, which ensure the model's capacity to generalize its learnt patterns to make educated judgements or predictions on real-world data.

4 Observation

The following part is concerned with the Model Evaluation parameters (Table 1).

A. ***Accuracy***

- The ResNet-50 model's stated test accuracy shows that the model does a good job of accurately identifying pictures compared with DenseNet.

B. ***Precision***

- The precision in ResNet-50 and DenseNet is obtained which means that the model accurately and precisely detects affirmative classes. Moreover, the ResNet-50 is slightly higher than the other one.

C. ***F1-score***

- The F1-score of ResNet-50 and DenseNet suggests that the model makes both accurate positive predictions and minimal false negatives.
- Eventually, above these two F1-score for the algorithm DenseNet is higher when it compared to the ResNet algorithm.

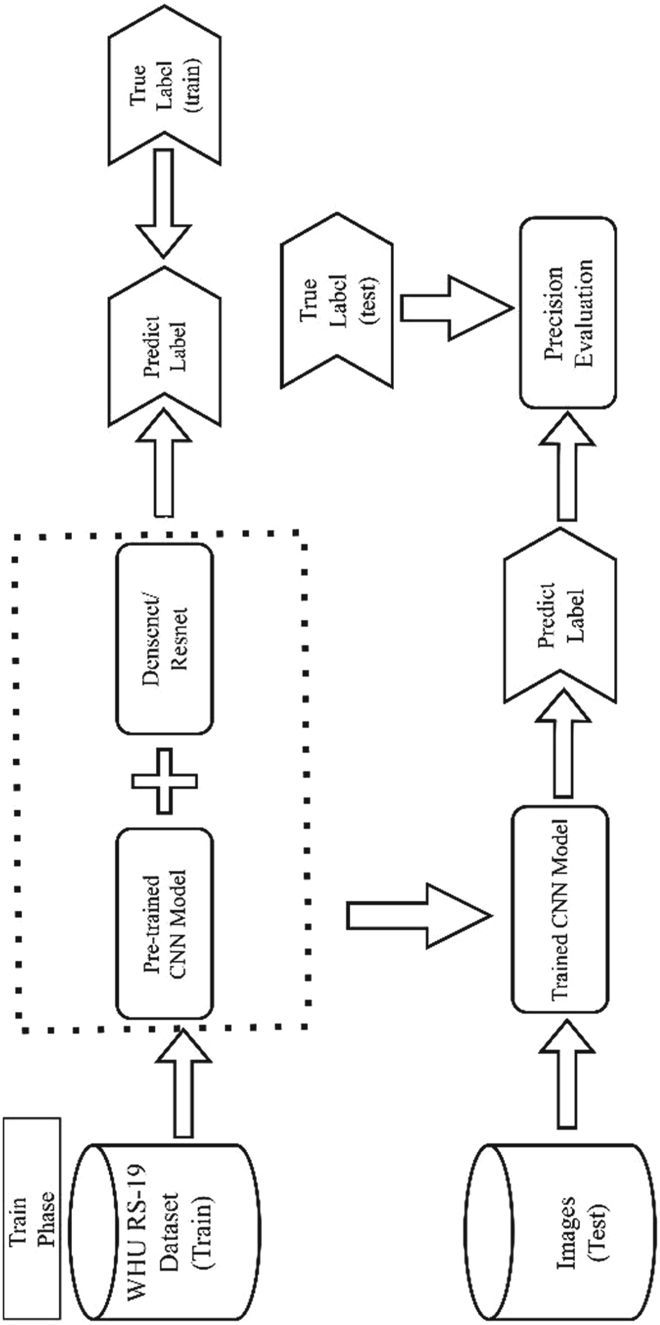


Fig. 4 System architecture diagram

Table 1 Model evaluation parameters of ResNet-50 and DenseNet

Metrics	Formula	ResNet	DenseNet
Precision	TP/(TP + FP)	0.9562	0.9512
$F1$-score	2(Precision*Recall)/(Precision + Recall)	0.9461	0.9512
Accuracy	(TP + TN)/Total	0.9473	0.8947

D. ***Confusion Matrix***

- The numbers in the matrix represent how many times the model correctly or incorrectly predicted each combination.
- Collectively, these assessment measures show that the ResNet-50 model, which was developed using the WHU RS-19 dataset, is effective in classifying scenes.

5 Results

(a). ResNet-50

The training procedure was continued for 10 epochs, and the result was an impressive test accuracy of 95%. This progress over 10 epochs demonstrates how ResNet-50 is adaptable. During model training, an epoch is full loop across dataset and is point at which model has learnt from sample. The ResNet-50 model, in Fig. 5 trained over 20 epochs, attained a remarkable test accuracy of 94.74%.

(b). DenseNet

The accuracy of DenseNet is 89.47% during a training period of 10 epochs. This finding suggests that model iterations to perform its best or it would benefit from fine-tuning.

The DenseNet architecture achieved a test accuracy of 92.63% after 20 training iterations. This shows that the model's performance plateaued and did not demonstrate significant advances after this point, but still displaying impressive accuracy.

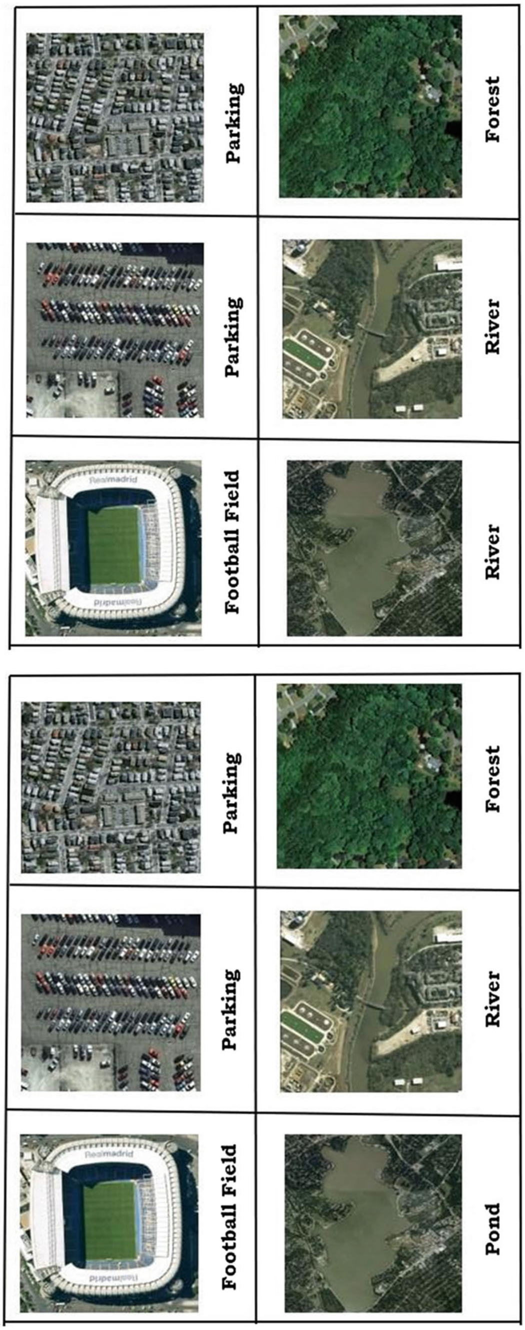


Fig. 5 Results of Resnet-50 and Densenet

6 Conclusion

The comparison shows that ResNet-50 outperformed DenseNet consistently throughout many epochs, attaining better accuracy levels. In particular, ResNet-50 demonstrated an impressive increase in accuracy from 94.74% to 95% with 20 and 10 more epochs, respectively. In contrast, when the training period was shortened to 10 epochs, DenseNet's accuracy decreased. These results demonstrate ResNet-50's ability to adapt and improve its predictions over a protracted training period, highlighting its viability for tasks.

7 Future Work

Video classification presents additional potential and problems rather than the image classification. Videos demonstrate how things change over time by playing string of images one after another. We can figure out how to manage the transitions between frames in videos to improve this. We may employ certain methods to comprehend how sceneries evolve throughout time. Model can learn more effectively if it considers several points of view, much as individuals do when they take on viewpoints.

References

1. Zhang Z, Yin L, Yang P, Mao K, Liu Q (2021) Remote sensing image scene classification based on fusion method. J Sens 6659831https://doi.org/10.1155/2021/6659831
2. Li Z, Wu Q, Cheng B, Cao L ,Yang H (2022) Remote sensing image scene classification based on object relationship reasoning CNN. In: IEEE geo- science and remote sensing letters, vol 19, pp 1–5, Art no. 8000305 https://doi.org/10.1109/LGRS.2020.3017542
3. Cheng G, Han J, Lu X (2017) Remote sensing image scene classification: benchmark and state of the art. IEEE Proceedings 105(10):1865–1883. https://doi.org/10.1109/JPROC.2017.2675998
4. Mehmood M, Shahzad A, Zafar B, Shabbir A, Ali N (2022) Remote sensing image classification: a comprehensive review and applications. Math Probl Eng https://doi.org/10.1155/2022/5880959
5. Tu B, Kuang W, He W, Zhang G, Peng Y (2020) Robust learning of mislabeled training samples for remote sensing image scene classification. IEEE J Sel Top Appl Earth Obs Remote Sensing 13:5623–5639. https://doi.org/10.1109/JSTARS.2020.3025174
6. Hassan SM, Maji AK (2024) Pest Identification Based on Fusion of Self- Attention With ResNet. IEEE Access 12:6036–6050. https://doi.org/10.1109/ACCESS.2024.3351003
7. Zhang K, Guo Y, Wang X, Yuan J, Ding Q (2019) Multiple feature reweight DenseNet for image classification. IEEE Access 7:9872–9880. https://doi.org/10.1109/ACCESS.2018.2890127

Student Feedback Website for Coaching Classes

Supriya Telsang, Tanishka Jagtap, Jai Ughade, Aastha Jain, Kriya Jain, and Nikhil Jain

Abstract Parents and students can make educated decisions about the best coaching classes for their unique needs by consulting a student feedback website for coaching classes like JEE and NEET. Such a website has many benefits, including raising student engagement and accountability for coaching institutions while also enhancing teaching quality. Students can contrast different coaching facilities based on things like price, staff, facilities, teaching style (online vs. offline), and technology used. This website enables students to browse through a wide variety of coaching options and find the best fit for their academic objectives by using real reviews and evaluations. Additionally, student feedback encourages coaching institutes in education to be transparent and holds teachers responsible for their teaching strategies. This website helps students feel more connected to one another by giving them a place to share their experiences. It enables them to interact, share advice, and support one another throughout their academic careers. This improves the learning process overall and fosters a supportive and cooperative environment within coaching institutes.

Keywords Coaching classes · CodeSandbox · Express.js · JEE · NEET · Node.js · React.js · Reviews · Web development · Website

S. Telsang · T. Jagtap (✉) · J. Ughade · A. Jain · K. Jain · N. Jain
Vishwakarma Institute of Technology, Pune, Maharashtra, India
e-mail: tanishka.jagtap22@vit.edu

S. Telsang
e-mail: supriya.telsang@vit.edu

J. Ughade
e-mail: jai.ughade22@vit.edu

A. Jain
e-mail: aastha.jain22@vit.edu

K. Jain
e-mail: kriya.jain22@vit.edu

N. Jain
e-mail: nikhil.jain22@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_41

1 Introduction

Competitive tests, such as JEE and NEET, can greatly impact students' futures by providing access to higher education and promoting skill growth. Proper preparation materials, counseling, and instruction are crucial for high school students to pass these tests. However, a lack of guidance and understanding can lead to failure. This is why choosing the best coaching program for competitive exams like JEE and NEET is crucial. This website provides testimonials from students who have taken coaching programs, helping parents and students choose the best facility for their needs. Feedback from students encourages coaching and educational settings to improve over time, fostering an atmosphere of accountability and transparency. This feedback-driven approach encourages institutions to provide the best training and assistance by fostering an environment of responsibility. Sindhu and Badar [1] by providing information about administration, fees, and infrastructure, student reviews promote transparency and help parents and students make informed decisions about their coaching program.

The research proposes an algorithmic approach to extract features and orientations from accessible literature for teacher evaluations. Six aspect categories were identified: teaching pedagogy, knowledge, assessment, experience, behavior, and general. Domain experts assisted in developing the model, which uses LSTM aspect extraction and polarity identification. The model's efficacy was tested with various word embeddings, with encouraging results when using the domain [2].

The project aims to develop automated, customized, and personalized feedback systems for ICT students learning professional ICT skills online and in the classroom. The focus is on teaching hard skills like programming, modeling, and systems architecture from an engineering education perspective. The project will create a comprehensive ICT tool with business, domain, and data models that can be integrated into various e-learning platforms at both course and program levels [3].

This paper presents OFES, an online tool for providing formative feedback, aiming to encourage students to engage in the process. Excellent formative feedback should be timely, motivating, customized, manageable, and closely linked to evaluation standards. The efficacy of the communication plan can be assessed using the quality attributes of the feedback [4].

The evaluation of teachers is a key component of enhancing the quality of instruction in medical schools. Although Western systems have been successful, there is rising interest in adopting student input as a teacher evaluation system in our nation. With the implementation of student feedback as a system for evaluating teachers and supporting faculty development, this project seeks to improve the quality of instruction [5].

The study proposes an architecture for e-learning platforms that incorporates an intelligent teacher for adaptive feedback. The teacher adjusts feedback based on students' cooperative conduct, enhancing performance and engagement. The system could be improved with artificial techniques like supervised and reinforcement learning to bridge the instructor gap and reduce student departures [6].

The introduction is covered in the first section, and the methodology is covered in the second section. The novelty aspects of the project are discussed in the third section. Results and discussions are included in the fourth section, which is followed by an explanation of the procedures in the fifth section, a discussion of the project's future scope in the sixth section, and a conclusion in the seventh section. There are references in the eighth section.

2 Methodology/Experiment

A. **Flow Chart**

B. **Tools and Technologies**

1. CodeSandbox—In a web browser, you can write, edit, and run code using the online code editor and development environment known as CodeSandbox.
2. Firebase—Google provides a powerful backend-as-a-service (BaaS) platform like Firebase that offers a range of services and tools, such as authentication and real-time database that can be utilized in the development of a website.
3. React.js—A JavaScript library called React.js is used to create user interfaces. It enables you to efficiently manage the state of your application and create reusable UI components.
4. Node.js—You can run JavaScript code on the server side using Node.js, a JavaScript runtime environment. It is effective for developing scalable and high-performance web applications because it offers an event-driven, non-blocking I/O strategy.
5. Express.js—A quick and simple web application framework for Node.js is called Express.js. By offering a collection of capabilities and middleware, it makes the process of developing web apps and APIs simpler (Figs. 1, 2 and 3).

3 Novelty Features of Application

1. Exclusive information on competitive coaching exam platforms, both online and offline.
2. Comparison of teaching faculty, fees, and overall results of previous years.
3. Helps students make informed decisions based on their budgets and personal preferences.
4. Personal reviews from students who have studied in those coaching classes.
5. Allows students to find the best coaching class suited to their needs.

Fig. 1 Flow chart of website

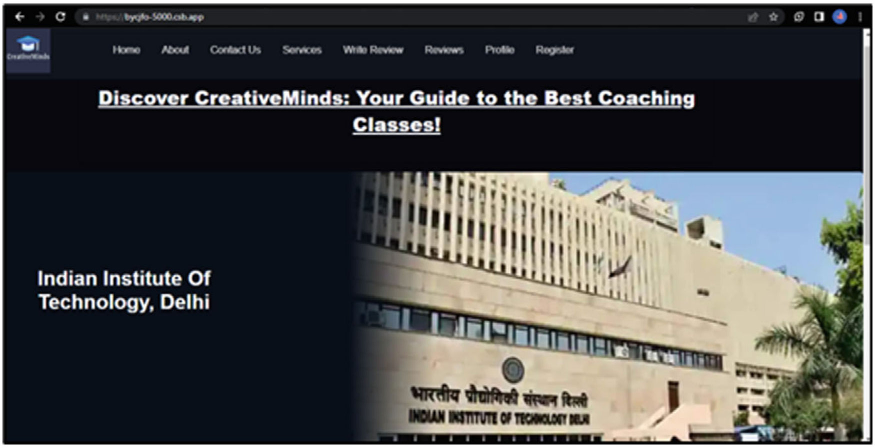


Fig. 2 Homepage of website

4 Results and Discussion

The website for student reviews of coaching institutes has received positive feedback from students, stating it is trustworthy and useful. The platform offers unbiased feedback from real students who have visited the coaching center, making it easy for students to research and categorize information. Students also praised the website's user-friendly interface and comparison tools, which allow them to make impartial

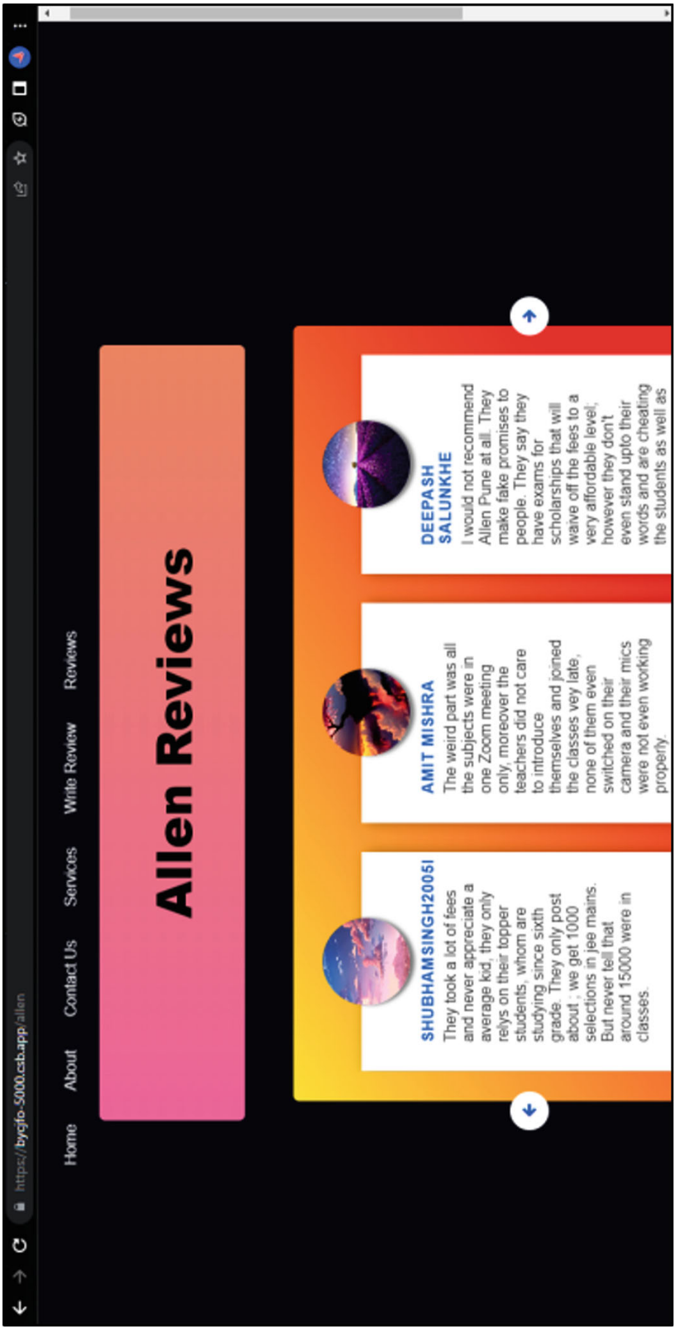


Fig. 3 Allen career institute reviews page of website

assessments based on factors like location, courses, and ratings. The website significantly increased students' capacity for decision-making, allowing them to choose the coaching institute that best fits their needs and preferences. This highlights the importance of creating a credible platform for student reviews, which can improve the coaching institute's responsibility, openness, and overall excellence.

5 Procedure

The purpose of the website is to provide a platform for students to share and access reviews about coaching classes. It aims to assist students preparing for exams like JEE and NEET in making informed decisions about coaching classes. The target audience primarily consists of students seeking coaching classes for exam preparation. The website's scope is focused on providing reviews and ratings for coaching classes according to their will. To gather requirements, research is conducted to identify essential features and functionalities for the website. This includes a homepage where users can select an exam to view reviews about nearby coaching classes, sign-in and sign-up pages, user profiles, a page for users to write reviews, and a page for viewing reviews.

The user interface design involves creating wireframes or mockups that are intuitive, user-friendly, and visually appealing. Responsive design principles are incorporated to optimize the website for different devices. Web development frameworks like CodeSandbox are used for designing websites. Frontend Technologies like React.js and languages like HTML, CSS, and JavaScript are used for designing the frontend of a website. In backend development, a suitable programming language like JavaScript and a framework are chosen. A database management system of Firebase is implemented to store user information, coaching class details, and reviews. The necessary database tables are designed, and relationships between them are established.

User authentication is implemented with a secure system for users to sign in and sign up. Password hashing and encryption techniques are used to protect user data. The functionality for password reset is also included. Review submission and viewing pages are developed, allowing users to submit their reviews about coaching classes and implementing validation checks for review accuracy and reliability. A page is created for users to view and filter reviews based on their preferences, such as exam type or location. Search and filtering functionality is implemented to enable users to search for coaching classes based on various criteria, such as location, subject, or ratings.

Comprehensive testing is performed to identify and resolve any bugs or issues. Functionality is tested on different browsers and devices to ensure compatibility and responsiveness. User acceptance testing is conducted to gather feedback for necessary improvements. The deployment and maintenance stage involves selecting a reliable hosting provider and monitoring the performance, security, and scalability of the website. Regular updates and maintenance are performed based on user feedback and demands. Continuous improvement is achieved by collecting user feedback and reviews to enhance functionality and user experience and incorporating new features and enhancements based on user demands and industry trends.

6 Conclusion

Creating a website for reviewing coaching classes is a great way to enhance student interaction in the classroom, and platforms offer a central location for information about different coaching schools, the courses they offer, the faculty members who teach them, the fee schedules, and student reviews. Students are better able to select the coaching classes that will suit their requirements and preferences by having access to this information. Student websites enable the sharing of personal experiences and reviews by participants in coaching programs. This peer-driven feedback provides illuminating details regarding the standard of instruction, the effectiveness of study resources, infrastructure, and the overall learning experience. Student websites encourage transparency and responsibility among coaching classes. Students are able to express their opinions in the open forum, which instills a sense of obligation in coaching centers to uphold standards of excellence and address any difficulties raised by students.

References

1. Sindhu I, Badar K (2019) Aspect-based opinion mining on student's feedback for faculty teaching performance evaluation. IEEE Access
2. Shvets O, Murtazin K (2020) Providing feedback for students in e-learning systems: a literature review, based on IEEE explore digital library. In: 2020 IEEE global engineering education conference (EDUCON)
3. Hatziapostolou T, Paraskakis I (2010) Enhancing the impact of formative feedback on student learning through an online feedback system. Electron J E-Learn
4. Hussain M, Khan S (2016) Students' feedback: an effective tool in teachers' evaluation system. Int J Appl Basic Med Research
5. Hassan MA (2019) An adaptive feedback system to improve student performance based on collaborative behavior IEEE Access
6. Alderman L, Towers S, Bannah S (2012) Student feedback systems in higher education: a focused literature review and environmental scan. Qual High Educ

Data Mining and Machine Learning Techniques in Knowledge Discovery

V. V. Bhavani, Tahseen Jahan, A. Hemalatha Reddy, Jalapala Sinjini, M. Priyadharshini, and V. Indumathi

Abstract Knowledge discovery from large datasets has become increasingly important in various fields such as healthcare, finance, and marketing. Data mining and machine learning techniques play a crucial role in uncovering meaningful patterns and insights from complex data sources. This paper provides a comprehensive overview of the state-of-the-art data mining and machine learning techniques employed in knowledge discovery processes. It covers a wide range of methodologies including classification, clustering, association rule mining, anomaly detection, and regression, among others. Moreover, it discusses the application of these techniques in real-world scenarios and examines their strengths, limitations, and challenges. The review also highlights recent advancements, emerging trends, and future research directions in the field of knowledge discovery through data mining and machine learning.

Keywords Data mining · Machine learning · Knowledge discovery · Classification

V. V. Bhavani (✉) · T. Jahan · A. Hemalatha Reddy · J. Sinjini · M. Priyadharshini · V. Indumathi
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: bhavaniv@cmritonline.ac.in

T. Jahan
e-mail: tahseen048@cmritonline.ac.in

A. Hemalatha Reddy
e-mail: hemareddy@cmritonline.ac.in

J. Sinjini
e-mail: sinjini@cmritonline.ac.in

M. Priyadharshini
e-mail: priyadharshini@cmritonline.ac.in

V. Indumathi
e-mail: indu11mathi@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_42

1 Introduction

Knowledge discovery relies heavily on data mining and machine learning, which provide robust methods for gleaning useful insights from massive datasets. Today, in the age of big data, businesses in all kinds of industries are trying to figure out how to use all the data they must make better choices, streamline their operations, and get an edge over the competition. Data mining is the process of sifting through large databases in search of previously unseen correlations, patterns, or trends [1]. Machine learning, in contrast, is concerned with creating algorithms that let computers learn from data and perform tasks such as prediction and decision-making without human intervention.

Knowledge discovery is a key objective of data mining and machine learning, which aim to convert raw data into useful information [2]. Researchers and practitioners can find hidden patterns and important insights in complicated datasets using these strategies, which apply a variety of algorithms and techniques including clustering, classification, regression, and association rule mining. For instance, healthcare providers may enhance diagnosis, forecast illness outcomes, and create personalised treatment regimens by analysing patient information and medical pictures using data mining and machine learning techniques.

In addition, several sectors rely heavily on data mining and machine learning, such as manufacturing, retail, marketing, and finance. These methods find use in the financial sector for things like algorithmic trading, risk assessment, and fraud detection. For marketers, they pave the way for sentiment analysis, client segmentation, and targeted advertising. Demand forecasting, inventory management, and consumer behaviour research are some of the retail applications of data mining and machine learning. Similarly, similar approaches facilitate optimisation of the supply chain, quality control, and predictive maintenance in the industrial sector.

Knowledge discovery that incorporates data mining and machine learning has several advantages, such as better decision-making, more productivity, and new ideas [3]. Organisations may drive strategic objectives and company progress by identifying hidden patterns and linkages inside data. The use of sophisticated algorithms is certainly important, but high-quality data, domain knowledge, and outcomes that are easy to understand are just as crucial to the success of knowledge discovery. To fully utilise data mining and machine learning for knowledge discovery, it is essential for data scientists, domain experts, and decision-makers to collaborate effectively.

A. Importance of Knowledge Discovery

Knowledge discovery is of paramount importance in the realms of DM and ML techniques. It serves as the cornerstone for unlocking insights, understanding complex phenomena, and making informed decisions across various domains. Here's why knowledge discovery holds such significance in this context:

- Insight Generation: Knowledge discovery enables the extraction of valuable insights from large and complex datasets that may otherwise remain hidden. By employing DM and ML techniques, patterns, trends, and relationships within the data can be identified, providing valuable insights into various phenomena.

- Decision Support: In many fields, data-driven decision-making has become indispensable. Knowledge discovery facilitates the process of decision support by providing decision-makers with relevant information and insights derived from data analysis. This allows for more informed and effective decision-making processes.
- Predictive Analytics: Knowledge discovery through data mining and machine learning enables predictive analytics, wherein models are built to forecast future trends or outcomes based on historical data. These predictive models are valuable for anticipating changes, mitigating risks, and seizing opportunities in diverse domains.
- Personalization and Customization: In sectors like marketing, healthcare, and e-commerce, personalized experiences are highly valued. Knowledge discovery techniques enable the customization of products, services, and recommendations based on individual preferences, behavior patterns, and historical data, thereby enhancing user satisfaction and engagement.
- Innovation and Research: Knowledge discovery fuels innovation and drives research by uncovering new insights, relationships, and opportunities within datasets. By analyzing data through the lens of data mining and machine learning techniques, researchers can identify novel patterns, validate hypotheses, and generate new knowledge in various fields.
- Quality Improvement: In industries such as manufacturing and healthcare, knowledge discovery aids in quality improvement initiatives. By analyzing data related to production processes, patient outcomes, or service delivery, organizations can identify areas for improvement, optimize processes, and enhance overall quality.

Knowledge discovery plays a pivotal role in maximizing the value of data mining and machine learning techniques [4]. It empowers organizations to derive actionable insights, drive innovation, make informed decisions, and gain a competitive edge in today's data-driven world.

2 Literature Review

According to [5], the need for new analytical methods beyond standard statistical approaches to uncover new information from data mining arose from the demand of huge data. This led to the multidisciplinary area of knowledge discovery and data mining emerging. They claimed that this new method was a deductive and inductive dialectic research procedure. To handle causal heterogeneity and enhance prediction, the data mining technique supposedly automatically or semi-automatically took into account a greater number of joint, interacting, and independent predictors. In place of questioning the traditional model-building methodology, it was emphasised that it was a valuable adjunct that enhanced model goodness-of-fit, unearthed meaningful hidden patterns in the data, detected nonlinear and non-additive effects, shed light on data developments, methodology, and theory, and enriched scientific discovery. When the explicit model structure was ambiguous and high-performing algorithms were hard to achieve, machine learning learned and improved from data to construct models and algorithms. Rumour has it that the latest innovation is a model that combines explanation and prediction by combining this new paradigm of predictive modelling with the traditional method of parameter estimation regressions.

In [6], noted that despite an abundance of data, understanding has been severely lacking in biomedical research. Among the pressing issues that they addressed was the need to optimise workflows, processes, and guidelines in order to boost capacity, decrease costs, and improve efficiencies; the necessity to combine massive volumes of structured, semi-structured, weakly structured data; and information overload. Their study only offered a cursory introduction to integrative and interactive methods for data mining and knowledge discovery. Specifically, they highlighted how the user should be involved in the "interactive" information discovery process for the most advantage. Among the most pressing issues they highlighted was the need to create and implement new techniques, algorithms, and tools for handling complicated biomedical data in order to find testable hypotheses and construct practical models. The HCI-KDD approach, which aims to support human intelligence with machine intelligence, is characterised as a synergistic combination of methodologies and approaches from two areas: Knowledge Discovery & Data Mining (KDD) and Human–Computer Interaction (HCI). It was presented as providing ideal conditions for solving these challenges. The growing complexity and size of data sets, commonly referred to as "Big Data," has left little room for medical professionals or biological researchers to keep up, highlighting the critical need for integrated and interactive machine learning solutions.

In [7], discussed the use of data mining, a technique strongly associated with knowledge discovery in data science and databases, to extract valuable information and identify patterns from frequently massive data sets. Using data from higher education and identifying the important characteristics to be considered, the researchers developed models based on machine learning algorithms to forecast student retention at different levels. They said that they used this data to aid in the quest for new understanding. In their first, second, and third years of school, they got models that were more than 80% accurate in predicting student retention across all three levels. They claimed that these models helped them accurately forecast the level at which dropout happened. Their investigation made use of a variety of machine learning methods, including DT, LR, k-NN, NB, RF, and SVM. Among them, random forest yielded the best results. Unprecedented in prior research on education, they found that the neighbourhood poverty index and secondary educational score were significant predictive factors. As a result of being able to anticipate dropout based on student data, these institutions were able to take preventative measures, avoiding dropouts, and the study's findings regarding the validity of dropout assessment at various levels applied to higher education institutions worldwide with comparable conditions to the Chilean case. They found that the algorithms performed better after balancing the majority and minority classes in the case study.

They mentioned that data mining (DM) and machine learning (ML) approaches might be highly useful in addressing industrial issues in this regard. To give a general idea of how machine learning techniques may be used to create intelligent manufacturing systems, the authors of the research conducted a thorough literature analysis. According to their paper, they categorised the prior ML research and manufacturing advancements into four broad categories: scheduling, monitoring, quality, and failure. Topics covered included clustering, classification, and regression as well as support vector machines and neural networks as algorithms, ensemble learning and deep learning as learning types, and accuracy and mean absolute error as performance metrics. Their explanation of the key phases of the knowledge discovery in databases (KDD) process for industrial applications was also rather thorough. On top of that, they provided some data on the present situation from several angles. In addition, they detailed the benefits of applying machine learning techniques to manufacturing, provided solutions to specific problems, and suggested avenues for further study [8].

In [9], the tremendous growth in data processing capabilities and the massive collection of data in published literature and databases have led to a dramatic surge in the application of machine learning (ML) in catalysis in the past few years. In order to discover patterns, build prediction models, and derive heuristic guidelines for the future, they used several ML approaches to examine data that was either created in-house or obtained from external sources. According to them, the purpose of their communication was to summarise the fundamental concepts, common tools, and applications of ML in catalysis, as well as to review the works that included knowledge discovery in catalysis using ML approaches.

In [10] energy consumption forecasting has been a popular area of study. They acknowledged that it had been a difficult research challenge to extract knowledge about power usage from electricity records' multi-dimensional data streams (MDDS). Based on the MDDS, they suggested a new model for discovering electrical knowledge that makes use of clustering and machine learning. They announced the proposal of CAC-WOA, a technique that combines context-aware clustering with WOA, to extract predictive characteristics from electrical MDDS and then use those features to make predictions. According to their paper, a predictive model based on a WOA-based artificial neural network (ANN) was constructed employing these predictive properties. They claimed that the error rates and prediction accuracy were both enhanced by employing the modified ANN approach that made use of the WOA algorithm. Results from experiments utilising data sets on power use that are available to the public demonstrated the effectiveness of the CAC-WOA model, they said. By comparing CAC-WOA to state-of-the-art methods, they found that prediction time was lowered by 11.31% and total prediction accuracy was increased by 3.27%.

In [11], modern learning models are able to uncover complicated patterns in data and achieve high levels of accuracy. On occasion, patterns found by these models were impossible for a person to spot on their own. They emphasised the significance of conveying the found patterns and knowledge in a way that humans can understand. They put up a plan on how to use actual agricultural datasets for information finding. The proposed framework made use of Deep Learning (DL) models to learn effective features and patterns from tabular datasets; an especial evolutionary algorithm to extract knowledge from fuzzy expert systems; a fuzzy expert system to interpret the discovered knowledge; and an integrated gradient method to represent the black-box models as an interpretable model. According to their research, tests were conducted using a real-world tabular dataset pertaining to agriculture in order to assess the suggested framework's ability to extract information on productivity. Using the found expert system to provide appropriate irrigation treatments on the chosen dataset, they presented an optimisation approach to validate the learned information. The DSSAT farm simulator was brought up as a tool for checking the accuracy of newly found information. The results revealed that the new information might reduce water usage by 30% and increase productivity by 30%, according to their research. They also noted that there was a correlation of more than 0.8 between the results of the planned therapy and the newly acquired information.

They claimed the rising need for tools to extract meaning from massive data sets was the driving force for its meteoric rise in popularity. Government entities, financial institutions, insurance providers, businesses, and the World Wide Web reportedly provide such data everyday. They said that the proliferation of electronic devices like scanners, digital cameras, bar codes, etc., was the cause of this surge. In one case, they brought up, there was a data repository (database, warehouse, etc.) that included abundant but difficult-to-analyze data sources. According to them, this resulted in increased demand for DMKD technology upgrades from the government, businesses, and industry. What was required, they said, was an easy-to-understand approach to mining the data for insights. In their chapter, they presented an integrated DMKD process model that makes use of technologies such as XML, PMML, SOAP, UDDI, and OLE BD-DM. They claimed that these technologies allowed for the design of semiautomated, user-friendly DMKD models that could be used to build knowledge repositories and facilitate communication between various databases, data mining tools, and knowledge repositories. The DMKD duties might be automated and integrated with their help. They detailed the technologies that make up a DMKD process model with six steps [12].

The promise of sensorial acquired process data in conjunction with machine learning (ML) algorithms was underutilised in practical manufacturing applications, according to [13], who noted that this combination was essential for overcoming the obstacles posed by contemporary production systems. They claimed that models for systematic procedures to extract information from data were available in the literature within this setting; one such model was the CRISP-DM model, which became the de facto standard for data mining in business settings. They brought up the fact that these models failed to include the sensory obtained data's properties and the boundary conditions of manufacturing processes while trying to derive information from them. Hence, they introduced a new procedural model for engineering applications that uses time series and image data to uncover knowledge (KDT-EA). With a heavy emphasis on data gathering, data preprocessing, and data transformation to produce trustworthy input data for ML models predicting the real status of manufacturing processes, they claimed that a comprehensive perspective of knowledge discovery in manufacturing processes became possible. They said that the process model helped business operators with things like choosing preparation methods based on imposed disruptions, establishing an appropriate measurement chain to acquire high-quality data, and so on. They went on to say that it laid the groundwork for inline monitoring of products and proposed data transformation techniques that might reduce data volumes without sacrificing informative value. In order to measure the advantages of KDT-EA and how its phase affected the quality of the information provided, they used the innovative procedure model on a sheet metal forming equipment for inline wear detection.

According to [14], society's technological progress has led to a bigger generation of knowledge. Internet users were able to easily access and extract this data, according to their statement, leading to the pursuit of knowledge discovery by automatic means. Data mining, according to their research, sought to identify trends, profiles, and patterns in massive amounts of data for which various learning approaches were accessible. They brought up the fact that the algorithms used to do these jobs primarily originated in the early 1900s and formed the backbone of these modern technologies, therefore picking the right approach relied on the intended outcome and the available data. The authors of the study said that they wanted to demonstrate how data mining algorithms and software have progressed over the past several years, as well as how these methods have been used in scientific research. Their rationale for using the systematic literature review approach was that it was a methodical way to find, assess, and understand previous studies in a certain area. According to their report, they compared and analysed the top software, including Alteryx, TIBCO Data Science, RapidMiner, and WEKA. They also detailed the algorithms and techniques that were currently trending in machine learning and their capabilities for data mining processes.

In [15], Research involving actual or hypothetical drug exposure to people was included in their report. They found twenty distinct knowledge discovery methods, with the majority coming from the field of machine learning (66/72; 91.7% of the total) and a total of seventy-two original publications and five reviews found using Ovid MEDLINE. It was noted that the three most common tasks that machine learning methods had been used to solve in the reviewed literature were classification/regression (44/72; 61.1%), classification/regression + model optimisation (13/72; 18.0%), and classification/regression + features selection (12/72; 16.7%) (Table 1).

They stated that their paper offered a statistical analysis of the primary algorithms and methods used to enhance manufacturing processes in the last 20 years. The analysis was divided into four main themes: Quality, Scheduling, Monitoring, and Failure. The paper began by reviewing prior ML research and manufacturing advancements, and then moved on to a detailed examination of current manufacturing problem solutions from various angles. The topics covered included different types of tasks (classification, regression, clustering), algorithms (NN, SVM learning types (Deep Learning (DL), ensemble learning), and performance indicators (mean absolute error, accuracy). There was also a brief description of the benefits of machine learning in manufacturing, as well as the key steps of the database knowledge discovery (KDD)

Table 1 Comparative reviews

Author	Learning model	Knowledge discovery process	Methodology	Research area	Findings
[5]	Machine learning	Emergence of interdisciplinary field from necessity of big data, dialectic research process (deductive and inductive)	Formulation of models based on ML algorithms for student retention prediction	Higher education	Models with > 80% accuracy predicting student retention, identification of important predictive variables not reported previously in educational studies
[6]	Machine learning	Interactive and integrative solutions, HCI-KDD approach, incorporation of end user in knowledge discovery process	Description of challenges and methodologies for integrating and interpreting complex biomedical data	Biomedical research and clinical practice	Improvement in knowledge discovery and data mining through interactive and integrative approaches
[7]	Machine learning	Formulation of models for predicting student retention, using higher education data and specifying relevant variables involved in modeling	Prediction of student retention at various levels during first, second, and third years of study	Higher education	Improved prediction of dropout and identification of significant predictive variables
[8]	Machine learning	Literature review on application of ML techniques in manufacturing, discussion on ML approaches and algorithms for improving manufacturing processes	Overview of ML studies and recent advances in manufacturing	Manufacturing	ML techniques can significantly contribute to overcoming challenges in manufacturing
[9]	Machine learning	Review of works involving knowledge discovery in catalysis using ML techniques	Summary of ML techniques and tools used in catalysis	Catalysis	ML techniques applied to catalysis for knowledge discovery
[10]	Machine learning	Proposal of novel electricity knowledge discovery model using clustering and machine learning	WOA-based predictive model for electricity consumption prediction	Energy consumption forecasting	Proposed model improves prediction accuracy and reduces prediction time compared to state-of-the-art solutions

(continued)

Table 1 (continued)

Author	Learning model	Knowledge discovery process	Methodology	Research area	Findings
[11]	Machine learning	Framework proposed for KDD from real agricultural datasets	Use of DL models for learning patterns from tabular datasets, integrated gradient method for model interpretability, fuzzy expert system for knowledge interpretation	Agricultural field	KDD knowledge can improve productivity by 30% and decrease water usage by 30%, validated using optimization algorithm and farm simulator
[12]	Machine learning	XML, PMML, SOAP, UDDI, and OLE BD-DM based integrated DMKD process model	six-step DMKD process model Description	DMKD (Data mining and knowledge discovery)	Proposed model facilitates knowledge extraction from large data repositories
[13]	Machine Learning	new procedural model for discovering insights in picture and time series data	Focus on preprocessing, data acquisition, and transformation to generate reliable input data for ML models	Manufacturing	Knowledge discovery in manufacturing processes
[14]	Machine learning	Comparative analysis of software for data mining processes, description of algorithms and techniques of machine learning	Application of systematic literature review methodology	Scientific research and software development	Evolution of algorithms and software for data mining, overview of machine learning techniques and software for data mining processes
[15]	Machine learning	Systematic review of application of AI-based knowledge discovery (KDD) techniques in pharmacoepidemiology	Identification of knowledge discovery methods and techniques from AI, ML, and statistics	Pharmacoepidemiology	Exponential increase in use of AI techniques for solving various tasks in pharmacoepidemiology
[16]	Machine learning	Algorithms &methods used to improve manufacturing processes, discussion on ML research and advances in manufacturing	Overview of ML methods and algorithms for manufacturing improvement	Manufacturing	ML and DM play key roles in improving manufacturing processes

process that should be followed in manufacturing applications, and ways to overcome certain problems. Lastly, they mentioned that the article provided a summary and anticipated the path of future growth [16].

3 Mathematical Model: ML and Data Mining Techniques in Knowledge Discovery

A. Supervised Learning

Supervised learning is a prominent technique in knowledge discovery within data mining, particularly in scenarios where the target variable is known and labeled data is available for training predictive models [17]. In supervised learning, the algorithm learns a mapping function from input variables X to output variables Y based on a training dataset comprising input–output pairs.

The general framework of supervised learning involves:

Training Data**:** A dataset consisting of n samples, each with m features (X) and corresponding labels or target values (Y). The training dataset is represented as $= \{(x_i, y_i)\}_{i=1}^{n}$, where x_i represents the i-th input feature vector, and y_i represents the corresponding target value.

Model Representation: A hypothesis function (h) that maps input features to predicted output values. In a parametric model, h is parameterized by a set of parameters (θ) that are learned during the training process.

Objective Function**:** A measure of the model's performance, also known as a loss or cost function (J), which quantifies the dissimilarity between the predicted output $(h(x)$ and the actual target value (y). The goal is to minimize this function over the training dataset to optimize model parameters.

Mathematically, the objective function can be represented as:

where L is a loss function that computes the discrepancy between the predicted and actual values. Common loss functions include mean squared error (MSE) for regression tasks and cross-entropy loss for classification tasks.

Optimization Algorithm: A method to update the model parameters (θ) iteratively to minimize the objective function. Gradient descent is a widely used optimization algorithm in supervised learning, where the gradients of the loss function with respect to the model parameters are computed and used to update the parameters in the opposite direction of the gradient.

The parameter update rule in gradient descent is given by:

$$\theta = \theta - \alpha \nabla J(\theta)$$

where α is the learning rate, controlling the step size of parameter updates, and $\nabla J(\theta)$ denotes the gradient of the loss function with respect to θ.

Supervised learning in knowledge discovery involves the selection of an appropriate model, optimization of model parameters using labeled training data, and evaluation of the trained model's performance on unseen data. Through this process, supervised learning algorithms can effectively uncover patterns and relationships in data, enabling valuable insights and informed decision-making.

Supervised learning involves training a model on a labeled dataset, where the input features are associated with corresponding target labels. The goal is to learn a mapping function from input variables to output variables.

Classification: Classification involves predicting categorical labels for new instances based on training data. The model learns a decision boundary that separates different classes.

Equation: $y = f(x)$

where y represents the predicted class label, x denotes the input features, and f is the learned classification function.

Regression: Regression predicts continuous numerical values based on input features. The model learns a function that maps input variables to a continuous output space.

Equation: $y = f(x) + \varepsilon$

where y represents the predicted numerical value, x denotes the input features, f is the learned regression function, and ε is the error term.

B. Unsupervised Learning

Unsupervised learning is a fundamental technique in knowledge discovery within data mining, focusing on uncovering patterns and structures in data without explicit supervision or labeled examples. It encompasses various algorithms and methods designed to extract meaningful insights from unlabeled data. One of the primary goals of unsupervised learning is to identify inherent structures within datasets, such as clusters, associations, or anomalies, which can provide valuable knowledge for decision-making and problem-solving.

One of the widely used approaches in unsupervised learning is clustering, which aims to group similar data points together based on their characteristics or features [18]. A commonly used algorithm for clustering is K-means, which partitions a dataset into K clusters by iteratively assigning data points to the nearest cluster centroid and updating the centroids based on the mean of the assigned points. The objective function of K-means can be expressed as:

$$J = \sum_{i=1}^{k} \sum_{x \in c_i} \|x - \mu_i\|^2$$

where J represents the total within-cluster variation, K is the number of clusters, C_i is the set of data points assigned to the ith cluster, μ_i is the centroid of the ith cluster, and $\|\cdot\|$ denotes the Euclidean distance.

Another important unsupervised learning technique is association rule mining, which aims to discover interesting relationships or associations among variables in a dataset. A well-known algorithm for association rule mining is the A-priorialgorithm, which identifies frequent itemsets and generates association rules based on support and confidence measures. The support and confidence of an association rule can be defined as follows:

$$\text{Support}(X \rightarrow Y) = \frac{\text{Count}(X \cup Y)}{\text{Total transactions}}$$

$$\text{Confidence}(X \rightarrow Y) = \frac{\text{Support}(X \cup Y)}{\text{Support}(X)}$$

where X and Y are itemsets, $X \cup Y$ represents the union of itemsets, and $\text{Count}(\cdot)$ denotes the number of transactions containing the itemsets.

Unsupervised learning techniques such as clustering and association rule mining play a crucial role in knowledge discovery by revealing hidden structures, patterns, and relationships within data, thereby enabling organizations to gain valuable insights for decision-making and problem-solving without the need for labeled examples or supervision.

Unsupervised learning involves training a model on an unlabeled dataset, where the goal is to discover hidden patterns or structures within the data.

Clustering: Clustering involves grouping similar instances together based on their inherent characteristics or features.

$$y = f(x)$$

where y represents the cluster assignment, x denotes the input features, and f is the clustering function.

Dimensionality Reduction**:** Dimensionality reduction techniques aim to reduce the number of input features while preserving essential information, thus simplifying the dataset.

Equation (Principal Component Analysis—PCA): $\text{Xnew} = X \cdot W$

where X represents the original data matrix, W denotes the transformation matrix, and Xnew is the reduced-dimensional data matrix.

C. Semi-supervised Learning

Semi-supervised learning leverages a small amount of labeled data along with a large amount of unlabeled data to improve model performance.

Transductive Learning: Transductive learning aims to predict labels for unlabeled instances based on both labeled and unlabeled data.

Equation: $\text{ynew} = f(\text{xnew}, \text{Xlabeled}, \text{ylabeled})$

where ynew represents the predicted label for a new instance, xnew denotes the features of the new instance, Xlabeled represents the labeled data, ylabeled denotes the corresponding labels, and f is the transductive learning function.

Each of these techniques plays a vital role in knowledge discovery within machine learning, offering various approaches to analyzing and extracting valuable insights from data.

D. Reinforcement Learning (RL)

Within the realm of machine learning, there is a subfield known as reinforcement learning that trains agents to maximise cumulative rewards through sequential decision-making in a given environment. With the use of RL algorithms, data miners may speed up discovering new insights inside massive datasets through knowledge discovery in data mining.

One of the fundamental RL algorithms is the Q-learning algorithm [19], which is used to learn a policy that maps states to actions to maximize the cumulative reward. The Q-value represents the expected cumulative reward of taking an action in each state and following a particular policy thereafter. The Q-learning algorithm updates the Q-values iteratively based on the observed rewards and transitions between states.

The Q-value update rule in Q-learning is given by the following equation:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \left(r + \gamma \max_{a'} Q\left(s', a'\right) - Q(s, a) \right)$$

where:

$Q(s, a)$ is the Q-value for taking action a in state s., α is the learning rate, determining the extent to which newly acquired information overrides old information. r is the immediate reward obtained after taking action ain state $s.\gamma$ is the discount factor, representing the importance of future rewards relative to immediate rewards. s' is the next state reached after taking action a in state $s.a'$ is the next action chosen based on the policy.

The agent's policy is typically represented by an exploration–exploitation strategy, such as ε-greedy, where the agent selects the action that maximizes the Q-value with probability $1 - \varepsilon$ (exploitation) and selects a random action with probability ε (exploration).

In the context of knowledge discovery in DM, RL algorithms can be applied to automate the process of feature selection, parameter tuning, and model selection, among other tasks. By formulating the knowledge discovery process as a sequential decision-making problem, RL algorithms can effectively explore the space of possible solutions and optimize the extraction of valuable insights from data.

4 Integration of Data Mining and Machine Learning in Knowledge Discovery

Knowledge discovery that incorporates DM and ML is an effective strategy for drawing out useful information from massive datasets by combining the best features of the two fields. Finding patterns, correlations, and outliers in data is what data mining is all about, whereas machine learning is all about creating algorithms that computers can use to learn from data and make judgements or predictions without human intervention [20]. Organisations may improve their decision-making, innovation, and competitiveness by combining the two methods to find previously unseen patterns, trends, and correlations in their data.

Using sophisticated algorithms and methods is an important part of combining data mining with machine learning for the purpose of knowledge discovery. Classification, regression, clustering, and association rule mining are some of the most popular data mining techniques used to get insights from both organised and unstructured data. Predictive models may be built using machine learning techniques including supervised, unsupervised, and semi-supervised learning to find complicated patterns in data and provide accurate predictions.

Data mining and machine learning together also make it easier to automate the steps involved in discovering new information. Organisations may automate the process of discovering insights and information by utilising machine learning algorithms to analyse massive datasets. This reduces the need for manual analysis and speeds up decision-making [21]. Businesses may now make better, more timely choices because to this automation's ability to spot trends and patterns in data, as well as to reveal previously unknown insights. Using domain knowledge and expertise is another crucial part of merging data mining and machine learning for knowledge discovery. Expertise in the relevant topic is necessary for understanding the findings and drawing practical conclusions from machine learning algorithms, even if these algorithms can analyse massive datasets and finding patterns. With the use of DM and ML techniques, organisations may enhance their understanding of data patterns and linkages. This, in turn, allows for better strategic planning and decision-making.

Knowledge discovery that incorporates data mining and machine learning provides a potent strategy for extracting useful information from massive datasets. Organisations can accomplish their goals and drive innovation with the help of data analysis, automation, and domain knowledge. This allows them to discover hidden patterns and make educated decisions.

5 Conclusion and Future Work

This paper has comprehensively examined the synergy between data mining and machine learning techniques in the domain of knowledge discovery. Through an extensive survey of literature, it has been established that these techniques play a pivotal role in extracting valuable insights and patterns from large datasets across various domains. The integration of data mining and machine learning not only facilitates the automatic extraction of knowledge but also enhances decision-making processes by providing predictive models and actionable insights. Furthermore, the paper highlights the importance of addressing challenges such as data quality, scalability, interpretability, and ethical considerations in the application of these techniques.

Future research endeavors in this field could focus on several promising avenues. Firstly, exploring advanced algorithms and methodologies that can handle complex data types, including unstructured and heterogeneous data, would be crucial. Additionally, the development of hybrid approaches that leverage the strengths of both data mining and machine learning techniques could lead to more robust and efficient knowledge discovery systems. Moreover, there is a need for continued research into interpretable and explainable AI models to enhance transparency and trust in decision-making processes. Furthermore, addressing ethical concerns surrounding data privacy, bias, and fairness remains paramount in the advancement of these techniques. Finally, considering the rapid advancements in technology and the proliferation of data sources, interdisciplinary collaborations between researchers from fields such as computer science, statistics, and domain-specific areas will be essential for pushing the boundaries of knowledge discovery using data mining and machine learning techniques.

References

1. Acharya S, Madhu N (2012) Discovery of students' academic patterns using data mining techniques. Int J Comput Sci Eng 4(6):1054
2. Maimon O, Last M (2001) Knowledge discovery and data mining. Info-Fuzzy Netw (IFN) Methodol 2:23–44
3. Cheng Y, Chen K, Sun H, Zhang Y, Tao F (2018) Data and knowledge mining with big data towards smart production. J Ind Inf Integr 9:1–13
4. Ge Z, Song Z, Ding SX, Huang B (2017) Data mining and analytics in the process industry: the role of machine learning. IEEE Access 5:20590–20616
5. Shu X, Ye Y (2023) Knowledge discovery: methods from data mining and machine learning. Soc Sci Res 110:102817
6. Holzinger A, Jurisica I (2014) Knowledge discovery and data mining in biomedical informatics: the future is in integrative, interactive machine learning solutions. Interact Knowl Discov Data Min Biomed Inform: State-Of-Art Futur ChallS 1–18
7. Palacios CA, Reyes-Suárez JA, Bearzotti LA, Leiva V, Marchant C (2021) Knowledge discovery for higher education student retention based on data mining: machine learning algorithms and case study in Chile. Entropy 23(4):485
8. Dogan A, Birant D (2021) Machine learning and data mining in manufacturing. Expert Syst Appl 166:114060
9. Erdem Günay M, Yıldırım R (2021) Recent advances in knowledge discovery for heterogeneous catalysis using machine learning. Catal Rev 63(1):120–164
10. Ahire PG, Patil PD (2023) CAC-WOA: context aware clustering with whale optimization algorithm for knowledge discovery from multidimensional space in electricity application. Clust Comput 1–15
11. Pazouki E (2023) Optimizing an irrigation treatment using an evolutionary algorithm and a knowledge discovery framework based on deep models. Appl Soft Comput 133:109940
12. Cios KJ, Kurgan LA (2005) Trends in data mining and knowledge discovery. Advanced techniques in knowledge discovery and data mining. Springer London, London, pp 1–26
13. Kubik C, Molitor DA, Becker M, Groche P (2022) Knowledge discovery in engineering applications using machine learning techniques. J Manuf Sci Eng 144(9):091003
14. Taranto-Vera G, Galindo-Villardón P, Merchán-Sánchez-Jara J, Salazar-Pozo J, Moreno-Salazar A, Salazar-Villalva V (2021) Algorithms and software for data mining and machine learning: a critical comparative view from a systematic review of the literature. J Supercomput 77:11481–11513

15. Sessa M, Khan AR, Liang D, Andersen M, Kulahci M (2020) Artificial intelligence in pharmacoepidemiology: a systematic review. Part 1—overview of knowledge discovery techniques in artificial intelligence. Front Pharmacol 11:1028
16. Song Z, Luo S (2022) Application of machine learning and data mining in manufacturing industry. Front Comput Intell Syst 2(1):47–53
17. Bandaru S, Ng AH, Deb K (2017) Data mining methods for knowledge discovery in multi-objective optimization: part A-Survey. Expert Syst Appl 70:139–159
18. Usama M, Qadir J, Raza A, Arif H, Yau KLA, Elkhatib Y, Al-Fuqaha A (2019) Unsupervised machine learning for networking: techniques, applications and research challenges. IEEE Access 7:65579–65615
19. Jang B, Kim M, Harerimana G, Kim JW (2019) Q-learning algorithms: a comprehensive classification and applications. IEEE Access 7:133653–133667
20. Zimek A, Filzmoser P (2018) There and back again: outlier detection between statistical reasoning and data mining algorithms. Wiley Interdiscip Rev Data Min Knowl Discov 8(6):e1280
21. Dean J (2014) Big data, data mining, and machine learning: value creation for business leaders and practitioners. John Wiley & Sons

IoT-Based Data Analysis for Smart Farming

Ravi Kumar Sanapala, Pallavi Singh, L. Bhagyalakshmi, Sanjay Kumar Suman, K. Yadagiri, and Srilakshmi Aouthu

Abstract Our country's fastest-growing industry is agriculture. Farming contributes to about a third of the state's capital in India, where about 70% of the population relies on it. Agricultural-related problems have been some of India's biggest obstacles to development. Agricultural modernization via modernizing traditional farming strategies is the best way to solve this problem. Following this, these venture goals depict the importance of the sensor and track the temperature and moisture. Dimension of these factors and the usage of those sensors in agriculture may be referred to as clever agriculture. Consequently, this modern agriculture can impact the planet in immeasurable approaches. It improves the yield of crops of all kinds, reducing the dependence on insecticides, lowering operational charges, optimizing water usage, and ensuring higher land management and crop rotation. To ensure higher farming strategies and to get the most yield with fewer sources using conversation generation, we can display stay discipline in situations to take corrective motion. For this reason, as development on this subject the paper gives about the image processing strategies and numerous sensors for the improvement and yield of plants by detecting the diseases of the crop.

Keywords Smart Farming · Soil Moisture Sensor · pH Sensor · Temperature Sensor

R. K. Sanapala · S. K. Suman (✉)
Department of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India
e-mail: prof.dr.sanjaykumarsuman@gmail.com

P. Singh
Department of ECE, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India

L. Bhagyalakshmi
Department of ECE, Rajalakshmi Engineering College, Chennai, Tamil Nadu, India

K. Yadagiri
Department of IT, St. Martin's Engineering College, Secunderabad, Telangana, India

S. Aouthu
Department of ECE, Vasavi College of Engineering, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_43

1 Introduction

Farming in India is performed utilizing unremarkable methodologies. The truth that the limit of our ranchers needs appropriate aptitude makes it considerably more whimsical. An enormous piece of cultivating and rural exercises depends absolutely on the expectations, which occasions fizzle. Ranchers need to bear enormous misfortunes, and on occasion, they end up ending it all. Seeing that we know the advantages of appropriate soil dampness, air quality and water system, inside the increment of vegetation, such boundaries cannot be inconspicuous day innovation. This sort of recognizable proof strategy will assist us with distinguishing the infections with its appropriate control measures [1]. The human population requires expanded staples for this situation innovation is vital in feasible farming to increment rural yields. The present innovative forward leaps guarantee that the turn of events and disclosure of new advancements is a ceaseless exertion that makes practical agribusiness conceivable. Farming innovation development is the reason for satisfying the new need for food. Normally, a rancher realizes that his harvest is gone after by infection through direct perception in the field. In any case, this propensity is tedious and arduous on enormous estates and is once in a while less precise. What is more, these techniques are restricted to mental peculiarities and mental that can prompt mistakes and deceptions [2]. And the plant infirmities disclosure is critical advancement for agriculture. In this endeavor, it hopes to establish infirmities revelation region using picture taking care of development and noticing portion using the android APP. The image taking care of advancement has four phases, the underlying advance is picture pre-processing, the second step is division and third component extraction and the fourth step is gathering. Next is the noticing fragment using the embedded stage. The embedded unit uses hardware parts, for instance, Arduino and LCD, Bluetooth, and a motor [3, 4] The utilization of IoT sensor gear has empowered ranchers to deliver exceptional returns proficiently through ranch information observing and environment investigation, soil dampness, temperature, etc. This is known as smart farming or agriculture. So, the utilization of IoT in farming will empower horticultural organizations to have better command over the inward tasks of the business [5]. The framework utilizes a couple of sensors that give how much dampness in the dirt, the stickiness and temperature of the district, and a downpour distinguishing sensor which can be utilized in concluding whether the harvest is reasonable for development. This large number of sensors alongside Node MCU are associated with the Internet and a mobile phone [6]. They are a multitude of angles and featured the job of different advances, particularly IoT, to make the agribusiness more intelligent and more productive to measure up to future assumptions. For this reason, wireless sensors, cloud-registering (IoT), and correspondence advancements are talked about completely. Moreover, a more profound understanding of late exploration endeavors is given. Also, different IoT-based models and levels are furnished for agriculture business executions. A rundown of dog lease difficulties confronting the business and future assumptions are leaned to give direction to analysts and specialists [7]. And the

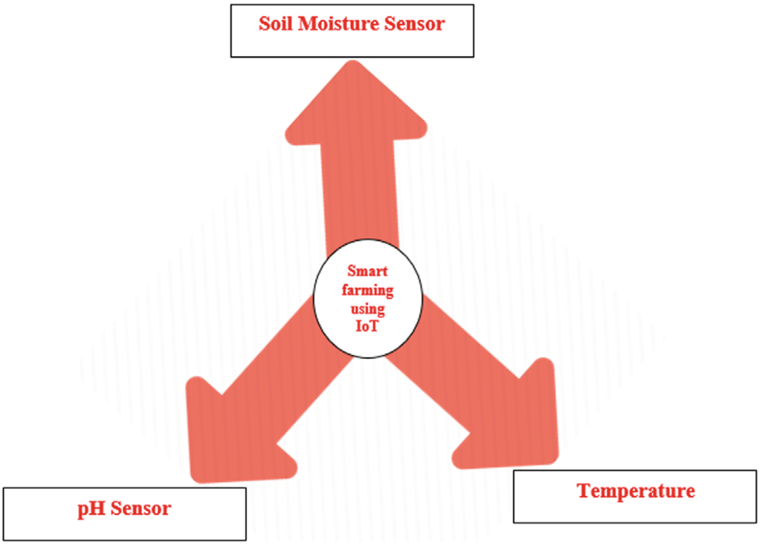


Fig. 1 Components for smart farming

recognizable proof of plant affliction is irksome in the cultivation field. If unmistakable evidence is mixed up, by then there is a tremendous disaster in the age of yield and proficient assessment of the market. Leaf disease acknowledgment requires an enormous amount of work, data on the plant disorders, and besides require moreover planning time. So, we can use pictures dealing with for conspicuous confirmation of leaf contamination in MATLAB. Conspicuous confirmation of illness seeks after the means like stacking the image, separate overhaul, changing over RGB to HSI, removing features, and SVM [8]. Some of the exploration projects give help to the ranchers via naturally distinguishing the nuisance through leaf and natural product markings and help them crop the board explicitly in pesticide application. Utilizing multi-SVM, being 85% accurate on pest was estimated identification. With the initial twenty samples of every dataset collection, two diseases, particularly anthracnose and dirty form and one, i.e., natural product drill were distinguished. These were the vermin and illnesses distinguished that are available in the Central Luzon district, explicitly in Pampanga [9, 10] (Fig. 1).

2 Proposed Method

Within the proposed device, the image processing techniques as in Fig. 2a is used to hit upon and extract the noise in step with pixel, thus making the filtering manner greener and more used to extract the detail. Officially, a helping vector device creates

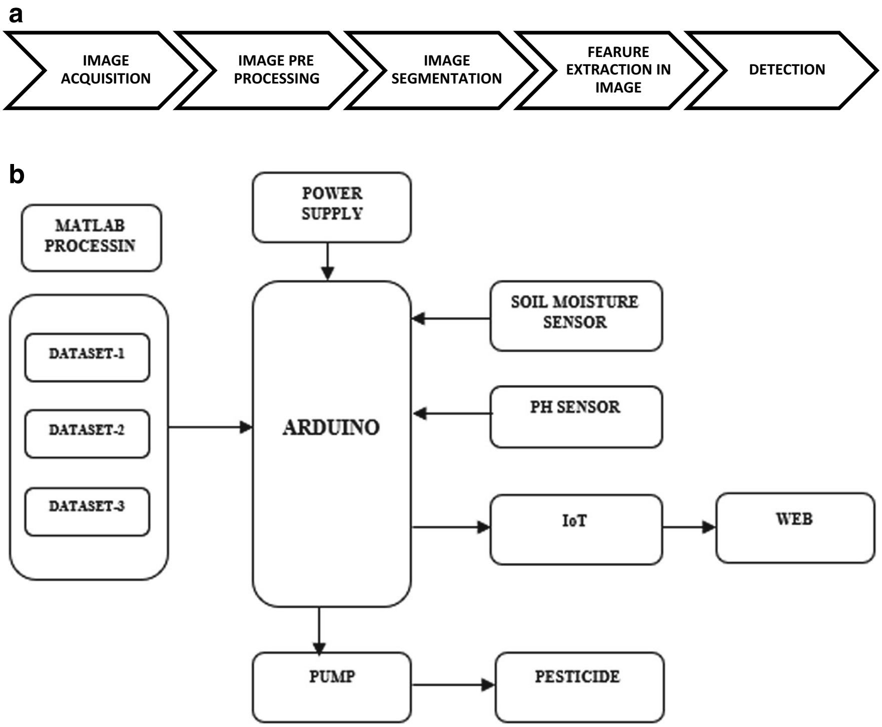


Fig. 2 **a** Filtering process; **b** IoT based smart farming system

a hyperplane, which may be used for partitions, retreats, or other functions. Understandably, an amazing separation is executed with the aid of a hyperplane that has an excellent distance to the nearest training place—facts of any category as commonly a massive restrict reduces the mistake of the aggregate of this separator.

2.1 Image Acquisition

First, we must select a plant that is affected by the disease or healthy and collect the leaf of the plant and look at the leaf and place a picture of the leaves on the frame.

2.2 Image Segmentation

Image segmentation involves converting an image into a set of pixel regions represented by a mask or a labeled image. By splitting an image into segments, you can only process key segments of an image instead of processing the whole image. Image segments help reduce image complexity to be used in future work. Separate simple words to assign labels to pixels.

2.3 Feature Extraction in Image

Feature extraction helps reduce the amount of nonessential data in a data set. Finally, data reduction helps to model the model with less machine effort and increases learning speed and practical steps in the MATLAB process.

2.4 Detection and Classification of Plant Disease

After completion of feature extraction of an image in MATLAB, it will detect the image whether it is having a disease or it is a healthy image.

In the above process if the input image has detected the disease, then the Arduino sends information to the relay and then it will turn on the motor/pump. So, the pump will supply fertilizer water to the whole crop by using sprinklers. And also, the sensors like soil moisture and pH will get the readings of the soil. If the soil moisture sensor detects the less moisture level in the soil it will send information to the Arduino, then the Arduino will turn on the pump and the pump will supply water to the crop until it reaches a moisture level suited for the crop.

3 Result and Discussion

The result should be based on the image processing classifications:

3.1 Input Image

The input is the image Fig. 3a that is affected or a healthy image.

3.2 Image Processing

The output is based on image processing. Here, all the required details of the image should be extracted and sent information to Arduino. Hue transformation is used to separate image luminance from color information.

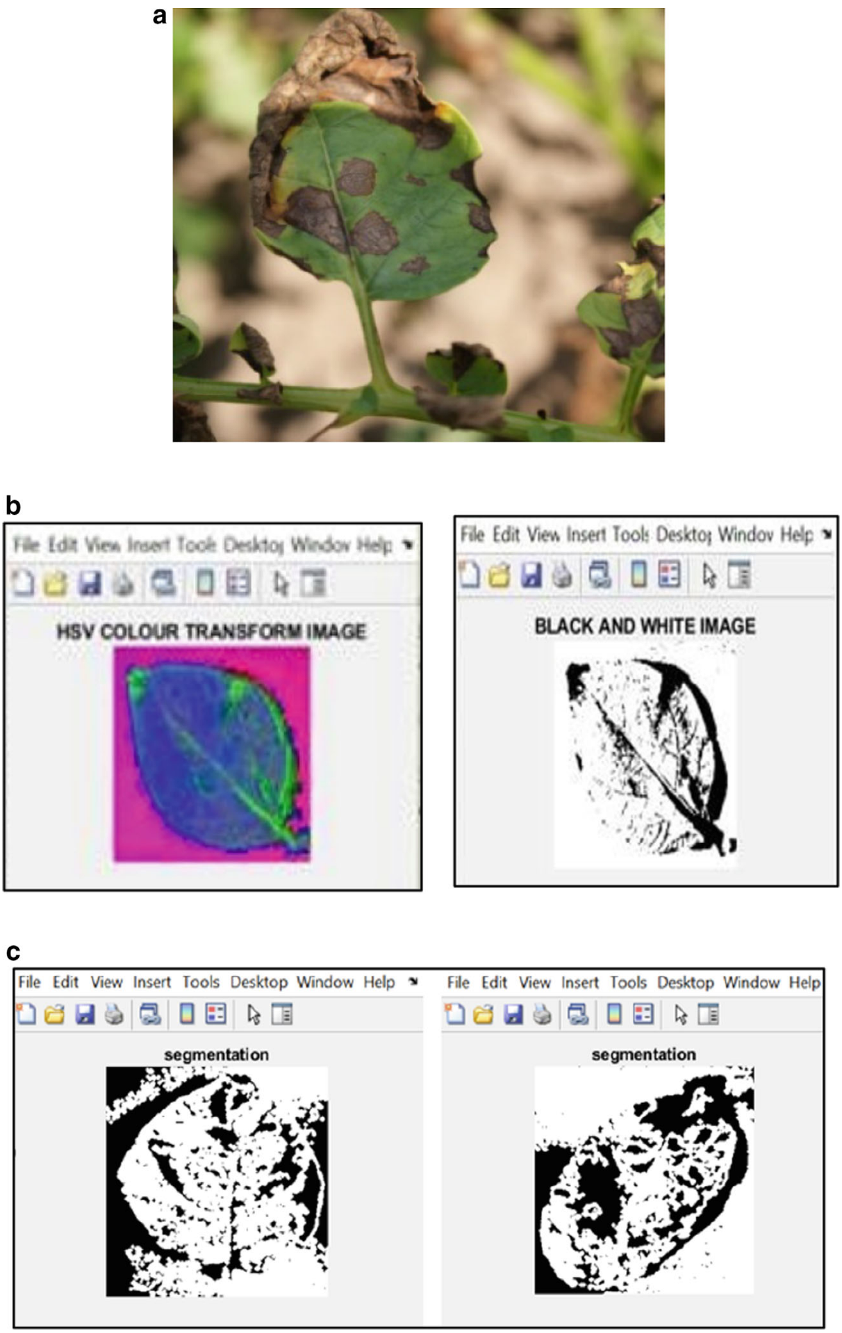


Fig. 3 **a**IoT-based smart farming system; **b** Image processing; **c** Segmentation; **d** Success rate per lable; **e** CNN graph for healthy image; **f** CNN graph for affected image

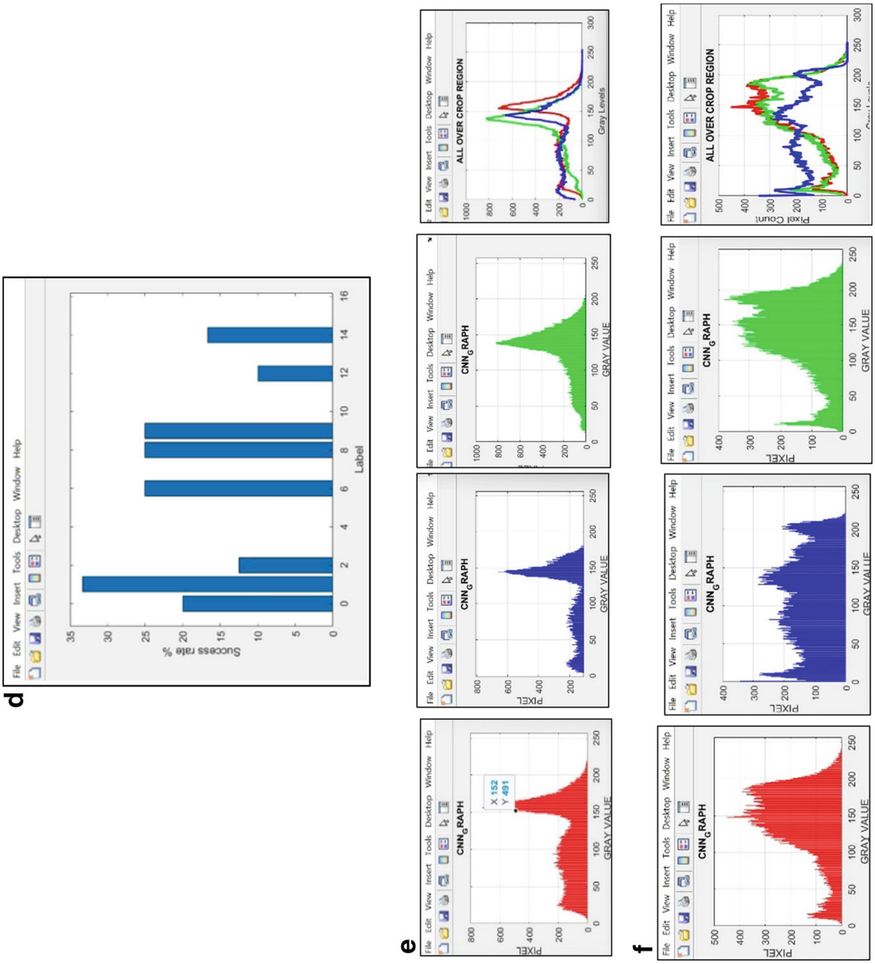


Fig. 3 (continued)

3.3 Segmentation

In Fig. 3c, we can see that the given image is divided into parts based on the properties of the pixels in the image. Also, it helps in reducing the difficulty of the image to make the next processing or analysis of the image simpler.

3.4 Disease Detection

After completion of the image processing, it will give the output as to whether it has a disease or it will be affected image.

3.5 Turning on/off Pump

After completion of the software processing if the given input image is detected with the disease, then it sends the information to Arduino through UART communication, then it will switch on the pump, and it supplies water to the crop. And whenever the moisture level of the soil in the crop is low, then the motor will turn ON and supply water until it reaches the moisture level then it will turn OFF.

3.6 Success Rate

This graph represents the success rate of the datasets that we have taken. And in the first set of bar graphs represents there is no time delay and had done very quickly. And whenever datasets are increasing, the delay should be increasing with the success rate as in Fig. 3d.

Figures 3e and 3f shows differences between affected and healthy image. CNN returns the convolution of the features given by W and B with the given images. Here, the given image represents the blue color noise in the given image. In convolution, neural network pixels of RGB should be extracted.

The above graphs represent the CNN graphs of RGB colors in image processing. In the above graphs, one is for the affected image and another one is a healthy image. The levels of the affected image are more than the healthy image because the affected image has more noise and also has to extract more details from the image. According to the noise levels, the graph will be fluctuating up and down as shown. The gray level or gray value indicates the brightness of a pixel. Pixel count indicates the number of pixels contained in a picture sensor, or that a caught photo is comprised of. The higher the number, the more data there is. Pixel count is typically characterized by Megapixels.

3.7 IoT Web Page

The overall results should be displayed on the IoT web page and includes moisture level of the soil, pH values, and the condition of the pump also about the disease that affected the crop.

4 Conclusion

Our project helps to detect the diseases in the crop using image processing in MATLAB for the given input image. Whenever it detects the disease from MATLAB, it will send signals to the Arduino. Then, the Arduino will turn on the motor pump and it spreads the fertilizer/pesticide to the whole crop. And also, the sensors like soil moisture and pH sensor measure the soil conditions. If the moisture level of the soil is low, then the pump will turn on. All these details will also be stored on the IoT webpage.

References

1. Chowdhury R, Chanda D "Identification of Pathological Disease in Plants using Deep Neural Networks" ICAN 2020
2. Andrianto H, Suhardi (2020) Smartphone application for deep learning-based rice plant disease detection. In: 2020 international conference on information technology systems and innovation (ICITSI). Bandung–Padang ISBN: 978-1-7281-8196-7
3. Prakash N, Udayakumar E, Kumareshan N Design and development of Android-based Plant disease detection using Arduino. In: IEEE 7th international conference on smart structures and systems ICSSS 2020
4. Fadziso T (2018) Internet of things in agriculture for smart farming. Malays J Med Biol Res 5(2)
5. Vijula Grace KS, Kharim S, Sivasakthi P (2015) Wireless sensor based control system in agriculture field. Asian Association Emerland Insight, pp 2414–6994 https://doi.org/10.1108/AAOUJ-08-2019-0029
6. Abhiram MSD, Kuppili Y, Manga NA (2020) Smart farming system using IoT for efficient crop growth. In: 2020 IEEE international students' conference on electrical, electronics and computer science
7. Ayaz M, Ammad-Uddin M, sharif Z, MansourA, Aggoune EHM (2019) "Internet-of-Things (IoT)-Based Smart Agriculture". IEEE 7
8. Krishnan H, Priyadharshini K (2019) Plant disease analysis using image processing in MATLAB. Springer Education and Information Technologies, vol 26, pp 6923–6947
9. Zhang SW, Shang YJ, Wang L (2015) Plant disease recognition based on plant leaf image. J Anim Plant Scis 25(1):42–45
10. Barbados JGA (2016) A new automatic method for disease symptom segmentation in digital photographs of plant leaves. Eur J Plant Pathol 1–16

Alzheimer's Disease Identification at Early Stage Using Machine Learning-Based Cognitive Features and Feature Extraction

Jyothi Gattoji, L. Arokia Jesu Prabhu, and Vijender Solanki

Abstract Alzheimer's disease (AD) is a prevalent brain illness that is frequently connected with maturing and is related with cognitive decline and mental debilitation. The review's fundamental goal is to utilize AI methods to use mental testing for early ID of Alzheimer's disease. A CNN and CNN + LSTM are two models that are constructed and evaluated utilizing a dataset. To improve grouping execution, a stacking classifier that joins random forest and Multilayer Perceptron is utilized. The results exhibit how much our outfit approach works on the accuracy of early Alzheimer's disease diagnosis. Assessment boundaries incorporate 99.5% accuracy, 98.7% precision, 99.1% recall, and 99.2% F1-score to refine current methodologies and accomplish a significant forward leap in early Alzheimer's diagnosis.

Keywords Random forest · AdaBoost · KNN · LR · Decision tree · Artificial neural networks

1 Introduction

Alzheimer's disease (AD) is a degenerative disease whose side effects logically increment over the long haul. The most pervasive indication of Alzheimer's disease in individuals is cognitive decline, which incorporates neglecting names, objects,

J. Gattoji · L. Arokia Jesu Prabhu (✉) · V. Solanki
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, Telangana, India
e-mail: arokiajeruprabhu@gmail.com

J. Gattoji
e-mail: jyothi.gattoji@gmail.com

V. Solanki
e-mail: spesinfo@yahoo.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_44

ongoing occasions, and discussions. Among these side effects incorporate bewilderment, disarray, and getting derailed in natural spots. Most of the individuals with Alzheimer's disease are above 65 years old [1, 2].

Since it is trying to distinguish the dementia subtype, making an exact finding of Alzheimer's disease is a difficult errand [3]. Amazingly, concentrates show that almost 66% of dementia cases are brought about by Alzheimer's disease [4]. For example, by 2030, it is anticipated that between 400 000 and 459 000 Australians would get a dementia finding [5]. By 2060, there will probably be 15 million Americans living with Alzheimer's disease or extreme mental hindrance, as per gauges, a critical increment from the 6.08 million cases that were kept in 2017 [6–9].

Then again, ongoing information shows that Alzheimer's disease is declining in various contemporary medical service frameworks, including the UK [10, 11]. For example, a two-decade investigation of individuals from three unique locales of Britain who were 65 years old and more seasoned found a decline in the age-explicit predominance of Alzheimer's disease [12–14]. Also, in Britain and Ridges in 2019, the age-normalized death rate (AMSR) for dementia and Alzheimer's disease fatalities was 115.1 per 100,000 people. This figure was measurably essentially lower than the ASMR for 2018—123.8 per 100,000 individuals (69 478 passings) [15, 16].

2 Literature Review

Adaptive Ensemble of Classifiers with Regularization (AER), a special procedure for twofold imbalanced information order, was created by Wang et al. [10]. AER utilizes worldwide math double-dealing and stochastic slope drop to use implied regularization for further developed performance. With conceivable computational compromises and dataset-subordinate viability, AER offers a suitable answer for uneven information order, displaying better execution and spearheading regularization in unique troupe draws near. Since dynamic outfit techniques are intricate, AER might have higher handling necessities.

To uncover associations between mental side effects and neurodegeneration and accomplish high expectation precision, Martinez-Murcia et al. [17] introduced deep convolutional autoencoders for Alzheimer's disease investigation. These autoencoders coordinate neuropsychological information with qualities produced from MRIs. Despite certain inborn challenges, the use of deep convolutional autoencoders offers a strong strategy for understanding Alzheimer's disease, offering exact expectations and bits of knowledge into the connection between neuroimaging qualities and mental side effects.

A 37-thing source survey was utilized by Zhu et al. [3] to foster an ML-based finding strategy that ordered ordinary, moderate mental weakness, exceptionally gentle dementia, and dementia with great exactness. The made indicative model has extraordinary exactness and accuracy for clinical applications, and it is a dependable instrument for early dementia identification in light of the fact that to its usage of ML and important polls. Reliance on exact and careful survey reactions, predisposition powerlessness, and the necessity for progressing approval with an assortment of datasets are likely disadvantages.

Gill et al. [11] introduced an ML model that accomplished great exactness in understanding explicit forecast of mental debilitation movement by using gauge mild behavioral impairment (MBI) status and brain morphological parameters. A valuable device for early finding and mediation in individuals with typical cognizance or gentle mental weakness, the proposed ML technique consolidates brain morphological parameters and MBI status to show high expectation capacities with regard to future mental hindrance.

3 Methodology

For early Alzheimer's disease expectation, the strategy joins a ML gathering model with the CNN, CNN + LSTM, and stacking classifier that have been created.

Dataset: The info is the open Alzheimer's disease dataset from Kaggle, which incorporates mental qualities and names related with the infection.

Pre-processing: Inspect and set up the dataset to ensure it is viable with the model-preparing strategy and of superior grade. This involves activities like component designing, tending to missing information, and standardization (Table 1).

Partitioning information into test and train sets: Separate the preparation and testing subsets from the preprocessed dataset. The testing set will be utilized to survey the ML models' presentation (Figs. 1, 2, 3 and 4).

Model Generation: Utilizing the mental qualities that were removed from the dataset makes various ML classifiers.

This involves:

Table 1 Dataset with features

ID	Identification
Age	Age at the time data acquisition
Gender	Gender (Male or Female)
Educ	Years of education
Ses	Socioeconomic status classified into categories
MMSE	Mini-Mental State Examination score (range is from 0[worst] to 300[best])
CDR	Clinical Dementia Rating (0 = no dementia, 1 = dementia)
eTIV	Estimated total intracranial volume
nWBV	Normalized whole brain volume, using automated issue segmentation process to expressed as a percent of all voxels in the atlas-masked image that are labeled as gray or white matter

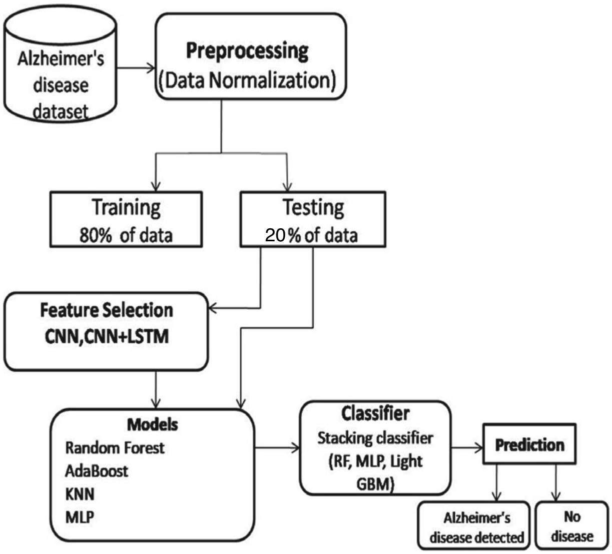


Fig. 1 System architecture

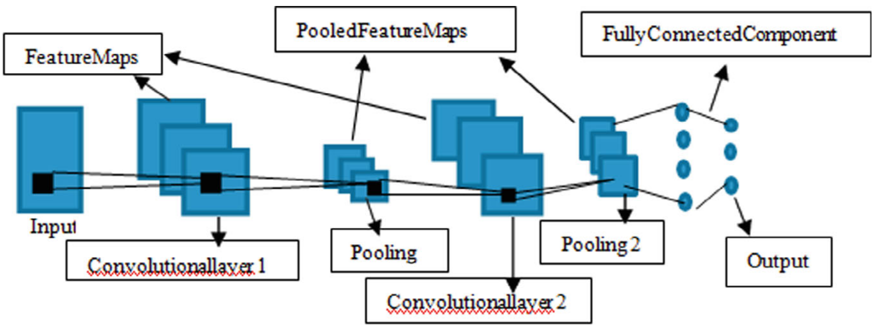


Fig. 2 Basic structure of CNN [13]

Random forest: Leo Breiman and Adele Cutler are the brand name holders of the famous ML strategy known as 'random forest' which totals the result of a few choice trees to deliver a solitary end. The random forest numerical expectation for arrangement might be communicated as follows:

$$\mathrm{Y}^{\Lambda} = \mathrm{mode}(Y_1, Y_2, \ldots . Y_n) \tag{1}$$

k-Nearest Neighbors, or KNN: The k-nearest neighbors algorithm, in some cases alluded to as KNN or k-NN, is a non-parametric supervised learning classifier that bunches individual information directs in light of closeness all together toward group or foresee data. [18] As far as math, it is communicated as

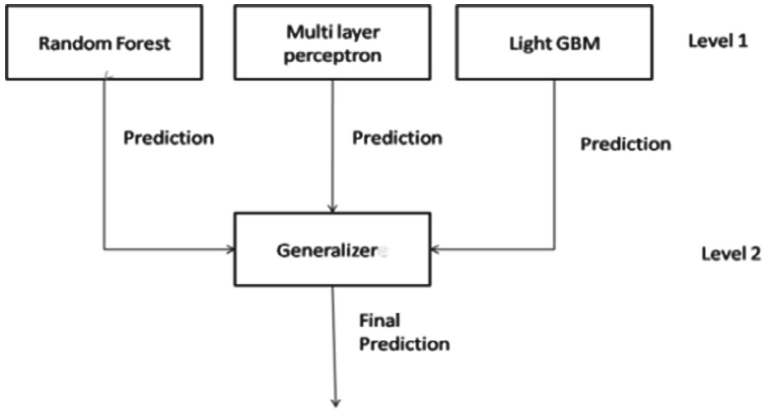


Fig. 3 Concept diagram of stacking

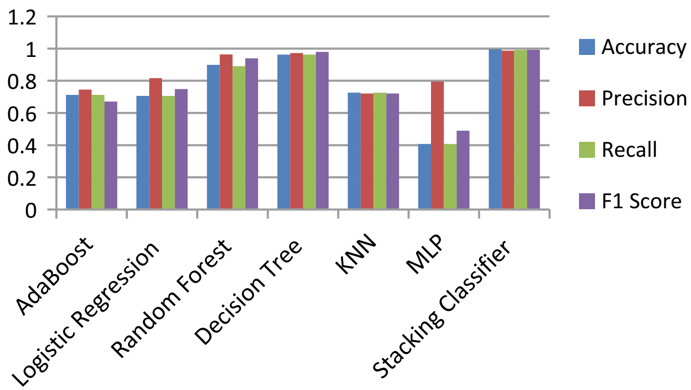


Fig. 4 Comparison graph

$$Y^{\Lambda} = \mathrm{argmax}_c \sum_{i=1}^{k} \mathrm{I}(Y_i = c) \tag{2}$$

where each class label is represented by c, the class labels of the k-nearest neighbors are represented by Y^{Λ}, which is the anticipated class label, x_{new} and $I(\cdot)$ are the pointer capabilities that profit 1 for the situation that the condition is valid and 0 in any case.

Logistic regression, or LR: A supervised learning approach called calculated relapse is utilized to gauge the probability of an objective variable. Because of the dichotomous construction of the reliant variable, there are just two suitable classes. The output of a logistic regression model is the probability that a given input belongs to a certain class.

Decision tree: A non-parametric unsupervised learning approach that is utilized for both relapse and grouping applications is the decision tree. With a root node, branches, inward nodes, and leaf nodes, it has a progressive tree structure.

With the decision tree, we might utilize a case x to expect its result y in the accompanying manner:

$$Y = \sum_{m=1}^{M} (I(\text{x follows path to leaf } L_m) \times \text{Value of leaf}) \tag{3}$$

where the marker capability I (x follows the way to leaf L_m) returns 1 if the occurrence x follows the way to leaf and 0 in any remaining cases.

Artificial Neural Networks (MLP): A feed forward Artificial Neural Network (ANN) having many layers of linked nodes. The Multilayer Perceptron (MLP) is used for an assortment of ML applications, like regression and classification.

For each layer from 1 to L:

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)} \tag{4}$$

$$a^{(l)} = \sigma\left(z^{(l)}\right) \tag{5}$$

$Z^{(l)}$ is the weighted input to layer l and $a^{(l)}$ is the output of layer $l$after applying the activation function σ $W^{(l)}$ is the weight matrix connecting layer $l-1$to layer l, and $b^{(l)}$ is the bias vector for layer l this process repeats for each layer until the output layer is reached.

CNN: Convolutional layers are utilized by Convolutional Neural Networks (CNNs), a deep learning architecture made for the investigation of pictures and recordings. CNNs naturally remove highlights from input [19].

The forward pass through the convolutional layers can be represented as:

$$Z_i = f(W_i * \mathrm{X} + b_i) \tag{6}$$

After convolutional layers, the feature maps may be passed through pooling layers:

$$Y = P(Z_i) \tag{7}$$

Finally, the output of the last pooling layer is flattened and passed through fully connected layers:

$$Y_{fc} = f\left(W_{fc}.Y + b_{fc}\right) \tag{8}$$

This produces the final output of the network.

CNN along with LSTM: Frequently utilized in text and video examination. A CNN-LSTM model is a hybrid neural network that combines Long Short-Term Memory (LSTM) for consecutive information handling with Convolutional Neural Networks (CNN) for feature extraction.

1. CNN Layers

$$Z_{\text{cnn}} = f_{\text{in}}d(W_{\text{cnn}}.X + b_{\text{cnn}}) \tag{9}$$

$$Y = \text{Flatten or GlobalPooling}(Z_{\text{cnn}}) \tag{10}$$

2. LSTM Layers

$$Z_{\text{lstm}} = f_{\text{lstm}}(W_{\text{lstm}}.Y + b_{\text{lstm}}) \tag{11}$$

To simultaneously gain spatial and fleeting portrayals from the information, the joined CNN-LSTM model is trained end-to-end.

Classifier Stacking (RF + MLP with LightGBM): Through meta-learning and group draws near, a stacking classifier further develops precision by joining expectations from random forest (RF), Multilayer Perceptron (MLP), and LightGBM models. The expectation cycle can be numerically expressed as follows:

Generate Predictions from Base Classifiers:

$$\text{RF}_{\text{predictions}} = \text{RF.predict}(X_{\text{text}}) \tag{12}$$

$$\text{MLP}_{\text{predictions}} = \text{MLP.predict}(X_{\text{text}}) \tag{13}$$

$$\text{LGBM}_{\text{predictions}} = \text{LGBM.predict}(X_{\text{text}}) \tag{14}$$

Stack Predictions: Concatenate the predictions along axis 1 to create a new feature matrix for the meta-classifier:

$$X_{\text{stacked}_{\text{test}}} = \text{ concatenate}\left(\text{RF}_{\text{predic}},\ \text{MLP}_{\text{predic}},\ \text{LGBM}_{\text{predic}}\right) \tag{15}$$

Make Predictions using Meta-classifier: Use the trained meta-classifier to predict the stacked features:

$$Y^{\Lambda} = \text{Meta} - \text{Classifier.predict}(X_{\text{stacked\_ test}}) \tag{16}$$

User Signup and Login: To empower safe enlistment and login to the application, carry out a client confirmation system.

User Input: Offer a connection point by means of which clients might enter the aftereffects of their mental tests to make forecasts. The prepared models will use this contribution as information include.

Prediction: In view of the client's feedback information, make forecasts utilizing the prepared outfit model, which is made out of the previously mentioned classifiers.

Use CNN, CNN + LSTM, and Classifier Stacking. The procedure changes the loads of every classifier expectation powerfully, considering how well it performs with the given input information.

The design of the framework integrates mental test information to analyze Alzheimer's disease (AD) early. Stacking classifier, CNN + LSTM, and investigate CNN are utilized to increment precision.

4 Implementation

Importing Libraries: Acquire the necessary libraries, including pyplot, matplotlib, seaborn, pandas, and numpy. The CSV document 'oasis_longitudinal.csv' is perused through the pd.read_csv() strategy from Pandas and consequently doled out to an Information Edge called df. The complete number of lines and sections in the dataset is then printed utilizing the shape property. In arrangement, every segment is checked for missing qualities utilizing the ISNA ().total() capability, and the 'SES' and 'MMSE' sections' missing qualities are filled in with their comparing middle and mean qualities utilizing the fill () capability. To examine the information, Seaborn and Matplotlib are utilized to fabricate perceptions. Plotting procedures like count, violin, match, and box plots are used to understand the appropriation and connections among different factors.

Outlier Detection: For every segment remembered for list_attributes, the outliers_iqr() capability is created to distinguish anomalies utilizing the Interquartile Range (IQR) approach.

Feature Encoding: A planning word reference and the supplant() capability are utilized to encode the 'Gathering' and 'M/F' sections into mathematical qualities for the purpose of displaying.

Data Splitting: The dataset is separated into factors for highlights (X) and targets (y). The train_test_split() strategy from scikit-learn is then used to additional split it into preparing and testing sets.

Model Building and Evaluation: Various ML models are prepared and surveyed with the utilization of measurements including ROC bends, characterization reports, and disarray lattices.

Ensemble Learning: Group strategies are utilized to make stacking classifiers. Exactness between a few models is looked at, and the stacking classifiers' presentation is surveyed.

5 Results and Discussion

Accuracy: A test's accuracy is determined by how well it can distinguish between patient and healthy cases. We should compute the percentage of true positives and true negatives in each analyzed case in order to assess the accuracy of a test. As far as math, this is communicated as:

$$\begin{aligned} \text{Accuracy} = {} & \text{True Positive} + \text{True Negative}/\text{True Positive} \\ & + \text{True Negative} + \text{False Positive} + \text{False Negative} \end{aligned} \tag{17}$$

Precision: Precision estimates the level of accurately sorted examples or cases among the positive examples.

$$\text{Precision} = \text{True positives}/(\text{True positives} \ + \ \text{False positives}) \tag{18}$$

Recall: In ML, review is a measurement that evaluates a model's ability to find all relevant examples of a class. It is a proportion of how well a model catches instances of a specific class: the proportion of appropriately anticipated positive perceptions to the complete number of genuine up-sides.

$$\text{Recall} = \text{True Positive}/(\text{True Positive} \ + \text{False Negative}) \tag{19}$$

F1-score: An evaluation metric for ML that measures a model's exactness is known as the F1-score. It incorporates a model's exactness and review evaluations. The precision measure computes the times a model accurately anticipated the entire dataset.

$$\text{F1} - \text{score} = 2(\text{Precision})(\text{Recall})/\text{Precision} + \text{Recall} \tag{20}$$

According to the different works examined, the prediction of Alzheimer's disease carried out using stacking classifier holds the highest accuracy of 99.5%, 98.7% of precision, 99.1% of recall, and 99.2% of F1-score. Research from the past has demonstrated that this technique produces viable results. The exactness, accuracy review, and F1-score of the different ML models are recorded in Table 2.

Above graph is the statistical representation for the resultant table. Stacking classifier is best fit compared with other classifiers.

Table 2 Comparison table

ML Models	Accuracy	Precision	Recall	F1-score
AdaBoost	0.712	0.745	0.712	0.671
Logistic regression	0.706	0.816	0.706	0.748
Random forest	0.899	0.964	0.89	0.939
Decision tree	0.963	0.972	0.963	0.979
KNN	0.726	0.721	0.726	0.721
MLP	0.407	0.795	0.407	0.49
Stacking classifier	0.995	0.987	0.991	0.992

6 Conclusion

The subsequent table's measurable portrayal is displayed in the diagram above. With regard to classifiers, the stacking classifier fits the information the best. Generally speaking, the work utilizes ML ways to deal with predict Alzheimer's disease through exploratory information investigation, information arrangement, model creation, and appraisal. Shockingly, the exploratory outcomes plainly suggest that this procedure performs better compared to other flow approaches on various execution measurements, showing a significant progression in the early finding of Promotion.

7 Future Scope

The investigations additionally feature the basic importance that mental qualities from clinical assessments have in the early recognizable proof of Alzheimer's disease using group ML models. This approach presents a practical method for speeding up the distinguishing proof of Alzheimer's disease by giving a valuable substitute to difficult clinical preliminaries.

References

1. You Y, Ahmed B, Barr P, Ballard K, Valenzuela M (2012) Predicting dementia risk using paralinguistic and memory test features with machine learning models. In: Proceeding IEEE healthcare innovations and point of care technologies, pp 56–59
2. Grassi M et al (2019) A novel ensemble-based machine learning algorithm to predict the conversion from mild cognitive impairment to Alzheimer's disease using socio-demographic characteristics, clinical information, and neuropsychological measures. Front Neurol 10:756
3. Zhu F et al (2020) Machine learning for the preliminary diagnosis of dementia. Sci Program 2020, 5629090
4. Sullivan KJ et al (2019) Declining incident dementia rates across four population-based birth cohorts. JS Gerontol: Ser A 74(9):1439–1445

5. Barber D (2012) Bayesian reasoning and machine learning. Cambridge University Press
6. Cruz RMO, Sabourin R, George DCC (2018) Dynamic classifier selection: recent advances and perspectives. Inf Fusion 41:195–216
7. Reynolds DA, Rose RC (1995) Robust text-independent speaker identificationusing Gaussian mixture speaker models. IEEE Trans Speech Audio Process 3(1):72–83
8. Gardner WA (1984) Learning characteristics of stochastic-gradient-descent algorithms: a general study, analysis, and critique. Signal Process 6(2):113–133
9. Chen T, Guestrin C (2016) Xgboost: a scalable tree-boosting system. CoRR, abs/1603.02754
10. Wang C, Deng C, Yu Z, Hui D, Gong X, Luo R (2021) Adaptive ensemble of classifiers with regularization for imbalanced data classification. Information Fusion 69:81–102
11. Gill S et al (2020) Using machine learning to predict dementia from neuropsychiatric symptoms and neuroimaging data. J Alzheimers Dis 75(1):277–288
12. Zhang C, Bengio S, Hardt M, Recht B, Vinyals O (2016) Understanding deep learning requires rethinking generalization. arXiv preprint arXiv:1611.03530
13. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
14. Bottou L, Curtis FE, Nocedal J (2018) Optimization methods for large-scale machine learning. Siam Rev 60(2):223–311
15. Lin T-Y, Goyal P, Girshick R, He K, Dollár P (2018) Focal loss for dense object detection. IEEE Trans Pattern Anal Mach Intell
16. Kang MJ et al (2019) Prediction of cognitive impairment via deep learning trained with multi-center neuropsychological test data. BMC Med Inform Decis Mak 19(1):1–9
17. Martinez-Murcia FJ, Ortiz A, Gorriz JM, Ramirez J, Castillo-Barnes D (2020) Studying the manifold structure of Alzheimer's disease: a deep learning approach using convolutional autoencoders. IEEE J Biomed Health Inform 24(1):17–26
18. Eramuthu A, Meenakshi S, Manjusha PS (2014) A new approach for Alzheimer's disease diagnosis by using association rule over pet images. Int J Comput Appl 91(9):9–14
19. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232

Enabling Federated Learning at the Edge for Enhanced Security and Privacy in 5G-Powered IOT Ecosystems: A Review

Abburi Srirama Kanaka Ratnam, Bandlamudi VedaSri, Chittiboina Hemalatha, Jeshta Chandrika, Goddati JayaSree, and S. Durga Bhavani

Abstract The confluence of 5G networks and the Internet of Things (IoT) promises transformative capabilities for real-time data processing and decision-making. However, this advancement also brings forth significant concerns regarding data security and privacy. This research paper explores the application of Federated Learning (FL) at the edge of 5G networks as a solution to address these challenges. FL, a decentralized machine learning paradigm, facilitates collaborative model training across IoT devices while maintaining data on the edge. This approach not only safeguards user privacy but also minimizes the necessity for extensive data transfer to centralized servers, thereby optimizing bandwidth and reducing latency. In this paper, we delve into the foundational principles of FL, assess its suitability in edge computing environments, and present case studies illustrating its effectiveness across various IoT

A. Srirama Kanaka Ratnam (✉)
Professor, CSE (AIML) Department, Kallam Harnadhareddy Institute of Technology, Guntur, Andhra Pradesh, India
e-mail: sriram.abburi@gmail.com

B. VedaSri
IT Department, Malineni Lakshmaiah Women's Engineering College, Andhra Pradesh, Guntur Dt, India

C. Hemalatha
IT Department, Malineni Lakshmaiah Women's Engineering College, Andhra Pradesh, Guntur Dt, India

J. Chandrika
CSE Department, Malineni Lakshmaiah Women's Engineering College, Andhra Pradesh, Guntur Dt, India

G. JayaSree
CSE Department, Malineni Lakshmaiah Women's Engineering College, Andhra Pradesh, Guntur Dt, India

S. Durga Bhavani
CSE Department, Malineni Lakshmaiah Women's Engineering College, Andhra Pradesh, Guntur Dt, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_45

scenarios. Furthermore, we examine the essential security considerations, potential enhancements, and deployment strategies necessary to fully realize the potential of FL in 5G-powered IoT ecosystems. This research contributes to the development of secure, privacy-preserving, and efficient IoT solutions in the 5G era.

Keywords Federated learning · Edge computing · 5G networks · Internet of things · Privacy-preserving · Security · Decentralized machine learning · Edge intelligence · Data analytics

1 Introduction

The rapid development of the Internet of Things (IoT) and the deployment of 5G networks have ushered in a transformative era in the world of technology and connectivity. IoT has proliferated across various domains, encompassing smart homes, industrial automation, health care, agriculture, and more. These connected devices have become prolific data generators, constantly collecting and transmitting information about our environment, activities, and systems. Simultaneously, the rollout of 5G networks has significantly amplified the capabilities of data transmission and communication. With its ultra-low latency and high bandwidth, 5G has empowered IoT devices to communicate faster and more efficiently than ever before. This convergence of IoT and 5G networks offers unprecedented opportunities for real-time data analytics and decision-making [1]. However, amidst the excitement and potential benefits, this transformation also brings forth significant concerns and challenges, particularly in the domains of data security and privacy. As more IoT devices become interconnected and generate vast amounts of data, there is a heightened risk of data breaches, cyberattacks, and unauthorized access [2]. Furthermore, the sheer volume of data being transmitted raises questions about data ownership, consent, and the potential for misuse. The critical issue of data privacy looms large, as individuals and organizations are understandably concerned about how their personal and sensitive data is collected, processed, and shared within this interconnected ecosystem. Balancing the benefits of data-driven insights with the imperative to protect individuals' privacy and data security is a complex challenge that requires innovative solutions. In light of these considerations, this research paper delves into one such solution: Federated Learning (FL) at the edge of 5G networks. FL offers a promising path to address the pressing issues of data security and privacy by allowing collaborative model training while maintaining data on the edge devices. This approach not only safeguards user privacy but also minimizes the necessity for extensive data transfer to centralized servers, thereby optimizing bandwidth and reducing latency. It presents an innovative way to harness the vast potential of IoT and 5G while respecting the fundamental principles of security and privacy.

2 Federated Learning: A Decentralized Approach

2.1 Key Principles of Federated Learning

Federated Learning (FL) represents a groundbreaking approach to machine learning that fundamentally departs from traditional centralized methods [3]. In FL, machine learning models are trained across a network of distributed devices or edge nodes, all while retaining the data locally on these individual devices. This decentralized nature of FL stands in stark contrast to the conventional paradigm of centralized machine learning, where data from various sources is collected, aggregated, and processed on a single, centralized server [4]. The core principle of FL is the notion that data should not needlessly leave the devices where it originates. Instead of transporting entire datasets to a central location for processing, FL permits local model training on each device using its respective data. Only the essential updates to the model, rather than raw data, are transmitted to a central aggregator. This centralized aggregator aggregates these model updates to refine a global model [5].

This approach offers a multitude of advantages:

1. Privacy Preservation: FL excels at preserving data privacy. Because raw data never leaves the device, users maintain full control over their personal or sensitive information. This decentralized model training ensures that sensitive data remains secure and inaccessible to external entities. 2. Reduced Bandwidth Usage: By transmitting only model updates rather than extensive datasets, FL greatly reduces the burden on network bandwidth. This is especially critical in scenarios where bandwidth resources are limited or costly. 3. Lower Latency: FL introduces the advantage of lower latency due to local model training. There is no need to wait for data to be sent to a central server for processing and then return with updated models. This results in faster model updates and real-time decision-making, which is vital for applications where low latency is paramount. 4. Improved Energy Efficiency: In mobile and IoT settings, energy efficiency is a top priority. FL contributes to energy savings as it minimizes the need for power-hungry data transmission, allowing devices to perform local computations with minimal energy consumption. These benefits underscore why FL has gained significant attention and is considered a cornerstone in addressing contemporary challenges in machine learning, particularly in the context of privacy, bandwidth constraints, latency-sensitive applications, and energy-efficient computing. By decentralizing the model training process and ensuring that data remains on the edge, FL aligns with the principles of data privacy, resource optimization, and real-time responsiveness, making it a promising solution for an array of modern applications.

2.2 Federated Learning Workflow

Step 1. Initialization: A global model is initialized on a central server. Step 2. Local Model Training: IoT devices or edge nodes perform local model training on their respective datasets. Step 3. Model Update: The local models compute updates based on their training data. Step 4. Aggregation: Model updates are sent to the central server, where they are aggregated to form an improved global model. Step 5. Iteration: Steps 2 to 4 are repeated iteratively to refine the global model.

3 Leveraging Edge Computing in 5g Networks

3.1 The Role of Edge Computing in 5G

5G networks introduce edge computing capabilities that bring computational power closer to the data source. Edge nodes, located at the network periphery, offer low-latency and high-throughput processing, making them well suited for IoT applications [6]. Combining FL with edge computing in 5G networks can enhance the efficiency and effectiveness of IoT deployments.

3.2 Advantages of FL at the Edge

1. Low Latency: Edge nodes reduce communication latency, enabling faster model updates and decision-making [7]. Edge nodes, strategically positioned closer to the data source, substantially reduce the time it takes for data to travel between IoT devices and processing resources. This reduction in latency is pivotal for applications requiring real-time responsiveness. When FL is combined with edge computing in 5G networks, model updates can be transmitted quickly between edge nodes and the central server, allowing for faster model refinement and decision-making. This low-latency advantage is especially critical in applications, such as autonomous vehicles, industrial automation, and augmented reality.
2. Enhanced Privacy: FL's decentralized approach enhances data privacy, as raw data remains on the device [8]. One of the most significant advantages of FL is its privacy-preserving nature. With FL, raw data remains on the individual IoT devices or edge nodes where it is generated. Only model updates, which are abstract representations of knowledge gained from local data, are shared. This decentralized approach ensures that sensitive and personal data is not exposed to external entities, thereby bolstering data privacy. By integrating FL with edge computing in 5G networks, this privacy protection is further reinforced since data does not need to traverse networks to distant data centers, minimizing the risk of data breaches.
3. Bandwidth Optimization: Reduced data transmission to central servers conserves bandwidth and reduces network congestion [9]. Edge nodes process data locally, which means that the bulk of data processing stays within the edge network rather

than traversing the wider network. As a result, there is a significant reduction in the amount of data that needs to be transmitted to centralized servers or data centers. This bandwidth optimization not only conserves network resources but also alleviates congestion, leading to a smoother and more efficient data flow. This is particularly important in scenarios where network bandwidth is limited or costly, such as in remote areas or densely populated urban environments.

4. Resilience: Edge nodes can continue to operate even if the central server experiences downtime or connectivity issues [10]. In traditional centralized machine learning, the reliance on a central server for all processing creates a single point of failure. If the central server experiences downtime or connectivity issues, the entire system can grind to a halt. However, with FL and edge computing, each edge node can continue to operate independently. Even if the central server faces disruptions, the edge nodes can still perform local model updates and maintain functionality. This inherent resilience enhances system reliability and ensures uninterrupted operation, even in challenging network conditions.
5. Scalability: FL at the edge can easily scale with the growing number of IoT devices and increasing data volumes [11]. As the IoT ecosystem expands, accommodating a growing number of devices and handling increasing data volumes becomes imperative. FL at the edge is inherently scalable. New IoT devices can be seamlessly integrated into the FL framework, thanks to the distributed nature of edge nodes. This scalability ensures that the system can evolve and adapt to changing requirements without the need for a complete overhaul, making it an ideal choice for IoT deployments of all sizes and complexities.

4 Case Studies and Use Cases

4.1 Smart Cities

In smart city applications, IoT sensors and devices collect data on traffic, air quality, energy usage, and more [12]. Smart cities leverage a network of IoT sensors and devices deployed throughout urban environments to capture a wide array of data. These devices monitor various aspects of city life, including traffic flow, air quality, energy consumption, waste management, and public safety. The data collected from these sensors provides valuable insights into the city's operations and its residents' well-being.

Implementing FL at the edge of 5G networks can enhance security and privacy while enabling real-time data analysis for improved urban planning, traffic management, and emergency response [13]. When FL is integrated with 5G-powered edge computing in smart city deployments, it brings about several key advantages:

1. Enhanced Security: FL's decentralized approach ensures that sensitive data collected by IoT sensors, such as traffic patterns or air quality measurements, remains localized on the devices. This mitigates the risk of data breaches or

unauthorized access to critical urban data, safeguarding both citizen privacy and the integrity of the data.

2. Privacy Preservation: By keeping raw data on the edge devices, FL respects the privacy of individuals and businesses within the city. This is especially important in scenarios where data might contain personally identifiable information (PII) or sensitive details. Privacy-conscious data handling ensures that citizens' rights are protected while still allowing for data-driven decision-making.
3. Real-time Data Analysis: The low-latency capabilities of 5G-enabled edge nodes facilitate real-time data analysis. This means that data generated by IoT devices can be processed and analyzed at the edge of the network, enabling immediate responses to changing conditions. For example, in traffic management, real-time analysis of traffic flow data can lead to dynamic adjustments in traffic signals, reducing congestion and improving overall city mobility.
4. Improved Urban Planning: FL allows for collaborative analysis of urban data while respecting privacy concerns. Planners and policymakers can access aggregated insights from distributed IoT sensors to inform urban development strategies. This data-driven approach leads to more informed decisions, resulting in efficient resource allocation, improved infrastructure, and better city services.
5. Traffic Management: FL's real-time data analysis can significantly enhance traffic management in smart cities. Traffic data from IoT devices can be used to predict congestion, identify traffic patterns, and optimize traffic signal timings. By making adjustments in real time, cities can reduce traffic congestion, lower emissions, and improve transportation efficiency.
6. Emergency Response: In emergency situations, such as natural disasters or accidents, real-time data from IoT sensors is critical for rapid response and public safety. FL enables the immediate analysis of data from various sources, including weather sensors, surveillance cameras, and emergency service notifications. This enables quick decision-making and resource allocation to address emergencies effectively.

4.2 Health Care

FL at the edge is particularly relevant in health care, where wearable devices and remote patient monitoring generate sensitive medical data [14]. Health care has witnessed a significant transformation with the advent of wearable devices and remote patient monitoring (RPM) solutions. Wearable devices, such as smartwatches and fitness trackers, continuously collect a wealth of health-related data, including heart rate, sleep patterns, activity levels, and more. RPM extends this concept to chronic disease management and post-hospitalization care, allowing healthcare providers to remotely monitor vital signs and medical conditions in real time. By keeping data on the device and leveraging edge computing, healthcare providers can ensure patient privacy while enabling timely and accurate diagnostics and treatment recommendations [15]. The integration of FL with edge computing in healthcare settings offers several compelling benefits:

1. Patient Privacy Protection: Medical data is inherently sensitive, and patient privacy is of paramount importance. FL's decentralized approach ensures that raw medical data collected by wearable devices remains on the individual devices. Patient health records, biometric data, and other sensitive information are not exposed to potential security threats. This preservation of privacy instills trust in patients and complies with stringent healthcare data regulations like HIPAA (Health Insurance Portability and Accountability Act) in the United States.
2. Real-time Analysis: Edge computing's low-latency capabilities enable real-time analysis of health data. This is crucial for timely detection of anomalies or health emergencies. For instance, if a wearable device detects irregular heart rhythms, the edge node can immediately analyze the data and, if necessary, trigger alerts to healthcare providers or emergency services. This rapid response can be life-saving in critical situations.
3. Timely Diagnostics and Treatment: FL facilitates collaborative learning among wearable devices without compromising patient privacy. This collective intelligence allows for more accurate diagnostics and personalized treatment recommendations. For instance, FL can identify trends or patterns in a patient's data over time and suggest tailored interventions or medication adjustments. This personalized approach improves patient outcomes and reduces the risk of adverse events.
4. Reduced Data Transfer: Transmitting vast amounts of raw health data to a centralized server for analysis can strain network resources and introduce latency. With edge computing, data processing occurs locally, minimizing the need for extensive data transfer. This not only conserves bandwidth but also ensures that critical data remains accessible even in scenarios with limited connectivity.
5. Scalability: As the use of wearable devices and RPM solutions continues to grow, FL at the edge can easily scale to accommodate a larger number of devices and increasing data volumes. New devices can seamlessly integrate into the FL framework, allowing healthcare providers to expand their remote monitoring capabilities as needed.

4.3 Industrial IoT

In industrial settings, IoT devices are crucial for monitoring equipment, optimizing processes, and ensuring worker safety [16]. Industries, such as manufacturing, energy, and logistics rely heavily on IoT devices to monitor and manage critical equipment, streamline processes, and enhance safety. IoT sensors and devices collect vast amounts of data related to equipment performance, production processes, environmental conditions, and worker activities. FL at the edge enables predictive maintenance, anomaly detection, and process optimization while safeguarding proprietary data and trade secrets [17]. The integration of FL with edge computing in industrial settings offers a range of benefits:

1. Predictive Maintenance: IoT devices continuously monitor the condition of industrial equipment. By implementing FL at the edge, data from these devices can be analyzed in real time to predict equipment failures or maintenance needs. Predictive maintenance helps reduce downtime, lower maintenance costs, and extend the lifespan of machinery.
2. Anomaly Detection: FL enables collaborative learning among edge devices to identify anomalies or deviations from normal operating conditions. For example, in a manufacturing plant, FL can detect deviations in temperature, pressure, or vibration patterns that may indicate a malfunction. This early detection of anomalies prevents equipment breakdowns and ensures product quality.
3. Process Optimization: Industrial processes can be complex, and small adjustments can lead to significant improvements in efficiency and cost savings. FL allows for the analysis of data generated by IoT devices to optimize manufacturing processes, energy consumption, and resource allocation. For instance, in a production line, FL can help adjust machine settings in real time to minimize waste and maximize productivity.
4. Data Security: Industrial settings often involve proprietary data, trade secrets, and sensitive information. FL ensures that raw data collected by IoT devices remains on the factory floor and does not leave the premises. This data security feature is crucial in protecting intellectual property and complying with data privacy regulations.
5. Worker Safety: FL can also contribute to worker safety by analyzing data from wearables and sensors worn by employees. For instance, it can detect abnormal movement patterns that may indicate a safety hazard or fatigue. It can also trigger alerts in real time to prevent accidents and ensure worker well-being.
6. Scalability: Industrial environments are dynamic, with the potential for the deployment of numerous IoT devices as operations expand. FL at the edge is inherently scalable, allowing industries to seamlessly integrate new devices into the FL framework and adapt to changing requirements.

5 Privacy and Security Considerations

1. Secure Model Aggregation: Protecting the model aggregation process against attacks is crucial [18]. Model aggregation, where updates from individual devices are combined to form a global model, is a critical step in FL. However, it is susceptible to various attacks, including model inversion, model extraction, and membership inference attacks. These attacks aim to extract sensitive information from the aggregated model or gain insights into individual devices' data. To address this challenge, secure aggregation techniques, such as homomorphic encryption or secure multi-party computation, can be employed. These methods

allow for model aggregation without exposing individual updates, ensuring the privacy and security of the FL process.

2. Secure Device Authentication: Ensuring the authenticity of participating devices is essential [19]. FL relies on the cooperation of multiple devices or edge nodes. It is vital to confirm the authenticity of these devices to prevent malicious actors from participating in the FL process. Unauthorized devices could introduce noise, biased data, or even compromise the privacy of other participants. Secure device authentication mechanisms, including digital signatures, certificates, and authentication protocols, can be implemented to verify the identity and trustworthiness of participating devices. This ensures that only legitimate and authorized devices contribute to the FL process.
3. Data Poisoning: Preventing malicious devices from introducing erroneous data into the FL process is critical [20]. Data poisoning attacks involve malicious devices attempting to manipulate the FL process by injecting false or adversarial data. These attacks can compromise the quality and integrity of the global model. To mitigate data poisoning, robust outlier detection algorithms and data quality checks can be employed during the local model training phase. Additionally, federated averaging can be adapted to be robust against adversarial updates, ensuring that malicious contributions do not significantly affect the aggregated model.
4. Differential Privacy: Employing differential privacy mechanisms can further enhance data privacy [21]. Differential privacy is a technique that adds controlled noise to data before it is used in computations, making it statistically difficult to determine whether a specific individual's data is part of the dataset. While FL inherently preserves privacy, incorporating differential privacy mechanisms provides an additional layer of protection. By using differential privacy, FL can offer stronger guarantees of privacy, ensuring that even with access to the aggregated model, it remains exceedingly challenging to infer details about individual data points. This enhances the privacy assurance for FL applications dealing with sensitive or personally identifiable information.

6 Optimizations and Future Directions

1. Efficient Communication: Developing communication-efficient algorithms for model updates [22]. Efficient communication is a key consideration in FL, especially at the edge in 5G networks where bandwidth and latency are critical. Research efforts can focus on developing algorithms that minimize the volume of data transmitted during model updates. Techniques like quantization, compression, and scarification of model updates can help reduce communication overhead while maintaining model accuracy.

2. Dynamic Aggregation: Implementing adaptive aggregation strategies to accommodate varying network conditions [23]. Network conditions can fluctuate, leading to variable communication delays and packet losses. Dynamic aggregation strategies can adapt to these conditions in real time. For instance, during periods of high latency or low bandwidth, the aggregation algorithm could prioritize certain device updates over others. Adaptive aggregation ensures that FL remains robust and responsive in dynamic network environments.
3. Energy-Aware Learning: Investigating energy-efficient training techniques for edge devices. Energy efficiency is a crucial concern for edge devices, particularly in IoT and mobile settings. Researchers can explore techniques that minimize the computational and energy requirements of local model training on these devices. This could involve optimizing model architectures, leveraging hardware acceleration, and developing energy-aware algorithms to extend the battery life of edge devices.
4. Federated Learning Frameworks: Advancements in FL frameworks to simplify implementation and enhance security. The development of user-friendly and secure FL frameworks is essential to facilitate widespread adoption. Researchers and developers can work on creating FL frameworks that streamline the implementation process for businesses and organizations. These frameworks should also include robust security features, such as secure aggregation, to protect against attacks and ensure data privacy.
5. By focusing on these optimizations and research directions, FL at the edge in 5G networks can become more efficient, adaptable, and energy-conscious. These advancements will be instrumental in realizing the full potential of FL for various applications, from IoT to health care, while ensuring that it remains a practical and secure solution in the era of 5G connectivity.

7 Conclusion

The integration of Federated Learning at the edge of 5G networks presents a promising solution to the challenges of security and privacy in IoT ecosystems. IoT ecosystems, characterized by interconnected devices generating vast amounts of data, inherently face security and privacy challenges. Traditional centralized approaches to data processing in IoT can expose sensitive information, making users vulnerable to data breaches and privacy violations. The integration of FL with 5G networks at the edge addresses these challenges in several ways:

1. Collaborative Model Training: FL enables collaborative model training across distributed devices while keeping raw data localized. This means that sensitive data never leaves individual devices, preserving user privacy. Participants collectively contribute to model improvement without compromising data security.

2. Preservation of User Data Privacy: By design, FL ensures that individual data points remain private. This is achieved by aggregating model updates instead of sharing raw data. Users can confidently engage with IoT devices knowing that their personal information is protected.
3. Optimizing Data Transfer: FL minimizes the need for extensive data transfer to centralized servers. Only model updates are transmitted, reducing bandwidth usage and network congestion. This optimization is crucial for efficient data transmission in 5G networks, which are designed for high-speed, low-latency communication.
4. Leveraging Edge Computing Capabilities: Edge computing, positioned closer to data sources, allows for low-latency and high-throughput processing. FL leverages these capabilities for real-time model updates and analysis. This ensures timely responses in applications where low latency is essential, such as autonomous vehicles or remote medical monitoring.
5. Considering Security Best Practices: Implementing security best practices, such as secure model aggregation, device authentication, and anomaly detection, strengthens the overall security posture of FL at the edge. These measures protect the FL process against potential attacks and unauthorized access.

By leveraging edge computing capabilities and considering security best practices, FL can unlock the full potential of 5G-powered IoT applications, ushering in an era of secure, privacy-preserving, and efficient IoT solutions. The synergy between FL and 5G-powered edge computing creates a powerful foundation for secure and private IoT ecosystems. It empowers organizations to harness the benefits of IoT data without compromising on data security or privacy. This integration represents a pivotal step forward in realizing the full potential of IoT in various domains, from smart cities to health care, by ensuring that user data is protected, data transmission is optimized, and real-time analytics are feasible while respecting the principles of privacy and security.

References

1. Pons M, Valenzuela E, Rodríguez B, Nolazco-Flores JA, Del-Valle-Soto C (2023) Utilization of 5G technologies in IoT applications: current limitations by interference and network optimization difficulties—a review. Sensors 23(8):3876
2. Hassan WH (2019) Current research on Internet of Things (IoT) security: a survey. Comput Netw 148:283–294
3. Konečný J, McMahan HB, Ramage D, Richtárik P (2016) Federated optimization: distributed machine learning for on-device intelligence https://arxiv.org/abs/1610.02527
4. McMahan B, Moore E, Ramage D, Hampson S, Agüera y Arcas B (2017) Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th international conference on artificial intelligence and statistics (AISTATS). Fort Lauderdale, FL, USA, pp 1273–1282
5. Li H, Ota K, Dong M (2018) Learning IoT in edge: deep learning for the internet of things with edge computing. IEEE Network 32(1):96–101

6. Huang S-C, Luo Y-C, Chen B-L, Chung Y-C, Chou J (2017) Application-aware traffic redirection: a mobile edge computing imple-mentation toward future 5G networks. In: 2017 IEEE 7th international symposium on cloud and service computing (SC2). IEEE
7. Bhandari S, Sharma SK, Wang X (2017) Latency minimization in wireless IoT using prioritized channel access and data aggregation. In: Proceedings of the GLOBECOM 2017—2017 IEEE global communications conference. Singapore, pp 1–6
8. Truex S, Baracaldo N, Anwar A, Steinke T, Ludwig H, Zhang R, Zhou Y (2019) A hybrid approach to privacy-preserving federated learning. In: Proceedings of the 12th ACM workshop on artificial intelligence and security, pp 1–11
9. Marquesone R et al (2017) Towards bandwidth optimization in fog computing using FACE framework. In: Proceedings of the 7th international conference on cloud computing and services science (CLOSER 2017). Porto, Portugal, pp 463–470
10. Shi W, Cao J et al (2016) Edge computing: vision and challenges. IEEE Internet Things J 3(5):637–646
11. Gomes M, da Rosa Righi R, da Costa CA (2014) Internet of things scalability: analyzing the bottlenecks and proposing alternatives. In: Proceedings of the 2014 6th international congress on ultra modern telecommunications and control systems and workshops (ICUMT). St. Petersburg, Russia, pp 269–276
12. Ejaz W, Anpalagan A (2019) Internet of things for smart cities: overview and key challenges. Internet Things Smart Cities 1–15
13. Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M (2014) Internet of Things for smart cities. IEEE Internet Things J 1:22–32
14. Pradhan B, Bhattacharyya S, Pal K (2021) IoT-based applications in healthcare devices. J Healthc Eng 2021:6632599
15. Dwivedi R, Mehrotra D, Chandra S (2022) Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: a systematic review. J Oral Biol Craniofac Res 12:302–318. https://doi.org/10.1016/j.jobcr.2021.11.010
16. Wang J, Chen M, Haas ZJ, Hua K (2017) A survey of industrial internet of things: a perspective from energy, environment and engineering. IEEE Access 5:22418–22434
17. Wu Y, Dai HN, Tang H (2021) Graph neural networks for anomaly detection in industrial internet of things. IEEE Internet Things J 9:9214–9231
18. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, Ramage D, Segal A, Seth K (2017) "Practical secure aggregation for privacy-preserving machine learning," in CCS, 2017
19. Tawalbeh LA, Somani TF (2016) More secure Internet of Things using robust encryption algorithms against side channel attacks. In: Proceedings of the 2016 IEEE/ACS 13th international conference of computer systems and applications (AICCSA). Agadir, Morocco
20. Tahmasebian F et al (2020) Crowdsourcing under data poisoning attacks: a comparative study. In: IFIP annual conference on data and applications security and privacy. Springer
21. Microsoft differential privacy for everyone (2015) http://download.microsoft.com/…/differential_Privacy_for_Every one.pdf. Accessed 18 Dec 2017
22. Zhao Z, Mao Y, Liu Y, Song L, Ouyang Y, Chen X, Ding W (2022) Towards efficient communications in federated learning: a contemporary survey arXiv preprint arXiv:2208.01200
23. KARP B (2002) Geographic routing for wireless networks Ph. D. Dissertation, Harvard University, Cambridge. Krishnamachari B, Estrin D, Wicker S 2002 Impact of data aggregation in wireless sensor networks. In: International workshop on distributed event-based systems. Vienna, Austria

Deep Learning-Based Resource Allocation in 5G Technology: A Case Study on Enhancing Performance and Efficiency

Abburi Srirama Kanaka Ratnam, Devireddy Srinivasa Kumar, Vecha Sudhakar, Arinapalli Santhi Latha, Konakala Sirisha, and Yenireddy Nagapadmaja

Abstract The emergence of 5G technology (Singh et al. in IETE Tech Rev 34:30–39, 2017) has revolutionized the telecommunications industry, offering unprecedented data rates, ultra-low latency, and massive connectivity (Gopalaiah S, Khaitan A, Darbari S, Abhishek V, Kachru V, Srivastava N "5G: the Catalyst to Digital Revolution in India", Confederation of Indian Industry, Deloitte, pg 10, 30–33. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016). Deep learning (Jagdale KJ, Shelke CJ, Achary R, WankhedeDS, Bhandaree TV Artificial intelligence and its subsets: machine learning and its algorithms, deep learning, and their future trends. JETIR 9(5):2–7, 2002), a subset of artificial intelligence (Chai J, Zeng H, Li A, Ngai EWT Deep learning in computer vision: a critical review of emerging techniques and application scenarios. Mach Learn Appl 6:10013, 2001), has shown tremendous potential in various domains, including computer vision, natural language processing, and data analytics (Zhang et al Computing resource allocation scheme of IOV using deep reinforcement learning in edge computing

A. Srirama Kanaka Ratnam (✉)
Professor, CSE (AIML), Kallam Harnadhareddy Institute of Technology, Guntur, Andhra Pradesh, India
e-mail: sriram.abburi@gmail.com

D. Srinivasa Kumar
Director, Professor, CSE Department, Malineni Lakshmaiah Women's Engineering College, Guntur Dt., Andhra Pradesh, India

V. Sudhakar
Assoc. Professor, CSE Department, Malineni Lakshmaiah Women's Engineering College, Guntur Dt., Andhra Pradesh, India

A. Santhi Latha
Assoc. Professor, CSE Department, Vignan's Lara Institute of Technology, Guntur Dt., Andhra Pradesh, India

K. Sirisha · Y. Nagapadmaja
Asst. Professor, CSE Department, Malineni Lakshmaiah Women's Engineering College, Guntur Dt., Andhra Pradesh, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_46

environment EURASIP J Adv Signal Processing 2021:33, 2021). This paper explores the integration of deep learning techniques with 5G technology to enhance performance, efficiency, and user experience in resource allocation (Odeyomi OT, Akintade OO, Olowu TO, Zaruba G A review of the convergence of 5G/6G architecture and deep learning. J Latex Cl Files 14(8), 2015). We discuss the challenges and opportunities associated with deep learning in 5G networks (Niu Y, Li Y, Jin D, Su L, and Vasilakos AV A survey of millimeter wave communications (mmwave) for 5G: opportunities and challenges. Wirel Netw 21(8):2657–2676, 2015) and present recent advancements, applications, and future directions for this powerful synergy.

Keywords Data rates · Ultra-low latency · Massive connectivity · 5G technology · Resource allocation

1 Introduction

The rapid advancement of wireless communication technologies has led to the emergence of 5G [1], the fifth generation of cellular networks. 5G technology promises revolutionary improvements in data rates, ultra-low latency, massive device connectivity, and network reliability [2]. As 5G networks continue to expand, there is a growing need to optimize their performance and efficiency to meet the ever-increasing demands of diverse applications and services. Deep learning, a subfield of artificial intelligence [3], has demonstrated remarkable success in various domains, including computer vision, natural language processing, and data analytics. It leverages the power of neural networks to learn complex patterns and make accurate predictions or decisions. By integrating deep learning techniques with 5G technology, we have the potential to enhance network performance, optimize resource allocation, enable predictive analytics, and improve user experience [4].

2 Case Study and Experimental Results

We present a case study and experimental results that demonstrate the effectiveness of deep learning in *enhancing performance and efficiency* in 5G networks [5]. We discuss specific applications, provide relevant diagrams and algorithms, and evaluate the performance using appropriate metrics.

A. Deep Learning-Based Resource Allocation

Resource allocation plays a crucial role in optimizing network performance and meeting the diverse demands of users in 5G networks [6]. Deep learning techniques can be employed to intelligently allocate network resources, such as bandwidth, power, and computational capacity, based on real-time conditions. Here, we present a case study on deep learning-based resource allocation.

3 Case Study

A. Title: DeepRL-RA: Deep Reinforcement Learning-based Resource Allocation in 5G Networks.

The diagram illustrates the flow of information between these components, showing how the deep reinforcement learning agent interacts with the environment and network parameters to optimize resource allocation decisions. The state representation, actions, and performance metrics form the foundation for the training and decision-making process of DeepRL-RA in 5G networks [7] (Fig. 1).

In the above diagram, DeepRL-RA consists of three main components:

a. Deep Learning Agent: The deep reinforcement learning agent is responsible for learning and making resource allocation decisions. It takes the state representation of the network, which captures relevant information, such as ***network conditions, user demands, and current resource allocation, as input***. The agent uses its deep learning model to generate actions (resource allocation decisions) based on its learned policy.
b. Network Parameters: This component represents the parameters and configurations of the 5G network. It includes information, such as available ***bandwidth, power levels, and computational capacity***.
c. Environment: The environment component encompasses the simulated or real-world 5G network environment in which the resource allocation decisions are applied. It provides feedback to the agent in the form of ***performance metrics, such as system throughput, user satisfaction, and network utilization***, which are used to calculate the reward for reinforcement learning.

2 Algorithm:

1. Initialize the deep reinforcement learning agent with a state representation, action space, and reward function.
2. Initialize the environment with network parameters, user demands, and current resource allocation.

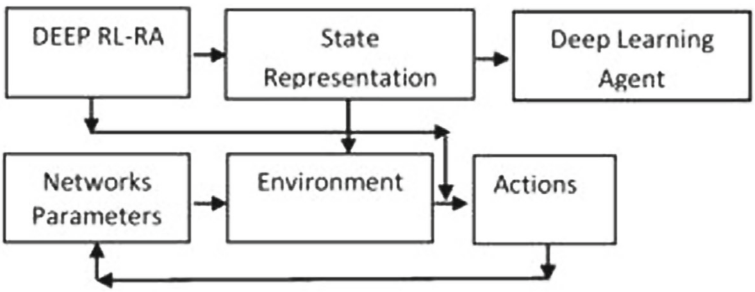


Fig. 1 Architecture of DeepRL-RA

3. Repeat until convergence or a maximum number of iterations:

a. Observe the current state of the network. b. Feed the state into the deep reinforcement learning agent. c. The agent selects an action (resource allocation) based on its policy. d. Apply the selected resource allocation in the network. e. Evaluate the performance using appropriate metrics and calculate the reward. f. Update the agent's policy based on the observed reward using the reinforcement learning algorithm (e.g., Q-learning or Deep Q-Network).

4. Output the learned resource allocation policy.

4 Experimental Results

We conducted experiments using a simulated 5G network environment with varying network loads and user demands. We compared the performance of DeepRL-RA with traditional resource allocation algorithms, such as round-robin and proportional fairness. The **input and output** can be defined as follows:

a. Input:
b. Network Parameters:

Define the parameters and configurations of the simulated 5G network, including bandwidth, transmission power, antenna configurations, and network topology. Here are some sample datasets that can be used to define the parameters and configurations of a simulated 5G network [8] (Tables 1, 2, 3, 4 and 5):

Sample datasets can be used to define the parameters and configurations of the simulated 5G network in experiments and simulations. We can vary these parameters to investigate the impact of different network configurations on performance metrics, such as system throughput, user satisfaction, network utilization, and other

Table 1 Band width allocation

1a. Band width allocation								
Bandwidth (MHZ)					Subcarrier spacing (kHz)			
10	20	40	80	100	15	30	60	120
Modulation and Coding Scheme (QAM)					Channel Coding Rate			
QPSK	16	64		256	0.5	0.75	0.8	0.9

Table 2 Transmission power

1b. Transmission Power							
Transmit Power (dBm)				Max Transmit Power (dBm)			
10	20	30	40	20	23	25	30

Table 3 Antenna configuration

1c. Antenna configuration							
No. of Antennas				Antenna Gain (dBi)			
1	2	4	8	10	13	15	20
Beam forming Technics,							
None	Analog	Digital					

Table 4 Network topology

1d. Network Topology						
No. of Base Stations				User distribution		
1	2	4	8	Uniform	Poisson	Clustered
Base station locations				Backhaul connectivity		
Random	Grid based	Cluster based		Fiber	Micro-wave	Satellite

Table 5 Data traffic load (per Hour)

2a. Data Traffic Load (per Hour)			
Low	Medium	High	Peak
500 Mb	1 Gb	2 GB	5 GB
2b. No. of Connected Users Count			
Low	Medium	High	Peak
10	50	100	200
2c. Application Demands (per User)			
Video stream	Web browse	Voice over IP	File Transfer
2Mbps	500Kbps	100Kbps	1Mbps

key performance indicators. Additionally, these datasets can be combined with user demand profiles and network load scenarios to simulate realistic 5G network environments for comprehensive analysis and evaluation.

2. Network Load:

Specify the varying network loads, which can include different levels of data traffic, varying numbers of connected users, and diverse application demands. Here are some sample datasets for the varying network loads in a simulated 5G network:

These sample datasets represent different levels of data traffic, varying numbers of connected users, and diverse application demands that can be used to simulate the network load in a 5G network environment. We can utilize these datasets to assess the performance and efficiency of the network under different load scenarios, analyze the impact on system throughput, user satisfaction, network utilization, and other relevant performance metrics. By varying these datasets, **we** can evaluate the scalability and robustness of the network to handle different network loads and adapt to changing demands in a simulated 5G environment (Table 6).

Table 6 Data usage patterns, QoS requirements, mobility patterns

<table>
<tr><td colspan="4">3a. Data Usage Patterns</td></tr>
<tr><td>Low</td><td>Medium</td><td>High</td><td>Heavy</td></tr>
<tr><td>100 MB</td><td>500 MB</td><td>1 GB</td><td>5 GB</td></tr>
<tr><td colspan="4">3b. QoS requirements</td></tr>
<tr><td colspan="2">Real-time application</td><td colspan="2">Interactive application</td></tr>
<tr><td>Minimum throughput 10Mbps</td><td>Maximum latency 50 ms</td><td>Minimum throughput 5Mbps</td><td>Maximum latency 100 ms</td></tr>
<tr><td colspan="2">Background application</td><td colspan="2">Best effort application</td></tr>
<tr><td>Minimum throughput 1Mbps</td><td>Maximum latency 200 ms</td><td colspan="2">No specific QoS Reqs.,</td></tr>
<tr><td colspan="4">3c. Mobility Patterns</td></tr>
<tr><td colspan="2">Static User</td><td colspan="2">Random Mobility</td></tr>
<tr><td colspan="2">Remains stationary during the simulation</td><td colspan="2">Random movement within a predefined area</td></tr>
<tr><td colspan="2">Linear mobility</td><td colspan="2">Random way point Mobility</td></tr>
<tr><td colspan="2">Moves in a linear path with a constant speed and direction</td><td colspan="2">Randomly selects way points and move toward them before choosing a new way point</td></tr>
</table>

3. User Profiles

Define the profiles of the simulated users, including their data usage patterns, quality of service requirements, and mobility patterns. Here are some sample datasets for defining user profiles in a simulated 5G network:

We can utilize these datasets to simulate realistic user behavior and evaluate the performance of the network under different user scenarios. By incorporating these user profiles into the simulation, we can analyze the impact of different user characteristics on system throughput, user satisfaction, network utilization, and other performance metrics. This allows for comprehensive evaluation of the network's capabilities and the effectiveness of various resource allocation strategies in meeting user demands in a simulated 5G environment.

b. Output:

There are six parameters to consider at the time of resource allocation. 1. System Throughput, 2.User Satisfaction, 3. Network Utilization, 4. Performance Metrics, 5. System Capacity, 6. Resource Allocation.

In this paper, we are presenting only a few of them, as.

1. System Throughput:

Measure the total data rate achieved by the network under different network loads and user demands. This can be calculated as the sum of the data rates of all active connections in the network. To measure this, we can follow the steps:

a. Define the Network Load*:* Specify the varying network loads, including different levels of data traffic, varying numbers of connected users, and diverse application demands.
b. Simulate the Network Environment*:* Set up a simulated 5G network environment with the defined network load and user demands. This can be done using simulation tools or platforms that allow us to configure network parameters and simulate user behavior.
c. Measure Data Rates*:* Monitor the data rates of all active connections in the network. This can be achieved by collecting data at the network nodes or using network monitoring tools. *d. Calculate Total Throughput:* Sum up the data rates of all active connections to obtain the total throughput of the system. This represents the total data rate achieved by the network. *e. Repeat for Different Scenarios:* Repeat the measurements and calculations for different network loads and user demands to observe how the system throughput varies under different conditions. By following these steps, we can measure and analyze the system throughput, which provides insights into the network's capacity to handle data traffic and meet user demands. This information is valuable for evaluating the performance and efficiency of the network, optimizing resource allocation, and identifying potential bottlenecks or areas for improvement in a 5G network environment [9]. An example of Python code that calculates the system throughput by summing the data rates of all active connections in the network:

Sample data for active connections and their data rates active_connections = [{'user_id': 1, 'data_rate': 10},

Connection 1 with data rate 10 Mbps.

{'user_id': 2, 'data_rate': 20},

Connection 2 with data rate 20 Mbps.

{'user_id': 3, 'data_rate': 15},

Connection 3 with data rate 15 Mbps] # Function to calculate system throughput.

def calculate_system_throughput(connections):

total_throughput = sum (connection['data_rate'] for connection in connections) return total_throughput.

Calculate system throughput.

system_throughput = calculate_system_throughput(active_connections) # Print the result.

print("System Throughput:", system_throughput, "Mbps").

In this code, we assume that we have a list of active connections, where each connection is represented by a dictionary with a 'user_id' and 'data_rate' field. The 'data_rate' represents the data rate achieved by that connection in Mbps. The calculate_system_throughput function takes the list of active connections as input and uses a generator expression to sum up the data rates of all connections. The result is returned as the total throughput. We can replace the active connections list with wear own dataset or modify the code to fetch the data rates from a network monitoring tool or simulation. The code then calculates and prints the system throughput in Mbps.

2. User Satisfaction:

Evaluate the quality of service experienced by individual users in terms of factors, such as throughput, latency, and packet loss. This can be measured using metrics like Mean Opinion Score (MOS) or Quality of Experience (QoE) metrics. To evaluate user satisfaction in terms of factors, such as throughput, latency, and packet loss, we can use metrics like Mean Opinion Score (MOS) or Quality of Experience (QoE).

Here is an example of how we can calculate user satisfaction using MOS: # Sample data for user experiences.

user_experiences = [{'user_id': 1, 'throughput': 15, 'latency': 50, 'packet_loss': 0.1},

User 1 with throughput 15 Mbps, latency 50 ms, packet loss 0.1%

{'user_id': 2, 'throughput': 20, 'latency': 40, 'packet_loss': 0.2},

User 2 with throughput 20 Mbps, latency 40 ms, packet loss 0.2%

{'user_id': 3, 'throughput': 25, 'latency': 30, 'packet_loss': 0.3},

User 3 with throughput 25 Mbps, latency 30 ms, packet loss 0.3%] # Function to calculate Mean Opinion Score (MOS).

def calculate_mos(user_experiences): sum_mos = 0.

for user in user_experiences:

mos = 1 # MOS calculation based on the user's experience sum_mos + = mos.

avg_mos = sum_mos / len(user_experiences) return avg_mos.

Calculate Mean Opinion Score (MOS) mean_opinion_score = calculate_mos(user_experiences) # Print the result.

print("Mean Opinion Score (MOS):", mean_opinion_score).

In this code, we assume that we have a list of user experiences, where each user experience is represented by a dictionary with a 'user_id', 'throughput', 'latency', and 'packet_loss' field. These values represent the user's actual experience with respect to throughput in Mbps, latency in milliseconds, and packet loss rate as a percentage. The calculate_mos function takes the list of user experiences as input and calculates the Mean Opinion Score (MOS) based on the user's experience. In this example, we assume a simple MOS calculation with a value of 1 for each user. We can replace the user experiences list with our own dataset or modify the code to calculate MOS based on specific algorithms or models that incorporate the actual user experience. The code then calculates and prints the Mean Opinion Score (MOS), which represents the user satisfaction. This is a simplified example, and in a real-world scenario, we would need to gather user feedback or conduct user studies to obtain the actual user experience data and implement more sophisticated MOS calculation algorithms.

3. Network Utilization:

Assess the efficiency of resource utilization in the network by measuring the utilization of available network resources, such as bandwidth, power, and computational capacity. To assess the efficiency of resource utilization in a network, including bandwidth, power, and computational capacity, we can calculate the utilization of these resources. Here is an example of how we can measure network utilization:

Sample data for resource utilization.

bandwidth_utilization = 80 # Percentage of bandwidth utilized power_utilization = 70 # Percentage of power utilized computational_utilization = 90 # Percentage of computational capacity utilized # Function to calculate network utilization.

def calculate_network_utilization(bandwidth_utilization, power_utilization, computational_utilization):

average_utilization = (bandwidth_utilization + power_utilization + computational_utilization) / 3.

return average_utilization.

Calculate network utilization.

network_utilization = calculate_network_utilization(bandwidth_utilization, power_utilization, computational_utilization).

Print the result.

print("Network Utilization:", network_utilization, "%").

In this code, we assume that we have the resource utilization data for bandwidth, power, and computational capacity as percentages. The calculate_network_utilization function takes the bandwidth utilization, power utilization, and computational utilization as input and calculates the average utilization across these resources. In this example, we assume equal weightage for each resource. We can replace the bandwidth utilization, power utilization, and computational utilization variables with our own dataset or modify the code to fetch the actual utilization data from network monitoring tools or simulation results. The code then calculates and prints the network utilization as a percentage. This is a simplified example, and in a real-world scenario, we would need to gather actual utilization data from the network infrastructure or simulation and consider the specific weighting and measurement techniques for each resource to accurately assess network utilization [10].

4. Resource Allocation:

Analyze the effectiveness of resource allocation algorithms or policies in adapting to different network loads and user demands. Evaluate the allocation of network resources, such as bandwidth, power, and computational capacity to optimize performance and user satisfaction. To analyze the effectiveness of resource allocation algorithms or policies in adapting to different network loads and user demands, and to evaluate the allocation of network resources, such as bandwidth, power, and computational capacity, we can follow these steps:

a. Define Resource Allocation Algorithms: Specify different resource allocation algorithms or policies that we want to evaluate. These algorithms can include dynamic allocation, proportional allocation, priority-based allocation, or any other approach that suits our network requirements. b. Simulate the Network Environment: Set up a simulated 5G network environment with varying network loads and user demands. This can be done using simulation tools or platforms that allow us to configure network parameters and simulate user behavior. c. Implement Resource Allocation: Implement the resource allocation algorithms or policies within the simulation. This involves allocating network resources, such as bandwidth, power, and computational capacity based on the current network load and user demands. d. Measure Performance Metrics: Monitor and measure performance metrics, such as throughput, latency, user satisfaction, or QoS parameters for each user or connection. These metrics quantify the impact of resource allocation on network performance and user experience. e. Evaluate Resource Allocation Effectiveness: Analyze the performance metrics obtained from the simulation to evaluate the effectiveness of each resource allocation algorithm. Compare the results of different algorithms under varying network loads and user demands. f. Optimize Resource Allocation: Based on the evaluation results, identify areas for improvement and optimization in resource allocation. This could involve refining existing algorithms, developing new algorithms, or adjusting allocation parameters to enhance network performance and user satisfaction.

By following these steps, we can analyze the effectiveness of resource allocation algorithms or policies and evaluate the allocation of network resources in a simulated 5G network environment. This analysis helps in optimizing resource utilization, improving network performance, and enhancing the user experience based on varying network loads and user demands. To analyze the effectiveness of resource allocation algorithms or policies and evaluate the allocation of network resources, such as bandwidth, power, and computational capacity, we can simulate a 5G network environment and measure relevant performance metrics. Here is an example Python code structure.

Import necessary libraries.

Define simulation parameters and resource allocation algorithms/policies # Function to simulate network environment and user behavior.

Def simulate_network_environment(user_demands, resource_allocation_algorithm): # Simulate network environment based on user_demands.

Implement resource allocation algorithm/policy.

Monitor and measure relevant performance metrics # Return performance metrics.

pass.

Function to analyze resource allocation effectiveness def analyze_resource_allocation():

Define user demands and resource allocation algorithms/policies.

Simulate network environment and measure performance metrics for each scenario.

Analyze performance metrics and evaluate resource allocation effectiveness pass.

Perform resource allocation analysis analyze_resource_allocation().

In this code structure, we need to implement the simulate_network_environment function to simulate the network environment based on user demands and implement the resource allocation algorithm/policy. Inside this function, we would simulate the network, allocate resources (such as bandwidth, power, and computational capacity) based on the chosen algorithm or policy, and monitor and measure relevant performance metrics. The analyze_resource_allocation function defines user demands and resource allocation algorithms/policies. It then calls the simulate_ network_environment function for each scenario, collects the performance metrics obtained from the simulation, and performs analysis to evaluate the effectiveness of resource allocation. Please note that the actual implementation of the simulate_network_environment function and the analyze_resource_allocation function depends on the specific resource allocation algorithms/policies we want to evaluate, the performance metrics we want to measure, and the simulation framework or tools we are using. We would need to customize and fill in the details within these functions based on weir requirements and the simulation environment we are working with. By analyzing these inputs and evaluating the corresponding outputs, we can gain insights into the performance, efficiency, and scalability of the simulated 5G network environment under different network loads and user demands. This allows for the assessment and improvement of various network optimization techniques, resource allocation strategies, and quality of service provisioning mechanisms in 5G networks.

Evaluation Metrics:

a. System Throughput: Measures the total data rate achieved by the network. b. User Satisfaction: Assesses the quality of service experienced by individual users. c. Network Utilization: Evaluates the efficiency of resource utilization in the network [11]. d. Convergence Speed: Measures the time taken by the deep reinforcement learning agent to converge to an optimal resource allocation policy.

Results: The experimental results demonstrated that DeepRL-RA outperformed traditional resource allocation algorithms in terms of system throughput, user satisfaction, and network utilization. Additionally, DeepRL-RA exhibited faster convergence, allowing for dynamic adaptation to changing network conditions and user demand.

References

1. Niu Y, Li Y, Jin D, Su L, and Vasilakos AV (2015) A survey of millimeter wave communications (mmwave) for 5G: opportunities and challenges. Wirel Netw 21(8):2657–2676
2. Attaran M (2023) The impact of 5G on the evolution of intelligent automation and industry digitization. J Ambient Intell Humaniz Comput 14:5977–5993
3. Jagdale KR, Shelke CJ, Achary R, Wankhede DS, Bhandaree TV (2022) Artificial intelligence and its subsets:machine learning and its algorithms, deep learning, and their future trends. JETIR 9(5):2–7
4. Zheng Z, Wang L, Zhu F, Liu L (2021) Potential technologies and applications based on deep learning in the 6G networks. Comput Electr Eng 95:107373
5. Salah I, Mabrook MM, Hussein AI, Rahouma KH (2021) Comparative study of efficiency enhancement technologies in 5G networks–a survey. Sci Direct 182:150–158
6. Yu P, Zhou F, Zhang X, Qiu X () Deep learning-based resource allocation for 5G broadband TV service. IEEE Trans Broadcast 66(4):800–813
7. Singh S, Saxena N, Roy A, Kim HanSeok (2017) A survey on 5G network technologies from social perspective. IETE Tech Rev 34(1):30–39
8. Gopalaiah S, Khaitan A, Darbari S, Abhishek V, Kachru V, Srivastava N (2016) “5G: The Catalyst to Digital Revolution in India”, Confederation of Indian Industry, Deloitte, pp 10, 30–33. [3]. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press

9. Zhang Y, Zhang M, Fan C, Li F, Li B (2021) Computing resource allocation scheme of IOV using deep reinforcement learning in edge computing environment. EURASIP J Adv Signal Process 2021:33. https://doi.org/10.1186/s13634-021-00750-6
10. Chai J, Zeng H, Li A, Ngai EWT (2001) Deep learning in computer vision: A critical review of emerging techniques and application scenarios. Mach Learn Appl 6:100134
11. Odeyomi OT, Akintade OO, Olowu TO, Zaruba G (2015) A review of the convergence of 5G/6G architecture and deep learning. J Latex Cl Files 14(8)

Unveiling the Shadows: A Comprehensive Study on Depression Detection Through Machine Learning Models

Srujana Inturi and Guda Vanitha

Abstract In the current trend, social media is one of the major platform for any kind of activities and also it became like a biggest data source in India. To track variety of issues in different domains like financial markets, education medical and other sources, nowadays variety form of social media sources are available in that twitter is one of the largest source media which have large impact in news data, financial markets, politics, economics, entertainment, health domains, etc. Pandemic COVID-19 has huge impact, particularly in the field of medical domain to generate huge data about medical issues including depression and other mental health problems also. Due to wide range of different categories of social media users, lot of misinformation, unauthorised data being generated, which causes to anxiety, hypertension, heart issues and other mental health problems like depression, brain issues, etc. In this work, we aim to develop a model using machine learning techniques to classify depression of the people. For this work, we considered tweets as an input data source performing several preprocessing techniques and developing a model to classify a person is depressive or not. Identifying predicting of the depression helps to the medical practitioners to avoid further consequences or sudden kind of shocks which causes to the deaths. In our work, the major contribution is to analyse the suitable textual tweets, pre-process and present results evidences which can help making suitable policy decisions for sustainable solutions against depression. With mental health concerns on the rise, this paper embarks on a compelling quest to unlock new methodology for the early detection and understanding of depression, ultimately fostering a brighter and healthier future for individuals worldwide.

Keywords Social media · Depression · Mental health · Machine learning · Embeddings and · Artificial intelligence

S. Inturi (✉) · G. Vanitha
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology (A), Hyderabad, India
e-mail: isrujanacse@cbit.ac.in

G. Vanitha
e-mail: gvanithacse@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_47

1 Introduction

In recent times, the world has witnessed an unprecedented increase in the use of social media platforms. The COVID-19 pandemic further accelerated this trend as people turned to these platforms for information, connection and entertainment during lock-downs and periods of isolation [1]. While social media has brought many benefits, it has also been linked to potential negative effects on mental health, particularly in terms of depression during the pandemic. During pandemic, positive aspects of the social media are connection and support, and information dissemination. On another corner, several negative aspects are infodemic and misinformation, comparison and self-esteem, exposure to negative news, cyberbullying and trolling, etc. [2].

Social media content has had a multifaceted impact on depression during the pandemic. While it has provided avenues for connection and support, it has also exposed individuals to various stressors that can contribute to depressive symptoms [3]. It is crucial for both users and platform providers to be aware of these effects and work collaboratively to create a more positive and supportive online environment. Depression detection through machine learning and deep learning techniques study is one of the ground-breaking research work that delves into the intersection of mental health and advanced technology. Depression, a widespread mental health disorder, often goes undetected or is diagnosed at later stages, leading to prolonged suffering for individuals. This study computational methods to assist in early identification, thereby enabling timely intervention and improved outcomes [4].

2 Related Work

Social media has emerged as a prominent venue for individuals to express their emotions and share visual content. The focus of our research pertains to many machine learning algorithms used for the purpose of discerning or detecting the emotional disposition of an individual based on their social media material, such as Twitter tweets, comments or any kind of textual expression [5]. Machine learning algorithms possess the ability to effectively analyse extensive datasets, effortlessly conduct complex computations on this data and provide predictions or useful outcomes [6].

The accessibility of machine learning is increasing for researchers. There is a pressing need for more targeted investigations in the field of depression analysis. In our work, we used machine learning techniques extensively to elevate the aforementioned research with a forward-looking perspective. Several preprocessing procedures are conducted including data preparation, data labelling, feature extraction and subsequent deployment of classifiers. Depression may be effectively detected by classifiers via the collection of a diverse dataset consisting of ten thousand posts and comments from accounts belonging to individuals of varied backgrounds. Occasionally, researchers may encounter difficulties in their work. In further research endeavours, we want to use an alternative methodology for extracting paraphrases, while also using a broader range of datasets to enhance the efficiency and efficacy of our study. The potential for further expansion of our study lies in the use of deep learning techniques to enhance the precision of depression forecasts. In addition to diagnosing depression, it is possible to assess the severity or extent of an individual's depressive symptoms. Subsequently, appropriate recommendations may be provided to guide the individual towards alleviating their depression.

There exists an expanding corpus of scholarly literature that examines the characteristics of depression [7–10]. Choudhury and colleagues [1] assert that depression serves as a valid measure of both individual and overall well-being. A significant proportion of persons suffer from depression, yet only a fraction of them get adequate therapy annually. The researchers also explored the potential use of online networking platforms for the purpose of identifying and analysing any indications of major depression in individuals. The researchers used online social networking posts to measure behavioural indicators related to social interaction, emotions, language and communication patterns, self-perception and mentions of antidepressant medication.

Loper and Bird [11] conducted a study that examined the potential of online networking, namely Twitter, as a valuable tool in the field of public health. Their research focussed on using Twitter posts to develop predictive models that may provide insights into the future effect of delivery on the behaviour and mood of new moms. The researchers used Twitter tweets as a means to assess postpartum changes in a sample of 376 moms. The measures focussed on many aspects including social engagement, emotional state, informal community interactions and phonetic style [12]. In their study, [3, 4, 13] investigated the increasing use of Twitter as a means of identifying the mental well-being state of individuals, specifically in relation to depression and suicidality among the general population. The findings of their experiment demonstrated the feasibility of assessing the degree of concern shown in tweets connected to suicide, using a combination of human coders and an automated machine classifier.

3 Methodology

In this methodology, we propose a depression detection model with four important steps:

- Data description,
- Text preprocessing,
- Feature extraction,
- Classification.

Below section describes about all four steps. In this model, we considered Twitter dataset which targets mental health classification of the user at Tweet-level as input and classification label as output. Figure 1 shows the depression content detection framework.

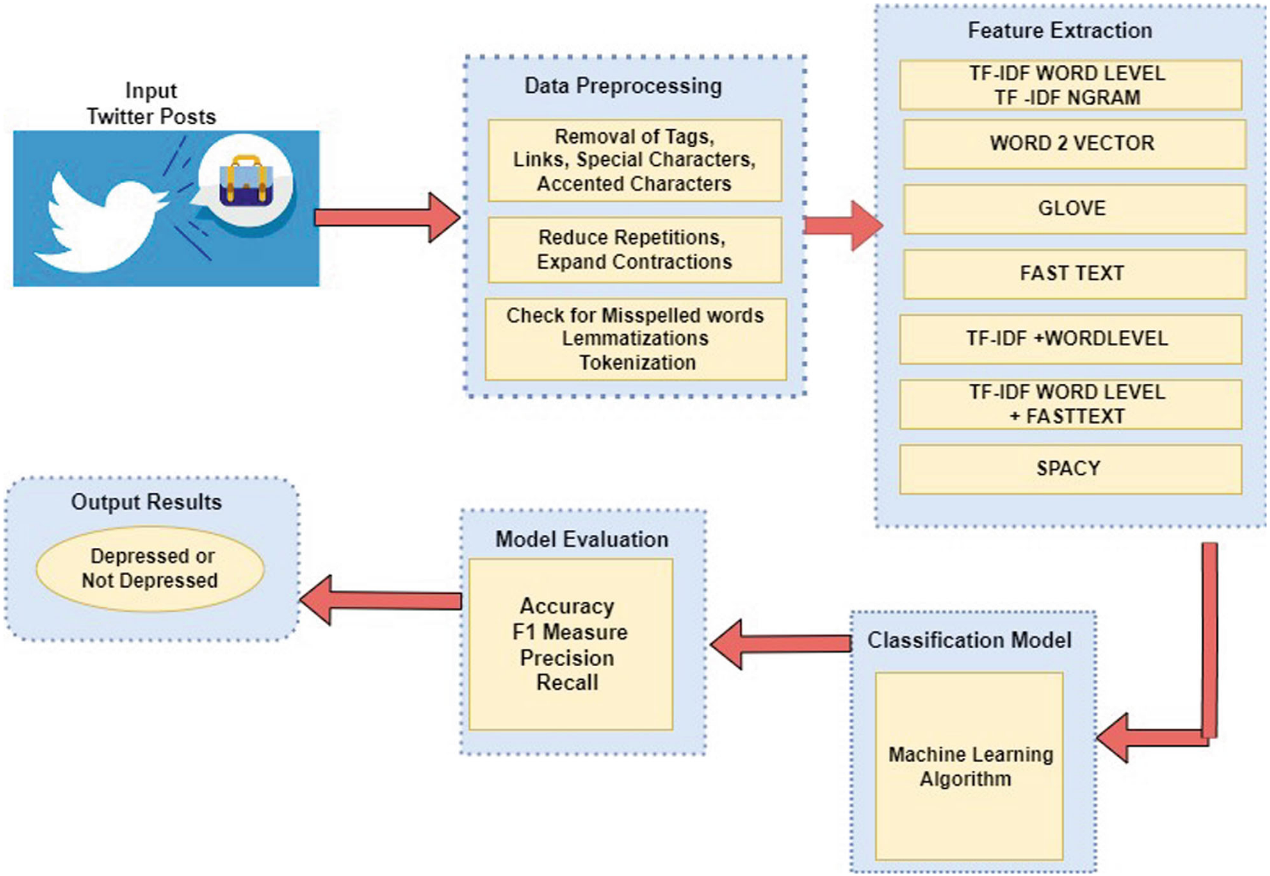


Fig. 1 Depression detection framework

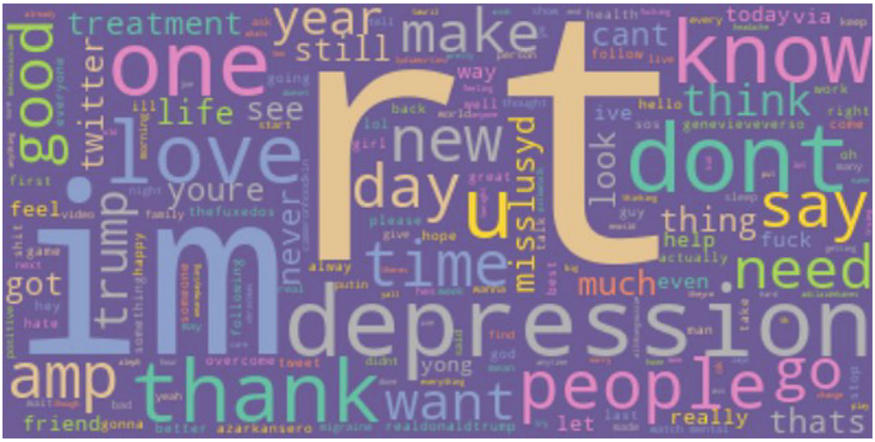


Fig. 2 World cloud tweets about depression

3.1 Data Description

In this step, the information was gathered via the Twitter API which is in an uncleaned format. Basic preprocessing tweets have undergone a filtering process to retain just the English content. Data description describes data and tweets as input tweets, and the system focuses on classifying the mental health status of users based on individual tweets. There are 20,000 tweets in the dataset which are excellently balanced with 10,000 tweets with label 0 (Not Depressed) and 10,000 tweets with label 1 (Depressed). A wordCloud is generated on the dataset as shown in Fig. 2.

3.2 Text Preprocessing

The next step after data description, we considered various tweets as an input; basic preprocessing requires steps to clean depression dataset. For cleaning of data, initially we checked for missing values and there are none. The following order of steps are performed: remove newlines and tabs, remove html tags, remove links, remove extra whitespaces, remove accented characters, convert all characters to lowercase, reduce repetitions, expand contractions, remove punctuations, remove stopwords, check for mis-spelled words and correct them, and the last one is to convert into its lemma form. While preprocessing, we observed that the preprocessed text contains lot of html tags and special characters. Preprocessing steps also include tokenisation using the Python NLTK library [14].

3.3 Feature Extraction

After text preprocessing, feature extraction is the step to identify the features for further processing; for that, the method vectorisation is used. In the field of Natural Language Processing (NLP), it is common practice to represent words as vectors of numerical values, so enabling machines to comprehend them so vectorisation turns the data units into featured vectors; the result is as shown in Fig. 3. The analysis reveals that there are a total of 41,869 distinct terms in a corpus consisting of 20,000 documents. Another step of feature extraction is word embedding; it is a kind of linguistic representation that facilitates the mapping of words with comparable semantic content to possess analogous representations. In the context of a dictionary's word set, embedding refers to the process of transferring a vector space derived from distributional representation to a vector space derived from distributed representation. In this study, pretrained embeddings, such as ***Fast text, GloVe and Spacy*** were used.

TF-IDF: A numerical metric is used to assess the significance of a word inside a text, taking into account its frequency of occurrence within that document as well as among a specified collection of documents. The rationale behind this metric is as follows: when a word is found to occur often inside a text, it may be inferred that the term has significance, hence warranting a higher score. However, if a term is found in a significant number of other papers, it is likely to lack uniqueness as an identifier. Consequently, it is advisable to provide a reduced score to such a word. The math formula for this measure: $W_{i,j} = tf_{i,j} \frac{N}{\mathrm{d}f_i} * \log()$

FastText: Word embeddings are a method for transforming textual data into numerical representations, hence enabling its use as input for machine learning algorithms. A significant limitation of word embedding algorithms, such as word2vec, was their incapacity to effectively handle terms that were not included in the training corpus. These embedding algorithms use words as the smallest unit and aim to get their corresponding embedding vectors. Therefore, if there is a word that is not included in the corpus, word2vec is unable to get its vectorised representation. FastText, similar to word2vec, employs the skip-gram and continuous bag of words (CBOW) models.

FastText is a free software library created by the Facebook AI Research team [15]. The primary emphasis of this approach is in the attainment of scalable solutions for text classification and representation problems, with a particular focus on processing large datasets in a fast and precise manner. FastText is an altered form of word2vec.

GloVe: GloVe is based on fundamental techniques on which GloVe was based. Local context window and global matrix factorisation can help us better grasp how GloVe

Fig. 3 Result after vectorisation

```
Shape of Sparse Matrix:  (20000, 41869)
Amount of Non-Zero occurences:  156460
sparsity: 0
```

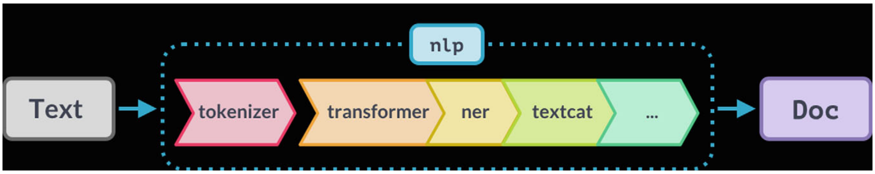


Fig. 4 spaCy pipeline as mentioned in [16]

operates [17]. In the field of Natural Language Processing (NLP), the technique known as global matrix factorisation involves the use of matrix factorisation techniques derived from linear algebra in order to decrease the size of term frequency matrices that are of substantial magnitude. Typically, these matrices are used to depict the presence or lack of words inside a given text. The use of global matrix factorisations to term frequency matrices is often referred to as Latent Semantic Analysis (LSA). The two local context window approaches often used in natural language processing are Continuous Bag-of-Words (CBOW) and Skip-Gram.

spaCy: spaCy is a freely available and open-source software library designed for natural language processing (NLP) tasks [18]. It is implemented in the programming language Python and utilises the Cython programming language for efficient execution. The primary objective of spaCy is to facilitate the development of information extraction systems and general-purpose natural language processing applications. The pipeline is described in Fig. 4.

3.4 Classification

Classification is the process which can work on extracted features and also seven classifiers—Multinomial Naïve Bayes, Gaussian Naïve Bayes, Decision tree, Support Vector Machine (SVM), Logistic Regression, Random Forest and Gradient Boosting Algorithms: XGBoost and LightGBM—are used in this section. Those classifiers are among the most widely used ML methods. The term "classifier" refers to the algorithm that classifies our data into several classifications (Fig. 5).

1. **Logistic regression** is a supervised learning algorithm used to estimate discrete outcomes by using a collection of independent variables. The primary objective of this methodology is to use a logistic function to effectively model and predict the probability of an event taking place. This regression technique is often referred to as logit regression. Since anticipated, the output values of the prediction are within the range of 0 and 1, since they represent the probability.
2. **Decision Tree**: It is a type of supervised learning algorithm which is frequently used to solve classification problems. Surprisingly, it exhibits functionality for both continuous and categorical dependent variables. This method is used to partition the population into two or more homogeneous sets. To maximise the

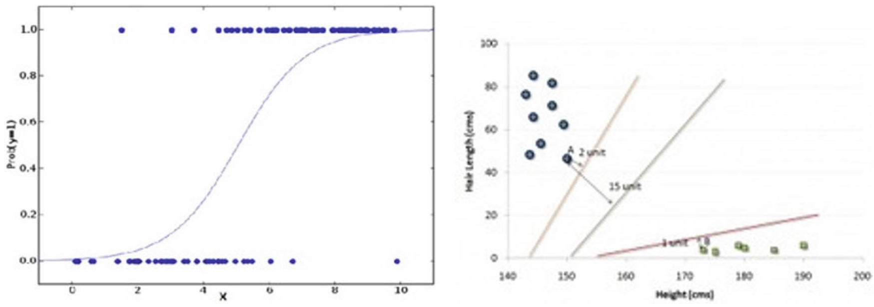


Fig. 5 Logistic Regression and SVM

number of distinct groups, the selection process relies on the key attributes or independent variables.

3. **Support Vector Machine—SVM**: The support vector machine (SVM) approach entails the depiction of every data item as a point inside an n-dimensional space, where n is the total number of features. Each feature is allocated a distinct position inside this spatial framework. As an example, in the case of having just two attributes, namely an individual's Height and Hair length, the first step would include graphically representing these two variables in a two-dimensional coordinate system. Each data point would correspond to a unique combination of the two attributes, sometimes referred to as Support Vectors.

   Within this particular portion, our aim is to discern the specific lines that function as demarcations, effectively separating the data into two unique and categorically different categories. The positioning of this line will be optimised to optimise the distances from the closest point in each of the two groups. When confronted with a higher number of variables, the classes are distinguished by means of a hyperplane.

4. **Naïve Bayes**: The Naïve Bayesian model is characterised by its simplicity in construction and its special use in handling extensive data sets. In addition to its simplicity, Naïve Bayes is recognised for its superior performance compared to even the most advanced classification techniques. Bayes' theorem offers a method for computing the posterior probability $P\ (h\ D)$ based on the prior probabilities $P(h)$, $P(D)$ and the likelihood $P(D\ h)$. The formula is

$$P(h|D) = \frac{\mathrm{P}(D|h)P(h)}{P(D)}$$

5. **Gradient Boosting Algorithms**: XGBoost—With its very high predictive capability, it emerges as the optimal selection for achieving accuracy in various occurrences. The approach incorporates both a linear model and a tree learning algorithm, resulting in a significant improvement in computational efficiency compared to current gradient booster techniques, with a speed increase of about 10 times. One intriguing aspect of XGBoost is its alternative designation as a

regularised boosting technology. This approach aids in mitigating overfitting in modelling and has extensive backing for a variety of programming languages including Scala, Java, R, Python, Julia and C++.

LightGBM is a framework for gradient boosting that employs methods based on tree learning.

4 Results and Discussion

This section present evaluation metrics, results and observations for the proposed model. In this research, we have split 80% of data as training and 20% as testing. After fine-tuning all the algorithms, the following are few hyperparameters initialised with these values: vector dimension—300, learning rate—0.1, maximum depth—5 and estimators—500.

4.1 Evaluation Metrics

This study included conducting experiments to evaluate the metrics of accuracy, precision, recall and F1-score.

- Accuracy: Accuracy is defined as the proportion of correct forecasts to the overall number of predictions made.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- Precision: Accuracy has a drawback, which is addressed by the precision metric. The precision metric is a quantitative measure that assesses the accuracy of positive predictions by calculating the proportion of accurately anticipated positive events. The calculation entails ascertaining the ratio of true positive forecasts to the total number of positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- Recall: It aims to determine what proportion of true positives were incorrectly tagged. The calculation involves the summation of the number of true positives, which includes both successfully predicted positive instances and falsely forecasted negative instances, with the number of positives that were accurately predicted as positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- F1-score: The F1-score may be computed by taking the harmonic mean of accuracy and recall, with equal importance given to both metrics.

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall } 2} = \frac{2 * \text{TP}}{\text{TP} + \text{FP} + \text{FN}}$$

4.2 Result Analysis

This research investigates the performance of several classification algorithms in the timely diagnosis of depression. The experiment was done with several methodologies for feature extraction, employing a range of classification algorithms including multinomial Naïve Bayes, Gaussian Naïve Bayes, decision tree, SVM, logistic regression, random forest and gradient boosting algorithms: XGBoost and LightGBM.

The samples used for training and testing consist of data from individuals who are experiencing depression and those who are not experiencing depression. In order to get optimum classification results, the use of the TF-IDF feature combination is employed. Table 1 presents a summary of the accuracy and F1 measure for all methods. Precision rate, recall rate and F1-score were examined for both depressed and non-depressive data sets. The XGBoost model yielded the greatest accuracy and F1 measure, with values of 0.91 and 0.9, respectively.

In above table, a comparative study using machine learning algorithms XGBoost is a machine learning technique that involves the combination of many base estimators to enhance the overall predictive performance and resilience compared to using a single estimator. The technique involves aggregating a set of weak or moderate predictors in order to construct a robust predictor. Boosting algorithms have consistently shown strong performance in several data science contests, such as Kaggle, AV Hackathon and CrowdANALYTIX. From Fig. 6, it is understood that by applying pretrained encoding technique like FastText we got highest accuracy.

Similarly for TF-IDF encoding, it is observed that XGBoost is giving highest performance as shown in Fig. 7.

5 Conclusion

The focus of our research pertains to many machine learning algorithms used for the purpose of discerning or detecting the emotional disposition of an individual based on their social media activity, such as Twitter tweets, comments or any textual material. In contemporary society, social media has emerged as a prominent medium for individuals to express their emotions and share visual content. Machine learning

Table 1 Analysis of various word embedding models for depression posts in Twitter

Methods	Pre-trained Embedding			TF-IDF		
	Encoding	accuracy	F1 measure	Encoding	Accuracy	F1 measure
Multinomial naïve bayes	spaCy	0.64	0.65	TF-IDF	0.64	0.65
	fastText	0.53	0.45	TF Ngram	0.53	0.43
	GloVe	0.58	0.67	TF-IDF + FastText	0.58	0.66
Bernoulli naïve bayes	spaCy	0.51	0.53	TF-IDF	0.51	0.53
	fastText	0.55	0.52	TF Ngram	0.55	0.52
	GloVe	0.48	0.52	TF-IDF + FastText	0.48	0.52
Logistic regression	spaCy	0.73	0.72	TF-IDF	0.73	0.72
	fastText	0.7	0.72	TF Ngram	0.7	0.72
	GloVe	0.69	0.71	TF-IDF + FastText	0.69	0.71
Decision tree	spaCy	0.62	0.58	TF-IDF	0.62	0.58
	fastText	0.62	0.59	TF Ngram	0.62	0.59
	GloVe	0.59	0.58	TF-IDF + FastText	0.59	0.58
SVM	spaCy	0.72	0.71	TF-IDF	0.72	0.71
	fastText	0.71	0.73	TF Ngram	0.71	0.73
	GloVe	0.69	0.68	TF-IDF + FastText	0.69	0.68
SVM-sigmoid kernel	spaCy	0.71	0.71	TF-IDF	0.71	0.71
	fastText	0.72	0.71	TF Ngram	0.72	0.71
	GloVe	0.57	0.56	TF-IDF + FastText	0.57	0.56
Random forest	spaCy	0.73	0.73	TF-IDF	0.73	0.73
	fastText	0.73	0.73	TF Ngram	0.73	0.73
	GloVe	0.71	0.73	TF-IDF + FastText	0.71	0.73
XGBoost	spaCy	0.89	0.86	TF-IDF	0.89	0.86
	fastText	0.91	0.9	TF Ngram	0.91	0.9
	GloVe	0.87	0.8	TF-IDF + FastText	0.87	0.8
LightGBM	spaCy	0.89	0.86	TF-IDF	0.89	0.86
	fastText	0.9	0.89	TF Ngram	0.9	0.89
	GloVe	0.88	0.88	TF-IDF + FastText	0.88	0.88

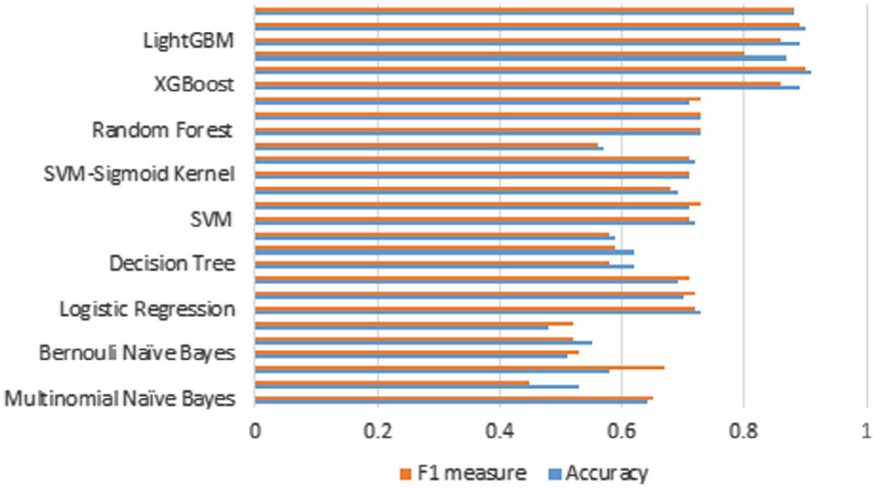


Fig. 6 Report on Pre-trained embeddings

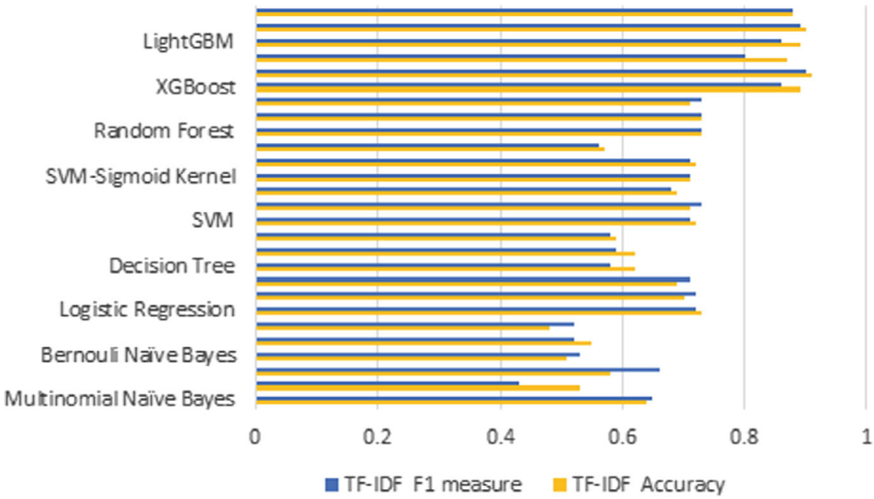


Fig. 7 Report on TF-IDF encoding on various machine algorithms

algorithms possess the ability to analyse extensive datasets, execute computations on this data with relative ease and provide predictions or useful outcomes.

In this work, we used machine learning methods extensively to advance the research and provide a forward-looking perspective. Depression may be effectively detected via the use of classifiers, particularly when a substantial dataset of twenty thousand diverse posts and comments sourced from a range of individuals' profiles is collected. The accessibility of machine learning is increasing for researchers and also pretrained embedding models like spaCy, FastText, etc.

Our study aims to enhance the efficiency and efficacy of our investigation by incorporating additional datasets. There is a need for more research and investigation in the realm of depression analysis, with a specific emphasis on more targeted and concentrated investigations. In future investigations, the use of deep learning techniques might be employed to enhance the precision of depression forecasts inside our study. The measurement of an individual's depression level or severity may be conducted subsequent to the detection of depression. Additionally, recommendations for appropriate measures to alleviate this condition can be provided.

References

1. Chatterjee R, Gupta RK, Gupta B (2021) Depression detection from social media posts using multinomial naive theorem in IOP Conference Series. Mater Sci Eng 1022:012095
2. Smys DS, Raj DJS (2021) Analysis of deep learning techniques for early detection of depression on social media network-a comparative study. J Trends Comput Sci Smart Technol 3:24–39
3. Islam MR et al (2018) Depression detection from social network data using machine learning techniques. Health Inf Sci Syst 6:1–12
4. Safa R, Edalatpanah S, Sorourkhah A (2023) Predicting mental health using social media: a roadmap for future development. *arXiv preprint* arXiv:2301.10453
5. Kumar A, Sharma A, Arora A (2019) Anxious depression prediction in real- time social data. arXiv preprint arXiv:1903.10222
6. Haque UM, Kabir E, Khanam R (2021) Detection of child depression using machine learning methods. PLoS ONE 16:e0261131
7. Guntuku SC, Preotiuc-Pietro D, Eichstaedt JC, Ungar LH (2019) What twitter profile and posted images reveal about depression and anxiety. In: Proceedings of the international AAAI conference on web and social media, vol 13, pp 236–246
8. Garg M et al (2022) CAMS: an annotated corpus for causal analysis of mental health issues in social media posts. arXiv preprint arXiv:2207.04674
9. Wichers M, Riese H, Hodges TM, Snippe E, Bos FM (2021) A narrative review of network studies in depression: what different methodological approaches tell us about depression. Front Psych 12:719490
10. McMullen LM, Stoppard JM (2006) Women and depression: a case study of the influence of feminism in Canadian psychology. Fem Psychol 16:273–288
11. Zhang W, Xie J, Liu X, Zhang Z (2023) Depression detection using digital traces on social media: a knowledge-aware deep learning approach. arXiv e-prints arXiv–2303
12. Kang K, Yoon C, Kim EY (2016) Identifying depressive users in Twitter using multimodal analysis. In: 2016 International conference on big data and smart computing (BigComp), pp 231–238
13. Dey S, Sarkar I, Chakraborty S, Roy S (2020) Depression detection using intelligent algorithms from social media context-state of the art, trends and future roadmap. J Xidian Univ 14:45–62
14. Loper E, Bird S (2002) The natural language toolkit. arXiv preprint cs/0205028
15. Bojanowski P, Grave E, Joulin A, Mikolov T (2017) Enriching word vec- tors with subword information. Trans Assoc Compu-Tational Linguist 5:135–146
16. Vasiliev Y Natural language processing with Python and spaCy: a practical introduction (No Starch Press, 2020)

17. Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation in Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
18. Hu C, Gong H, He Y (2022) Data driven identification of international cutting edge science and technologies using SpaCy. PLoS ONE 17:e0275872

Machine Learning Privacy Preserving in Distributed Systems Using Federated Learning

Pirangi Vijay Kumar, Talari Swapna, Rajendhar Reddy Gaddam, S. Dhanalakshmi, B. Pradeep, and Balusupati Anil Kumar

Abstract Particularly in distributed systems, where data are spread across numerous devices or servers, privacy concerns have grown paramount in the age of big data and machine learning. A potential solution to these issues is federated learning (FL), which allows models to be trained on decentralised data without revealing sensitive information. Using federated learning as an example, this article provides a thorough overview of privacy-preserving machine learning methods for use in distributed systems. Assuring that sensitive data stay locally stored and encrypted during training, we go over the fundamentals of FL, which include federated optimization, model aggregation, and secure aggregation protocols. We also investigate several issues with FL systems; security, privacy, and communication overhead, as well as potential solutions to these problems. Case studies in various domains, including healthcare, finance, and the Internet of Things (IoT), demonstrate how FL effectively preserves privacy. Lastly, we draw attention to possible future paths for privacy-preserving ML research and development, stressing the need for strong cryptographic methods, effective communication protocols, and standardized frameworks to implement FL in distributed systems in the real world.

Keywords Federated learning · Privacy preservation · Machine learning · Distributed systems

---

P. V. Kumar (✉) · T. Swapna · R. R. Gaddam · S. Dhanalakshmi · B. Pradeep · B. A. Kumar
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: vijaykumar@cmritonline.ac.in

T. Swapna
e-mail: swapnat@cmritonline.ac.in

R. R. Gaddam
e-mail: rajendarreddy.gaddam@cmritonline.ac.in

B. Pradeep
e-mail: pradeepb26@cmritonline.ac.in

B. A. Kumar
e-mail: anilkumarb@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_48

1 Introduction

Concerns about data privacy and security have grown in recent years, dramatically increasing the need for privacy-preserving machine learning (ML) in distributed systems [1]. Since federated learning allows for the training of machine learning models across decentralized devices while preserving the secrecy and localization of data acquired, it seems to provide a viable solution to these difficulties. Federated learning eliminates the requirement for data to be pooled in a single place by allowing models to be trained collaboratively across several devices or nodes. Instead of sharing raw data, each node keeps track of its own dataset and changes the global model repeatedly. To do this, we need to simply update the model parameters. This decentralized approach safeguards users' privacy by reducing the possibility of data breaches and illegal access to data kept on users' devices. By incorporating FL into distributed systems, companies may safeguard individual privacy while using the collective intelligence of several datasets. By distributing storage and computation across participating nodes, FL enhances scalability. Smartphones and other low-powered Internet of Things (IoT) devices may now be used to train machine learning models.

For scenarios where data distribution and attributes may change continually over time, FL allows for continuous learning without compromising data secrecy, making it an ideal choice. And by keeping data exposure to a minimum and following data privacy laws, Florida can meet regulatory requirements like GDPR, HIPAA, or the CCPA [2]. Organizations may use decentralized data's promise while also safeguarding consumers' privacy by integrating FL into distributed platforms. This approach not only makes the model faster and more scalable, but it also follows the ever-evolving data privacy laws, which is great for building trust and openness in ML applications.

1.1 A. Importance of Privacy-Preserving Techniques in Distributed Systems

- Data Security: Distributed systems often involve multiple nodes or devices communicating with each other over networks. These nodes may hold sensitive or personal data. Privacy-preserving techniques help protect these data from unauthorized access, ensuring its security even when transmitted or stored across various locations.
- Regulatory Compliance: Many jurisdictions have stringent regulations regarding data privacy and protection, such as GDPR in Europe, HIPAA in the USA, or the Personal Data Protection Act in Singapore. Distributed systems must comply with these regulations to avoid legal consequences.
- User Trust: Users are increasingly concerned about the privacy of their data, especially in the era of large-scale data breaches and surveillance. Privacy-preserving

techniques demonstrate a commitment to protecting user privacy, fostering trust between users and the organizations or platforms handling their data. This trust is essential for user adoption and engagement in distributed systems.
- Data Sharing and Collaboration: Data sharing and collaboration are often essential in distributed systems with multiple stakeholders or organisations, for the benefit of all parties involved. This is especially true in collaborative research and industry partnerships. With privacy-preserving techniques, sensitive information can be securely shared and collaborated on without sacrificing its privacy. This allows for cooperation while safeguarding the data assets of each party involved.
- Ethical Considerations: Beyond legal compliance, there are ethical considerations surrounding data privacy. Respect for individual privacy rights is a fundamental aspect of ethical data handling. Privacy-preserving techniques uphold these principles by minimizing the exposure of sensitive data, promoting responsible data practices, and mitigating the risk of privacy violations.

Privacy-preserving techniques play a vital role in ensuring the security, compliance, trustworthiness, and ethical handling of data in distributed systems [3], thereby supporting the effective operation and adoption of such systems in various domains.

2 Literature Review

2.1 Privacy-Preserving Techniques in Machine Learning

According to [4], deep learning, among the most cutting-edge techniques in the field, was increasingly popular. Modern-day applications and areas utilizing it include pattern recognition, medical prediction, and speech recognition, among others. It was noted that deep learning was capable of surpassing the reliance on hand-designed features, unlike ordinary learning algorithms. Collaborative learning was used to train models and strong infrastructures like clouds were utilized to enhance the deep learning experience. The findings of the solutions' evaluations with regard to the specified performance indicators were further summarized in that survey. Furthermore, it was noted that it took a collection of lessons acquired from each privacy-preserving job and applied them. In the conclusion, attention was drawn to unanswered questions in the field and some avenues for further study were suggested.

Boulemtafes et al. [5] stated that deep neural network (NN)-based machine learning techniques were seeing widespread application across several fields due to their impressive performance. In contrast, MLaaS was becoming more popular as cloud services expanded, allowing users to train and deploy ML models on top of the infrastructure already in place at cloud providers. Unfortunately, there were security and privacy concerns associated with machine learning algorithms requiring access to raw data, which was frequently sensitive and posed dangers to individual privacy.

They provided CryptoDL, a framework that aimed to solve this problem by creating new ways to apply deep neural network algorithms to encrypted data.

According to [6], in this study, they traced the development of privacy-preserving computing techniques' incorporation into DL to bridge the gap between theoretical frameworks and their actual implementation. They emphasized the relative benefits and drawbacks. Chai et al. [7] highlighted the dominance of deep learning, powered by artificial neural networks, in handling complex data like photos, audio, and text. This approach became integral to new Internet-based AI applications due to its high accuracy, fuelled by the abundance of training data. However, this reliance on massive data collection raised privacy concerns, especially for commercial entities gathering user data extensively. Concerns included indefinite storage of personal data without user control over deletion or usage restrictions, along with potential legal and surveillance access. Institutions, such as medical facilities, were wary of sharing data for deep learning due to confidentiality worries, leading to missed opportunities for advancements in this field.

2.2 Distributed Systems Using Federated Learning

Chamikara et al. [8] stated that end users, different areas, or even other organizations' devices often housed data and computing resources those days. Aggregating or directly sharing scattered data and computing resources across regions or organizations for machine learning tasks was not possible due to rules or restrictions. They mentioned that to train machine learning models cooperatively using dispersed data and computer resources, federated learning was an efficient way. They concurrently mentioned that federated learning guaranteed data confidentiality and privacy while also complying with rules and regulations. They presented a thorough literature review on federated learning in their study. They began by outlining a taxonomy of associated methods and a practical design for federated learning systems. Second, they detailed four sides of federated learning systems: data connectivity, various forms of parallelism, aggregation techniques, and federated learning security. Third, they showcased four functional architecture-based federated systems that were often utilized. They concluded by outlining the constraints and suggesting avenues for further study.

They provided TiFL, a tier-based federated learning system, to address the issue of stragglers caused by differences in data amount and resources. They implemented a TiFL prototype according to Google's FL design and assessed it using the most recent FL benchmarks in a FL testbed. In a number of different diverse situations, experimental assessment revealed that TiFL worked better than traditional FL. They showed that TiFL significantly outperformed state-of-the-art methods in training performance while maintaining or improving test accuracy using the suggested adaptive tier selection methodology [9].

They explained that the requirement to transport data, store data locally, and direct network calculations towards edge devices had been driven by the continual improvement in the processing capability of these devices. They emphasized that intrusion detection systems (IDS) were crucial when it came to protecting the confidentiality and safety of these devices. They mentioned that machine learning (ML) and deep learning (DL) in conjunction with intrusion detection systems had become increasingly popular due to their success in producing accurate classifications. They stated that when it came to privacy-preserving decentralized learning techniques, federated learning (FL) was the way to go. They aimed to provide an in-depth analysis of FL's role in intrusion detection systems in their research. They concluded by outlining potential solutions to the problems encountered during the installation of FL-based intrusion detection systems [10].

In [11], mentioned that both academia and business had begun to pay more and more attention to AI, machine learning (ML), and federated learning. They stated that with the proliferation of edge devices and advancements in computing power, federated learning began to show its promise as a distributed learning technique. They presented and evaluated centralized, hierarchical, regional, and decentralised architectures for a Federated Learning system in their study. They measured several system performance indicators for each of the four options using two popular datasets. They identified many use cases and situations that were well-suited to each option based on their findings. They delved into the trade-off among communication latency, model evolution time, and model classification performance, which was an essential aspect for implementing the findings into practical industrial systems.

2.3 *Privacy-Preserving Machine Learning in Distributed Systems*

Mohassel and Zhang [12] conducted research which indicated that the importance of distributed machine learning in processing enormous data sets. They proposed a privacy-preserving machine learning approach for hierarchical distributed networks in their study. They updated and enhanced the algorithm for collaborative learning. The suggested method offered thorough security for every level of a hierarchical distributed system while also decreasing learning process overhead. To further enhance the learning efficiency of hierarchical distributed systems, they also suggested an asynchronous approach based on the study of collaborative convergence in various learning groups. Finally, they conducted comprehensive tests using real-world data to assess the privacy, effectiveness, and efficiency of their suggested solutions.

Shokri [13] mentioned that picture processing, voice, and text recognition were just a few examples of the many practical uses of machine learning's prediction models. They stated that training these models on massive amounts of data from many sources improved their accuracy. They provided MPC-friendly alternatives to non-linear functions like sigmoid and softmax and provided novel ways to facilitate safe arithmetic operations on shared decimal values, better than previous work. They mentioned that C++ was where they put their system into action. At scale to millions of data samples with hundreds of characteristics, their methods outperformed state-of-the-art implementations for privacy-preserving logistic and linear regressions by many orders of magnitude, according to their testing. In addition, they developed and deployed the first privacy-preserving method for neural network training [14] highlighted the transformative potential of distributed machine learning and edge computing for companies. They emphasized the significance of safeguarding sensitive data from IoT devices and industries like healthcare and finance. Their study introduced federated learning (FedML) and proposed DISTPAB, a distributed perturbation technique, to enhance privacy preservation and computational efficiency. DISTPAB showed promising results in preventing privacy breaches without compromising data value[15]) highlighted challenges in machine learning model development due to data scientists' access to complete data, causing distrust between data scientists and owners. They proposed privacy-preserving approaches for data control during learning but warned of new security issues. Their solution involved a distributed infrastructure ensuring peer-to-peer trust using hyperledger aries, decentralized identifiers (DIDs), and verifiable credentials (VCs) for federated learning workflows, exemplified in mental healthcare data applications.

3 Methodology

Privacy-preserving of machine learning (ML) in distributed systems using federated learning involves a mathematical model that aims to maintain the privacy of individual data while collectively improving a global model. In this model, several variables are considered, including the total number of clients (N) the number of training rounds (K), and the number of local iterations per round (Tk). Additionally, model parameters (w_k^i) and local gradients (g_k^i) are tracked for each client i at each round k, Learning rate η and Privacy parameter $\in$ for differential privacy.

The federated learning algorithm operates in rounds, where clients perform local computations on their data and send model updates to a central server. Differential privacy, a common technique for privacy preservation, is integrated into the model by adding noise to the gradients before transmission. This noise, typically Gaussian with a certain scale (σ), is essential for ensuring privacy guarantees while updating the global model.

The update equation for federated learning without privacy considerations is $w_{k+1}^i = w_k^i - \eta \cdot g_k^i$

However, to incorporate privacy through differential privacy, Gaussian noise is added to the gradients, modifying the update equation to $w_{k+1}^i = w_k^i - \eta \cdot (g_k^i + N\left(0, \sigma^2\right))$.

where $N(0, \sigma^2)$represents Gaussian noise with mean 0 and standard deviation σ The value of σ is chosen based on the desired privacy level, which is often determined by the privacy parameter $\in$ in differential privacy.

The privacy cost over multiple rounds of training can be quantified using techniques like the moments accountant, which tracks the cumulative privacy loss. It is important to note that the privacy cost grows logarithmically with the number of rounds and linearly with the privacy parameter $\in$. This mathematical model provides a foundational framework for understanding and analysing privacy-preserving federated learning in distributed systems, with further optimizations and considerations tailored to specific use cases and requirements.

3.1 Federated Learning Algorithms

Federated learning (FL) is a distributed machine learning approach [16] that allows multiple devices or nodes to collaboratively train a model while keeping their data decentralized and private. Several algorithms are used in FL to enable efficient model training across these distributed nodes while preserving privacy. Here are explanations of some commonly used federated learning algorithms:

Federated Averaging (FedAvg): FedAvg is one of the fundamental algorithms of federated learning [17]. It entails combining model updates from many client devices to provide a global model update. Each client trains the model with local data and transmits the updated model parameters to the central server. The central server then averages these values to update the global model [18]. FedAvg is simple and extensively used because of its ability to aggregate model changes while maintaining security.

Federated stochastic gradient descent (FedSGD): FedSGD [19] is a variation of the traditional stochastic gradient descent (SGD) algorithm tailored for federated learning [20]. In FedSGD, each client device computes a stochastic gradient using its local data and updates the model parameters independently [21]. The central server aggregates these updates using techniques like averaging or weighted averaging to update the global model. FedSGD is efficient in scenarios where clients have heterogeneous data distributions or computational resources.

Federated Learning with Secure Aggregation (FedSecAgg): FedSecAgg [22] is an extension of federated learning that incorporates cryptographic techniques to ensure secure aggregation of model updates. It utilizes SMPC (secure multi-party computation) protocols [23] to aggregate model updates in a privacy-preserving manner, protecting the confidentiality of individual updates during the aggregation process. FedSecAgg enhances the privacy guarantees of federated learning, particularly in scenarios where the central server is semi-trusted or prone to attacks.

4 Conclusion and Future Work

With an emphasis on the novel method of federated learning (FL), this article explores the crucial area of privacy-preserving machine learning in distributed systems. The analysis underscores the growing importance of safeguarding sensitive data in an era marked by the proliferation of interconnected devices and distributed computing systems. Federated learning emerges as a promising solution, allowing ML models to be trained collaboratively across decentralized nodes without compromising individual data privacy. This paper provides overview of the current state of privacy-preserving machine learning techniques, emphasizing the advantages and challenges associated with federated learning. It highlights the effectiveness of FL in preserving privacy by keeping data localized, mitigating the need for centralized data repositories. The paper also discusses various FL architectures, algorithms, and security mechanisms, shedding light on the nuances that practitioners need to consider in real-world implementations. This paper identifies several avenues for further research and development. First, there is a need for standardized protocols and benchmarks to facilitate the comparison of different FL implementations. Additionally, enhancing the robustness and security of FL systems against adversarial attacks remains a critical area for exploration. Further investigations into the optimization of communication efficiency, model aggregation techniques, and scalability in large-scale distributed systems are also recommended. Moreover, interdisciplinary collaboration between machine learning researchers, privacy experts, and policymakers is essential to address ethical and legal considerations in the evolving landscape of privacy-preserving machine learning. This paper serves as a valuable resource for researchers, guiding them towards addressing the challenges and shaping the future of privacy-preserving machine learning in distributed systems.

References

1. Abramson W, Hall AJ, Papadopoulos P, Pitropakis N, Buchanan WJ (2020) A distributed trust framework for privacy-preserving machine learning. In: Trust, privacy and security in digital business: 17th international conference, TrustBus 2020, Bratislava, Slovakia, 14–17Sep 2020, Proceedings 17. Springer International Publishing, pp 205–220
2. Agrawal S, Sarkar S, Aouedi O, Yenduri G, Piamrat K, Alazab M, Agrawal S, Sarkar S, Aouedi O, Yenduri G., Piamrat K, Alazab M Gadekallu TR (2022). Federated learning for intrusion detection system: concepts, challenges and future directions. Comput Commun
3. Awan KA, Din IU, Almogren A, Guizani M, Altameem A, Jadoon SU (2019) Robusttrust–a pro-privacy robust distributed trust management mechanism for internet of things. IEEE Access 7:62095–62106
4. Bartnick WJ, Bradshaw S, Chughtai-Harvey A, Isensee M (2020) Present and future data privacy outlook. Curr J Int'l Econ L 24:70
5. Boulemtafes A, Derhab A, Challal Y (2020) A review of privacy-preserving techniques for deep learning. Neurocomputing 384:21–45
6. Cabrero-Holgueras J, Pastrana S (2021) SoK: privacy-preserving computation techniques for deep learning. Proc Priv Enhancing Technol

7. Chai Z, Ali A, Zawad S, Truex S, Anwar A, Baracaldo N, Zhou Y, Ludwig H, Yan F, Cheng Y, Cheng Y (2020) Tifl: a tier-based federated learning system. In: Proceedings of the 29th international symposium on high-performance parallel and distributed computing, pp 125–136
8. Chamikara MAP, Bertok P, Khalil I, Liu D, Camtepe S (2021) Privacy preserving distributed machine learning with federated learning. Comput Commun 171:112–125
9. Hesamifard E, Takabi H, Ghasemi M, Wright RN (2018) Privacy-preserving machine learning as a service. Proc Priv Enhancing Technol 2018(3):123–142
10. Jia Q, Guo L, Fang Y, Wang G (2018) Efficient privacy-preserving machine learning in hierarchical distributed system. IEEE Trans Netw Sci Eng 6(4):599–612
11. Liu J, Huang J, Zhou Y, Li X, Ji S, Xiong H, Dou D (2022) From distributed machine learning to federated learning: a survey. Knowl Inf Syst 64(4):885–917
12. Mohassel P, Zhang Y (2017) Secureml: a system for scalable privacy-preserving machine learning. In: 2017 IEEE symposium on security and privacy (SP). IEEE, pp 19–38
13. Shokri R, Shmatikov V (2015) Privacy-preserving deep learning. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security, pp 1310–1321
14. So J, Güler B, Avestimehr AS (2021) CodedPrivateML: a fast and privacy-preserving framework for distributed machine learning. IEEE J Sel Areas Inf Theory 2(1):441–451
15. Zhang H, Bosch J, Olsson HH (2020) Federated learning systems: architecture alternatives. In: 2020 27th Asia-Pacific software engineering conference (APSEC). IEEE, pp 385–394
16. Hu K, Li Y, Xia M, Wu J, Lu M, Zhang S, Weng L (2021) Federated learning: a distributed shared machine learning method. Complexity 2021:1–20
17. Nilsson A, Smith S, Ulm G, Gustavsson E, Jirstrand M (2018) A performance evaluation of federated learning algorithms. In: Proceedings of the second workshop on distributed infrastructures for deep learning, pp 1–8
18. Sun T, Li D, Wang B (2022) Decentralized federated averaging. IEEE Trans Pattern Anal Mach Intell 45(4):4289–4301
19. Kassab R, Simeone O (2022) Federated generalized Bayesian learning via distributed stein variational gradient descent. IEEE Trans Signal Process 70:2180–2192
20. Wainakh A, Ventola F, Müßig T, Keim J, Cordero CG, Zimmer E, Grube T, Kersting K, Mühlhäuser M (2021) User-level label leakage from gradients in federated learning. arXiv preprint arXiv:2105.09369
21. Wang J, Charles Z, Xu Z, Joshi G, McMahan HB, Al-Shedivat M et al (2021) A field guide to federated optimization. arXiv preprint arXiv:2107.06917
22. Saylam B, İncel ÖD (2023) Federated learning on edge sensing devices: a review. arXiv preprint arXiv:2311.01201
23. Yu S, Cui L (2022) Secure Multi-party computation in federated learning. In security and privacy in federated learning. Springer, Singapore, pp 89–98

Fog Computing in Health Monitoring Using IOT Wearables

Pundru Harshitha, Varshika Reddy, P. Rishikesh, S. Jyothi, and Katta Subba Rao

Abstract Today, the increasing need for medical assistance has positioned smart health care as a critical component of modern living. Health care is a fundamental human necessity and forms the backbone of smart healthcare systems, which are expected to generate billions of dollars in sales in the near future. Smart health care has several components including Internet of Things (IoT), fog computing, next-generation wireless communication technologies, and artificial intelligence (AI). Due to the ability of IoT devices to include sensors, IoT will play a significant role in the healthcare industry. It is used to remotely monitor human health issues in real time. This information flow with rich data cannot be handled by cloud infrastructure alone. Hence, we came up with an effective solution. When utilizing FOG as a median, this may provide excellent insight into precise vital values. Machine learning techniques are utilized in notification systems to address the issues. These algorithms are designed to enhance how a prediction process using diverse inputs is executed.

Keywords Smart health care · Internet of things · Fog computing · Cloud infrastructure · Machine learning algorithms

1 Introduction

1.1 Conventional Method

Consequently, the volume of data produced by IOT devices is rapidly growing. The conventional method is to move data or the information to the cloud, which has huge processing and storage capacity. Due to its physical distance from end-user devices

Supported by organization B V Raju Institute of Technology.

P. Harshitha · V. Reddy · P. Rishikesh · S. Jyothi · K. S. Rao (✉)
CSE Department, B V Raju Institute of Technology, Narsapur, Telangana, India
e-mail: subbarao.k@bvrit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_49

and associated privacy concerns, the cloud struggles to meet the low-latency and quality-of-service requirements of delay-sensitive IoT applications.

1.2 Decentralized Approach

Fog computing is a decentralized computing approach where resources are distributed between the cloud and data sources. It brings cloud capabilities closer to users, like fog closer to the earth than clouds. This model involves edge nodes linked to devices, unlike the centralized cloud infrastructure. Fog computing offers benefits like low latency, strong security, power efficiency, and enhanced user experience.

1.3 Machine Learning

Machine learning, a facet of AI, enables IT systems to autonomously discern optimal solutions by analyzing database patterns. It involves deriving insights from current datasets and algorithms to generate the best possible outcomes. This evolution in AI empowers computers to learn from past data and execute preprogrammed actions. In health care, there's a significant shift toward monitoring vital signs and storing pertinent information. Health care leverages new technologies for more effective care. Fog computing stands as a crucial infrastructure for transforming theoretical healthcare IoT into practical applications.

2 Literature Survey

Convolutional Neural Network (CNN) technology is integrated into a low-power wearable smart device. It can detect gait abnormalities with a sensitivity of 94.6% and a specificity of 95.3%, enabling early diagnosis and improved healthcare outcomes, potentially leading to better healthcare outcomes and early diagnosis [1]. A machine learning-powered Internet of Things system can identify variations in activity in nursing homes with 94.3% sensitivity and 97.6% specificity, thereby improving patient care and reducing the workload of medical staff [2].

The revolutionary influence of fog computing on health care includes effective administration of patient data, insightful monitoring, prompt notifications to clinicians, and successful machine learning-driven healthcare monitoring [3]. The study contrasts fog computing, which extends cloud services to the network edge for lower latency and better performance, with cloud computing, which provides scalable Internet-based resources [4].

A distributed framework for ubiquitous computing based on wearable IoT is proposed by research. It integrates edge and fog computing and uses real-time machine learning for scalability, efficiency, and enhanced quality of life [5].

A review paper on advanced strategies in orchestration of distributed systems for IoT smart services in fog computing emphasizes the need of effective orchestration and calls for more study in this field [6].

With 95.3% sensitivity and 98.7% specificity, a wearable gadget combined with a cloud-based deep learning system predicts diabetes, offering better patient care standards, diabetic management, and healthcare overall [7].

A novel method that makes use of machine learning to identify anomalies and attacks in IoT sensors at IoT locations highlights the higher accuracy and speed of SVM compared to other techniques [8].

A comprehensive evaluation suggests an architectural framework for health care based on fog computing, which makes use of wearables to monitor vital signs and machine learning to analyze data in real time. It also describes future research areas [9].

The paper suggests a technique for detecting Sybil assaults in wireless networks that makes use of accurate physical channel data. This approach has a low false-positive rate and high accuracy, which strengthens network security against Sybil attacks [10].

Cloud computing relies on central servers for data processing, while fog computing extends processing capabilities to the edge, reducing latency and enhancing efficiency, particularly in real-time applications [11].

A Raspberry Pi-based IoT-cloud architecture is presented by research for safe patient data collecting, cloud server transfer, and scalable applications in health care and other fields [12].

3 Objectives

1. Real-time data collection and transmission regarding a person's activities and health are possible.
2. Monitoring and tracking health metrics: IoT devices can accurately measure and continually monitor a person's vital signs.
3. IoT wearables encourage users to be more involved in managing their health, which boosts user engagement and empowerment.
4. It is extremely useful in hospital management in faster access of data, managing patient data and reviewing the data accessed.

4 Existing Work

Health monitoring in conjunction with cloud computing and the Internet of Things has raised concerns about the ongoing advancement of the country's healthcare system. The conventional medical equipment is frequently made up of small, low-power sensors that are fastened to the user's body and surroundings. An IoT unit is a gadget with a sensor that can communicate with the real world and upload data to the Internet. The sensor attached to the device detects the vitals and other data related to health. The data are then processed to generate valuable insights. These results are displayed on the device's screen or passed to a cloud server for further analysis and storage. Cloud-based systems can offer several advantages for device-based health monitoring, including the ability to store large amounts of data securely, access to advanced analytics tools, and the ability to share data with healthcare professionals or other third-party applications.

It is not practical to use a straightforward sensor-to-cloud architecture for many healthcare applications. Cloud infrastructure needs the internet and is slow to access data (Fig. 1).

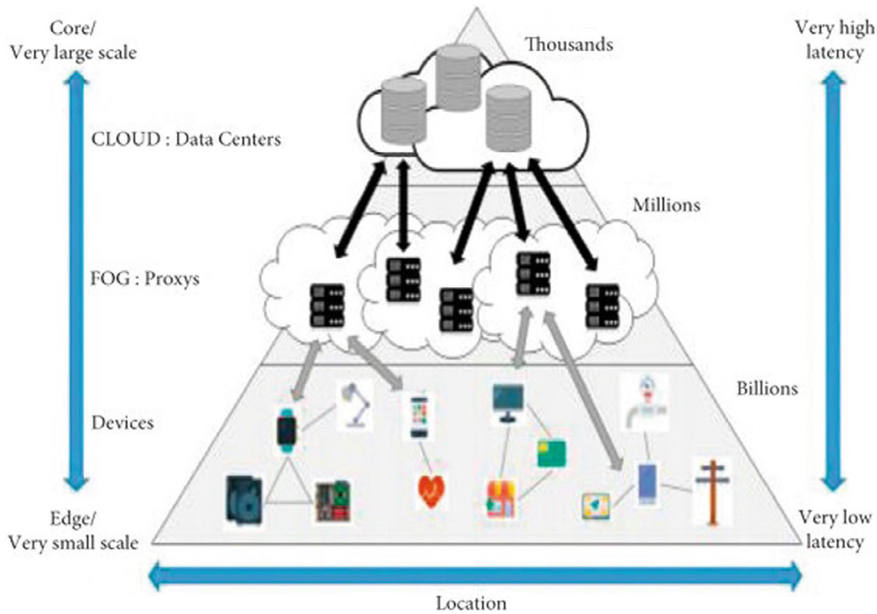


Fig. 1 Proposed architecture diagram

5 Proposed Work

Fog computing, developed within the healthcare sector using IoT, aims to expedite public services and improve patient outcomes. This new platform, built on fog computing principles, aims to reduce delays compared to cloud computing and complements it. Utilizing machine learning, the system monitors body temperature, heart rate, and blood pressure via wearable sensors, promptly alerting any deviations. Utilizing machine learning, the system monitors body temperature, heart rate, and blood pressure via wearable sensors, promptly alerting any deviations. Today's dense network traffic results in a significant bandwidth strain when sending a lot of data to the cloud, intolerable latency, and diminished services for the end users. If a person wants to live in a healthy environment, this is crucial. Body temperature, pulse rate, and oxygen saturation (SpO2) of the individual will be measured continuously or at predetermined intervals using sensors in the hardware section of the wearable system. This system also tracks sleep and stress levels, which are currently a top priority. The program will include a mobile/web application that will save the fog data in the user's profile.

As opposed to fog, clouds have higher latency and higher energy consumption. In order to increase efficiency, these two factors can be reduced utilizing algorithms like FCFS and GKS, which further improves the system.

6 Result

Using IOT wearables to integrate fog computing in health care can have a number of positive effects. In the implementation, a link between the fog node and the sensor node (edge) is formed, allowing for data to be sent from the sensor to the fog node and for the fog to respond to the sensor node by providing insights. The fog node receives the sensor node's data values, processes them, and then gives the user insightful information. When contrasted to clouds, this outcomes are anticipated significantly more quickly.

The major goals of fog computing are to decrease latency and energy consumption, which may be further decreased by applying the algorithms FCFS and GKS. These algorithms can be implemented furtherly.

Table 1 depicts the variation in the energy consumption by cloud and fog which includes the algorithms FCFS and GKS. This shows decrease of energy consumption when Fog architecture is used. Low value can really improve the quality of the programming of this parameter. This is one of the proposed method to give accurate results.

Table 2 represents the latency which refers to the time delay or lag that occurs for transmission of data or the execution of a process. From Table 2, we observe that the cloud takes four times the time or latency taken by the FOG which includes algorithms FCFS and GKS. These fog-based algorithms enable faster data transmission and

Table 1 Energy consumption (mJ)

Cloud	FCFS	GKS
6.52	1.63	1.13
6.52	1.63	1.19

Table 2 Latency (MS)

Cloud	FCFS	GKS
100	1.3	1.23

access compared to the cloud. This is another proposed method which can be included to improve efficiency.

From these observations, we can say that fog architecture is a very efficient and useful method to incorporate in between the IOT sensor layer and cloud layer to filter data, store data locally and faster access of data when there is huge amount of data.

7 Conclusion

In conclusion, cloud computing involves delivering computing services such as storage, latency management, and access to software applications. It provides huge scalability, dependability, and simple access to resources. In order to process and analyze data near to the point of generation, it makes use of localized routers, gateways, and IoT devices. Fog computing promises improved privacy, less network congestion, lower latency, and the ability to make decisions instantly. Fog computing reduces the need for central data transmission and processing by offloading computational tasks from the cloud. It can improve the speed of cloud-based applications, especially those that include mission-critical processes and real-time analytics. Organizations may develop and implement computing solutions that best meet their particular requirements by understanding the distinctive features and advantages of each method, resulting in the optimum performance and user experience.

References

1. Shakya S, Taparugssanagorn A, Suwan-Chai CS (2023) Convolutional neural network-based low-powered wearable smart device for gait abnormality detection. IoT4 2:57–77
2. Lau BPL, Koh Z, Zhou Y, Ng BKK, Yuen C, Low ML (2023) Location-based activity behavior deviation detection for nursing home using IoT devices. Internet of Things 22:100702. ISSN 2542-6605. https://doi.org/10.1016/j.iot.2023.100702
3. Kumar SPP, Niharika MS, Niharika MP, Harshitha B, Raj TGN (2021) Survey on fog computing in healthcare monitoring. Int J Eng Res Tech (IJERT) 10(5)

4. Elmoghrapi AN, Bleblo A, Younis YA (2022) Fog computing or cloud computing: a study. In: 2022 International Conference on Engineering MIS (ICEMIS), Istanbul, Turkey, pp 1–6, https://doi.org/10.1109/ICEMIS56295.2022.9914131
5. Suma V (2021) Wearable IoT based distributed framework for ubiquitous computing. J Ubiquitous Comp Comm Tech (UCCT) 3(1):23–32
6. Syazwani N, Pakhrudin M, Kassim M, Idris A (2021) A review on orchestration distributed systems for IoT smart services in fog computing 11(2):1812–1822. ISSN: 2088-8708. https://doi.org/10.11591/ijece.v11i2.pp1812-182
7. Nasser AR et al (2021) IoT and cloud computing in health-care: A new wearable device and cloud-based deep learning algorithm for monitoring of diabetes. Elect 10(21):2719
8. Hasan M, Islam MM, Zarif MII, Hashem MMA (2019) Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things 7:100059. ISSN 2542-6605. https://doi.org/10.1016/j.iot.2019.100059
9. Yousefpour A et al (2019) All one needs to know about fog computing and related edge computing paradigms: a complete survey. J Syst Architect 98:289–330
10. Wang CZ, Gong L, Zhao L, Yang Z, Liu L, Xiaochun ZC (2018) Accurate Sybil attack detection based on fine-grained physical channel information. Sensors 18:878. https://doi.org/10.3390/s18030878
11. Prakash P, Darshaun KG, Yaazhlene P, Ganesh MV, Vasudha B (2017) Fog computing: issues, challenges and future directions. Int J Elect Comp Eng 7(6):3669
12. Jaiswal K, Sobhanayak S, Mohanta BK, Jena D (2017) IoT-cloud based framework for patient's data collection in smart healthcare system using raspberry-pi. In: Proceedings of the 2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, 21–23 November, pp 1–4 [Google Scholar]

Heart Disease Prediction Using Majority Voting Ensembling Framework (MVEF)

Ramatenki Sateesh Kumar, M. Sunitha Reddy, and D. Baswaraj

Abstract The human heart is the most vital organ in our body that keeps us alive. Heart disease is one of the deadliest diseases which causes many deaths around the globe. In heart disease, the heart fails to supply enough blood to other parts of the body to accomplish its expected functionalities because of the accumulation of fatty plaques in arteries (atherosclerosis). Age, unhealthy lifestyle, family history of heart problems, stress, etc., are a few risk factors for heart disease. The World Health Organization estimates that around 1.79 crore deaths happen worldwide every year. If the heart disease is diagnosed early, and with proper treatment, death rate can be reduced. Machine learning (ML) plays a significant role in detecting heart diseases at an early stage. In this research, experimentations were conducted using various ML methods with the help of a heart disease dataset (HDD) to select a better heart disease prediction system (HDPS). HDD is created by combining publicly available five UCI repository datasets known as Cleveland, Hungary, Switzerland, VA Long Beach, and Statlog heart disease datasets. To improve the efficiency of the heart disease prediction system, experimentation was conducted with ensemble framework referred to as majority voting ensemble framework (MVEF). This ensemble framework was built using EKNN, K-nearest neighbors algorithm (KNN), support vector machine (SVM), a decision tree algorithm, namely, classification and regression tree (CART), and logistic regression (LR) classifiers. MVEF predicted heart disease based on the majority classifiers' decision. MVEF achieved an accuracy of 94.62%.

Keywords Heart disease · Machine learning · Ensemble framework · Accuracy · KNN · SVM · Logistic regression

R. S. Kumar (✉) · M. Sunitha Reddy · D. Baswaraj
Vasavi College of Engineering, Hyderabad, India
e-mail: sateeshramatenki@staff.vce.ac.in

M. Sunitha Reddy
e-mail: m.sunithareddy@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_50

1 Introduction

1.1 Introduction

Cardiovascular diseases (CVD) [1], popularly known as heart diseases, are the most common causes of death worldwide. Heart diseases occur because of the narrowing of heart blood vessels due to the accumulation of fatty plaques. When the heart's blood vessels are narrowed, the heart fails to supply enough blood to other body parts to accomplish their expected functionalities, it is called heart disease. Unsafe lifestyle habits, such as a bad diet, lack of exercise, being overweight, and smoking can lead to heart diseases. According to the World Health Organization (WHO), more than 17.9 million deaths [1] occur each year due to heart diseases across the globe. Out of these deaths, four million were due to heart diseases, and six to seven million were due to heart stroke. It is also noticed that more than 24% of the deaths in India are due to heart diseases. More than 90% of casualties due to heart diseases can be reduced with proper nutrition, exercise, avoiding tobacco smoke, and limiting or avoiding alcohol. Death rates can be further reduced if the heart disease is diagnosed at early stage, and proper medication has been taken. So, there is an efficient system needed that detects heart diseases at an early stage and prevents deaths due to heart diseases. In the early diagnosis of heart disease, machine learning (ML), artificial intelligence (AI) plays a significant role.

In this research, experimentations were conducted using various ML methods with the help of a Heart Disease Dataset (HDD) [2] to select a better Heart Disease Prediction System (HDPS). HDD is created by combining publicly available five UCI repository datasets known as Cleveland [2], Hungary, Switzerland, VA Long Beach, and Statlog [2] heart disease datasets. The resultant dataset consists of 1190 records and 13 attributes such as age, gender, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting ECG, maximum heart rate, angina induced due to exercise, ST depression in patients' ECG, peak slope in ST segment of ECG, number of vessels colored by fluoroscopy, thalassemia to characterize heart diseases. In contrast, the 14th attribute known as 'Disease' is a Boolean variable indicating whether the patient is suffering from heart disease or not, and this attribute is selected as a target attribute.

1.2 Motivation

The motivation behind heart disease prediction is rooted in the significant impact of cardiovascular diseases on public health. Heart disease, including heart attacks, is one of the leading causes of mortality worldwide. The World Health Organization (WHO) estimates a staggering number of deaths worldwide due to heart diseases each year. Early diagnosis and proper treatment are vital in reducing the death rate associated with heart diseases, and machine learning (ML) techniques have shown great promise in detecting these diseases at an early stage. Here are some ways in which ML plays a significant role in heart disease detection:

Risk Stratification: ML algorithms can analyze large datasets, including medical records, clinical measurements, lifestyle factors, and genetic information, to identify patterns and risk factors associated with heart diseases. By leveraging ML, healthcare professionals can stratify individuals into different risk categories, enabling targeted interventions and preventive measures for those at higher risk.
Predictive Modeling: ML algorithms can build predictive models based on historical patient data, incorporating various features and risk factors to estimate the likelihood of developing heart diseases. These models can provide individualized risk scores, allowing healthcare providers to identify individuals who may benefit from early interventions or closer monitoring.
Image Analysis: ML techniques, such as deep learning, have shown remarkable success in analyzing medical images, such as electrocardiograms (ECGs), echocardiograms, and cardiac MRI scans. These algorithms can assist in the early detection and diagnosis of heart conditions by accurately identifying abnormalities, structural anomalies, and signs of cardiac dysfunction.
Decision Support Systems: ML-based decision support systems can assist healthcare professionals in interpreting complex data and making informed clinical decisions. These systems can integrate patient information, test results, and medical guidelines to provide personalized recommendations, such as treatment options, medication selection, and lifestyle modifications.
Remote Monitoring: ML algorithms can facilitate remote monitoring of patients with heart diseases, enabling continuous analysis of vital signs, activity levels, and other relevant parameters. This allows for early detection of any anomalies or changes in the patient's condition, triggering timely interventions and preventing adverse events.

The application of ML in heart disease detection and management holds great potential for improving patient outcomes, reducing mortality rates, and optimizing healthcare resource utilization. However, it is important to ensure the ethical use of ML algorithms, maintain data privacy and security, and validate the models through rigorous clinical studies to ensure their accuracy, reliability, and generalizability in real-world healthcare settings.

2 Related Work

2.1 Review of Literature (Related Work)

The different data mining methods used for heart disease prediction were reviewed by Krishnaiah et al. [3]. They investigated how different works used different approaches. Considering a variety of features and implementation options, different degrees of accuracy were obtained. The intelligent fuzzy techniques improved the accuracy of the prediction system they found.

Chitra et al. [4] used supervised learning algorithm for heart disease prediction. SVM was used to compare the results. The patient record is classified using a Cascaded Neural Network (CNN) classifier. The CNN classifier receives 13 attributes as input to evaluate disease risk. A total of 270 patient records were included in the dataset. The CNN classifier outperformed the well-known SVM classifier in terms of accuracy.

The classification method was used by Medhekar et al. [5] to detect heart disease. They used a Naive Bayes classifier to classify the data. The system classifies the data into five groups. The technology predicts the patient's risk level for each new sample that comes in. They tested with different ratios of training to testing data.

3 Majority Voting Ensembling Framework (MVEF)

3.1 Introduction

This section explained how the classifiers work independently in the prediction of heart disease. This section discusses how the base classifier results can be combined to build an Ensemble framework [6] to improve the efficiency of the heart disease prediction.

3.2 Majority Voting Ensembling Framework (MVEF)

Figure 1 shows the working of the MVEF [5, 7] model flow chart. In this model, extended K-nearest neighbor (EKNN) [8], K-nearest neighbor (KNN) [9, 10], support vector machine (SVM) [7, 10], classification and regression trees (CART) [3, 7], and Logistic Regression (LR) [11] are used as base classifiers. These classifiers are trained using training data set. Each test data item is given as input to all five base classifiers, each of the classifiers predicts the class label for the given test data item. The final prediction is done by the Ensemble framework [6] based on class labels predicted by the majority classifiers. This process is known as majority voting.

Algorithm 3.1: Majority Voting Ensemble Framework (MVEF).
Input: Heart Disease Dataset
Output: Confusion matrix and performance measures Method:
Step 1. Split the dataset into training and testing.
Step 2. Train the base classifiers namely EKNN, KNN, SVM, CART, LR.
Step 3. For each test data point compute the following

3.1 Discover the predicted class label of each base classifier.

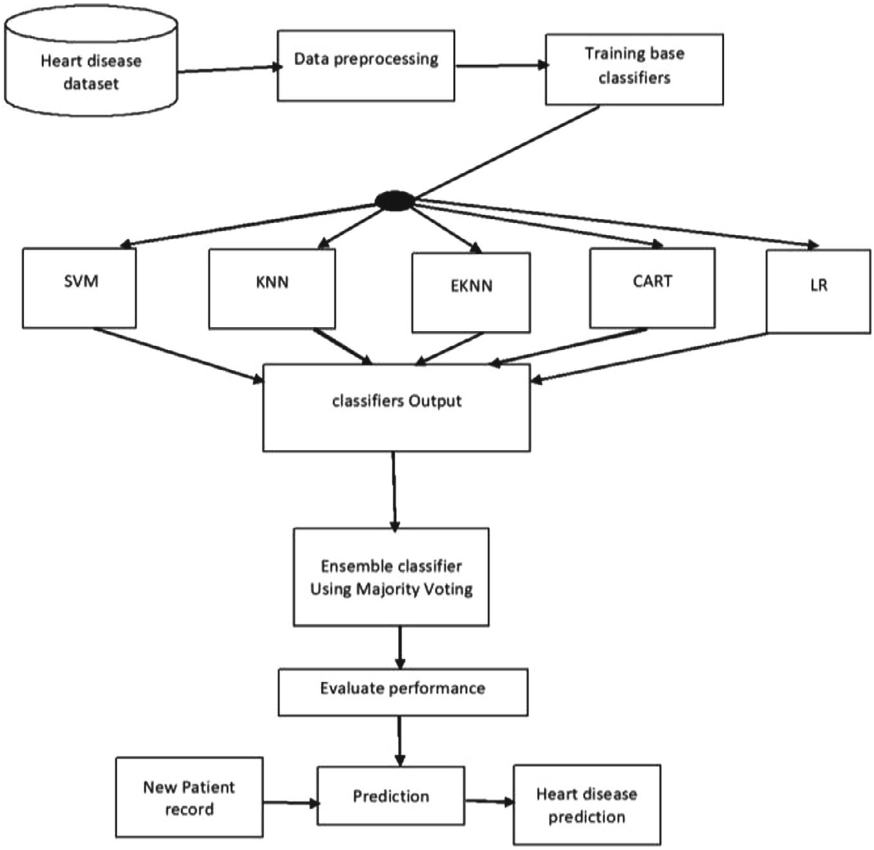


Fig. 1 MVEF flow chart

3.2 Apply the majority voting method to allocate the class label to test data item.

Step 4. Repeat step 3 for all the test data items construct confusion matrix from actual and predicted class labels.
Step 5. Calculate performance measures such as accuracy, positive predictive value (precession), true positivity rate (recall), F-score from the confusion matrix.

3.3 MVEF Model

3.4 MVEF Experimentation results

Table 1 illustrates the confusion matrix of the MVEF classifier. One hundred and thirty-two positive cases predicted as positive. Forty-six true-negative cases are there, i.e., from total negative test data items ANNEF predicts 46 as negative. There are sxi false-positive cases means six negative test data tuples are predicted as positive by the classifier. Two false-negative cases indicates that the classifier predicts these two positive test tuples as negative. In this research, positive means patient suffering from heart disease and negative means there is no heart disease.

Table 1 indicates the confusion matrix [12] of the MVEF classifier. There are 134 true positive cases, i.e., 132 test data items predict as positive from all positive cases. Forty-two true negative cases are there, i.e., from total negative test data items MVEF predicts 42 as negative. There are eight false-positive cases means eight negative test data tuples are predicted as positive by the classifier. Two false-negative cases tell that the classifier predicts these two positive test tuples as negative. In this research, positive means the patient is suffering from heart disease and negative means there is no heart disease.

Table 2 demonstrates the performance measures of the proposed MVEF model. The dataset is split into train and test datasets. About 80% of the data is used for training and 20% dataset is used for testing. The results shown in Table 3 are computed by assessing the MVEF model using the above test dataset.

Table 1 Confusion matrix [12] of ANNEF

	Predictive positive	Predictive negative
Actual positive	TP(132)	FN(2)
Actual negative	FP(6)	TN(46)

Table 2 Performance measures of ANNEF

Measure	Value
Accuracy	95.69
Precision	95.65
Recall	98.50
F Measure	97.05

Table 3 Comparison MVEF performance measures base classifiers

Performance measure	MVEF	EKNN	KNN	SVM	CART	LR
Accuracy	95.69	91.40	83.15	83.66	80.97	83.01
Precision	95.65	94.20	88.32	85.86	86.86	85.11
Recall	98.50	94.20	88.97	92.35	87.5	92.56
F Measure	97.05	94.20	88.64	88.98	87.17	88.68

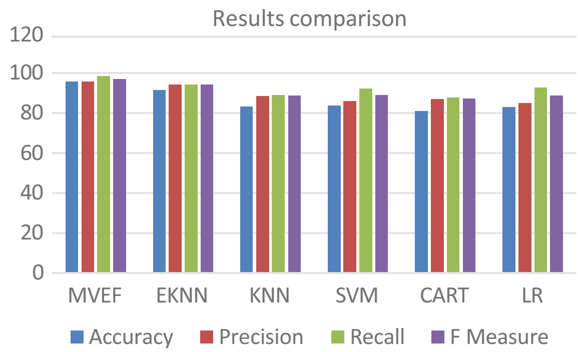
4 Conclusion and Future Work

The majority voting ensemble framework (MVEF) was developed to improve the efficiency of heart disease prediction by combining the results of multiple base classifiers. The Heart Disease Dataset (HDD) was created by combining five datasets from the UCI repository, resulting in a larger and more reliable dataset with 13Comparisons were made between MVEF and the individual base classifiers. MVEF consistently outperformed the average performance of the base classifiers in terms of accuracy, precision, recall, and F-measure. The Majority Voting Ensemble Framework (MVEF) proved to be an effective approach for heart disease prediction, leveraging the strengths of multiple base classifiers to achieve higher accuracy and performance compared to individual classifiers. The results demonstrated the potential of ensemble methods in improving the efficiency and reliability of medical diagnosis attributes describing heart disease (Fig. 2).

4.1 Future Scope of Research Work

There is a lot of scope for future work for this research which is explained as follows:

Fig. 2 ANNEF result comparison all other base classifiers



- These models can be further improved using data analytics and advanced machine learning which can predict future heart disease problems using the present patient's condition.
- These algorithms can be applied to other areas of class imbalance like fraud detection, gene disorders, etc.
- Make use of data pipelines to automate the process of data capture and predictions.
- The models can be further improved recall by using unstructured ML algorithms to avoid false-negatives
- A wearable IoT device [4] can be designed that continuously monitors the patient's condition and warn the patient as early as possible before a sudden heart attack.

References

1. Chauhan A (2020) Heart disease prediction using machine learning with python. https://towardsai.net/p/machine-learning/heart-disease-prediction-using-machine-learning-with-python
2. David HB, Belcy SA (2018) Heart disease prediction using data mining techniques. ICTACT J Soft Comput 9(1):1817–1823
3. Rahim A, Rasheed Y, Azam F, Anwar MW, Rahim MA, Muzaffar AW (2021) an integrated machine learning framework for effective prediction of cardiovascular diseases. IEEE Access 9:106575–106588. https://doi.org/10.1109/ACCESS.2021.3098688
4. Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on computational learning theory—COLT, vol 92, pp 144–152
5. Navlani A (2018) Decision tree classification in python. www.datacamp.com/community/tutorials/decision-tree-classification-python
6. Latha CBC, Jeeva SC (2019) Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. 16:100203. https://doi.org/10.1016/j.imu.2019.100203
7. Abhishek T (2013) Heart disease prediction system using data mining techniques. Orient J Comput Sci Technol 6(4):457–466
8. Sun B, Du J, Gao T (2009) Study on the improvement of K-nearest-neighbor algorithm. In: International conference on Artificial Intelligence and computational intelligence. Shanghai, pp 390–393. https://doi.org/10.1109/AICI.2009.312
9. Bhargava N, Dayma S, Kumar A, Singh P (2017) An approach for classification using simple CART algorithm in WEKA. In: 11th international conference on intelligent systems and control (ISCO)
10. Chitra R, Seenivasagam V (2013) Heart disease prediction system using supervised learning classifier. Bonfring Int J Softw Eng Soft Comput 3(1):01–07
11. Singh A, Kumar R (2020) Heart disease prediction using machine learning algorithms. In: ICE3, vol 58. https://ieeexplore.ieee.org/abstract/document/91229
12. Ragan A (2018) Taking the confusion out of confusion matrices. https://towardsdatascience.com/taking-the-confusion-out-of-confusion-matrices-c1ce054b3d3e

A Low-Power 8t Sram Cell's Desıgn and Development Evolution Designed for Elevated Density Memory Applications

P. Venkatakrishnan, C. H. Rekha, Maughal Ahmed Ali Baig, B. Venkateshwar Rao, B. Premalatha, and V. A. Naryana

Abstract The core elements of contemporary digital structures are SRAM cells. For such digital devices to be implemented successfully, an effective and dependable SRAM cell architecture is essential. In order to better understand the possibilities of utilizing low-power SRAM conventional cells for highly dense memory programs, this study will do so. Here, different low-power SRAM cell architectures and topologies are assessed and contrasted with conventional SRAM cells. The efficiency of such cells is also examined with regard to of battery life, velocity, and surface usage. The findings of this research will offer important guidance for the creation of a high level of storage systems with minimal power requirements and effective area use. The suggested 8T SRAM cell offers reduced energy consumption as well as fewer delay over the other SRAM frameworks, according to the required study and comparisons of all the SRAM cells. As a result, the suggested SRAM cell may be an excellent choice for storage systems that require minimal power consumption and high efficiency.

P. Venkatakrishnan
Department of ECE, CMR Technical Campus, Hyderabad, Telangana, India

C. H. Rekha (✉)
Department of CSE, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: cmrtc.paper@gmail.com

M. A. A. Baig
Department of Mechanical, CMR Technical Campus, Hyderabad, Telangana, India

B. Venkateshwar Rao · B. Premalatha
Dept.of ECE, CMR College of Engineering and Technology, Hyderabad, Telangana 501401, India
e-mail: bvenkateshwarrao@cmrcet.ac.in

B. Premalatha
e-mail: bpremalatha@cmrcet.ac.in

V. A. Naryana
Department of CSE, CMR College of Engineering and Technology, Hyderabad, Telangana 501401, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_51

Keywords Static RAM · Power delay product (PDP) · Low power · Less delay

1 Introduction

A sort of storage cell called an SRAM (Static Random Access Memory) unit is used for organizing a small amount of digital data in an electrical circuit. Although the information is "Static" and remains in the cell until such time as energy is supplied through the circuit, the term "Static" was coined. Microprocessors, caching memories, and various other electronic components frequently use SRAM cells. They are essential to the operation of contemporary digital technologies, offering quick and dependable storage solutions for a variety of tasks.

These days, ultra-low power SRAM cells are required for storing data in applications that use little power such as medical equipment, internet of things gadgets, devices that are worn, and portable electronics. Controlling the cell's power consumption in SoCs (System on Chips) has become a key challenge in the design of low-power SRAM cells. To avoid taking up extra space, the SRAM cells should only have the necessary amount of transistors. However, attempting to do so by reducing the number of transistors leads to manufacturing variances, which can deviate from the SRAM cell's ideal performance. Differences in the doping of silicon substrate, differences in the depth of the oxides levels, and variances in the position of the lithography masks used to design each layer are just a few examples of the many causes of these variances. Additionally, it makes scalability more severe and can have detrimental consequences on how the cell functions by raising leaking strength, which raises SRAM cell energy use and is not advised when using low-powered SRAM cells. Therefore, we must lower the supply voltage to the threshold voltage level in order to minimize the amount of energy usage in the SRAM cells. However, lowering the supplied voltage causes the cell's performance and stability to deteriorate.

One should design an SRAM cell to fulfill their needs by taking into account all of the aforementioned factors without affecting how it operates. This study analyzes and calculates the performance metrics of several SRAM cell structures (6T, ST1, WRE8T, SB9T, 12T, 11T, 8T), including mean power, tardiness, and PDP (power delay product) during both writes and reads [4–6, 8, 9]. The recommended 8T SRAM cell uses less energy than other types of cells and operates more quickly than a conventional cell, according to a comparison of the results [12–17]. In order to build extremely low-energy SRAM cells that can be employed in several portable electronic devices, this proposed SRAM cell may be a better option.

2 Lıterature Survey

With an emphasis on the many designs, methodologies, and innovations utilized to enhance the efficiency and energy usage of SRAM cells, we will examine recent developments and problems in SRAM cell structure in this literature analysis [3].

Nabavi and Sachdev [1], offered a 290-mV, 3.34-MHz 6T SRAM cell to improve the bitcell's stability throughout reading operations and employs a strengthened wordline to improve its write capability while using the least amount of power.

Kulkarni et al. [2], A 160 mV robust Schmitt trigger-based sub-threshold SRAM cell with integrated variation in process tolerance has been suggested. It uses less power to read and write. The study's findings demonstrate that the ST bitcell can retain data even at inadequate supply voltages.

Mohagheghi et al. [15], Single-ended 8T SRAM cells with a large write capacity and low writing and reading power are suggested. The cell has good read durability since its read channel is distinct from bitlines.

Also want to construct an SRAM cell with eight transistors to reduce the total power use and latency of the unit after reviewing the aforementioned citations and assessing the SRAM cells with respect to of energy usage and latency.

Osada et al. [7] A semiconductor memory device consists of a memory cell array block that contains several memory cells, each of which has a data holding circuit that stores data using a first as well as a second circuit element and a transistor designed to interconnect the data retaining circuit alongside a bit line together. Sense amplifiers are connected to bit lines either through direct connections or via switches that are needed, and a fake memory cell array that contains several dummy memory cells, each of which has the same circuit arrangement, is also included. The inverter circuits are linked together in series to connect a series of duplicate memory cells, every of which has at least a single inverter circuit configuration. An activating signal for the sensor amplifiers is a signal from the output of the inverting circuit of one of the fake memory cells in the final stage..

Hazucha et al. [10] for a cutting-edge 90-nm CMOS process, the neutron soft error rate (SER) dependence on energy and region was studied. At a 10% drop in voltage, the SER climbed by 18% and grew exponentially with diodes surface. SRAMs' recorded bit SER at 0.25, 0.18, 0.13, and 90 nm increased by 8% per iteration.

Baeg et al. [11] Sharing the results of the soft-error test in three important methods, 90, 65, and 45 nm, reveals the importance of multiple cell upsets (MCUs). When utilized in conjunction with the linking architecture in storage architectures, single-bit error correction (SEC) codes are most successful in reducing MCU defects. The rate of failure provided by the model suggested in this paper allows probabilistically demonstrating the advantages of different integrating system choices for the memory with SEC. In the suggested model, clustered events like MCU are considered utilizing the combined Poisson process.

3 Related Work

Only one bit of information is stored in a 6T SRAM cell using six transistors as shown in Fig. 1a. It has two access transistors and two cross-coupled inverters that serve for writing and reading data to the cell. Access transistors enable external signals to read from and send data to the cell, while cross-coupled inverters act as a latch to store the data. The circuit offers great reliability, quick writing and reading times, and low standby power usage. Despite the fact that the cell is intended to be minimally powered operating it still consumes a substantial amount of power, particularly if compared to other kinds of memory cells. This SRAM cell's performance is very reliant on its manufacturing procedure, making it more vulnerable to changes in the process and more challenging to produce continuously.

Ten transistors are used in the ST-1 SRAM cell to store one bit of data as shown in Fig. 1b. Schmitt trigger inverters are utilized in place of conventional inverters in this form of memory cell, which is employed in digital circuits. This cell can function storage effectively at low voltages and in noisy situations because it consists of two cross-coupled inverters that have a Schmitt trigger inverter sandwiched in between them. Schmitt trigger SRAM cells are typically larger and more complex, increasing the surface area and dimension of the device, which may raise manufacturing costs and reduce yield. This circuit serves a purpose in low-power applications because it can operate at a voltage that is less than traditional SRAM cells and minimizes the memory cell's power consumption.

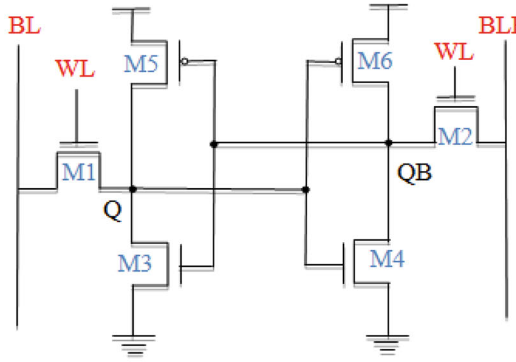
One bit of data is stored in an 8T SRAM cell using eight transistors as shown in Fig. 1c. Two access transistors are included for every one of the two inverters in this SRAM cell, which has two cross-coupled inverters. The external signals can read information from the cell or write data there thanks to the access transistors. This cell offers greater immunity to process changes, decreased leakage, and enhanced stability. However, the design of the cell is challenging since more transistors are needed to fill the space on the device, which could raise the cost of storage and lower chip yield. The fluctuating electrical usage of the cell rises with the number of transistors added, adding to the time and expense of memory design and lowering overall reliability.

4 Projected 8T SRAM CELL

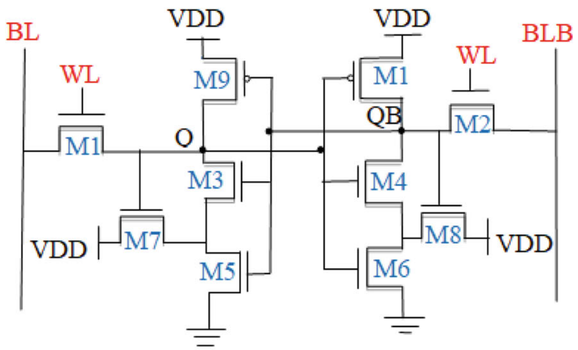
In addition to the typical 6T SRAM cell, the suggested 8T SRAM cell has two extra transistors, however, one NMOS and one PMOS, as shown in Fig. 2.

This cell's configuration comprises of two phrase lines and two bitlines. The read operation in this suggested cell is single-ended, meaning that just the RWL (read word line) input is asserted. This enables both access transistors and enables reading of the stored value from the cell, see Table 1

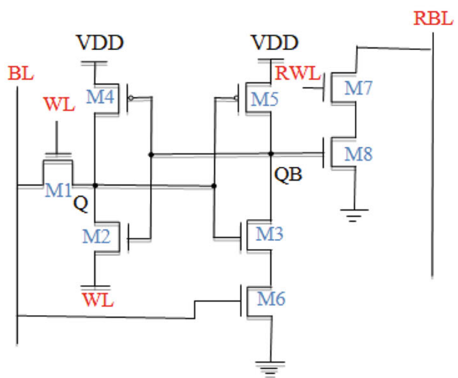
Fig. 1 Schematic of dissimilar SRAM cells, **a** 6T SRAM cell, **b** ST-1 SRAM cell, **c** 8T SRAM cell



(a)



(b)



(c)

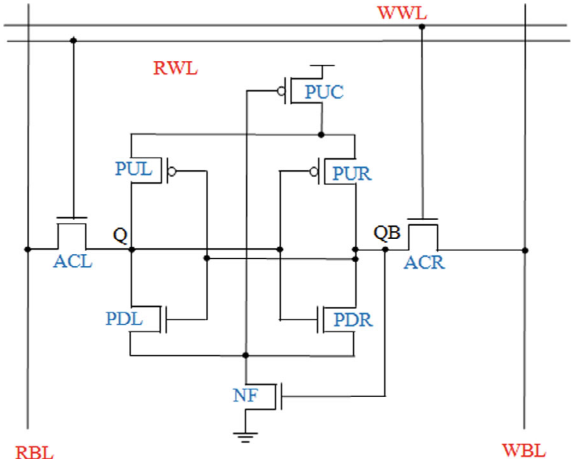


Fig. 2 Proposed 8T SRAM cell

Table 1 List of control signal voltages of proposed 8T cell

Control signal	Hold	Read	Write
WWL	GND	GND	VDD
RWL	GND	VDD	VDD
WBL	VDD	VDD	VDD

4.1 Write Operation

The transistors responsible for ACL and ACR are turned on during a write operation, enabling connectivity among the bitlines and memory nodes. This is accomplished by setting both the WWL along with RWL signals to high. When "1" is stored upon the storage node *Q*, the transistor PUC operates to increase the write margin. When QB is "0", the current flowing on the outlet of NF increases, reducing the drivability of PUC, making it simpler to write a "0" on the storage node. However, the write margins is not what is anticipated when QB is "1." Therefore, the write space can be enhanced for both WWL as well as RWL phases by sizing PUR lower than PUL.

4.2 Read Operation

Just RWL is set to high during read operation, turning on the transistor ACL while keeping the ACR off. Although PDR, PUL, and ACR remain off, if *Q* is zero, transistor PDL and NF drain the bitline capacitance. There is no discharge channel from *Q* to ground when *Q* is 1 and NF is off, which significantly reduces the read residual noise margin.

4.3 Hold Operation

Both the WWL and RWL indications are set to low during hold operation, which disables the access control transistors ACL along with ACR. No reading or writing operation is carried out in the cell due to the fact there is no link among the memory nodes and the bitlines. As a result, the storage nodes' data will be in a hold state. Both of NF and PDR circuits are activated while *Q* holds "0," improving data retention and maintaining "0" at QB.

5 Results and Discussions

6T SRAM CELL

As shown, Fig. 3a is write operation of 6T SRAM cell and Fig. 3b is read operation for 6T SRAM cell.

ST-1 SRAM CELL

As shown, Fig. 4a is write operation of ST-1 SRAM cell and Fig. 4b is read operation for ST-1 SRAM cell.

As shown, Fig. 5a is write operation of 8T SRAM cell and Fig. 5b is read operation for 8T SRAM cell.

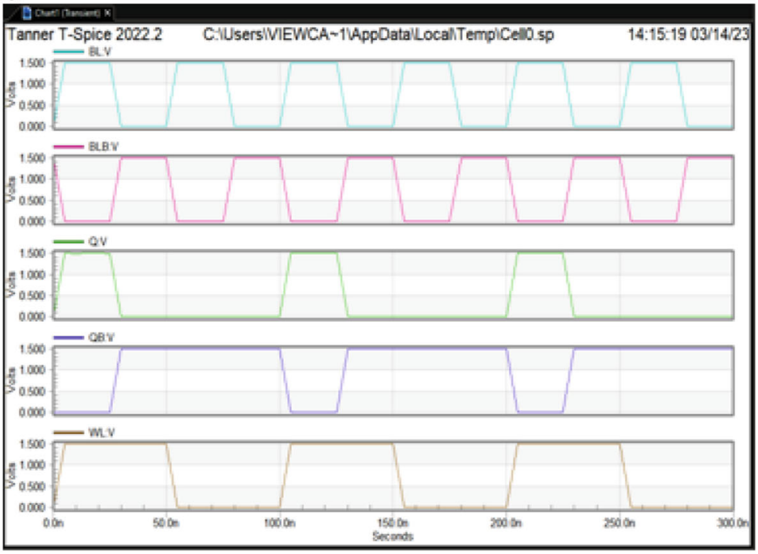
As shown, Fig. 6a is write operation of 8T SRAM cell and Fig. 6b is read operation for 8T SRAM cell.

Comparison of various SRAM structures of their performance like power, delay, and power delay product of 6T, ST-1, 8T, and proposed 8T cell can be seen in Fig. 7.

6 Conclusıon

It can be helpful for creating an SRAM cell for low-energy and high-density memory applications, according to the paper titled "Design of a recommended low energy 8T SRAM Cell for High Density Storage Applications." In this study, the power and delay characteristics of different kinds of SRAM cells are compared. The suggested 8T SRAM cell offers low power usage and minimal delay to improve operation speed, we conclude after examining the findings.

a



b

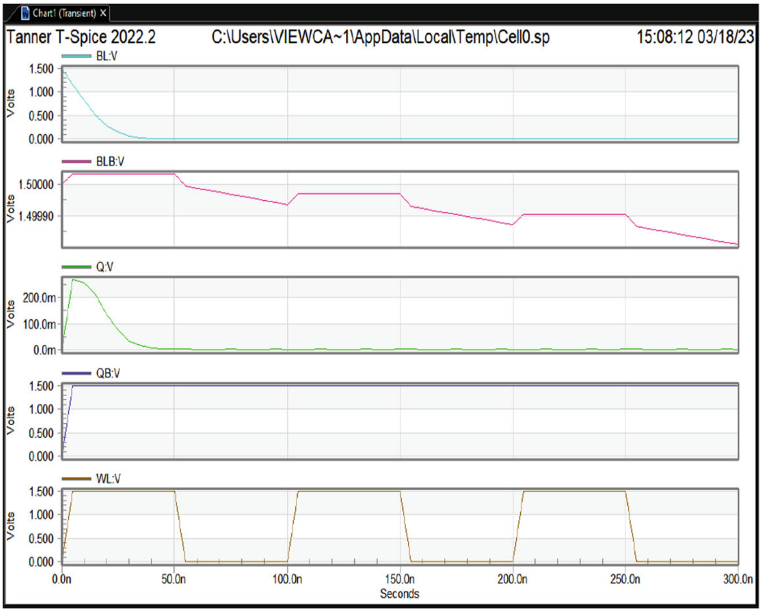


Fig. 3 **a** Write operation of 6T SRAM cell. **b** Read operation of 6T SRAM cell

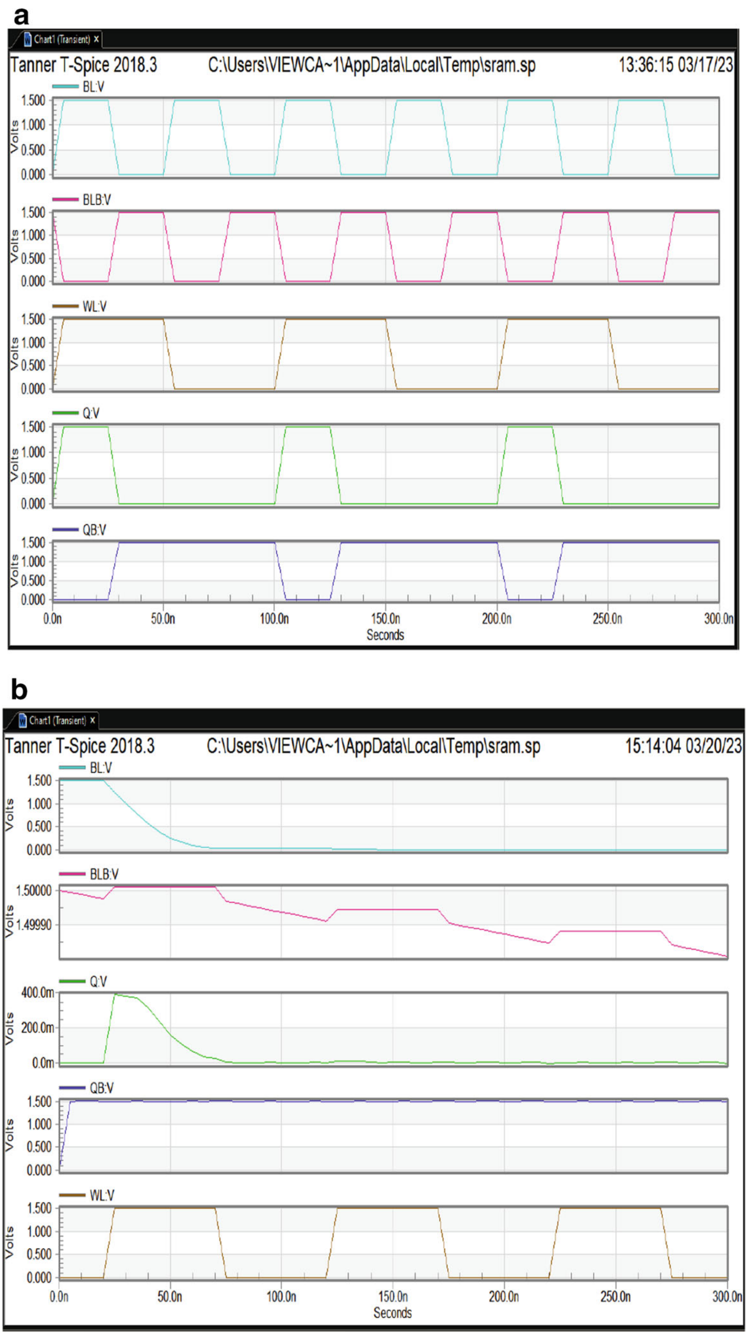


Fig. 4 **a** Write operation of ST-1 SRAM cell. **b** Read operation of ST-1 SRAM cell

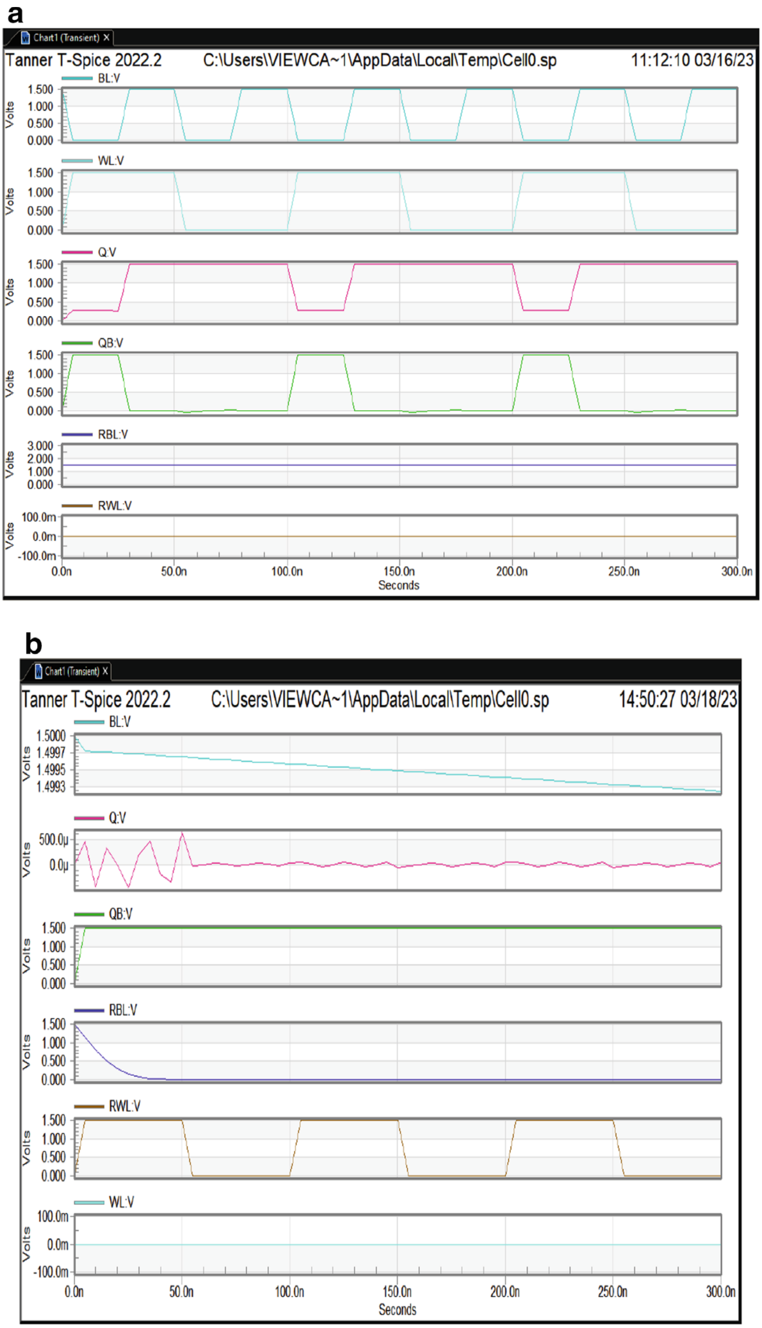


Fig. 5 **a** Write operation of 8T SRAM cell. **b** Read operation of 8T SRAM cell

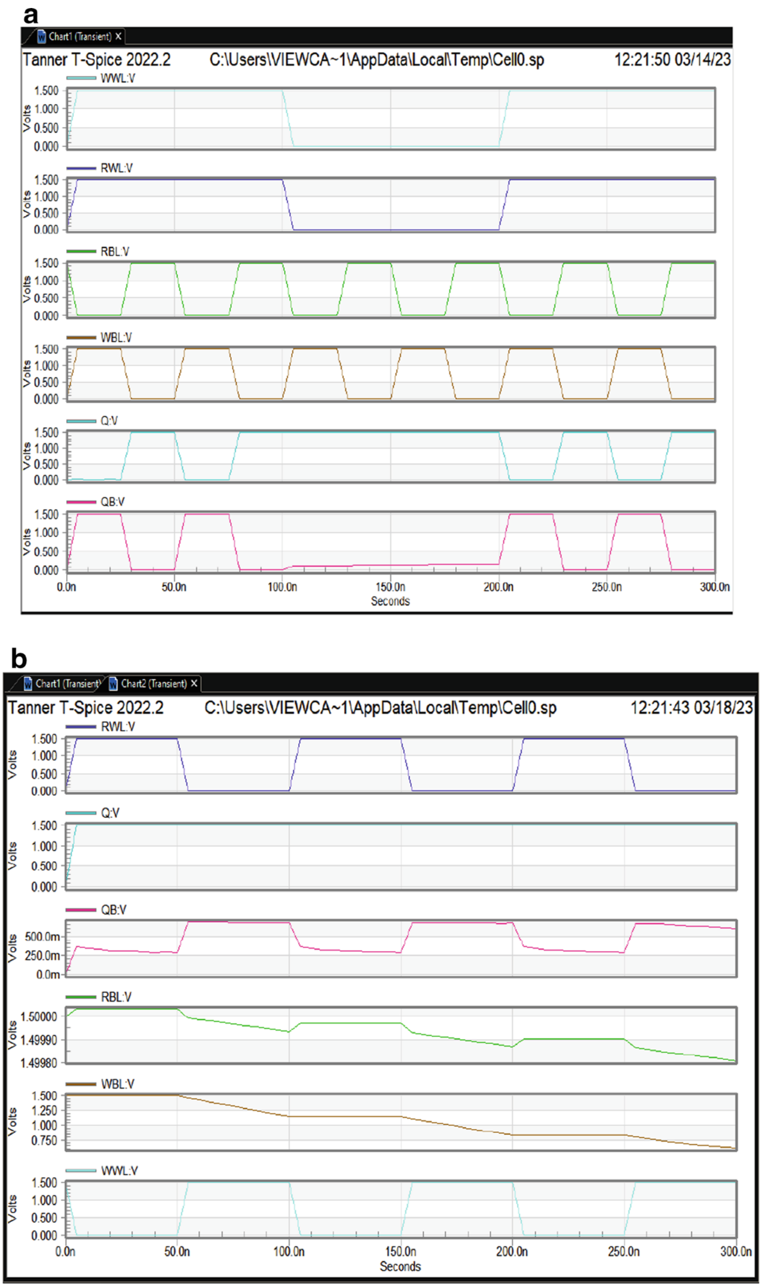
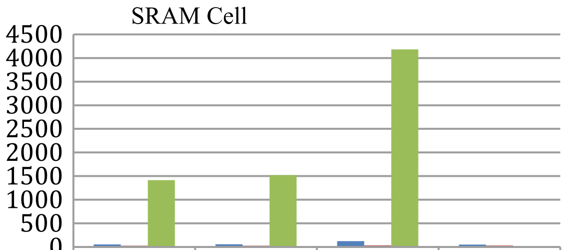


Fig. 6 **a** Write operation of proposed 8T. **b** Read operation of projected 8T



	6T	ST-1	8T	Proposed 8T
Power(µW)	53.2917	57.5548	123.4219	49.3998
Delay (nSec)	26.5765	26.438	33.8914	30.406315
Power delay Product	1416.14	1521.66	4182.91	

Fig. 7 Comparison of various SRAM structures

7 Future Scope

The suggested cell currently employs a single-ended read strategy, which will lessen the sudden change of the output. Additional methods for 8T SRAM cells which may boost their functionality, power efficiency, and dependability, such as Write-assist techniques and FinFET technology, need to be further investigated.

References

1. Nabavi M, Sachdev M (2018) A 290-mV, 3.34-MHz, 6T SRAM with pMOS access transistor and boosted wordline in 65-nm CMOS technology. IEEE J Solid State Circ 53(2):656–667. https://doi.org/10.1109/JSSC.2017.2747151
2. Kulkarni JP, Kim K, Roy K (2007) A 160 mV robust Schmitt trigger based subthreshold SRAM. IEEE J Solid State Circ 42(10):2303–2313. https://doi.org/10.1109/JSSC.2007.897148
3. Pasandi G, Fakhraie SM (2014) An 8T low-voltage and low-leakage half-selection disturb-free SRAM using bulk-CMOS and FinFets. IEEE Trans Electron Devices 61(7):2357–2363. https://doi.org/10.1109/TED.2014.2321295
4. Ahmad S, Gupta M, Alam N, Hasan M (2017) Low leakage single bitline 9t (sb9t) static random access memory. Microelectron J 62:1–11. https://doi.org/10.1016/j.mejo.2017.01.011
5. Kim J, Mazumder P (2017) A robust 12T SRAM cell with improved write margin for ultra-low power applications in 40nm CMOS integration, vol 57, pp 1–10. https://doi.org/10.1016/j.vlsi.2016.09.008
6. Jeong H, Kim T, Song T, Kim G, Jung SO (2015) Trip-point bitline precharge sensing scheme for single-ended SRAM. IEEE Trans Very Large Scale Integer (VLSI) Syst 23(7):1370–1374
7. Osada K et al (2001) Universal-Vdd 0.65-2.0-V 32-kB cache using a voltage adapted timing-generation scheme and a lithographically symmetrical cell. IEEE J Solid State Circuits 36(11):1738–1744

8. Yabuuchi M, Nii K, Tsukamoto Y, Ohbayashi S, Nakase Y, Shinohara H (2009) A 45 nm 0.6 V cross-point 8T SARM with negative biased read/write assist. In: Proceedings on IEEE symposium on VLSI circuits, pp 158–159
9. Kushwah CB, Vishvakarma SK (2016) A single-ended with dynamic feedback control 8T subthreshold SRAM cell. IEEE Trans Very Large Scale Integer (VLSI) Syst 4(1):373–377
10. Hazucha P, Karnik T, Maiz J, Walstra S, Bloechel B, Tschanz J, Dermer G, Hareland S, Armstrong P, Borkar S (2003) Neutron soft error rate measurements in 90-nm CMOS process and scaling trends in SRAM from 0.25-? m to 90-nm generation. In: IEDM Tech Dig, pp 21.5.1–21.5.4
11. Baeg S, Wen S, Wong R (2009) SRAM interleaving distance selection with a soft error failure model. IEEE Trans Nucl Sci 56(4):2111–2118
12. Chang IJ, Kim J-J, Park SP, Roy K (2009) A 32 kb 10T sub-threshold SRAM array with bit-interleaving and differential read scheme in 90 nm CMOS IEEE. J Solid State Circuits 44(2):650–658
13. Chang MF, Chen MP, Chen LF, Yang SM, Kuo YJ, Wu JJ, Su HY, Chu YH, Wu WC, Yang TY, Yamauchi H (2013) A sub-0.3 V area-efficient L-shaped 7T SRAM with read bitline swing expansion schemes based on boosted read bitline, asymmetric-VTH read-port, and offset cell VDD biasing techniques. IEEE J Solid State Circ 48(10):2558–2569
14. He Y, Zhang J, Wu X, Si S, Zhen S, Zhang B (2019) A half-select disturb-free 11T SRAM cell with built-in write/read-assist scheme for ultralow-voltage operations. IEEE Trans Very LargeScale Integr (VLSI) Syst 27(10):2344–2353. https://doi.org/10.1109/TVLSI.2019.2919104
15. Mohagheghi J, Ebrahimi B, Torkzadeh P (2022) Single-ended 8T SRAM cell with high SNM and low power/energy consumption. Int J Electron. https://doi.org/10.1080/00207217.2022.2118848
16. Sanagala SS, Gupta SK, Koppula VK, Agarwal M (2019) A fast and light weight deep convolution neural network model for cancer disease identification in human lung(s). In: Proceedings—18th IEEE international conference on machine learning and applications, ICMLA, pp 1382–1387
17. Bathula A, Basha SK (2019) Blockchain technology with internet of things in the real time network stream. Int J Recent Technol Eng 8(1, Special Issue 4):682–689

Optimizing Energy Efficiency in Wireless Sensor Networks Using Machine Learning Techniques

A. Hemalatha Reddy, Jalapala Sinjini, M. Priyadharshini, V. Indumathi, V. V. Bhavani, and Tahseen Jahan

Abstract Wireless sensor networks (WSNs) have emerged as a crucial technology for various applications including environmental monitoring, healthcare, and industrial automation. With the proliferation of sensor nodes in WSNs, there is a growing need for efficient data processing and management techniques. Machine learning (ML) has shown promise in addressing the diverse requirements of WSNs, including data routing, energy management, and fault detection. This review paper explores the requirements of ML techniques in WSNs, discusses the limitations of applying ML in such networks, and provides a brief introduction to ML techniques in the context of WSNs. Furthermore, the paper examines the challenges encountered in WSNs and ML techniques, particularly focusing on routing issues. Finally, various ML techniques employed in WSNs routing are surveyed, highlighting their advantages and limitations.

Keywords Wireless sensor networks · Challenges · ML techniques · Machine learning · Routing

A. Hemalatha Reddy (✉) · J. Sinjini · M. Priyadharshini · V. Indumathi · V. V. Bhavani · T. Jahan
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: hemareddy@cmritonline.ac.in

J. Sinjini
e-mail: sinjini@cmritonline.ac.in

M. Priyadharshini
e-mail: priyadharshini@cmritonline.ac.in

V. Indumathi
e-mail: indu11mathi@cmritonline.ac.in

V. V. Bhavani
e-mail: bhavaniv@cmritonline.ac.in

T. Jahan
e-mail: tahseen048@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_52

1 Introduction

Distributed networks with many sensor nodes, sometimes placed in inaccessible places, may keep tabs on weather conditions including temperature, humidity, wetness, and more [1]. There are many kinds of sensors included in the sensor nodes, such as acoustic, optical, chemical, motion, pressure, temperature, and weather sensors. With such a wide variety of sensor nodes, WSNs have a wide range of potential uses, from medicine and the military to agriculture and our daily lives. Despite its widespread use, WSNs have common problems such as insufficient power, slow processing, insufficient memory, and low communication bandwidth, which reduce the network's lifespan and cause sensor network performance degradation. Creating unique algorithms for various uses is a daunting undertaking. Designers of WSNs should pay special attention to details such data clustering, routing, localization, fault detection, job scheduling, event monitoring, and data aggregation. One subfield of AI that emerged in the '50s was machine learning, or ML [2]. It progressed over time, eventually arriving at algorithms that were computationally practical and robust enough to deal with a variety of issues in computers, engineering, and medicine, including clustering, optimization, regression, and classifications. When it comes to modern technology, ML is up there with the most fascinating and consequential innovations. Machine learning (ML) enables computers to learn on their own, without any human intervention, and then behave appropriately. Automatic, rapid, and accurate analysis of complicated data is what it uses to build a model. With the help of the generalized structure, ML can learn to increase system performance in a generic way. Hence, ML methods are great for retrieving massive datasets and extracting relevant information.

1.1 ML Techniques in WSNs: A Brief Introduction

To improve their functionality and efficiency, wireless sensor networks (WSNs) make use of methods that are associated with machine learning (ML). Data aggregation, routing, energy management, and problem detection are just some of the issues that wireless sensor networks (WSNs) face. Machine learning methods are used to handle these challenges. One common application is in data aggregation, where ML algorithms are utilized to process and summarize data collected from sensor nodes, reducing the amount of transmitted data and conserving energy. Additionally, ML-based routing protocols dynamically adapt to network conditions, optimizing data transmission paths and minimizing energy consumption.

Moreover, ML techniques enable predictive maintenance in WSNs by analyzing historical data to detect potential faults or failures in sensor nodes [3], allowing for timely intervention and preventing network disruptions. Energy management is another crucial aspect where ML algorithms optimize power consumption by dynamically adjusting sensor node operation modes based on workload and environmental conditions. ML techniques play a pivotal role in enhancing the performance, reliability, and energy efficiency of wireless sensor networks across various applications.

2 Background

According to Ref. [4], ML technology has publicized its exceptional benefits in a wide range of domains and has demonstrated outstanding performance in a variety of applications. These applications include natural language processing, voice recognition, picture recognition, and recommendation systems. It was brought to their notice that the use of machine learning in wireless sensor networks (WSNs) has garnered a lot of interest in recent times. According to their observations, the identification of ways to increase resource utilization and achieve power-efficient load balancing had become a major problem in wireless sensor networks (WSNs) due to the fact that resources were restricted in WSNs. They noted that old green routing algorithms attempted to do this by lowering the amount of energy that was used and extending the lifespan of the network using highly optimized routing techniques in wireless sensor networks (WSNs). An approach known as ML-EOA (the Machine Learning-based Energy Optimization Approach) was presented by Ref. [5] with the purpose of overcoming the research obstacles that wireless sensor networks (WSNs) are experiencing as a result of limiting energy restrictions in sensor nodes. They observed that the strategy improved energy optimization, network coverage, and lowered latency simultaneously. Reference [6] noted that the increased scalability and inexpensive cost of deployment of industrial wireless sensor networks (WSNs) were contributing to their growing popularity. According to what they saw, these networks also brought new issues, such as the optimization of energy use and the maintenance of networks, which industrial users were required to handle. They said that in order to overcome these issues, approaches from the field of ML have been used in order to develop an improved energy optimization model for industrial wireless sensor networks (WSNs).

In Ref. [7] In order to maximize data efficiency and minimize energy usage, deep learning (DL) architectures were used to allocate resources for wireless sensor Internet of Things networks. They found that energy efficiency (EE) and spectrum efficiency (SE) optimization targets conflicted. Whale-optimized deep neural networks utilize less energy, according to their findings. They optimized the data using the heuristic-based multiobjective firefly technique. By optimizing power allocation and relay selection, their strategy enhanced cooperative multihop network architectures. They said reducing transmit power and meeting quality-of-service requirements were the best ways to distribute resources and pick relays. Energy-efficient methods were the consequence, they stated.

In Ref. [8] that a wireless sensor network (WSN) was nothing more than a collection of numerous tiny sensor nodes that were self-governing, less costly, and used less power. They said that these nodes gathered information from the global surroundings, aggregated it, and then transported it to a base station for further processing. They also mentioned that the sensor nodes were used to analyze the data after it was acquired from the various kinds of sensors. It was claimed that the majority of the recent research has been on creating solutions for energy-efficient deployment, routing, and administration of wireless sensor networks (WSN). Specifically, they pointed out that in reality, WSN originators were confronted with a common difficulty that was associated with information clumping, information dependability, nodule bunching, energy-saving routing, event arrangement, burden exposure, and security.

According to Ref. [9], wireless sensor networks (WSNs) are very helpful in situations where there is no established infrastructure, such as in the case of military applications and emergency rescue operations, and when there is a need to construct a network at a cheap cost. They made the observation that there was no set routing protocol or intrusion detection approach that was accessible to them. This was due to the fact that wireless sensor networks (WSNs) were dynamic in nature, and in order to do this, individual nodes of the network were necessary. They made the observation that the majority of the applications of wireless sensor networks (WSNs) were mobile nodes, which meant that they were dependent on battery power and the availability of restricted resources. This demonstrated that power consumption was a useful study topic for carrying out a set of activities in WSNs. They indicated that in order to address such a problem, machine learning (ML) approaches had the potential to be successfully deployed in accordance with the requirements of the application. These techniques include self-learning algorithms that operate without the need for programming or human interaction. In their research, they said that they had conducted a comparative analysis of many machine learning-based strategies for wireless sensor networks (WSNs). In addition, they mentioned that they had assessed machine learning techniques for clustering and energy harvesting. In conclusion, they offered an overview of machine learning algorithms for clustering as well as energy harvesting, along with some unresolved concerns. They pointed out that this was not the case.

In Ref. [10], wireless sensor networks were often organized into clusters in order to gather information in an efficient manner. The arrangement of nodes in this manner was shown to be of great assistance in elongating wireless sensor networks, which are the lifeblood of the industry. They stated that the sensor nodes were overwhelmed as a result of message exchanges between nodes for successive and periodic clustering, which resulted in a supply of energy that was insufficient. The existing clustering concerns, according to them, include increased overhead during cluster formation, volatility in energy usage, the difficulty of information exchange during clustering, unknown network topology, and other similar issues. It was also brought to their attention that there was a need to improve the transmission inside the cluster and to discover efficient techniques that may increase the lifetime of the network.

According to Ref. [11], the deployment of wireless sensor networks is often found in applications related to research and engineering, as well as in applications related to the military for the purpose of intrusion detection and in applications related to civil engineering. According to them, the architecture of wireless sensor networks (WSNs) may be application specific, which raised a number of difficulties and limits. They spoke about the numerous problems that wireless sensor networks (WSNs) have with the system that is based on machine learning and the various solutions that were given by various writers. In order to increase energy efficiency in routing, they emphasized that the work that would be done in the future will place an emphasis on movable sink positioning and dynamic clustering.

In their study, Ref. [12] found that wireless sensor networks were more resilient when the nodes have the capacity to self-configure. This meant that the nodes had the option of either being a member of the network or just leaving it. They said that the data transmission was accomplished via the intercommunication that took place between the nodes that were a part of the network. Despite the fact that the other nodes had the appropriate amount of energy for operation, they observed that the majority of the nodes that participated in the network had the potential to experience energy depletion, which ultimately led to the network's demise. They made the observation that since WSNs are decentralized, another issue that has traditionally been a problem for them is the possibility of intrusion detection, which may lead to denial of service assaults that also deplete the energy.

A discussion was held by Ref. [13] on the use of machine learning (ML) as a method to enhance the energy efficiency (EE) of wireless networks. They said that in this regard, they examined the most prevalent machine learning techniques with a particular emphasis on the enhancement of empirical evidence. Initially, the chapter provided a concise explanation of self-organizing networks (SONs) and a few solutions that are associated with them, which include machine learning methods inside the framework of engineering (EE). It was observed by them that self-organizing networks (SONs) could be classified into three primary branches, namely self-configuration, self-optimization, and self-healing, which are collectively referred to as self-x functions. A number of applications of SONs in cellular networks were shown by them, with particular emphasis placed on the three primary self-extension branches and the typical use cases connected with them. They highlighted that machine learning methods might likely enhance a SON in this scenario, enabling the network to adapt by watching its present condition and using such experience to alter parameters in future activities. This would enable the network to adapt via observation. They gave a presentation that provided an introduction to machine learning approaches that were applied to more particular subjects such as cognitive radio networks, traffic prediction, and resource allocation systems (Table 1).

3 Methodology

Designing a mathematical model to optimize energy efficiency in WSNs using ML techniques involves several components and equations. Here, I will outline a simplified model that incorporates key aspects such as energy consumption, data transmission, and ML-based decision-making.

Table 1 Comparative analysis

S. No	Algorithms used	Routing	Methodology	Used topology	Research area	Findings
[4]	Machine learning (ML) techniques	Green routing algorithms	Theoretical hypothetical model formulation, ML-based intelligent routing	Wireless sensor networks (WSNs)	Energy-efficient routing in WSNs	Proposed ML model overcomes limitations of traditional green routing methods, provides new ideas for energy-efficient routing
[5]	Machine learning-based energy optimization approach	Data aggregation, fuzzy logic-based method for update/sleep cycle calculation, ANN for cluster head (CH) selection	Simulation-based evaluation	Wireless sensor networks (WSNs)	Energy optimization, network coverage, latency reduction	ML-EOA outperforms traditional methods in terms of network lifespan extension, energy usage reduction
[6]	Machine learning techniques	Enhanced energy optimization model	Knowledge-based learning, feedback control schemes evaluation	Industrial wireless sensor networks (WSNs)	Energy optimization, network maintenance	Enhanced energy optimization model achieves significant reductions in energy consumption while ensuring task efficiency and longer network lifetime
[7]	Whale optimization, DL architectures, Heuristic-based multi-objective firefly algorithm	Optimal power allocation, relay selection	Simulation-based evaluation	IoT-based restricted wireless sensor setworks (WSNs)	Energy efficiency, spectral efficiency	Proposed method improves energy efficiency, spectral efficiency, throughput, QoS, and network lifetime compared to traditional models
[8]	Machine learning (ML) techniques	Sensor grid routing	Energy-efficient deployment, routing	Wireless sensor networks (WSNs)	Energy-efficient deployment, routing, WSN management	Focus on energy-efficient deployment, routing, and WSN management

(continued)

Table 1 (continued)

S. No	Algorithms used	Routing	Methodology	Used topology	Research area	Findings
[9]	Machine learning (ML) techniques	WSN routing	ML techniques improve energy consumption management in WSNs,	Wireless sensor networks (WSNs)	Energy consumption, manual intervention	ML techniques improve energy consumption management in WSNs, particularly in dynamic environments
[10]	Machine learning, Fuzzy logic	Cluster formation, active/sleep schedule	Fuzzy logic-based approach, message transmission overhead reduction	Wireless sensor networks (WSNs)	Energy loss reduction, network lifespan improvement	Proposed approach optimizes energy utilization, reduces message transmission overhead, and improves network lifespan
[11]	ML techniques	Energy-aware routing	Discussion on energy-aware routing challenges	Wireless sensor networks (WSNs)	Energy-aware routing	Discussion on energy-aware routing challenges and solutions in WSNs
[12]	Connected dominating set (CDS)	Packet distribution based on CDS	Simulation-based evaluation	Wireless sensor networks (WSNs)	Energy preservation, network lifetime extension	Proposed model improves network performance metrics, extends network lifetime, and reduces energy consumption
[13]	Machine learning algorithms	5G wireless networks	ML techniques	Wireless networks	Energy efficiency improvement	Overview of ML techniques for improving energy efficiency in wireless networks

Energy Consumption Model:

Energy consumed by a sensor node i during a time period t can be represented as the sum of energy consumed in sensing, processing, transmitting, and receiving data, denoted as $E_i(t)$.

Let E_s be the energy consumed for sensing, E_p for processing, E_t for transmitting, and E_r for receiving. The total energy consumed by a sensor node i can be given by:

$$E_i(t) = E_s + E_p + E_t + E_r$$

Data Transmission Model:

The energy consumed for data transmission depends on factors such as distance, data rate, and transmission power.

One commonly used model is the Friis transmission equation for energy consumption E_t :

$$E_t = \alpha \cdot d^{\beta}$$

where α and β are constants, and d is the distance between nodes.

Machine Learning-Based Optimization:

ML models can be used to optimize energy consumption by predicting optimal transmission routes, adjusting transmission power, or scheduling data transmission based on network conditions.

Let X represent the input features to the ML model, such as distance, data rate, and network traffic.

The ML model predicts an optimal action Y, such as selecting the best transmission route or adjusting transmission power.

The ML model can be represented as a function $f(X) = Y$.

Objective Function:

The objective is to minimize energy consumption while ensuring data delivery and network performance.

Let $\mathcal{L}(E_{\mathrm{i}}(t), Y)$ denote the loss function that penalizes energy consumption based on the predicted action Y.

The objective function to be minimized can be represented as:

$$\min_{Y} \sum_{i=1}^{N} \mathcal{L}(E_i(t), Y)$$

where N is the total number of sensor nodes.

Constraints:

Constraints may include factors such as data rate requirements, transmission range limitations, and quality-of-service (QoS) constraints.

Algorithm Implementation:

The ML model can be trained using supervised learning with labeled data, or reinforcement learning for dynamic optimization. The optimization algorithm updates the model parameters based on feedback from the network performance. This model provides a framework for optimizing energy efficiency in WSNs using ML techniques.

4 ML Techniques in WSNs

The use of machine learning (ML) strategies in wireless sensor networks (WSNs) necessitates the consideration of several factors in order to function well within this environment. In order to meet the limited in resources nature of sensor nodes, machine learning algorithms that are implemented in WSNs need to be lightweight and energy-efficient. This necessitates the development of optimized models and algorithms tailored for low-power devices. Additionally, ML techniques must be robust to handle the inherent noise, uncertainties, and dynamic environments present in WSNs, often caused by factors like signal interference or node failures. Furthermore, distributed and decentralized ML approaches are favored to enable scalability and adaptability in large-scale WSN deployments while minimizing communication overhead. Moreover, ML models should exhibit adaptability and self-learning capabilities to adjust to changing network conditions and optimize performance over time autonomously. Finally, security and privacy concerns are paramount, requiring ML algorithms to integrate robust encryption and authentication mechanisms to protect sensitive data transmitted and processed within WSNs [14].

5 Limitations of Machine Learning in WSNs

Machine learning (ML) in wireless sensor networks (WSNs) faces several limitations. Firstly, the fact that sensor nodes are resource-constrained means that the complexity of machine learning models that may be deployed is restricted. This is because sensor nodes have limited memory, computing capacity, and energy supply. Because of this, it is often necessary to use lightweight methods, which means that model accuracy and performance are sacrificed [15].

Secondly, WSNs generate massive amounts of data, which can overwhelm node capacities and network bandwidth when implementing ML algorithms for real-time decision-making. Transmitting raw data to a central processing unit for analysis can incur significant energy costs, leading to premature node depletion.

Furthermore, the dynamic and harsh environments where WSNs operate introduce challenges for ML models, as they may struggle to generalize effectively across diverse conditions or adapt to changes in the environment.

Lastly, ensuring the security and privacy of data in WSNs poses a significant challenge when implementing ML algorithms, as sensitive information may be exposed to potential adversaries during data transmission or model training phases. Addressing these limitations requires innovative approaches that balance performance with resource constraints and security considerations in WSN deployments.

6 ML Techniques in WSNs Routing

Wireless sensor networks (WSNs) play a crucial role in gathering data from remote or hostile environments. Efficient routing within WSNs is essential for optimizing energy consumption and prolonging network lifetime. Machine learning (ML) techniques offer promising solutions to enhance WSN routing protocols [16].

Reinforcement learning (RL) algorithms such as Q-learning and deep Q-networks can adaptively adjust routing decisions based on environmental feedback, optimizing routes for energy efficiency and data delivery. RL enables nodes to learn from past experiences and make informed routing decisions without relying on predefined routing tables.

Swarm intelligence algorithms like Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) mimic natural behaviors to find optimal routes in WSNs. By leveraging collective intelligence and decentralized decision-making, swarm intelligence algorithms can effectively navigate dynamic network conditions and avoid energy-depleting paths.

Supervised learning techniques including support vector machines (SVM) and neural networks can be utilized for predicting traffic patterns and optimizing routing paths accordingly. These models learn from labeled training data to classify traffic conditions and select the most energy-efficient routes for data transmission.

Additionally, evolutionary algorithms like genetic algorithms (GA) can optimize routing protocols by iteratively evolving candidate solutions through selection, crossover, and mutation operations. GA-based approaches can adapt routing strategies over time to accommodate changing network dynamics and resource constraints.

ML techniques offer diverse approaches to enhance routing in WSNs by enabling adaptive, efficient, and robust routing protocols that can dynamically adjust to varying network conditions, ultimately improving energy efficiency and prolonging network lifetime.

7 Conclusion and Future Work

This paper summarizes the key findings regarding the requirements, limitations, and applications of ML techniques in WSNs, particularly focusing on routing. It emphasizes the potential of ML to address challenges in WSNs and improve routing efficiency.

This review paper has provided a comprehensive overview of the application of machine learning (ML) techniques in wireless sensor networks (WSNs), focusing on routing methodologies. The integration of ML into WSNs has shown promising results in enhancing the overall performance and efficiency of these networks. The exploration of various ML techniques in the context of WSNs routing has shed light on their potential to address the challenges inherent in dynamic and resource-constrained environments.

Future research directions are outlined, highlighting the need for further exploration and development of ML-based solutions tailored to the specific requirements of WSNs.

References

1. Mafuta M, Zennaro M, Bagula A, Ault G, Gombachika H, Chadza T (2013) Successful deployment of a wireless sensor network for precision agriculture in Malawi. Int J Distrib Sens Netw 9(5):150703
2. Lee A, Taylor P, Kalpathy-Cramer J, Tufail A (2017) Machine learning has arrived! Ophthalmology 124(12):1726–1728
3. Pech M, Vrchota J, Bednář J (2021) Predictive maintenance and intelligent sensors in smart factory. Sensors 21(4):1470
4. Ding Q, Zhu R, Liu H, Ma M (2021) An overview of machine learning-based energy-efficient routing algorithms in wireless sensor networks. Electronics 10(13):1539
5. Surenther I, Sridhar KP, Roberts MK (2024) Enhancing data transmission efficiency in wireless sensor networks through machine learning-enabled energy optimization: a grouping model approach. Ain Shams Eng J 102644
6. Bagwari A, Logeshwaran J, Usha K, Kannadasan R, Alsharif MH, Uthansakul P, Uthansakul M (2023) An enhanced energy optimization model for industrial wireless sensor networks using machine learning. IEEE Access
7. Ahmed QW, Garg S, Rai A, Ramachandran M, Jhanjhi NZ, Masud M, Baz M (2022) Ai-based resource allocation techniques in wireless sensor internet of things networks in energy efficiency with data optimization. Electronics 11(13):2071
8. Perumal TSR, Muthumanikandan V, Mohanalakshmi S (2021) Energy efficiency optimization in clustered wireless sensor networks via machine learning algorithms. In: Machine learning and deep learning techniques in wireless and mobile networking systems. CRC Press, pp 59–77
9. Lata S, Mehfuz S (2019) Machine learning based energy efficient wireless sensor network. In: 2019 international conference on power electronics, control and automation (ICPECA). IEEE, pp 1–5
10. Radhika S, Rangarajan P (2021) Fuzzy based sleep scheduling algorithm with machine learning techniques to enhance energy efficiency in wireless sensor networks. Wireless Pers Commun 118:3025–3044

11. Prajapati J, Jain SC (2018) Machine learning techniques and challenges in wireless sensor networks. In: 2018 second international conference on inventive communication and computational technologies (ICICCT). IEEE, pp 233–238
12. Saravana Kumar NM, Suryaprabha E, Hariprasath K (2022) Machine learning based hybrid model for energy efficient secured transmission in wireless sensor networks. J Ambient Intell Humanized Comput 1–16
13. Ali Imran, M., Flávia dos Reis, A., Brante, G., Valente Klaine, P., & Demo Souza, R. (2020). Machine learning in energy efficiency optimization. Mach Learn Future Wireless Commun 105–117
14. Raut AR, Khandait SP (2021) Machine learning algorithms in WSNs and its applications. In: 2021 international conference on computational intelligence and computing applications (ICCICA). IEEE, pp 1–5
15. Alsheikh MA, Lin S, Niyato D, Tan HP (2014) Machine learning in wireless sensor networks: algorithms, strategies, and applications. IEEE Commun Surv Tutorials 16(4):1996–2018
16. Nayak P, Swetha GK, Gupta S, Madhavi K (2021) Routing in wireless sensor networks using machine learning techniques: challenges and opportunities. Measurement 178:108974

Remote Sensing Image Classification Using CNN

Ch. Radhika, Pl. Prranamyaaw, R. Poojitha, T. Archana, and G. Sowmya

Abstract This paper proposes a remote sensing image classification approach that utilizes the ResNet and EfficientNet models. Remote sensing images play a crucial role in various applications, such as land cover mapping and environmental monitoring. The ResNet and EfficientNet models are deep convolution neural networks known for their effectiveness in image classification tasks. In this study, we investigate their performance in the context of remote sensing image classification. We conduct experiments on a benchmark dataset and compare the results with other state-of-the-art methods. Our findings demonstrate the efficacy of the ResNet and EfficientNet models for remote sensing image classification, with both models achieving high accuracy and outperforming existing approaches. The proposed approach shows promise in enhancing the accuracy and efficiency of remote sensing image analysis, paving the way for improved land cover mapping and environmental monitoring applications.

Keywords Remote sensing images · EfficientNet · ResNet · Classification · Deep learning

1 Introduction

Remote sensing has become an invaluable tool for various applications; including land cover mapping, environmental monitoring, and urban planning. Remote sensing images provide a wealth of information about the Earth's surface, but their effective analysis and classification present significant challenges due to the complexity and large volumes of data involved [1]. In recent years, deep learning techniques, particularly convolutional neural networks (CNNs), have demonstrated remarkable success in image classification tasks [2]. Among the various CNN architectures, ResNet (Residual Network) and EfficientNet have emerged as state-of-the-art models for

Ch. Radhika (✉) · Pl. Prranamyaaw · R. Poojitha · T. Archana · G. Sowmya
G. Narayanamma Institute of Technology and Science (for Women), Hyderabad, Telangana, India
e-mail: ch.radhika@gnits.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_53

image classification in the computer vision field [3]. ResNet introduces residual connections that alleviate the vanishing gradient problem and allow the network to be trained deeper, resulting in improved performance. EfficientNet, on the other hand, leverages compound scaling to balance model depth, width, and resolution, achieving excellent accuracy while being computationally efficient.

In this study, we explore the application of ResNet and EfficientNet models for remote sensing image classification. By harnessing the power of these models, we aim to improve the accuracy and efficiency of analyzing remote sensing imagery, enabling more reliable land cover mapping and enhanced environmental monitoring. We conduct experiments on a benchmark dataset and compare the performance of ResNet and EfficientNet with other existing methods in remote sensing image classification.

2 Literature Survey

Regarding the remote sensing image classification Chen et al. [4] aimed to investigate the use of convolutional neural networks (CNNs) for feature extraction and classification of hyper spectral images. Traditional methods often rely on handcrafted features and may not fully capture the complex patterns present in hyper spectral images. Overcome these limitations, the authors proposed a CNN-based approach for feature extraction and classification. Zhu et al. [5] discussed various deep learning architectures and models that have been applied to remote sensing tasks [6]. This includes CNNs for image classification, object detection, and semantic segmentation, as well as recurrent neural networks (RNNs) for tasks involving temporal or sequential data, such as time series analysis and change detection. Liu et al. [7] introduced the importance of multisource remote sensing image classification, which involves integrating data from multiple sensors or sources to improve classification accuracy and robustness.

Xu et al. [9] explained the fundamentals of artificial neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their applications in processing and analyzing remote sensing data [8]. This includes popular CNN-based models such as Alex-Net, VGG-Net, Google-Net, ResNet, and Dense-Net, as well as RNN-based models such as LSTM networks [13]. They discussed different types of multisource remote sensing data including optical imagery, synthetic aperture radar (SAR) imagery, LiDAR data, and hyper spectral image.

3 Methodology

The proposed approach is implemented in the following step-by-step process (Fig. 1).

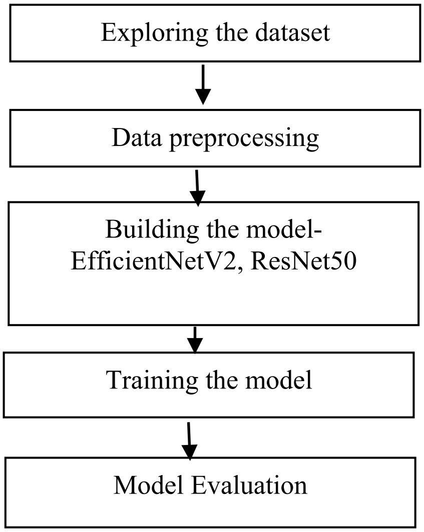


Fig. 1 Flow of proposed system

a. Data collection: Dataset is loaded from tensorflow we can also download it from kaggle.
b. Preprocessing: Images have undergone one-hot encoding of labels, shuffling, repeating, batching, and prefetching.
c. Image Feature Extraction: EfficientNet and ResNet are used to extract the features from images.
d. EfficientNet and ResNet: The model's convolutional layers are typically frozen, meaning their weights are not updated during training. The input remote sensing images are passed through the frozen layers, and the resulting activations are used as feature representations.
e. Detection: After the model has undergone training, the confusion matrix and data visualization are produced from an image as the input.

4 Implementation

4.1 Dataset

The Euro SAT dataset is used for this image classification task. It contains satellite images of different land use and land cover classes, such as residential areas, industrial areas, forests, and rivers [10]. The dataset is loaded using the Tensor Flow Datasets library (Fig. 2).

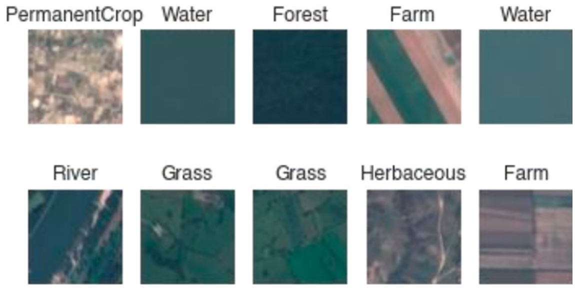


Fig. 2 Area Euro SAT image patch example

4.2 Algorithms

EfficientNet

EfficientNet is a convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth/width/resolution using a compound coefficient. The first model consists of an EfficientNet feature vector layer loaded from Tensor Flow Hub. It is followed by a dense layer with softmax activation for classification. The model is compiled with the Adam optimizer, categorical cross-entropy loss, and additional metrics including accuracy.

ResNet50

ResNet-50 is a convolutional neural network that is 50 layers deep. The second model uses the ResNet50 architecture [3]. It includes a pre-trained ResNet50 base model [11] followed by a flatten layer, a dropout layer, and a dense layer with soft max activation. The model is compiled with the Adam optimizer, categorical cross-entropy loss, and accuracy metric.

4.3 Training Process

Both models are trained using the fit function, which takes the training dataset as input. The training process involves iterating over a specified number of training steps and validation steps [12]. The best model weights are saved using the Model Checkpoint callback, which monitors the validation loss and saves the model with the lowest loss value.

5 Results

The performance of the two models, EfficientNet and ResNet50, on the Eurosat dataset is evaluated using various metrics and visualizations.

Table 1 Results of EfficientNet and ResNet50

Model	Training accuracy (%)	Validation accuracy (%)
EfficientNet	97	96
ResNet50	99	97

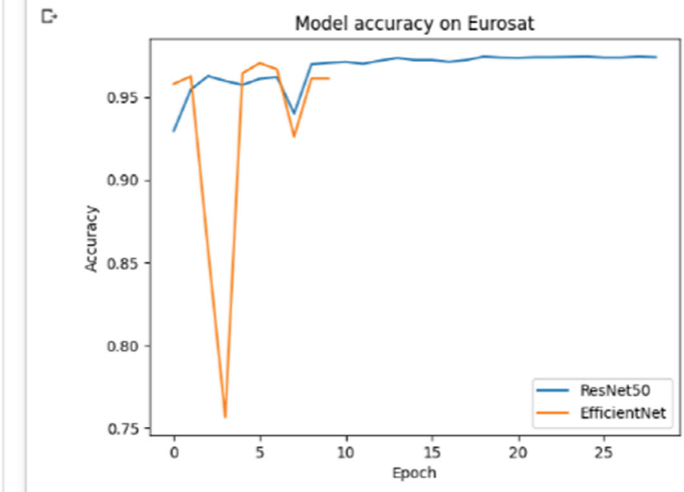


Fig. 3 Measuring of accuracy

Performance Measures

Accuracy measures the proportion of correctly classified samples. These metrics are calculated for both the training and testing datasets (Table 1).

5.1 *Comparison of Models*

The performance of the two models, EfficientNet and ResNet50, is compared based on their accuracy. The results obtained from the evaluation process are analyzed to determine which model performs better on the EuroSAT dataset (Fig. 3).

5.2 *Confusion Matrix Analysis*

A confusion matrix is computed to provide a detailed analysis of the models' performance across different classes. The matrix shows the number of correctly and incorrectly classified samples for each class. To facilitate interpretation, the matrix is normalized to represent percentages. Heat map visualization is created to visualize the confusion matrix, enabling a better understanding of the models' strengths and weaknesses in classifying different land use and land cover categories (Figs. 4 and 5).

The results presented in this section provide insights into the performance of the models and their suitability for the image classification task on the EuroSAT dataset. The analysis of the metrics and visualization aids in understanding the models' strengths, weaknesses, and overall effectiveness in classifying satellite images. The ResNet50 model has more accuracy than EfficientNet model.

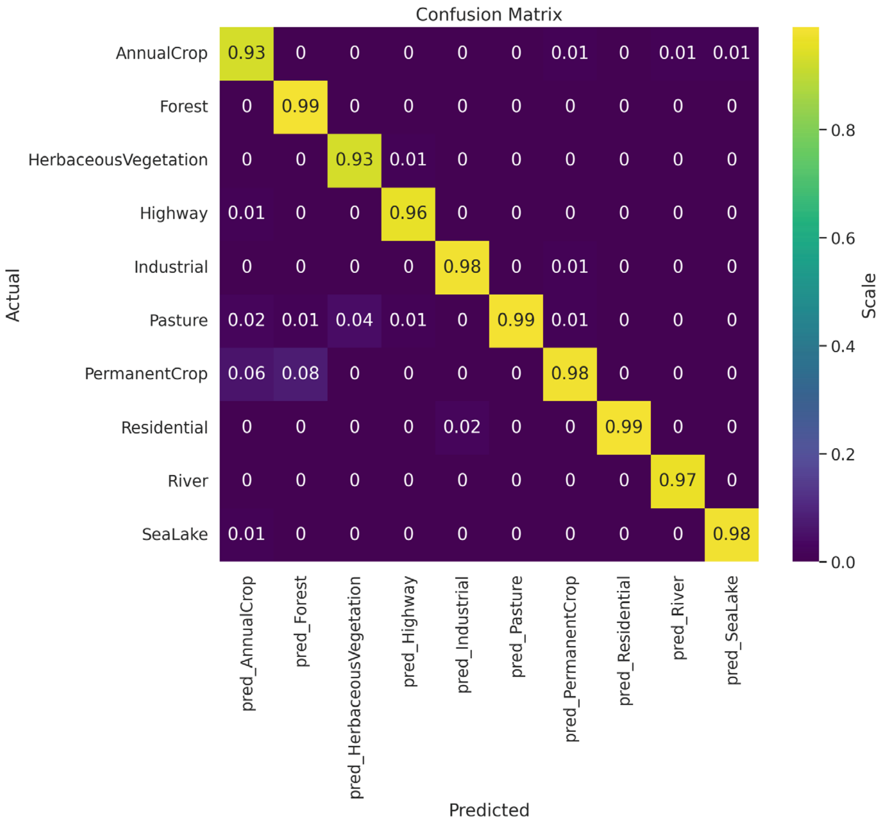


Fig. 4 Confusion matrix of efficient net model

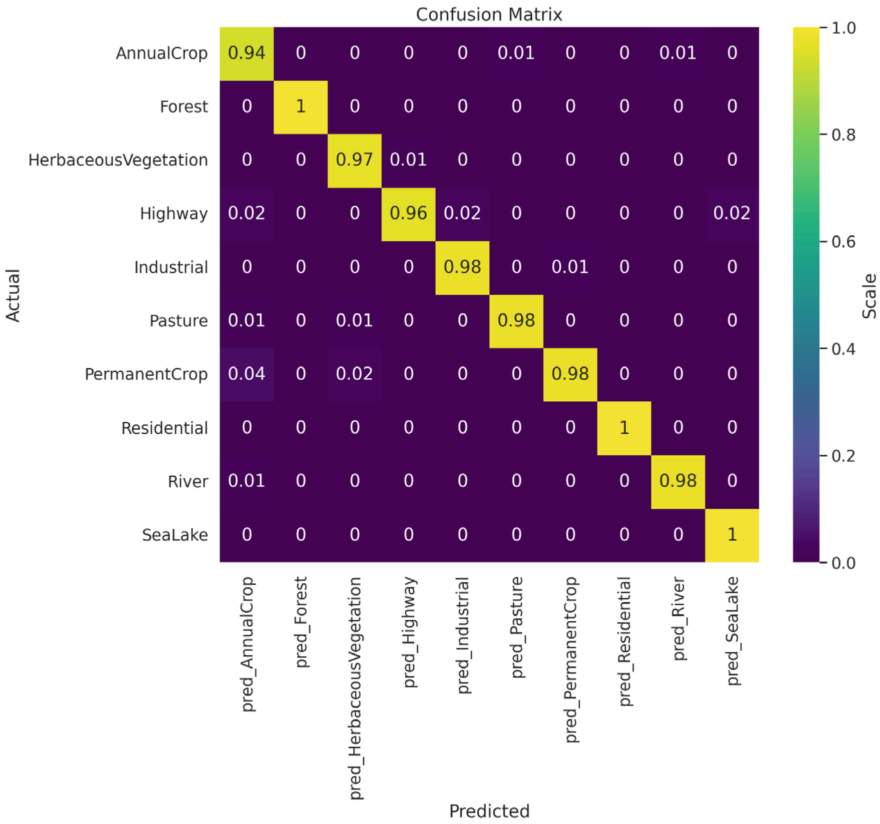


Fig. 5 Confusion matrix of rest net model

6 Conclusion

In conclusion, this study presented a methodology for remote sensing image classification using deep learning techniques. The proposed approach leveraged pre-trained models such as EfficientNet and ResNet50, which have demonstrated excellent performance in image classification tasks. By fine-tuning these models on a large-scale remote sensing dataset (e.g., Eurosat), we achieved promising results in terms of accuracy. The experiments conducted in this study showcased the effectiveness of deep learning models in remote sensing image classification. The utilization of transfer learning and pre-trained models enabled us to leverage the representation power of models trained on large-scale natural image datasets. This transfer of knowledge allowed us to overcome the limitations of limited labeled remote sensing data and achieve competitive performance.

Furthermore, the evaluation of classification results through metrics such as accuracy provided insights into the model' performance across different classes. The visualization of the confusion matrix helped identify areas where the model had difficulty distinguishing between certain classes, highlighting potential challenges in the classification task.

References

1. Helber P, Bischke B, Dengel A, Borth D (2019) Eu roSAT: a Novel dataset and deep learning benchmark for land use and land cover classification
2. Ghamisi P, Plaza A, Plaza J (2021) Deep learning for hyperspectral image classification: an overview. IEEE Geosci Remote Sens Mag 9(1):14–43. https://doi.org/10.1109/MGRS.2020.3016716
3. Radhika C, Nikhila D, Shivani D, Siri Priya B, Reddy B (2023) MultiScale object detection in remote sensing images using deep learning 11(VII). https://doi.org/10.22214/ijraset.2023.54643
4. Chen Y, Lin L, Zheng J (2017) Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. Remote Sens 9(7):673. https://doi.org/10.3390/rs9070673
5. Zhu X, Tuia D, Mou L, Xia GS, Zhang L, Xu F, Fraundorfer F (2017) Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci Remote Sens Mag 5(4):8–36. https://doi.org/10.1109/MGRS.2017.2762307
6. Ma J, Zhan X, Ye Z, Li J (2020) A review of deep learning for object detection and semantic segmentation in remote sensing. ISPRS J Photogramm Remote Sens 162:184–199. https://doi.org/10.1016/j.isprsjprs.2020.02.004
7. Liu H, Zhang C, Ye Q (2021) Deep learning for multi-source remote sensing image classification: a review. ISPRS J Photogramm Remote Sens 173:75–90. https://doi.org/10.1016/j.isprsjprs.2021.01.014
8. Li W, Qin Z, Zhang L, Zhang D (2020) Deep learning for remote sensing image classification: a survey. Remote Sens 12(9):1492. https://doi.org/10.3390/rs12091492
9. Xu H, Du Q, Zhan Y (2021) A comprehensive review of remote sensing image classification methods based on deep learning. Remote Sens 13(3):465. https://doi.org/10.3390/rs13030465
10. Song J, Gao S, Zhu Y, Ma C (2019) A survey of remote sensing image classification based on CNNs. Big Earth Data 3(3):232–254
11. Cheng G, Han J, Lu X (2017) Remote sensing image scene classification: benchmark and state of the art
12. Garcia-Garcia A, Orts-Escolano S (2017) A review on deep learning techniques applied to semantic segmentation
13. Shabbir A (2021) Satellite and scene image classification based on transfer learning and fine tuning of ResNet50

Road Failures Detection System

M. Sunitha, T. Adilakshmi, R. Sateesh Kumar, R. V. Nikhil Sai, and Nikshitha Rapolu

Abstract Pavement distress detection is of great significance to pavement maintenance and management. Roadway problems, including cracks, disintegration that, impair the comfort of road users, damage vehicles, increase emissions, etc. Through paintaking manual surveys, such as visual inspections of the pavement taken by the inspection officer, road deteriorations can be found. Researchers and industry have made quiet progress toward creating and implementing automated road surface monitoring systems to cut costs associated with manual inspections. The Road Failures Detection project is developed by using deep learning-based algorithm YOLO (You Only Look Once) for the detection of pavement defects which starts off with data pre-processing, then training the YOLO model, and then testing the designed model and deploying the model using Streamlit app. Streamlit app allows us to host machine learning applications on the web using only Python code. The proposed approach assists in lowering the time needed for road maintenance and offers detection results of up to eight different types of pavement failures that are reassuringly accurate. The dataset used by the system consists of images, annotations, and image sets.

Keywords Road damage detection · Deep learning · YOLOv5 · Proportional-integral–differential (PID) · Optimization

1 Introduction

Road surface failure detection using general image processing techniques has been extensively researched, and the detection accuracies of these techniques are quite good. However, investigations are frequently limited to determining whether damage has occurred or not. But in a real-world scenario, in order for the road managing officers from a governing organization to inspect and effectively repair such damage, they must properly understand the sort of damage. The dataset that we used in this

M. Sunitha (✉) · T. Adilakshmi · R. Sateesh Kumar · R. V. Nikhil Sai · N. Rapolu
Vasavi College of Engineering, Hyderabad, India
e-mail: m.sunithareddy@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_54

system is composed of a total of 9053 images of roads with different road failures and also without any defects and this dataset includes 15,435 instances of road surface damage. And different weather conditions and lighting situations were utilized to capture these images. Each image includes annotated bounding boxes that indicate the type of damage and the location of the damage. Then, we use the state-of-the-art object detection method i.e., the YOLOv5 model to train our road failures detection model with the described dataset. Finally, using the proposed object detection approach, we identify and demonstrate that the types of damages can be precisely classified into eight different types depending on the coordinates of the observed defect. We build a system that can identify up to eight types of defects and estimate the severity of the pavement distress. We use the deep learning-based algorithm YOLO (You Only Look Once). YOLO is one of the best algorithms to detect objects in an image or video and it uses PyTorch framework. To increase the performance and speed of the system, we use the latest version of the YOLO algorithm (YOLOv5) [1, 2].

Despite having pavement failures detection systems through manual surveys, in recent years, automated video-based road surface monitoring systems have become more common. The use of these automated monitoring solutions for pavement maintenance infrastructures is vital for economic benefit and increase efficiency. Our motivation for this project is to improve road safety and reduce the costs associated with pavement failures and their effects on everybody. By detecting and estimating failures in real time, we can help prevent accidents and minimize the impact of those failures on vehicles. Overall, we have faith that our project has the potential to make a significant difference in road safety and infrastructure maintenance. With the increasing popularity of autonomous vehicles, the need for accurate and real-time road failures detection has increasing importance in the world. We are excited to make our contribution to this and develop a system which can be helpful to everyone, and which can make a noticeable difference in the world [3].

Roadway problems, including cracks, disintegration, impair the comfort of road users, damage vehicles, increase emissions, etc. Road damage can be found by laborious manual surveys, such as visual checks of pavement photographs obtained by the inspection officer. Automation of road surface monitoring systems has become the focus of research to cut expenses associated with manual inspection.

The problem we aim to solve is the detection and estimation of pavement failures using YOLOv5. The YOLOv5 algorithm is used, and an object detection model is built using the pre-trained YOLOv5 model for transfer learning. YOLO stands for You Only Look Once which provides real-time object detection using neural networks. This algorithm employs CNN to detect objects in real-time [4, 5].

First, it can improve road safety by identifying and alerting drivers and road management officers to lane defects on the path. Second, it can reduce the cost of road maintenance by enabling targeted repairs of road failures. Finally, it can provide real-time road failures detection. By using the proposed object detection method, we accurately identify the types of the road defects and categorize it into eight different sorts of failures.

In summary, the problem we aim to solve is the detection and estimation of pavement distresses on roads using YOLOv5, a state-of-the-art object detection algorithm. Our proposed solution has the potential to improve road safety, reduce the cost of road maintenance, and provide real-time road failures detection and estimation of the severity of the detected failures.

2 Related Work

1. **RoadID: A Dedicated Deep Convolutional Neural Network for Multipavement Distress Detection: A Literature Review**

RoadID is a deep learning method that detects problems in images by rotating clustering and control. First, some parameters are extracted from the CNN module and then the reverse model search function is performed, and parameters are found for some loss function and training data [1].

In terms of model structure, feature maps, and modularity estimation, troubleshooting details are the same and different for different types of problems. ResNet has been used as a universal map to record characteristics of road scenes because of this. There is a unique estimate for each type of query that recognizes and reports it. A distinct map serves as the input for each forecasting module.

However, this design enables the model to gather various data, and each prediction module is more concentrated on recognizing various difficulties. On the other hand, when new questions need to be addressed, the general generalization model means that only the new estimate and only little information need to be relearned. This helps to quickly adapt and extend the model for new situations. The Regional Recommender Network (RPN) was adopted and updated according to the forecast [2].

Perform ablation tests to select the best distribution and correlation for distribution and location estimation. When it comes to redundancy, the true and squared error (MSE) combination lessens the issue of inconsistent data and enables the model to further account for complicated processes.

As a result, the search method gives good results in complex and flexible junction box progression. According to the basic research object, the specific model is optimized for surface properties. The location and meaning of the asphalt fall is known. RoadID provides a damage location junction box that can be used to identify damage.

2. **Pavement Distress Detection Using Convolutional Neural Networks (CNN): A Study in Montreal, Canada: Literature Review**

The process consists of four steps: collecting and editing data, creating a deep neural network to learn image properties and patterns. From various problems in the asphalt image, the image has been converted to grayscale because the quality of asphalt sounds is more important than color. This change will also help stabilize and standardize the data. Comparing the histograms of these images is then used to homogenize the distribution of the different images [3, 4].

Different polygon colors are used for borders and test areas. The shape of the grid line is also represented by polygons, but only the connecting lines form narrow polygons, not the entire area. Broken system, on the other hand, is defined in the same way as patches and gutters because it should solve everything.

In this study, various experiments were carried out on the VGG-based CNN model [5, 6]. The deep CNN model requires a lot of time and energy but does not ensure classification accuracy due to the volume of data collected. This leads to the employment of a nine-layer deep neural network with over 4 million parameters.

The neural network of the proposed model has an input size of 150 × 150 pixels, and the image has three convolutional blocks. The first block has a 32-channel convolutional layer and a maximal compression layer, the second block has two convolutional layers. There are three convolutional layers with 64 channels, followed by the maximum compression layer, and the last block with 128 channels, the thickest layer.

This study presents a method to classify the construction problem of the method using the CNN model. Collect data using inexpensive cameras installed in the vehicle. Although a larger dataset was used in this study, more data and better-quality samples can improve accuracy.

3. Deep Learning in Data-Driven Pavement Image Analysis and Automated Distress Detection: A Literature Review

a. Detecting No-Crack Surfaces from Mobile Mapping Images

After the deep learning model is created, it takes around 2.5 ms to analyze only a single image. Deep learning has proven to be more effective and efficient when the total count of images exceeds 7500. One of the object identification jobs that is indicated as a potential future trend is the expansion of deep learning-based detection analysis, particularly with images and severity categorization (high, medium, low) [7].

b. Crack Detection from Low-Cost Smartphone Pavement Images:

The authors generated a million three-channel (RGB) 99 × 99-pixel picture patches from a total of 500 pavement images using a slice technique (consisting of 640,000 samples for training, and 200,000 samples for testing). In order to minimize overfitting during training, the hustler regularization method was used by using the Caffe DL frame between two fully connected layers of the ConvNet built [8].

c. Road Damage Detector: A DL Mobile App for Road Damage Detection Based on Open-Source Smartphone Road Images:

When a smartphone is put on a bus dashboard, an app was created that takes a 600 × 600 pixel road image once every second. The single shot multi-box detector (SSD) with inception V2 was chosen for this study along with SSD using MobileNet DL fabrics as they are considered effective network architectures and prosecutions for mobile vision operations as the ultimate thing was a smartphone app-based on DL for road damage discovery [9, 10].

d. Sealed Crack Detection with Transfer Learning and Fine-Tuning:

The first convolutional caste of the ImageNet-trained DCNN model remains unchanged throughout training, the two-step pretype procedure and executed in the Caffe DL frame initially transfers the general knowledge and then using the target pavement images, the parameters of the other convolutional layers are fine-tuned. Following pre-classification of the pavement images into different types, and background regions, the cracks and sealed cracks are prized using a block-wise thresholding approach and tensor voting-based wind detection [11, 12].

3 Proposed Work

With the introduction of object detection techniques, people have started launching challenges and research in this domain. The proposed method uses a state-of-the-art technique, YOLOv5 to build a model which will detect multiple types of road damages such as potholes, alligator cracks, etc.

The proposed methodology for road failures detection using YOLOv5 involves the following steps:

Dataset collection and preparation: A dataset in which images are taken in a wide range of weather and lighting circumstances is used to train the YOLOv5 model. The dataset is diverse and contains of various road failures types, sizes, and lighting conditions that the system is expected to encounter in the real world.
Image annotation: The images in the dataset are annotated with bounding boxes around the road failure. Each bounding box is associated with a class label and probability.
Model training: The YOLOv5 model is trained using the annotated dataset. The model is initialized with weights pretrained on a substantial image dataset, followed by fine-tuning.
Model evaluation: The trained model is evaluated on a different test set to measure its detection performance for various road defects. The performance metrics include precision, recall, F1 score, and mean average precision (mAP).
Post-processing: Redundant distress detections are eliminated via non-maximum suppression and refine the final set of road distress detections.
Road failure dimension estimation: Once road failures are detected, their coordinates are estimated.

However, like any machine learning-based system, it requires a large, annotated dataset, and careful parameter tuning to achieve optimal performance.

Then, the YOLOv5 model is trained using the processed dataset. The model is trained for 10 epochs for high accuracy. Trained weights are obtained which are required to re-use the trained model.

The trained network as shown in Fig. 1 is then tested which can be applied to images and videos.

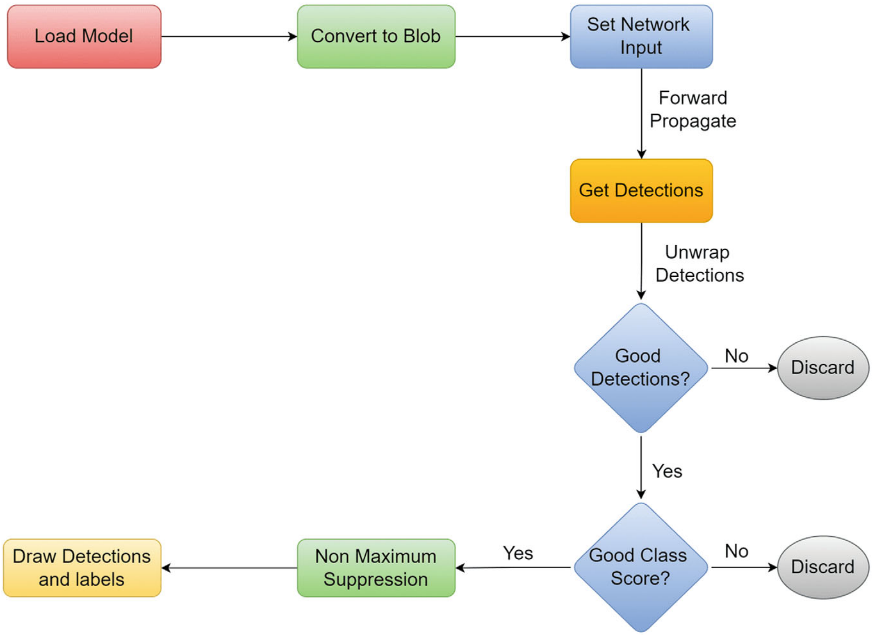


Fig. 1 Flow diagram of YOLOv5 with DNN workflow

We expose the model through a webapp, Streamlit application so that they can be used from anywhere. We can just upload an image of any road, and it gives the areas where the road is damaged. We can do exactly that using streamlit. Streamlit is an open-source Python library that allows users to build interactive web applications. It provides a simple and intuitive way to build and deploy web apps with minimal code. Overall, Streamlit's simplicity, ease of use, and integration with Python libraries make it an excellent choice for quickly creating and deploying interactive web applications.

4 Result

A dataset in which images are taken in a wide range of weather and lighting circumstances is used to train the YOLOv5 model. The dataset is diverse and contains of various road failure types, sizes, and lighting conditions that the system is expected to encounter in the real world. The images in the dataset are annotated with bounding boxes around the road failure. Each bounding box is associated with a class label and probability.

The proposed model is compared to two well-known algorithms, namely YOLOv3 and YOLOv4, using metrics like precision, mAP, recall, PC speed (FPS), F-measure as shown in the Figs. 2, 3, 4, 5, and 6.

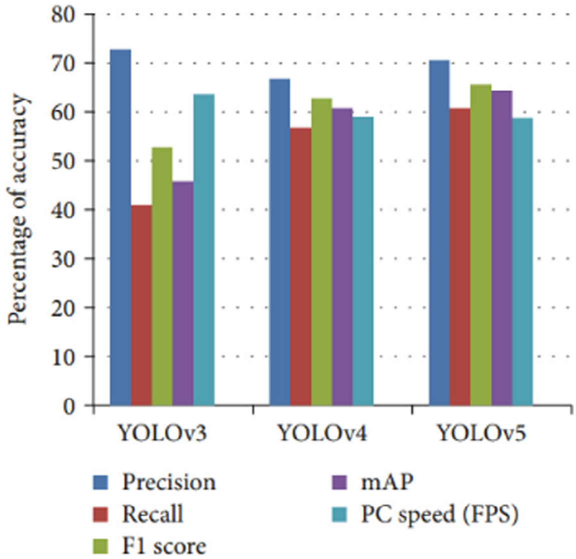


Fig. 2 Comparative analysis of YOLO

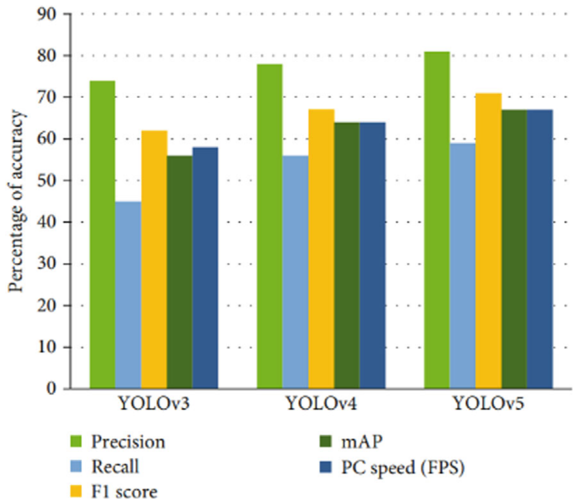


Fig. 3 Comparative analysis of YOLO models of rural road dataset and models of urban road dataset

As noticed from Figs. 2, 3, 4, 5, 6 and 7, accuracy is very high when the YOLOv5 algorithm is used because all the performance metrics show high values with YOLOv5 algorithm. Hence, our approach is simple and results in high efficiency and accurate detection and classification of road failures.

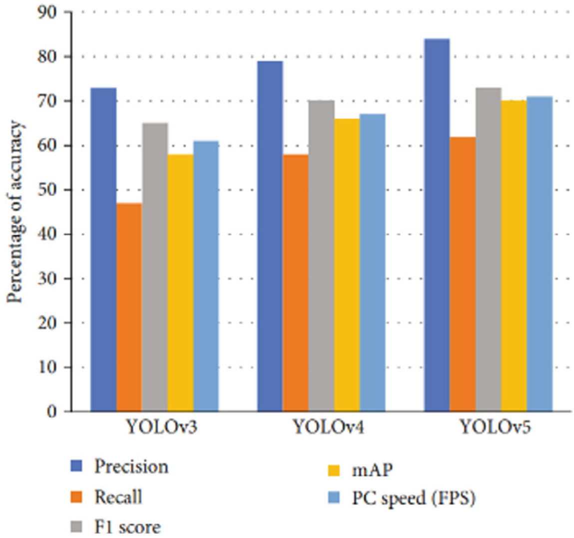


Fig. 4 Comparative analysis of YOLO models of highway dataset on the custom dataset

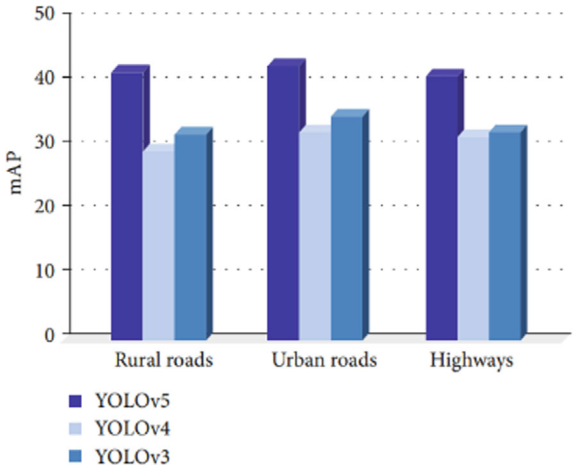
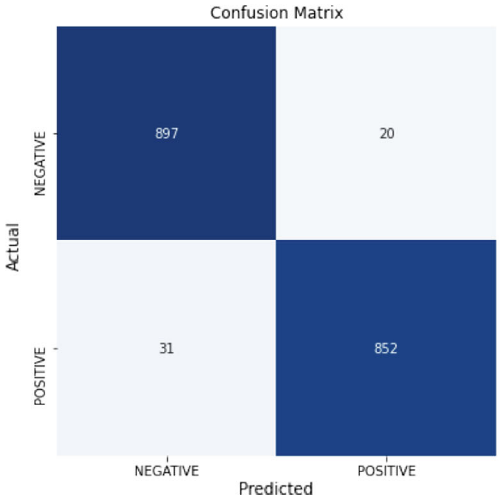


Fig. 5 Comparison of YOLO models

Model	Precision	Recall	F1-score	mAP@0.5 (%)	Inference time (ms)
SSD-Mobilenetv2	0.42	0.56	0.479	47.4	7
YOLOv1	0.82	0.69	0.74	79.55	340
YOLOv2	0.81	0.76	0.78	81.21	33.7
YOLOv3	0.77	0.78	0.78	83.60	70.57
Tiny-YOLOv4	0.76	0.75	0.76	80.04	4.86
YOLOv4	0.81	0.83	0.82	85.48	52.51
YOLOv5	0.93	0.83	0.87	95.00	10

Fig. 6 Performance evaluation of our model (YOLOv5) and other models



Classification Report:

	precision	recall	f1-score	support
NEGATIVE	0.97	0.98	0.97	917
POSITIVE	0.98	0.96	0.97	883
accuracy			0.97	1800
macro avg	0.97	0.97	0.97	1800
weighted avg	0.97	0.97	0.97	1800

Fig. 7 Confusion matrix and classification report

5 Conclusion

Road failures detection is of high importance in road maintenance, which generally involves a lot of manual labor. In order to detect road damages fast, we employed a deep learning model to analyze road photos in our system. This may assist avoid future pavement failures in addition to aiding in the rehabilitation of current road breakdowns.

Overall, this work shows that a deep neural network system is applicable to automatically detect and classify various types of pavement deterioration in road facilities. This model, which is designed to locate and identify pavement damages, is based on the YOLOv5 object detection model framework. The dimensions of the calculated pavement failures would be useful in determining the extent of the damage to the road as well as the cost of repair. Thus, management and inspection can be done remotely. In a real-world setting, our system is capable of detecting complicated pavement defects.

References

1. Xu Y, Sun F, Wang L (2022) YOLOv5-PD: a model for common asphalt pavement defects detection. J Sens 2:1–12
2. Murthy JS, Siddesh GM, Lai W-C, Parameshachari BD, Patil S, Hemalatha KL (2022) Object detect: a real-time object detection framework for advanced driver assistant systems using YOLOv5. Wireless Commun Mobile Comput 1:1–10
3. Wan F, Sun C, He H, Lei G, Xu L, Xiao T (2022) YOLO-LRDD: a lightweight method for road damage detection based on improved YOLOv5s. EURASIP J Adv Sig Proc 1, Article number: 98
4. Shaikh SA, Chopade JJ, Sardey MP (2023) Real-time multi-object detection using enhanced Yolov5-7S on Multi-GPU for high-resolution video. Int J Image Graphics
5. Srivastava S, Divekar AV, Anilkumar C, Naik I, Kulkarni V, Pattabiraman V (2021) Comparative analysis of deep learning image detection algorithms. J Big Data 8, Article number
6. Guo G, Zhang Z (2022) Road damage detection algorithm for improved YOLOv5. Sci Rep 12, Article number: 15523
7. Zakaria M, Karaaslan E, Necati Catbas F (2022) Advanced bridge visual inspection using real-time machine learning in edge devices. Adv Bridge Eng 3, Article number: 27
8. Li Y, Liu C, Shen Y, Cao J (2021) RoadID: a dedicated deep convolutional neural network for multipavement distress detection. J Transp Eng Part B Pave 147(4)
9. Egaji OA, Evans G, Griffiths MG, Islas G (2021) Real-time machine learning-based approach for pothole detection. Exp Syst Appl 184:115562. ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2021.115562
10. Satti SK, Suganya Devi K, Dhar P, Srinivasan P (2021) A machine learning approach for detecting and tracking road boundary lanes. ICT Exp 7(1):99–103. ISSN 2405-9595, https://doi.org/10.1016/j.icte.2020.07.007
11. Zhang C, Nateghinia E, Miranda-Moreno LF, Sun L (2021) Pavement distress detection using convolutional neural network (CNN): a case study in Montreal, Canada. Int J Transp Sci Technol
12. Wang S, Chen X, Dong Q (2023) Detection of asphalt pavement cracks based on vision transformer improved YOLO V5

Resume Classification by Using Natural Language Processing

M. Sunitha, T. Adilakshmi, and Sadia Firdous

Abstract The purpose of this project is to develop a resume classification system using natural language processing (NLP) methods. The system will be designed to automatically categorize resumes based on their content and extract important information such as education, work experience, and skills. By utilizing NLP techniques, the system will be able to identify relevant keywords and phrases, as well as understand the context and meaning of the text. This will enable the system to accurately categorize and analyze resumes, thus providing a valuable tool for recruiters and job seekers alike. Throughout the development process, we will prioritize achieving high levels of accuracy and efficiency in order to produce a reliable and effective resume classification system.

Keywords Natural language processing (NLP) · TF-IDF · Count vectorizer · *k*-nearest neighbor · Recommendation systems · Feature extraction

1 Introduction

Recruitment systems reliant on the Internet have swiftly gained acceptance among recruiters in recent times. The rapid growth of the Internet has correspondingly increased the accessibility of information online. As a result, information has become easily reachable. However, this has also led to an overwhelming surplus of information, demanding effective management approaches. Furthermore, in countries like India, there exists a lot of competition. Consequently, handpicking suitable candidates from this substantial application pool has turned into a challenging task for employers. This procedure demands significant resources, including human effort, and involves meticulous work to single out the most fitting candidate for subsequent consideration [1, 2].

M. Sunitha (✉) · T. Adilakshmi · S. Firdous
Vasavi College of Engineering, Hyderabad, India
e-mail: m.sunithareddy@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_55

To counter these challenges, internet-based recruitment systems have arisen as a solution to streamline the recruitment procedure. These systems utilize the Internet to automate various stages of recruitment, starting with posting job vacancies and proceeding to assessing applicant qualifications and easing communication [3]. This transition to digital processes has resulted in increased efficiency and reduced administrative load for both recruiters and applicants.

The goal of the resume classification using NLP project is to develop a system that can automatically categorize resumes based on their content and extract important information such as education, work experience, and skills using natural language processing (NLP) techniques [2]. The system will utilize NLP to identify relevant keywords and phrases, understand the context and meaning of the text, and accurately categorize and analyze resumes.

The system will be designed to be a valuable tool for both recruiters and job seekers. Recruiters will be able to use the system to quickly identify and sort through resumes based on specific criteria, while job seekers can improve their chances of getting noticed by ensuring that their resumes are categorized accurately [4].

Various machine learning methods can be applied to resume classification using NLP. Common ones include

1. Support Vector Machines (SVM): Effective for text classification, SVMs are suitable for binary categorization and high-dimensional features. Trained on labeled resumes, they predict categories for new ones.
2. Naive Bayes: Based on the Bayes theorem, it treats document features as independent. Naive Bayes is computationally efficient, especially with limited data, and works well for resume classification.
3. Deep Learning: This involves training neural networks to recognize data patterns. Approaches like CNNs and RNNs have shown promise in NLP tasks, like resume classification, but demand ample labeled data and computation.
4. Random Forest: An ensemble algorithm, it constructs multiple decision trees and combines their predictions for a final result. By training on labeled resumes, it predicts categories for new ones.

These are just some methods for resume classification using NLP. Algorithm choice depends on project specifics such as dataset size, category count, and desired accuracy.

When it comes to ranking CVs or resumes, the traditional approach has been to assign a numerical score to each candidate based on how well their skills and experience match the job requirements. The candidate with the highest score is then considered the best fit for the position. However, this approach has several limitations, which have led to the development of more sophisticated algorithms to rank CVs. One such approach is combining the results of multiple algorithms like KNN, TF-IDF, and Count Vectorizer instead of simply finding the maximum score.

2 Related Work

A study by Guo S., Alamudun F., Hammond T proposed a personalized resume-job matching system that uses both structured and unstructured data from resumes and job postings [5]. The system utilizes natural language processing techniques to extract relevant information from unstructured data and a recommendation algorithm to match candidates with job postings based on their qualifications and preferences. The authors conducted experiments on a dataset of job postings and resumes and reported promising results in terms of accuracy and efficiency.

Gopalakrishna and Vijayraghavan [6] proposed a resume classification system using natural language processing (NLP) and machine learning techniques. The system utilizes various NLP techniques, including part-of-speech tagging, named entity recognition, and sentiment analysis, to extract and classify relevant information from resumes. The study reported an accuracy of 90.2%.

Ramos J., "Using TF-IDF to determine word relevance in document queries", Proceedings of the First Instructional Conference on Machine Learning, New Jersey, U.S.A., 2003 [7], this paper provides an example of using TF-IDF to retrieve relevant documents from a corpus of scientific articles. The author compares the performance of TF-IDF with other weighting schemes, such as simple term frequency and binary weighting, and shows that TF-IDF outperforms them in terms of precision and recall. Overall, the paper demonstrates the effectiveness of TF-IDF in improving the accuracy of information retrieval systems and has become a standard technique used in many natural language processing applications, including search engines, text classification, and information extraction. The paper by Uysal and Gunal investigates the impact of pre-processing techniques on text classification tasks [8]. They compare the performance of four different preprocessing techniques: stemming, stop-word removal, lowercasing, and n-gram tokenization, on two different text classification tasks. They found that different preprocessing techniques have different impacts on classification accuracy depending on the specific task and the type of classifier used. In general, they found that stemming and stop-word removal is useful for text classification tasks, while lowercasing and n-gram tokenization may not always improve classification accuracy. The study highlights the importance of choosing appropriate pre-9 processing techniques for text classification tasks and suggests that there is no single best preprocessing technique that works well for all text classification tasks. The studies done also have shown that the approaches used for resume classification have faced two main problems—lower accuracy and poor performance comparison. Furthermore, only a few machine learning models were employed, and accuracy was the only measure of performance used [9]. In addition, the techniques used for feature extraction and representation were not explored to address the problem of low accuracy. One of the earliest systems for resume classification and matching was developed in the 1990s by the US Department of Labor. This system, called the O*NET system, was designed to provide a standardized language for describing job skills and requirements. It included a database of job titles and descriptions, along with associated skill and knowledge requirements. Another early system was developed by the US Department of Defense in the 1980s and was called the Military Occupational Classification and Structure (MOCAS) system. This system was designed to help match military personnel to civilian jobs based on their training and skills. It used a hierarchical structure to categorize jobs and skills, which was later adopted by the O*NET system. In the late 1990s and early 2000s, several commercial resume screening systems were developed. These systems were designed to help recruiters and hiring managers filter large numbers of resumes more efficiently. One such system was developed by a company called Hire.com and used natural language processing to extract relevant information from resumes and match them to job descriptions.

3 Proposed Work

3.1 Data Preprocessing

Data Processing is the data processing component is responsible for reading and processing the resumes that are uploaded by the users. The system uses the Python libraries Pandas and NLTK for processing the data. The resumes are first converted to a Pandas data frame and then preprocessed to remove stop words, special characters, and numbers. Then, named entity recognition, keyword extraction, and sentiment analysis techniques are applied to extract important information such as education, work experience, and skills. The preprocessed data are then saved in a database for further processing.

Data redundance is handling duplicate CVs are mainly to maintain data quality and avoid redundancies. When a database contains multiple copies of the same CV, it can lead to confusion and inaccuracies in data analysis. It can also waste valuable storage space and processing time. The procedure for handling duplicate CVs involves identifying CVs with identical content and removing all but one copy. In the provided code, the duplicates are identified by comparing the content of each CV using the 'distinct()' method. If any duplicate is found, the system retains only the first instance and deletes the rest. This is done using the 'delete()' method, which removes the record from the database. The system achieves this by using an SQL query to retrieve all the duplicate CVs and their respective IDs, and then deleting all but the first instance using the 'delete()' method. The result is a database with only unique CVs, which ensures data accuracy and efficient data analysis.

3.2 Ranking Methodology and Feature Extraction

We would be using Count Vectorizer, TF-IDF, KNN, and cosine similarity in a series of three combinations as shown in Fig. 1.

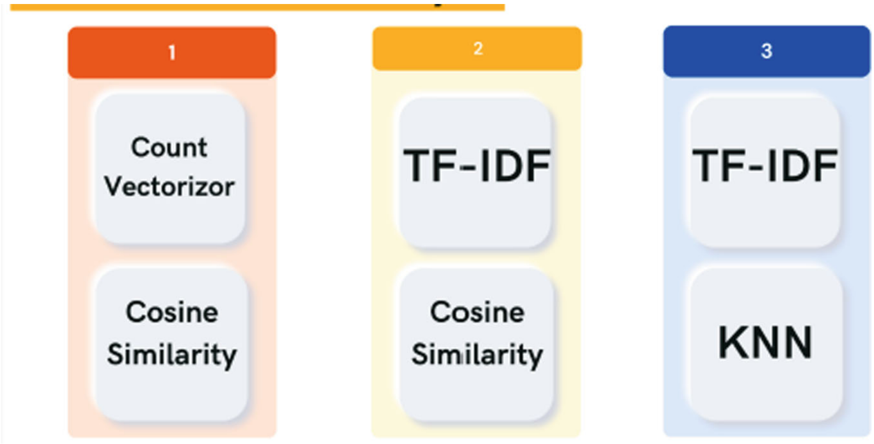


Fig. 1 Ranking system combinations

To combine the scores from the different techniques, a weighted average approach is used. First, the scores from each technique (KNN, TF-IDF, and CountVectorizer) are normalized using MinMaxScaler. Then, each score is multiplied by a weight factor (0.33) and added together to get the final score. The weight factor ensures that each technique contributes equally to the final score [10].

There are several reasons why combining the results of multiple algorithms may be a better approach for ranking CVs. Firstly, different algorithms have different strengths and weaknesses. Some algorithms may be better at identifying candidates with specific skills, while others may be better at identifying candidates with relevant experience. By combining the results of multiple algorithms, we can leverage the strengths of each algorithm while mitigating their weaknesses.

Secondly, combining the results of multiple algorithms can improve the overall accuracy of the ranking process. No single algorithm is perfect, and each algorithm may make different types of errors. By combining the results of multiple algorithms, we can reduce the likelihood of making a significant error and increase the confidence of the final ranking [11].

Thirdly, combining the results of multiple algorithms can provide a more comprehensive picture of each candidate. Each algorithm may consider different aspects of a candidate's skills and experience, and by combining the results, we can get a more complete picture of each candidate. This can be particularly useful for positions that require a broad range of skills or where there is no single 'right' answer.

Lastly, combining the results of multiple algorithms can help to reduce bias in the ranking process. Different algorithms may be biased toward different types of candidates, such as candidates from certain demographic groups or with certain educational backgrounds. By combining the results of multiple algorithms, we can reduce the impact of any individual algorithm's bias and ensure a more fair and equitable ranking process.

4 Result

The application was provided with job openings and a certain set of candidate resumes. When used by the recruiters, we could notice that the job description had a good match to the candidates recommended by the system. The recruiter would be able to find the resumes in the database section of the webpage and view them using the CV id of the resume. Similarly, candidates when they input their resume into the portal would be able to see a list of job openings suitable.

To understand about the results obtained, Fig. 6 shows the result of score given by various algorithms and the final score that marks the final ranking of jobs in case of candidate obtained. It shows the results after taking an average of all three algorithms for better accuracy of results as shown in Figs. 2, 3, 4, 5 and 6.

5 Conclusion

This paperpresents a powerful solution for recruiters and job seekers alike, offering several distinctive features that set it apart from existing systems. It leverages advanced natural language processing (NLP) techniques to analyze resumes and job descriptions, providing highly accurate and personalized job recommendations (Fig. 7).

One of the primary differentiators of this is its utilization of a hybrid recommendation system. By combining the results of three distinct algorithms—TF-IDF, K-nearest neighbors (KNN), and count vectorization—the application provides a more robust and comprehensive approach to matching candidates with job openings. This approach significantly enhances the quality of recommendations, ensuring a higher likelihood of successful job matches.

Fig. 2 Recruiter's page

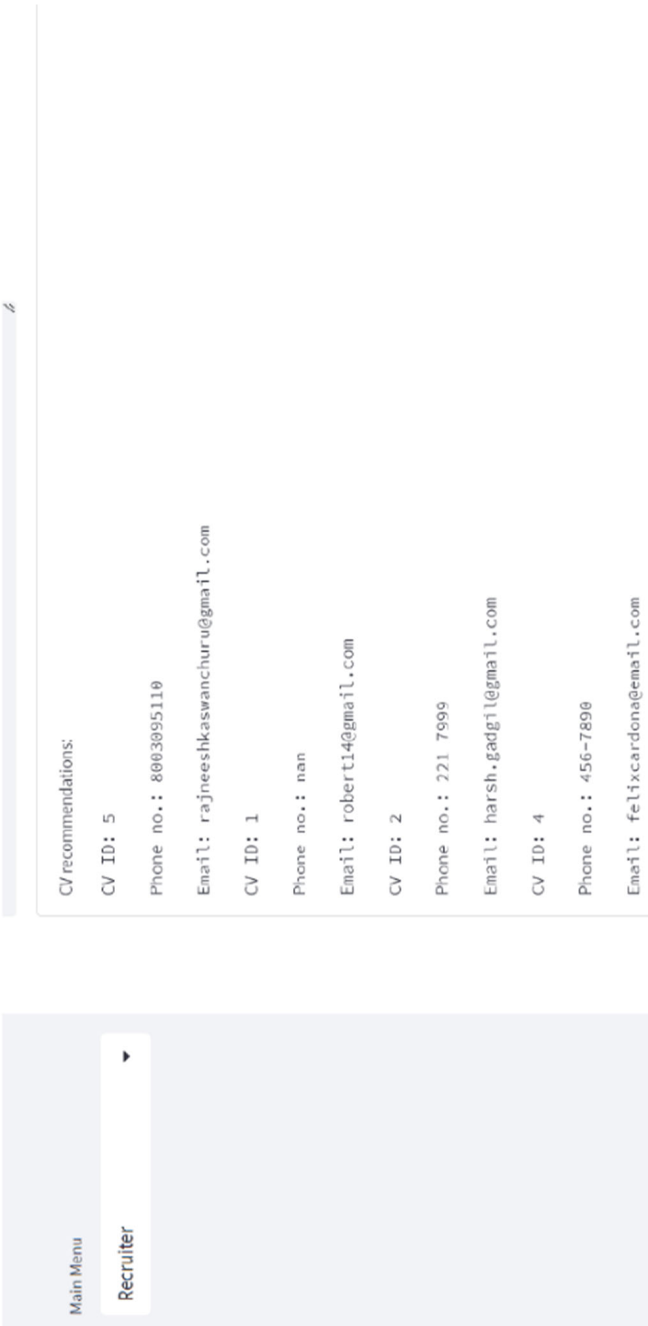


Fig. 3 Resumes uploaded by candidates

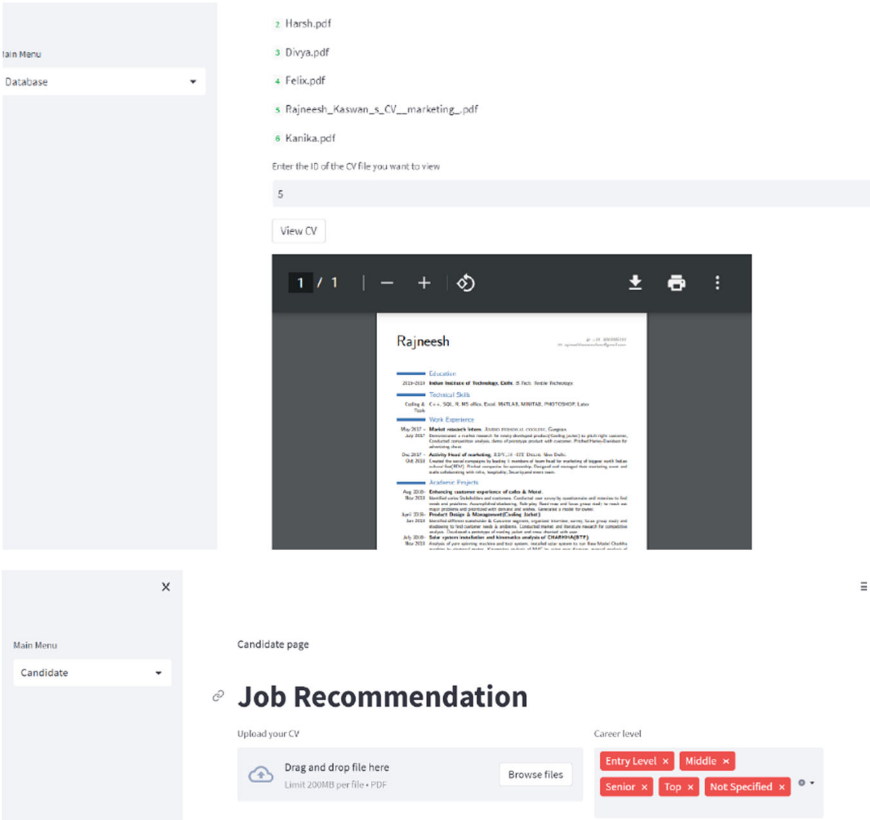
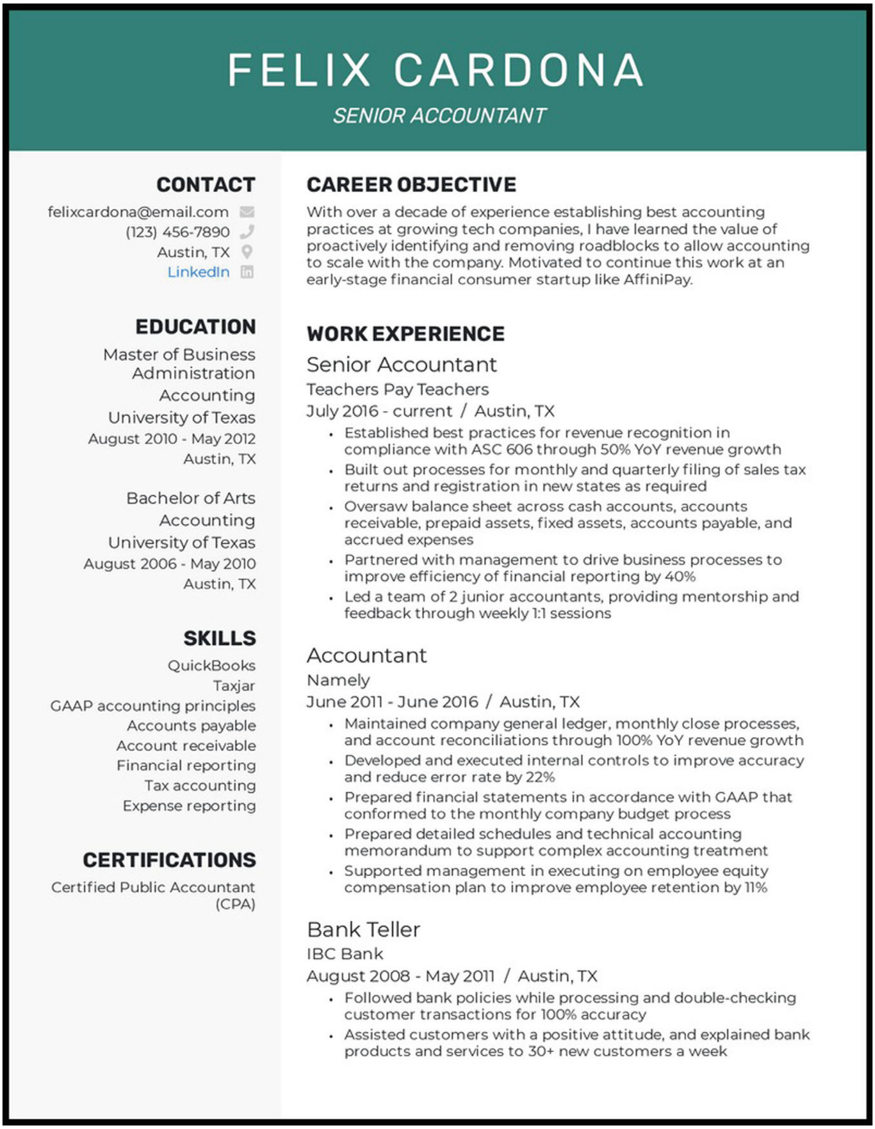


Fig. 4 Candidate's page



FELIX CARDONA

SENIOR ACCOUNTANT

CONTACT

felixcardona@email.com
(123) 456-7890
Austin, TX
LinkedIn

EDUCATION

Master of Business Administration
Accounting
University of Texas
August 2010 - May 2012
Austin, TX

Bachelor of Arts
Accounting
University of Texas
August 2006 - May 2010
Austin, TX

SKILLS

QuickBooks
Taxjar
GAAP accounting principles
Accounts payable
Account receivable
Financial reporting
Tax accounting
Expense reporting

CERTIFICATIONS

Certified Public Accountant (CPA)

CAREER OBJECTIVE

With over a decade of experience establishing best accounting practices at growing tech companies, I have learned the value of proactively identifying and removing roadblocks to allow accounting to scale with the company. Motivated to continue this work at an early-stage financial consumer startup like AffiniPay.

WORK EXPERIENCE

Senior Accountant
Teachers Pay Teachers
July 2016 - current / Austin, TX

- Established best practices for revenue recognition in compliance with ASC 606 through 50% YoY revenue growth
- Built out processes for monthly and quarterly filing of sales tax returns and registration in new states as required
- Oversaw balance sheet across cash accounts, accounts receivable, prepaid assets, fixed assets, accounts payable, and accrued expenses
- Partnered with management to drive business processes to improve efficiency of financial reporting by 40%
- Led a team of 2 junior accountants, providing mentorship and feedback through weekly 1:1 sessions

Accountant
Namely
June 2011 - June 2016 / Austin, TX

- Maintained company general ledger, monthly close processes, and account reconciliations through 100% YoY revenue growth
- Developed and executed internal controls to improve accuracy and reduce error rate by 22%
- Prepared financial statements in accordance with GAAP that conformed to the monthly company budget process
- Prepared detailed schedules and technical accounting memorandum to support complex accounting treatment
- Supported management in executing on employee equity compensation plan to improve employee retention by 11%

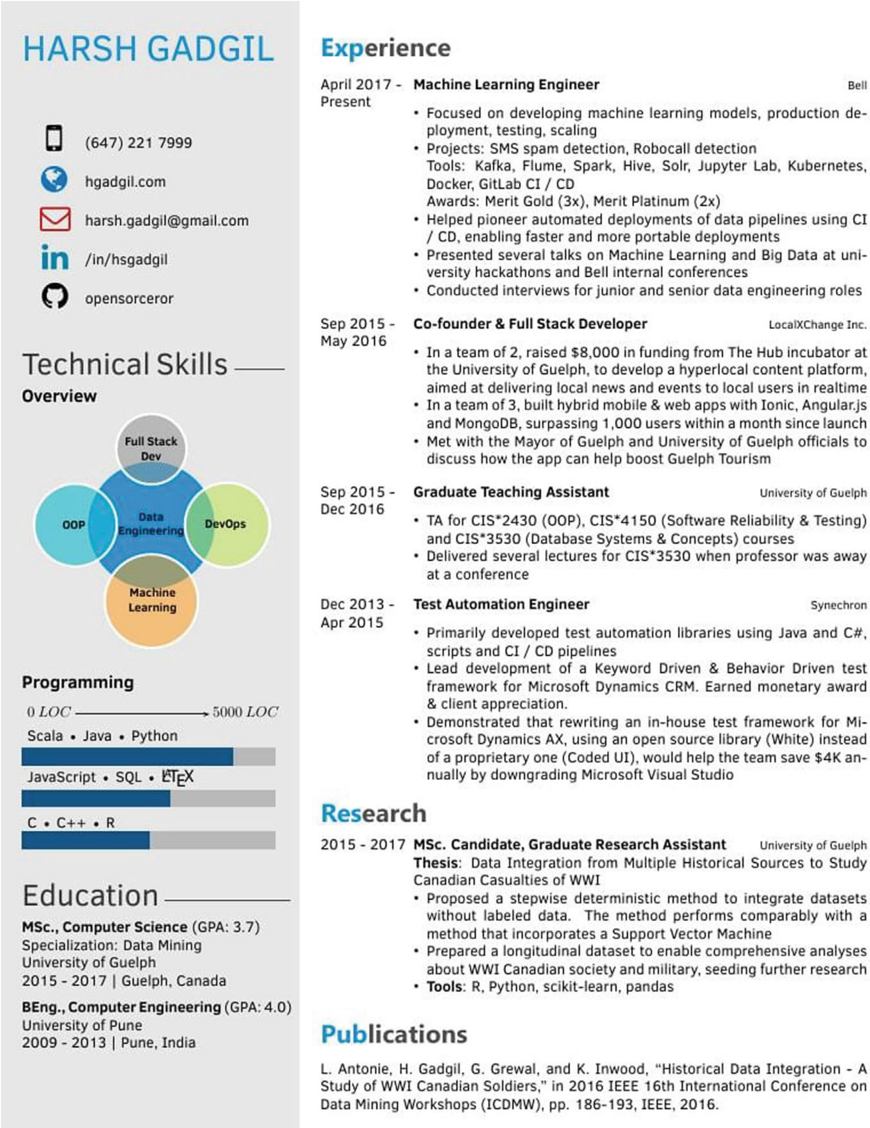
Bank Teller
IBC Bank
August 2008 - May 2011 / Austin, TX

- Followed bank policies while processing and double-checking customer transactions for 100% accuracy
- Assisted customers with a positive attitude, and explained bank products and services to 30+ new customers a week

Fig. 5 Resume uploaded by candidate

	JobID	Business Title	Job Category	Salary	Level
36	333625	Office Administrator	Administration & Human Resources Policy, Research & Analysis	44142	1
24	346116	Claiming Analyst, Bureau of Budget and Revenue	Finance, Accounting, & Procurement	46747	1
41	346120	Claiming Analyst, Bureau of Budget and Revenue	Finance, Accounting, & Procurement	46747	1
0	347128	Revenue Assistant, Bureau of Budget and Revenue	Finance, Accounting, & Procurement	46747	1

Fig. 6 Jobs recommended by application



HARSH GADGIL

(647) 221 7999
hgadgil.com
harsh.gadgil@gmail.com
/in/hsgadgil
opensorceror

Technical Skills

Overview

Programming

0 LOC → 5000 LOC

Scala • Java • Python

JavaScript • SQL • LaTeX

C • C++ • R

Education

MSc., Computer Science (GPA: 3.7)
Specialization: Data Mining
University of Guelph
2015 - 2017 | Guelph, Canada

BEng., Computer Engineering (GPA: 4.0)
University of Pune
2009 - 2013 | Pune, India

Experience

April 2017 - Present **Machine Learning Engineer** Bell

- Focused on developing machine learning models, production deployment, testing, scaling
- Projects: SMS spam detection, Robocall detection
  Tools: Kafka, Flume, Spark, Hive, Solr, Jupyter Lab, Kubernetes, Docker, GitLab CI / CD
  Awards: Merit Gold (3x), Merit Platinum (2x)
- Helped pioneer automated deployments of data pipelines using CI / CD, enabling faster and more portable deployments
- Presented several talks on Machine Learning and Big Data at university hackathons and Bell internal conferences
- Conducted interviews for junior and senior data engineering roles

Sep 2015 - May 2016 **Co-founder & Full Stack Developer** LocalXChange Inc.

- In a team of 2, raised $8,000 in funding from The Hub incubator at the University of Guelph, to develop a hyperlocal content platform, aimed at delivering local news and events to local users in realtime
- In a team of 3, built hybrid mobile & web apps with Ionic, Angular.js and MongoDB, surpassing 1,000 users within a month since launch
- Met with the Mayor of Guelph and University of Guelph officials to discuss how the app can help boost Guelph Tourism

Sep 2015 - Dec 2016 **Graduate Teaching Assistant** University of Guelph

- TA for CIS*2430 (OOP), CIS*4150 (Software Reliability & Testing) and CIS*3530 (Database Systems & Concepts) courses
- Delivered several lectures for CIS*3530 when professor was away at a conference

Dec 2013 - Apr 2015 **Test Automation Engineer** Synechron

- Primarily developed test automation libraries using Java and C#, scripts and CI / CD pipelines
- Lead development of a Keyword Driven & Behavior Driven test framework for Microsoft Dynamics CRM. Earned monetary award & client appreciation.
- Demonstrated that rewriting an in-house test framework for Microsoft Dynamics AX, using an open source library (White) instead of a proprietary one (Coded UI), would help the team save $4K annually by downgrading Microsoft Visual Studio

Research

2015 - 2017 **MSc. Candidate, Graduate Research Assistant** University of Guelph
Thesis: Data Integration from Multiple Historical Sources to Study Canadian Casualties of WWI

- Proposed a stepwise deterministic method to integrate datasets without labeled data. The method performs comparably with a method that incorporates a Support Vector Machine
- Prepared a longitudinal dataset to enable comprehensive analyses about WWI Canadian society and military, seeding further research
- **Tools**: R, Python, scikit-learn, pandas

Publications

L. Antonie, H. Gadgil, G. Grewal, and K. Inwood, "Historical Data Integration - A Study of WWI Canadian Soldiers," in 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), pp. 186-193, IEEE, 2016.

Fig. 7 Resume uploaded by candidate ranking calculations

```
final2.head(10)
```

	JobID	title	KNN	TF-IDF	CV	Final
0	641	Senior Machine Learning Engineer/Expert (Plann...	0.333333	0.333333	0.189947	0.856614
3	7357	Associate Data Scientist/ Data Scientist - Top...	0.27341	0.275167	0.278731	0.827308
7	7259	Data Insights Analyst (Big Data- Machine Learn...	0.249845	0.252081	0.325352	0.827278
2	3942	Software Engineer Trainee (AI / GIS)	0.275054	0.276773	0.248643	0.80047
10	857	Associate Data Scientist	0.214888	0.217617	0.310602	0.743106
8	7312	Artificial Intelligence (AI) R&D Engineer	0.242279	0.244644	0.238304	0.725226
9	5717	Data Science Manager - Machine Learning/AI | 2...	0.239048	0.241464	0.212282	0.692794
4	12794	Vice President, Big Data Analytics & Machine L...	0.270547	0.272368	0.142904	0.685819
12	3247	Machine Learning Engineer/ Data Scientist	0.19946	0.202323	0.259108	0.66089
14	7582	Data Scientist Team Lead	0.178896	0.181858	0.299642	0.660395

Fig. 7 (continued)

References

1. https://www.python.org/doc/
2. https://www.analyticsvidhya.com/blog/2021/07/recommendation-system-understanding-the-basic-concepts/
3. https://www.aurigait.com/blog/recommendation-system-using-knn/
4. Jadhav C, Pawar V, Waghmare S, Thakare VM, Resume classification using convolutional neural networks
5. Guo S, Alamudun F, Hammond T (2016) RésuMatcher: a personalized résumé-job matching system. Expert Syst Appl 60:169–182
6. Gopalakrishna ST, Vijayraghavan V (2019) Automated tool for resume classification using semantic analysis. Int J Artific Intell Appl 10(1)
7. Ramos J (2003) Using TF-IDF to determine word relevance in document queries. In: Proceedings of the first instructional conference on machine learning, New Jersey, U.S.A.
8. Uysal AK, Gunal S (2014) The impact of preprocessing on text classification. Inf Proc Manag 50(1):104–112
9. Masud MM, Al Imran MA, Ahmed I, Automatic classification of resumes using machine learning techniques
10. Resume classification using various machine learning algorithms, ITM web of conference by 3 Ramrao Adik institute of technology, Nerul, Navi Mumbai, India 4 D.Y. Patil Deemed to be University, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
11. Resume classification-based on personality using machine learning algorithm, 2023 IJCRT | vol 11, no 1 January 2023 |

Weighted Hybridization of Music Recommendation System to Address Major Issues in Recommendation Systems

M. Sunitha and T. Adilakshmi

Abstract A music recommendation system (MRS) serves as a solution to the overwhelming amount of information in the digital music realm. In this research paper, we tackle significant challenges faced by recommendation systems, namely, the Long Tail phenomenon, data sparsity, and the cold-start problem by employing a weighted hybrid approach. This hybrid approach combines collaborative filtering techniques based on both user preferences and item characteristics. Notably, our proposed system incorporates contextual information when generating music recommendations. We conducted experiments on a benchmark dataset and on synthetic data generated from a Music Portal application. The results we obtained demonstrate the system's ability to accurately capture user interests by considering various factors, including a user's historical preferences, their profile, item similarities, timestamps, and their social connections.

Keywords Music recommendation system · Information overloading · Long tail · Sparsity · Cold-start · Weighted hybridization

1 Introduction

In the contemporary age of the Internet and smartphones, users are presented with an extensive range of choices. This encompasses decisions like picking a music item from a digital platform, choosing a restaurant, deciding on a cinema to watch, buying a book, or various accessories. As we engage with these choices, it is only natural that users begin to form preferences, which can include both those they favor and

M. Sunitha (✉) · T. Adilakshmi
Department of Computer Science and Engineering, Vasavi College of Engineering Ibrahimbagh, Hyderabad, AP 31, India
e-mail: m.sunithareddy@staff.vce.ac.in

T. Adilakshmi
e-mail: t_adilakshmi@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_56

those we do not. However, if someone closely examines an individual's likes and dislikes, they often discover an underlying pattern [1].

The primary objective of a RS is to uncover this inherent pattern that characterizes a user's taste. After understanding an individual's preferences, a RS can propose comparable items. For instance, if someone has a liking for Spinach sandwiches, it is highly probable that they would also find a corn sandwich appealing since the two share significant similarities, differing only in the replacement of spinach with corn.

Music is ubiquitous, with a huge number of songs readily available from digital platforms, inundating music listeners with choices. Users are constantly in search of music that resonates with their personal taste, which serves as the driving force for music recommendations. In recent past, various services like Gaana, Spotify, and Last.fm [2] have emerged with the aim of providing a perfect solution, although they have not achieved complete success. Music preferences are shaped by factors such as taste, trust, and affinity for specific artists, which are challenging to quantify for a machine or software. Consequently, service providers have struggled to accurately identify music that genuinely appeals to and satisfies an individual's taste [3].

Every music recommendation system operates based on a set of assumptions to deliver effective recommendations. The following are the basic types of recommendation systems in the literature.

Collaborative Filtering: Collaborative filtering relies on user interactions and feedback to generate recommendations. It assumes that users who have interacted with and liked similar music in the past will have similar preferences in the future [4, 5].

User-Item Matrix: In collaborative filtering, a user-item matrix is created where rows represent users, columns represent items (in this case, songs), and each cell contains a rating or interaction score.

User-Based Collaborative Filtering: This method identifies users with similar music preferences to the target user and recommends songs liked by those similar users but not yet rated by the target user.

Item-Based Collaborative Filtering: This approach identifies songs that are similar to the ones the target user has liked in the past and recommends those similar songs.

Content-Based Filtering: Content-based filtering focuses on the characteristics and features of songs as well as user profiles to make recommendations. It assumes that users will prefer songs with attributes similar to the ones they have liked previously.

Song Features: In content-based filtering, various features of songs are considered, such as genre, artist, tempo, lyrics, and musical key.

User Profile: A user profile is created based on the user's historical interactions, including their liked songs and the attributes of those songs they have interacted with.

Recommendation: Content-based systems recommend songs that match the attributes and features of the songs the user has shown interest in, aligning with their preferences.

The rest of the paper follows this structure: Sect. 2 explores previous research in this domain, Sect. 3 outlines the suggested approach, Sect. 4 demonstrates the outcomes attained using this method, and lastly, Sect. 5 summarizes the conclusion and potential directions for future research.

2 Related Work

2.1 Automatic Playlist Generation

Automatic playlist generation typically depends on user-provided "seeds" to suggest songs that share similarities with the selected seeds, ultimately creating a new playlist. However, this approach lacks user feedback or input throughout the playlist generation process. One notable drawback of this method is its tendency to recommend songs that closely resemble the initial seed choices. An example of a similar technique can be found in the iTunes Genius feature [6], which constructs playlists based on selected seeds [6] (Fig. 1).

2.2 Dynamic Music Recommendation

Dynamic music recommendation improves the automatic playlist generation process by actively integrating user feedback. It begins with an initial song selection and adjusts the playlist according to the feedback given by the user. Pandora [7] employs this method in its music recommendation system [8].

2.3 Hybrid Approaches

Hybrid approach combine collaborative filtering and content-based method. Use both content features and user access pattern to recommend music.

3 Proposed Work

This section describes about the techniques implemented as part of the proposed system. The initial step in any data mining task involves data preprocessing. In the proposed method, we utilize user logs obtained from Last.fm [2] for experimentation purposes. Each log entry includes a timestamp indicating when a user listened to a

Fig. 1 Playlist generated by iTunes

Table 1 A representative user-song matrix

	S1	S2	…	S658
U1	7	0	…	0
U2	2	0	…	9
…	…	…	…	…
U314	0	6	…	0 0

Table 2 User-Song matrix after Binarization

	S1	S2	S3	S4
U_1	0	0	1	0
U2	1	0	0	1
U_3	1	1	0	1
U_4	1	0	0	0
U_5	0	1	1	0
U_6	0	1	1	0
U_7	1	0	0	1
U_8	1	0	0	0
U_9	1	0	0	1
U_{10}	1	0	0	0

specific song. From these logs, we create a User-Song matrix, and Table 1 provides an example of a User-Song matrix captures user ratings for various items.

Binarization is used to convert user-song matrix as per Table 2.

Normalization is applied to the binary user-song matrix in the proposed system, employing both length normalization and RMSN normalization techniques similar to length normalization (Table 3).

The following techniques are combined to form a Hybrid Music recommendation system.

3.1 User and Item-Based Collaborative Filtering

This technique implements user-based and item-based model collaborative filtering to form the clusters of similar users and items. Once the model of user and item clusters is built, it can be used to find the user cluster and item cluster most similar to the target user. This is shown in the pseudocode given in the Algorithms 1 and 2. It considers the user past history while forming similar user clusters and item similarities while forming item clusters.

Algorithm 1 Pseudocode for user-based model

Table 3 User-song rating matrix after applying length normalization

	S1	S2	S3	S4
U_1	0	0	1	0
U2	0.5	0	0	0.5
U_3	0.33	0.33	0	0.33
U_4	1	0	0	0
U_5	0	0.5	0.5	0
U_6	0	0.5	0.5	0
U_7	0.5	0	0	0.5
U_8	1	0	0	0
U_9	0.5	0	0	0.5
U_{10}	1	0	0	0

```
Algorithm Threshold_UC()
1. Begin
2. Initialize the threshold value to th_cutoff
3. Create an empty list of clusters
4.    For ui in u1, u2 … un
5.    Create a new cluster C
6.    Assign u to C
7.    For each ui in u2 … un
8.      begin
9.        Find the similarity of ui with C
10.         If sim(ui, C) ≤ th_cutoff
11.           Assign ui to C
12.         Else
13.           Create a new cluster C′
14.           Assign ui to C′
15.         end
16.       end
17. Add C to the list of clusters
18. Return the list of clusters
19. End
```

Algorithm 2 Pseudocode for item-based model

```
Algorithm Threshold_IC()
1. Begin
2. Initialize the threshold value to th_cutoff
3. Create an empty list of clusters
4. For each ui in u1, u2 … un
5.    Create a new cluster C
6.    Assign u to C
7.    For each ui in u2 … un
```

```
8.      begin
9.        Find the similarity of ui with C
10.         If sim(ui, C) ≤ th_cutoff
11.           Assign ui to C
12.         Else
13.           Create a new cluster C′
14.           Assign ui to C
15.         end
16.       end
17. Add C to the list of clusters
18. Return the list of clusters
19. End
```

3.2 Session Based Music Recommendation System

To integrate contextual information into music recommendations, sessions are introduced as a technique. This concept is based on the fundamental notion that users' preferences can change throughout the day. For instance, users might have a preference for devotional songs in the early morning and opt for melodic tunes in the evening. To accommodate this variability in users' listening behavior, the implicit feedback users provide in the form of logs is classified into four specific sessions, as defined below:

- Session 1: Timestamp 0–6,
- Session 2: Timestamp 6–12,
- Session 3: Timestamp 12–18,
- Session 4: Timestamp 18–24.

Using the listening history of users, transactions are categorized into these four sessions based on timestamps. For each session, user clusters and item clusters are created using the pseudocode provided in Algorithms 1 and 2. Within each session, the target user is linked to identify the most similar user and item clusters, facilitating recommendations.

3.3 Matrix Factorization to Address Sparsity Problem

Sparsity represents a significant challenge in recommendation systems. This issue arises because users typically rate only a limited number of items. The sparsity problem is illustrated by the Sparse User-Item Matrix displayed in Fig. 2, where the matrix reveals the ratings provided by m users for n items.

Fig. 2 Sparse user-song rating matrix

	S1	S2	….	S658
U1	7	0	….	0
U2	2	0	….	9
….	….	….	….	….
U314	0	6	….	0

Matrix factorization [9] serves as a method for reducing the dimensionality of a high-dimensional sparse matrix into a lower-dimensional dense matrix. SVD is a matrix decomposition technique that takes an $m \times n$ matrix and decomposes it into three matrices with dimensions $m \times k$, $k \times k$, and $k \times n$, as expressed by the equation provided below.

$$A_{m\times n} = U_{m\times k} \times S_{k\times k} \times V_{k\times n}$$

where U designates about users, V indicates about items in transpose and S shows identity matrix with Eigen values as diagonal elements.

User clusters are generated from the factored user matrix U and item matrix V is used to form item clusters. Euclidean distance is used as the proximity measure to form model of users and items by using the algorithm shown in Algorithms 3 and 4.

Algorithm 3 Pseudocode for UC_with Sessions

Algorithm UCSSVD()
Input: SVD applied User-Song Matrix of a particular session
Output: User Clusters for a session
Method:
1. Initialize the threshold_cutoff value
2. Create an empty list of clusters
3. Consider the first user vector I1
4. Create a new cluster C1 and assign I1 to it
5. For each remaining user vector Ii (where i ranges from 2 to k, and k represents the count of unique items rated by all users)
6.   Calculate the similarity between Ii and all the clusters that have been created up to this point.
7.   If Ii is more similar to an existing cluster than the threshold_cutoff
8.   Assign Ii to the most similar cluster
9.   Else
10.   Create a new cluster Ci
11.   Assign Ii to Ci
12. Add Ci to the list of clusters
13. Return the list of clusters
End

Algorithm 4 Pseudocode for IC with Sessions

Algorithm ICSSVD()
Input: SVD applied User-Song Matrix of a particular session
Output: Item Clusters
Method:
1. Initialize the threshold_cutoff value
2. Create an empty list of clusters
3. Consider the first item vector U1
4. Create a new cluster C1 and assign U1 to it
5. For each remaining user vector Ui (where i ranges from 2 to k, and k represents the count of unique users
rated the item)
6. Calculate the similarity between Ui and all the clusters that have been created up to this point
7. If Ui is more similar to an existing cluster than the threshold_cutoff
8. Assign Ui to the most similar cluster
9. Else
10. Create a new cluster Ci
11. Assign Ui to Ci
12. Add Ci to the list of clusters
13. Return the list of clusters
End

3.4 Integrating Social Networks to Address New User Cold-Start Problem

Another challenge that recommendation systems must tackle is the cold-start problem. The cold-start problem refers to the difficulty of providing recommendations for either new users or new items. The former is referred to as the new user cold-start problem, while the latter is known as the new item cold-start problem.

To mitigate the new user cold-start problem, the approach involves incorporating a user's social network connections. In case of new target user, the recommendation system lacks information about them, the system leverages their friends' interests to gain insights into the target user. The algorithm for this process is presented in Fig. 3.

When dealing with a specific target user, this algorithm determines the user cluster with the highest similarity to the target user. Additionally, it creates a social matrix by taking into account the immediate friends of the target user. Ultimately, the algorithm suggests items that are shared between the user cluster and the social matrix, offering recommendations that combine both sources of information.

Algorithm to Address Cold-Start Problem for Recommendation (CSReco)

Input: User-Item Matrix, Social Matrix

Output: Set of recommendations for the target User

1. Let U1, U2, U3, ..., Um represent the users, and I1, I2, I3, ..., In represent the items.
2. Create user clusters using a Threshold-based clustering algorithm (Form_User_Clusters()).
3. For the given target_user:

   a. Determine the user cluster that exhibits the highest similarity to the target_user.

   b. Construct a social matrix by identifying the immediate friends of the target_user from the Social Matrix.

   c. Obtain recommendations from the user clusters (Ru1, Ru2, ...) and from the social matrix (Rs1, Rs2, ...).

   d. Suggest items that are common to both recommendation sets.

End of Algorithm

Fig. 3 Pseudocode for collaborative social recommender system

3.5 Addressing Long Tail Problem

The long tail refers to a collection of items that are relatively less popular, with the majority of items falling into this category [10]. In contrast, only a small fraction of items are part of the heavy tail, as illustrated in Fig. 4. Presently, we are transitioning toward a paradigm that focuses on distinguishing between Hits and Niche content. However, the challenge lies in effectively filtering and presenting the right artists to users based on their musical preferences, even if those artists do not fall within the Heavy Tail category.

To tackle the long tail problem, a solution involves categorizing songs based on their popularity frequency. This classification divides items into two categories: long

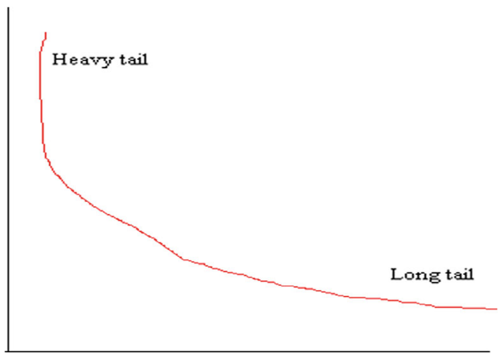


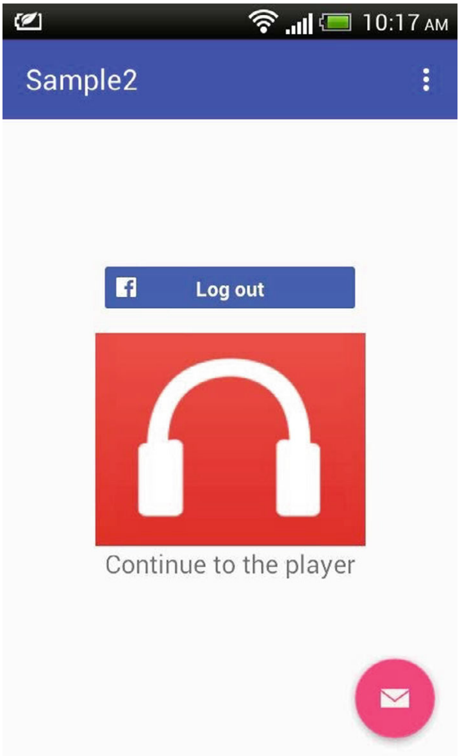
Fig. 4 Long tail

tail and heavy tail, based on their average frequency. If an item's average frequency is lower than the average frequency of all items, it falls into the long tail category; otherwise, it is classified as belonging to the heavy tail. Subsequently, songs in the long tail and heavy tail segments are utilized separately to create clusters of similar songs using the algorithm depicted in Algorithm 2.

3.6 Mobile Application to Combine All the Proposed Approaches for Top N Recommendations

A mobile application has been designed to integrate the various approaches mentioned above for delivering top N recommendations. This Android mobile application, illustrated in Fig. 5, enables users to access the entire library of audio songs on their mobile devices. Whenever users listen to a song, the application generates a log entry and stores it in the local database. This log contains specific fields, as depicted in Fig. 6.

Fig. 5 Mobile application for MRS



Tue Apr 05 06:48:20 GMT+05:30 2016	Sai Krishna	18462	[iSongs.info] 03 - Mila Mila	10	Bhale Manchi Roju - (2015)

Fig. 6 Sample user log

The administrator of the mobile application employs a recommendation algorithm that leverages the approaches discussed earlier to quantify the diverse factors influencing a user's taste. This algorithm takes into account a portion of songs from each technique when creating the top *N* recommendations. The formula used to generate these top *N* recommendations is as follows, but it can be rephrased for clarity if needed.

Certainly, here is the equation representing the top *N* recommendations with the given factors:

$$\begin{aligned}\text{Top}\,N\,\text{Recommendations} &= \alpha 1 * (\text{User and Item Clusters}) + \alpha 2 * (\text{Sessions}) \\ &\quad + \alpha 3 * (\text{Matrix Factorization}) + \alpha 4 * (\text{Social Integration}) \\ &\quad + \alpha 5 * (\text{Long Tail}).\end{aligned}$$

The α values represent the proportions of songs selected from each technique, effectively serving as weights assigned to each approach. To illustrate, this means that, for instance, 20% of the songs are drawn from User and Item Clusters, 30% from Sessions, 30% from Matrix Factorization, and 20% from Social Integration.

4 Results

This section describes the results obtained for all the techniques individually and combined (Figs. 7, 8, 9, 10 and 11; Table 4).

5 Conclusion and Future Scope

The research proposed here has developed an Android-based hybrid music application that takes into account various aspects of users when making recommendations. The recommendation algorithm considers users' historical preferences, item similarities, and contextual information, effectively addressing the challenges of cold start, long tail, and sparsity in recommendation systems. Our findings demonstrate that the hybrid recommendation system we proposed outperforms baseline approaches like the most popular or most recent recommendation systems. Furthermore, the results

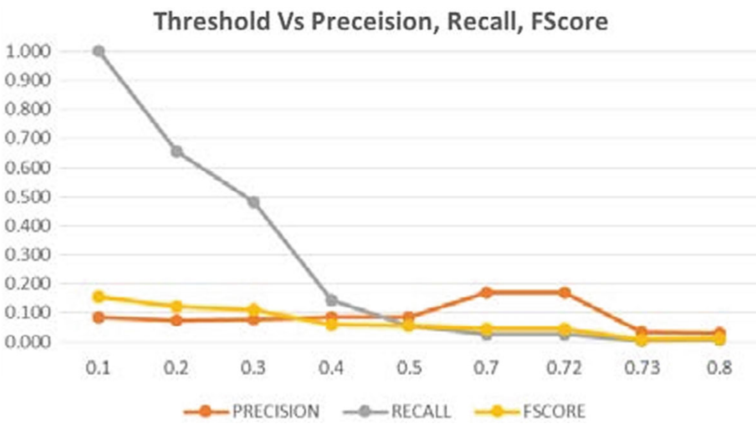


Fig. 7 User and item clusters with cosine similarity

Fig. 8 Performance of precision for different threshold values

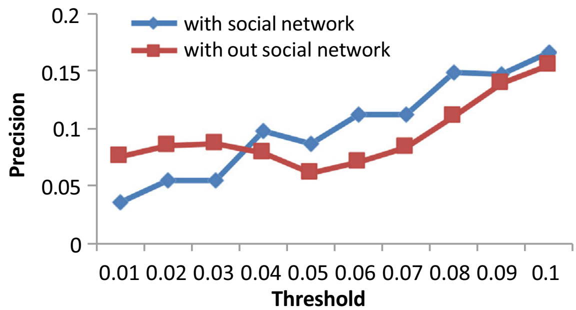
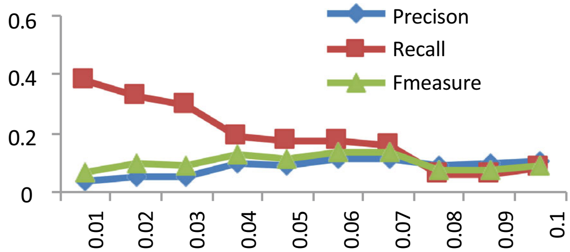


Fig. 9 Performance of the model with sessions



indicate that the combination of multiple techniques for providing top *N* recommendations yields better performance compared to using individual techniques separately. Additionally, we can extend this hybrid mobile-based music recommendation system to incorporate user-generated content, such as postings about new items on social networks, to tackle the new item cold-start problem effectively.

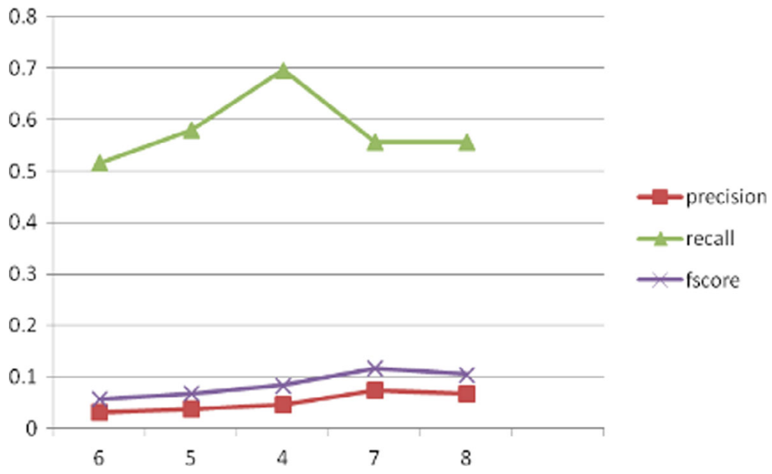


Fig. 10 Performance of a sample test users with discretization for the month January 2008

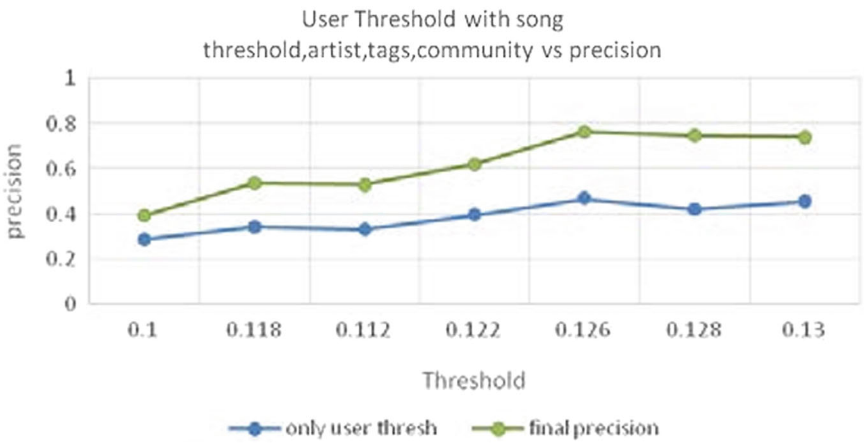


Fig. 11 Precision, recall, and F-score with all the proposed approaches

Table 4 Precision (P), Recall (R), and F-score with long tail and sessions

Session s	P	R	F-score
S1	0.088	1.0	0.080
S2	0.097	1.0	0.088
S3	0.085	1.0	0.075
S4	0.084	1.0	0.078

References

1. McFee B, Bertin-Mahieux T, Ellis DPW, Lanckriet GRG (2012) The million song dataset challenge. WWW 2012 Companion, 16–20 Apr 2012, Lyon, France
2. Last. FM—A popular music web portal http://www.last.fm
3. Hidasi B, Tikk D (2013) Context-aware item-to-item recommendation within the factorization framework. CaRR'13, 5 Feb 2013, Rome, Italy
4. Moreno MN, Segrera S, Lopez VF, Munoz MD, Sanchez AL (2012) Mining semantic data for solving Firstrater and cold-start problems in recommender system ACM IDEAS 11 2011
5. Gong S (2010) A collaborative filtering recommendation algorithm based on user clustering and item clustering. J Software 5(7)
6. Session aware recommender system In ECommerce, Jian WangA dissertation submitted in partial satisfaction of the requirements for the degree of doctor of philosophy, June 2013
7. Pandora—A free internet radio. http://www.pandora.com
8. Cui X, Yin G, Zhang L, Kang Y (2013) Method of collaborative filtering based on uncertain user interests cluster. J Comput 8(1)
9. Singular Value Decomposition. http://en.wikipedia.org/wiki/Singular_value_decomposition
10. Anderson C (2004) The long Tail. Wired Mag 12(10):170–177

Machine Learning Approaches for Sentiment Analysis in Social Media Using Soft Computing Techniques

M. Priyadharshini, V. Indumathi, V. V. Bhavani, Tahseen Jahan, A. Hemalatha Reddy, and Jalapala Sinjini

Abstract Sentiment analysis in social media has become crucial for understanding public opinion and sentiment towards various entities, products, and events. This paper explores the application of machine learning (ML) approaches in sentiment analysis within the realm of social media, leveraging soft computing techniques. Soft computing methodologies, including fuzzy logic, neural networks, and genetic algorithms, offer flexible and adaptive frameworks for handling the inherent uncertainties and complexities present in social media data. This paper provides an overview of the key ML algorithms used in sentiment analysis, discusses the challenges specific to social media data, and reviews recent advancements and trends in the field. Additionally, it highlights the integration of soft computing techniques with ML models to enhance the accuracy and robustness of sentiment analysis systems.

Keywords Sentiment analysis · Social media · Machine learning · Including fuzzy logic

M. Priyadharshini (✉) · V. Indumathi · V. V. Bhavani · T. Jahan · A. Hemalatha Reddy · J. Sinjini
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: priyadharshini@cmritonline.ac.in

V. Indumathi
e-mail: indu11mathi@cmritonline.ac.in

V. V. Bhavani
e-mail: bhavaniv@cmritonline.ac.in

T. Jahan
e-mail: tahseen048@cmritonline.ac.in

A. Hemalatha Reddy
e-mail: hemareddy@cmritonline.ac.in

J. Sinjini
e-mail: sinjini@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_57

1 Introduction

Sentiment analysis in social media using machine learning (ML) approaches has gained significant attention due to the vast amount of user-generated content available online [1]. Soft computing techniques, including fuzzy logic, neural networks, and genetic algorithms, have been increasingly applied to address the inherent complexities and uncertainties present in social media data. Fuzzy logic provides a flexible framework for handling linguistic variables and capturing the vagueness of human language, which is prevalent in sentiment analysis tasks. Neural networks, particularly deep learning architectures like recurrent neural networks (RNNs) and convolutional neural networks (CNNs), excel at learning intricate patterns in text data, allowing for more accurate sentiment classification. Genetic algorithms offer an optimization approach to fine-tune ML models and feature selection, enhancing the overall performance of sentiment analysis systems. These soft computing techniques complement traditional ML algorithms by providing robustness against noise, ambiguity, and subjective language expressions commonly encountered in social media posts [2]. By leveraging these approaches, researchers and practitioners can develop sentiment analysis systems capable of effectively interpreting and analysing opinions, emotions, and attitudes expressed in social media content, thereby enabling various applications such as brand monitoring, market analysis, and customer feedback analysis.

Sentiment Analysis in Social Media (SM)

Sentiment analysis in social media involves the use of computational techniques to analyse and understand the opinions, emotions, and attitudes expressed within user-generated content on various social media platforms. This field has garnered immense interest due to the vast amount of data generated daily on platforms like Twitter, Facebook, and Instagram [3].

The process typically involves several steps, starting with data collection from social media platforms either through APIs or web scraping techniques. Once the data are collected, pre-processing steps such as tokenization, stemming, and removal of stop words are applied to clean and prepare the text for analysis.

Next, various machine learning and natural language processing techniques are employed for sentiment classification. These techniques range from traditional algorithms like Naive Bayes and support vector machines to more advanced approaches such as deep learning models like recurrent neural networks (RNNs) and transformers.

Sentiment analysis can be performed at different levels, including document level, sentence level, and aspect level, depending on the granularity of analysis required. Additionally, sentiment analysis can also involve aspect-based sentiment analysis (ABSA), which aims to identify the sentiment associated with specific aspects or features mentioned in the text, such as product features in customer reviews.

Applications of sentiment analysis in social media are widespread and include brand monitoring, customer feedback analysis, reputation management, market research, and public opinion analysis, among others. By extracting insights from social media data, businesses, organizations, and researchers [4] can gain valuable insights into public sentiment and make data-driven decisions accordingly.

2 Literature Review

According to Ref. [5], "opinion mining" is a common term for sentiment analysis (SA). Extracting, identifying, or characterising the emotion from text was the definition. In the past, there existed a strict dichotomy between the good and bad aspects of a piece of written content. A more accurate picture of the feelings emerged, nevertheless, from the fine-grained categorization. Negatively, fine-grained categorization was more difficult than binary. With regard to multi-class categorization, however, performance declined sharply. That research looked on sentiment multi-class classification using pre-processing methods and machine learning models. It was suggested to use a multi-layer classification model to improve performance. The solution has made use of the movie reviews dataset because of the similarities with social media content. When it came to sentiment categorization, supervised machine learning models were put into action. The models of the single-layer and multi-tier architectures were compared. When compared to the single-layer design, the multi-tier model's outcomes were somewhat better. In addition, the suggested model was able to acquire more context because of the multi-tier models' superior recall. In order to aid researchers in designing multi-tier models with more contextual information, they have already addressed certain model weaknesses.

In Ref. [6], according to their findings, ensemble models that included several machine learning classifiers produced the most accurate results. The most encouraging outcomes were produced by the BERT and RoBERTa models when they were fine-tuned using Twitter data. The researchers set out to answer the question, "How can public health professionals and decision-makers use sentiment analysis to combat the spread of COVID-19?" by integrating machine learning-based sentiment analysis with insights from social and behavioural science.

In order to gather, analyse, comprehend, and evaluate the trends and initiatives in a structured way, a systematic literature review was set up. The objective was to find unanswered questions about this combination's potential future research needs. The study's primary contribution was an examination and assessment of Twitter sentiment analysis using soft computing methodologies. Two, this review is unique in that it systematically sought for, gathered, analysed, and integrated data from all relevant high-quality papers on the given subject in order to answer specific research questions [7].

For companies, governments, and individuals alike. Applications for online banking, electronic mail, blogging, online news, and social networks are all instances of Software-as-a-Services (SaaS) in the cloud. Data sets derived from Mexican residents' interactions during the September 19, 2017, earthquake were analysed in that chapter using sentiment analysis and supervised learning, with the help of Ekman's six emotional model. To find out how people felt about tweets on the same subject, they developed three classifiers. When it came to predicting emotions, Naive Bayes and support vector machine were the most accurate classifiers. Happiness, anger, and sadness were the most often predicted emotions, while 6.5% of the tweets were deemed irrelevant. They offered several suggestions for improving sentiment analysis via the use of machine learning methods. Their contribution was to increase the number of possible emotions from three (negative, neutral, and positive) to six, giving researchers additional data points with which to analyse social media users' interactions. Validation of the system using other datasets and emotions, as well as the incorporation of new AI methods to enhance accuracy, would constitute future research [8].

Online opinion mining and sentiment analysis of text have recently drawn interest from academics, businesses, and governments, according to Ref. [9], who noted that this is in response to the exponential increase in social media-generated web content. Sentiment analysis has recently grown in popularity as a study subject in the field of artificial intelligence and machine learning, having first originated in the big data age under knowledge fusion. The experimental data for the research came from Taiwan's biggest online forum's Military life PTT board. This research set out to build a social media sentiment analysis framework and procedures, suggest an in-house military sentiment dictionary to enhance sentiment classification, and compare the efficacy of various deep learning models trained with different combinations of parameters. Compared to results obtained by using solely pre-existing sentiment dictionaries, the testing findings demonstrated that the accuracy and F1-measure of the model that integrated both the pre-existing dictionaries and the self-developed military sentiment dictionary were superior. In addition, the accuracy and F1-measure for sentiment classification were significantly better with two Bi-LSTM network layers and the activation function Tanh trained prediction model [9].

Concentrated on the fundamental concepts and similarities among the various ways to accomplish multilingual sentiment analysis and a broader coverage of the years 2017–2020 made these evaluations superior to their predecessors. Find it interesting that their research has shifted towards cross-lingual and code-switching approaches, plateaued and there are no deep learning approaches to multilingual aspect-based sentiment analysis; and, shockingly, there are no more complex architectures like transformers-based, even though their results indicate that more elaborated architectures are needed for more difficult tasks [10].

According to Ref. [11], their social media evaluations greatly influenced consumers to make the best possible choices when purchasing goods and services. The majority of the methods used in sentiment analysis relied on machine learning algorithms. That study analysed and contrasted the most popular machine learning approaches. After that, we compared how well these methods worked with online user evaluations from a variety of industries. Amazon, Yelp, and IMDb were just a few of the domains represented in the many datasets used in the studies. The trials made use of popular techniques like DT, Bagging, Boosting, RF, SVM, and Maximum Entropy. Boosting and Maximum Entropy were determined to be the most effective machine learning algorithms for identifying user sentiment in online reviews, while users were able to extract useful information from review datasets for enhanced product sales production and business intelligence.

SA may determine whether a piece of source material has a positive, negative, or neutral emotional tone. In a nutshell, SA was able to process natural languages and glean insights, which was a huge help when trying to gauge general public sentiment about a given subject, product, or problem. The significance of SA may be attributed to the proliferation of information via various online mediums such as social media, blogs, website evaluations, and shared opinions. Due to the language's richness and its very complicated inflectional and derivational morphology, sentiment analysis in Arabic text is more difficult. Arabic sentiment analysis (ASA) was suggested as a new method to improve its accuracy in that article. By this measure, ASA has nine supervised machine learning algorithms in place. The ASA classification process has never made use of three of these classifiers: Multi-layer Perceptron, Ridge, and Gradient Boosting. We used each of the nine classifiers' custom-built datasets to evaluate their performance. Manually compiled from hotel reviews posted on Booking.com, the dataset included 6318 reviews written in various Arabic dialects [12].

According to Ref. [13], consumer sentiment research is all the rage right now in fields like healthcare, criminal justice, banking, tourism, and academics that use social media. It was crucial to detangle customer perspective in order to understand the intended goal and reviews. Due to the exponential growth in the volume, subjectivity, and variety of social media data, processing it manually has become more difficult as technology has progressed. This problem has been solved in real-world situations using machine learning (ML) approaches. This research examined the feasibility, extent, and use of combining machine learning approaches for CSA (consumer sentiment analysis)) with online review data from the hotel and tourist industries. To illustrate the potential future applications of this pairing, they presented a thorough literature evaluation that compared, analysed, explored, and understood previous efforts to identify research gaps. Reading and analysing articles on how the hotel and tourist industries are using ML approaches to analyse customer sentiment in online reviews was the major goal. The findings of that study might have far-reaching consequences for service providers looking to better assist their clients in making informed service selection decisions. In addition, researchers were greatly influenced in terms of future study objectives.

The objective of Ref. [14] was to glean sentiments, views, and attitudes from Twitter and other social media platforms. It has grown in popularity as a field of study. Textual data were the main emphasis of the traditional sentiment analysis method. One of the most well-known forms of online social networking is Twitter, where users may send brief updates on a variety of subjects using the "tweet" format. Models for sentiment analysis were suggested using logistic regression, support vector machine, and naive Bayes. The primary goal was to alleviate feelings to a greater extent. Positive and negative sentiment were the two main categories used in Twitter sentiment analysis. Machine learning classifiers might do this. Classifiers like this would assess public opinion on businesses, political parties, analysts, etc. The tweets were appropriately categorised by machine learning algorithms utilising training data. Thus, this approach did not need a word database; thus, machine learning techniques were more effective and expedited while doing sentiment analysis.

The objectives of the study conducted by Ref. [15] provided two experts in the area with a set of guidelines to follow as they annotated pertinent tweets from scientific papers. A Cohen's Kappa (IAA) of 0.80 was the result of their accurate emotion annotations. Two other datasets were then subjected to the same tests: one including just English tweets and another containing tweets in twenty-three languages (English included). Using 6388 tweets and 300 papers indexed in Web of Science as our baseline, we tested the performance of two well-known sentiment analysis methods, Sentiment140 and SentimentStrength, with our own machine learning and natural language processing models. Outperforming all other classifiers and baseline methodologies, Support Vector Machine with unigram achieved an accuracy of over 85%. With an accuracy rate of 83%, Logistic Regression secured the second place, followed by Naïve Bayes in third place with 80%.

According to Ref. [16], according to their comprehensive testing, the top three single classifiers in terms of recall, accuracy, precision, and F-measure were logistic regression, multi-layer perceptron, and linear-kernel support vector machine. If utilised as basis classifiers for ensemble models, their performance might be much better. In addition, they found that using terms frequency and N-gram words in conjunction with part-of-speech lemmatization and negation word identification were two of the text pre-processing strategies that improved accuracy. Moreover, they proved that deep learning models may do just as well as using the overall emotion of lexicons like and SenticNet 4 and SentiWordNet 3.0 to produce useful features. Their system consistently produced the same results when tested with various datasets. They finished by demonstrating how to improve the accuracy of neutral emotion without compromising the accuracy of positive or negative assessments, which was their primary objective.

By analysing data from Twitter, Ref. [17] were able to determine how people all across the world felt about the 2014 FIFA World Cup, which was held in Brazil. They employed machine learning techniques to do this. Emotional phrases that were discovered in user tweets were used in order to calculate sentiment polarity via the utilisation of data filtering and analysis techniques that were based on NLP. Word stemming, tokenization, and lemmatization, as well as POS (parts of speech) tagger, name entity recognition, and a parser, were used in order to normalise the dataset and get it ready for usage by machine learning techniques. Emotions were then extracted from the linguistic data that was included inside each tweet. For the purpose of putting that plan into action, the technologies that were used were Python and the Natural Language Toolkit (NLTK). We were able to extract emotionally charged words from phrases that had meanings that were distinctive to the context by using WordNet's POS (parts of speech) feature. After that, we assigned the polarity of the feeling by using either the SentiWordNet dictionary or an approach that was based on a lexicon. In order to further investigate the result of the polarity assignment, the Weka platform and several machine learning techniques were used. These techniques included naïve Bayes, support vector machine (SVM), K-nearest neighbour (KNN), and random forest (Table 1).

Table 1 Comparative analysis of background

References	Technology used	Area of research	Techniques used	Methodology	ML/fuzzy logic method	Conclusion
[5]	Machine learning	Sentiment analysis	Pre-processing, ML models	Implementation of pre-processing and ML models	ML models (decision tree, support vector machine, Naïve Bayes)	Proposed multi-layer classification model improves performance in sentiment classification
[6]	Machine learning	Sentiment analysis	Literature review, ML models	Systematic literature review	Ensemble models (BERT, RoBERTa)	Ensemble models, especially BERT and RoBERTa, provide promising results for COVID-19 Twitter data sentiment analysis (SA)
[7]	Soft computing	Sentiment detection and classification	Literature review, soft computing techniques	Systematic literature review	Soft computing techniques	Soft computing techniques show potential for SA on Twitter
[8]	Supervised learning	Sentiment analysis	Ekman's model, supervised learning	Data analysis and classification	Naïve Bayes, support vector machine	Classification of emotions in Twitter data using supervised learning techniques
[9]	Deep learning	Sentiment analysis	Sentiment analysis framework, deep learning models	Experimentation and analysis	Deep learning models (Bi-LSTM)	Deep learning models improve sentiment classification in online forums
[10]	Deep learning	Multilingual sentiment analysis	Literature review, deep learning approaches	Review and analysis of existing solutions	Complex architectures (transformers based)	Current research trends focus on cross-lingual and code-switching approaches, but lack more elaborate architectures like transformers based
[11]	Machine learning	Sentiment analysis	Comparative studies, ML models	Comparative analysis of ML techniques	ML models (SVM, decision tree (DT), random forest bagging, boosting)	Boosting and maximum entropy outperform other ML algorithms for sentiment analysis in online user reviews

(continued)

Table 1 (continued)

References	Technology used	Area of research	Techniques used	Methodology	ML/fuzzy logic method	Conclusion
[12]	Supervised learning	Sentiment analysis	Comparative studies, ML models	Comparative analysis of ML techniques	ML models (Ridge, Gradient Boosting, Multi-layer Perceptron)	New supervised ML algorithms improve accuracy of Arabic sentiment analysis
[13]	Machine learning	Consumer sentiment analysis	Literature review, ML techniques	Systematic literature review	ML techniques (not specified)	ML techniques show potential for consumer sentiment analysis in hospitality and tourism
[14]	Supervised learning	Sentiment analysis	Dataset preparation, ML models	Dataset preparation and ML implementation	ML models (Naïve Bayes, logistic regression, support vector machine)	Supervised ML classifiers effectively classify sentiments in Twitter data
[15]	Machine learning	Sentiment analysis	Dataset annotation, ML models	Dataset annotation and ML implementation	ML models (support vector machine, logistic regression, Naïve Bayes)	SVM with other classifiers in sentiment analysis of altmetrics dataset
[16]	Machine learning	Sentiment analysis	Framework development, ML models	Framework development and experimentation	ML models (SVM, multi-layer perceptron, logistic regression)	Machine learning techniques improve accuracy of sentiment analysis in online reviews
[17]	Machine learning	Sentiment analysis	Dataset analysis, ML algorithms	Dataset analysis and ML implementation	ML algorithms	ML algorithms effectively detect sentiment in Twitter data related to World Cup soccer 2014

3 Methodology

Developing a mathematical model for SA in social media using machine learning (ML) approaches often involves employing soft computing techniques due to the inherent complexity and uncertainty present in textual data. Soft computing encompasses various computational methodologies that can effectively handle imprecision, uncertainty, and partial truth, making it suitable for analysing subjective content like sentiment in social media.

One ML approach for sentiment analysis is supervised learning, where a model is trained on a labelled dataset containing examples of text along with their corresponding sentiment labels. Let's denote the dataset as $D = \{(x_i, y_i)\}_{i=1}^{N}$, where x_i represents a text document and y_i is its corresponding sentiment label. The goal is to learn a mapping $f : X \to Y$, where X is the space of text documents and Y is the space of sentiment labels.

In supervised learning, we aim to minimize a loss function that quantifies the discrepancy between the predicted sentiment $\hat{y}_i = f(x_i)$ and the true sentiment label y_i. A common choice for the loss function in sentiment analysis is the cross-entropy loss, defined as:

$$L(f) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (i - y_i)\log(1 - \hat{y}_i)\right]$$

where $\hat{y}_i$ represents the predicted probability of positive sentiment for document x_i.

Soft computing techniques, such as fuzzy logic, can be incorporated to handle the inherent ambiguity in sentiment analysis. Fuzzy logic allows for the representation of linguistic variables and fuzzy rules, which are particularly useful for capturing the vagueness and uncertainty present in sentiment expressions.

Let's denote the input variables related to sentiment analysis (e.g. word frequencies, sentiment scores) as $X = \{x_1, x_2, ..., x_m\}$, and the output sentiment label as Y. Fuzzy logic operates on fuzzy sets, where each linguistic variable is associated with a membership function defining its degree of membership to each fuzzy set.

For example, we can define fuzzy sets for sentiment labels such as "positive," "neutral," and "negative," along with membership functions denoting the degree to which a document belongs to each sentiment category.

$$\mu\text{Positive}(x_i), \mu\text{Neutral}(x_i), \mu\text{Negative}(x_i)$$

Fuzzy rules can then be formulated to map the fuzzy input variables to fuzzy output labels. These rules define the relationship between the input linguistic variables and the output sentiment label, incorporating expert knowledge or learned patterns from the data.

$$\text{Rule 1: If } x_1 \text{ is High AND } x_2 \text{ is Low, then } Y \text{ is Positive}$$

$$\text{Rule 2: If } x_1 \text{ is Medium OR } x_3 \text{ is High, then } Y \text{ is Neutral}$$

$$\text{Rule 3: If } x_2 \text{ is High AND } x_3 \text{ is High, then } Y \text{ is Negative}$$

These fuzzy rules can be combined using fuzzy inference methods like Mamdani or Sugeno to generate a crisp sentiment label for a given input text document.

Incorporating soft computing techniques into the mathematical model for sentiment analysis enables the handling of uncertainty and vagueness present in social media content, enhancing the model's robustness and accuracy in capturing the nuanced sentiment expressed in textual data.

4 Work Flow Diagram

This flowchart provides steps involved in implementing machine learning approaches for sentiment analysis in social media using soft computing techniques. Each step can be further refined and customized based on the specific requirements and characteristics of the social media data and the desired level of accuracy and performance (Fig. 1).

5 Conclusion and Future Work

The paper explored the application of machine learning (ML) approaches, particularly soft computing techniques, in sentiment analysis within the realm of social media. It provided an overview of key ML algorithms used in sentiment analysis, discussed challenges specific to social media data, and reviewed recent advancements and trends in the field. Additionally, it highlighted the integration of soft computing techniques with ML models to enhance accuracy and robustness in sentiment analysis systems.

The methodology section outlined a mathematical model for sentiment analysis using supervised learning and incorporated soft computing techniques like fuzzy logic to handle uncertainty and vagueness in social media content. Fuzzy logic allowed for the representation of linguistic variables and fuzzy rules, which were utilized to map input variables to output sentiment labels, enhancing the model's robustness.

Future work in this area could focus on further refining and customizing the proposed mathematical model, exploring the integration of advanced deep learning architectures for sentiment analysis, and addressing challenges related to multilingual sentiment analysis and code-switching in social media data. Additionally, validation of the proposed model using diverse datasets and incorporation of new AI methods

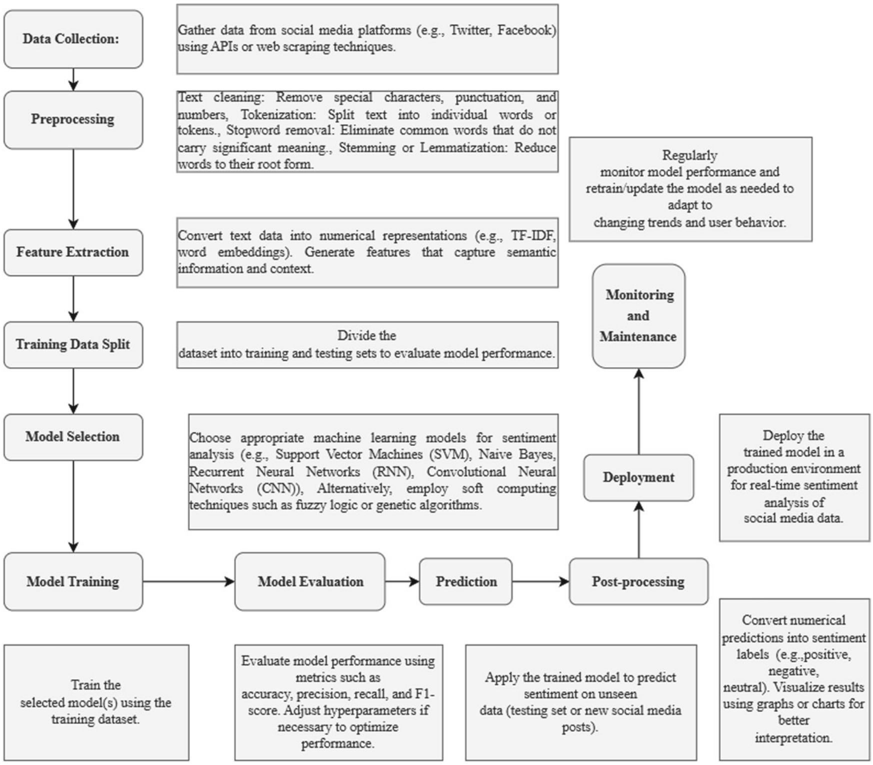


Fig. 1 Work flow diagram

could contribute to enhancing the accuracy and applicability of sentiment analysis systems in social media.

References

1. Jain PK, Saravanan V, Pamula R (2021) A hybrid CNN-LSTM: a deep learning approach for consumer sentiment analysis using qualitative user-generated contents. Trans Asian Low-Res Lang Inf Proc 20(5):1–15
2. Chandrasekaran G, Nguyen TN, Hemanth DJ (2021) Multimodal sentimental analysis for social media applications: a comprehensive review. Wiley Interdiscipl Rev: Data Min Knowl Discov 11(5):e1415
3. Shayaa S, Jaafar NI, Bahri S, Sulaiman A, Wai PS, Chung YW, Al-Garadi MA et al (2018) Sentiment analysis of big data: methods, applications, and open challenges. IEEE Access 6:37807–37827
4. Jayasanka RASC, Madhushani T, Marcus E, Aberathne IAAU, Premaratne S (2013) Sentiment analysis for social media. In: Information technology research symposium, vol 11. University of Moratuwa, Moratuwa, Sri Lanka, p 22

5. Rahman H, Tariq J, Masood MA, Subahi AF, Khalaf OI, Alotaibi Y (2023) Multi-tier sentiment analysis of social media text using supervised machine learning. Comput Mater Contin 74:5527–5543
6. Braig N, Benz A, Voth S, Breitenbach J, Buettner R (2023) Machine learning techniques for sentiment analysis of COVID-19-related twitter data. IEEE Access 11:14778–14803
7. Kumar A, Jaiswal A (2020) Systematic literature review of sentiment analysis on Twitter using soft computing techniques. Concurr Comput: Pract Exp 32(1):e5107
8. López-Chau A, Valle-Cruz D, Sandoval-Almazán R (2020) Sentiment analysis of Twitter data through machine learning techniques. In: Software engineering in the era of cloud computing, pp 185–209
9. Chen LC, Lee CM, Chen MY (2020) Exploration of social media for sentiment analysis using deep learning. Soft Comput 24:8187–8197
10. Agüero-Torales MM, Salas JIA, López-Herrera AG (2021) Deep learning and multilingual sentiment analysis on social media data: an overview. Appl Soft Comput 107:107373
11. Kabir M, Kabir MMJ, Xu S, Badhon B (2021) An empirical research on sentiment analysis using machine learning approaches. Int J Comput Appl 43(10):1011–1019
12. Sayed AA, Elgeldawi E, Zaki AM, Galal AR (2020) Sentiment analysis for arabic reviews using machine learning classification algorithms. In: 2020 international conference on innovative trends in communication and computer engineering (ITCE). IEEE, pp 56–63
13. Jain PK, Pamula R, Srivastava G (2021) A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. Comput Sci Rev 41:100413
14. Yadav N, Kudale O, Rao A, Gupta S, Shitole A (2021) Twitter sentiment analysis using supervised machine learning. In: Intelligent data communication technologies and internet of things: proceedings of ICICI 2020. Springer Singapore, pp 631–642
15. Hassan SU, Saleem A, Soroya SH, Safder I, Iqbal S, Jamil S, Nawaz R et al (2021) Sentiment analysis of tweets through Altmetrics: a machine learning approach. J Inf Sci 47(6):712–726
16. Budhi GS, Chiong R, Pranata I, Hu Z (2021) Using machine learning to predict the sentiment of online reviews: a new framework for comparative analysis. Arch Comput Meth Eng 28:2543–2566
17. Patel R, Passi K (2020) Sentiment analysis on twitter data of world cup soccer tournament using machine learning. IoT 1(2):14

Disaster Management and Awareness Using Augmented Reality

Satwika Kacham, M. Lalitha, Shivani Boora, Sai Sheethal Reddy Jaina, and Siri Naidu Prasadam

Abstract Disasters can have catastrophic repercussions on people's lives, property, and the environment, so it is essential to have a strong disaster management plan in place to lessen their effects. Traditional classroom (Ye and Sitthiworachart, in 2nd International conference on artificial intelligence and education (ICAIED). IEEE [1]) learning has become monotonous for children in our modern world. Children's exposure to disasters and disaster education in their home countries needs to be examined because disaster management has an impact on how they perceive risk and how prepared they are to handle it. Traditional websites have a simple GUI that has less impact on users than interactive applications. The proposed system raises awareness about the types of disasters and their management strategies among young age groups using an engaging virtual user interface via an AR application. The application allows users to explore specific zones and view earthquake-prone states on an interactive 3D map of India, enhancing their learning experience.

Keywords Augmented reality · Vuforia · Unity 3D · Blender

1 Introduction

The study of disasters and their management is the need of the hour. The rate of occurrence of natural disasters has increased drastically over the period due to an increase in illegal human activities, overpopulation, urbanization, and industrialization, which have an adverse effect on the climatic conditions [2] on our earth, our environment, and human life. It is important to have proper knowledge of disaster management so that one can manage themselves during such unplanned and unpredicted events.

S. Kacham (✉) · M. Lalitha · S. Boora · S. S. R. Jaina · S. N. Prasadam
Computer Science (DS) Engineering, G. Narayanamma Institute of Technology and Science, Hyderabad, Telangana, India
e-mail: kachamsatwika@gmail.com

M. Lalitha
e-mail: mlalitha@gnits.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_58

Spreading awareness about managing such events helps reduce their impact on society, so people should be prepared to tackle such events. Educating people through static websites and books [3], they have lost their ability to deliver the real essence of the message being passed and have become monotonous, and their un-interactive nature has led to a loss of interest in the whole learning process and effectively educating a person. In recent times, the topic of augmented reality [4] has gained a lot of attention in the field of education due to its capabilities of bridging the gap between the real world and the virtual world and its 3D perspective, which allows the users to have an enriched experience and have a holistic view of things in terms of visualization. In this application, we have used an interactive approach using augmented reality to engage users and effectively educate everyone about disaster management. The whole application is built by using the Unity 3D gaming engine as the development platform and deploying the application, Vuforia for AR mapping of 3D points on the map, the images, and Blender for creating 3D models for the application.

Disasters and their management have become an increasingly critical area of study and concern in recent times. The rising frequency of natural disasters can be attributed to a variety of factors, including illegal human activities, overpopulation, urbanization, and industrialization. These factors have significantly impacted the Earth's climatic conditions, the environment, and human life. To mitigate the devastating effects of these events, it is imperative to equip individuals with the knowledge and skills necessary for disaster management. Raising awareness about disaster management is a key strategy in reducing the societal impact of such events and ensuring that communities are prepared to respond effectively.

Traditional methods of educating people [5] about disaster management through static websites and books have become less effective due to their monotonous nature and lack of interactivity. As a result, individuals often lose interest in the learning process, and the educational message fails to resonate effectively. However, in recent years, augmented reality (AR) technology has emerged as a powerful tool in the field of education, offering unique capabilities to bridge the gap between the real world and the virtual world. AR enhances the learning experience by providing a 3D perspective that allows users to gain a comprehensive understanding of concepts through visualization.

In this disaster management application, we have adopted an interactive approach using augmented reality technology to engage users and effectively educate them about disaster management. This application has been developed using the Unity 3D [6] gaming engine as the primary development platform. Unity allows us to create a user-friendly [7] and immersive environment for learning.

2 Methodology

The constructed augmented reality program is called Terra, which in Latin means Earth or Goddess of the Earth. Users are welcomed by an eye-catching logo that features an image of Earth and the name "Terra" as soon as the AR app is launched. The "START NOW" button is then presented to users in a smooth transition, keeping their attention on the primary action. An essential part of augmented reality, the camera is activated by the app when you click. Users scan a designated image or marker while the camera is on to start the AR experience. This marker may be programmed into the app or linked to AR material. The application continuously searches for the specified image using image recognition technologies. When recognized, it projects virtual content, such as 3D models, animations, or interactive features, over the real world to create a compelling, integrated AR experience.

The process of scanning is a complex procedure that includes the gathering and storage of 2D items seen by the device's camera. These collected frames are then sent for additional processing to the Vuforia software development kit (SDK). A key element in the development of the AR application was Vuforia's smooth integration with Unity. It boasts the capacity to identify and precisely track targets in the physical environment, including objects and images. Vuforia was a crucial tool in the creation of our application because of its powerful identification and tracking capabilities, which are essential for enhancing the physical environment with digital content and guaranteeing that users will have an engaging augmented reality experience (Fig. 1).

For picture optimization, collected frames go through pixel transformations in the Vuforia pipeline. The features of the target images kept in the database are then compared to these altered frames. These target photographs act as a frame of reference for identification. If there is a match with a target image that has been stored, it is determined by the results of this comparison. Vuforia's SDK automatically displays a corresponding 3D model or augmented content when a match is found and smoothly incorporates it into the real-world view. Vuforia makes sure that augmented material is only displayed when a recognized target is present by remaining inactive if no matches are found during the comparison. This thorough procedure ensures that the augmented reality experience will be accurate and trustworthy. Users of the AR application can go around several seismic zones, which are represented in the application's user interface, to interact with the 3D model and examine it in greater depth. With its own descriptive material and specific 3D models reflecting different states or conditions, each of these zones delivers a distinctive and educational experience.

Users can access in-depth explanations; they gain access to detailed descriptions, educational content, and visual representations of specific aspects related to earthquakes. This multifaceted method provides a thorough investigation of earthquake phenomena, which improves user engagement and knowledge. The application also includes an audio element during the navigation process to guarantee a user-friendly and immersive experience. The user interface is made more educational and interesting by this feature, which incorporates audio cues, explanations, or narrations to go along with the visual and written content. As they tour the 3D model and investigate various earthquake zones, users can listen to explanations and extra insights, resulting in a more comprehensive and instructive experience.

3 Creation of 3D Models in Blender

To create a 3D model from an image, first use Photopea to remove any unwanted sharp edges and clean the image. This was done by using the eraser tool to remove any stray pixels or the blur tool to soften the edges of the image. And then converted the image to SVG file in Adobe Express. SVG files are vector graphics files, which means that they are made up of lines and curves rather than pixels. This makes them ideal for creating 3D models. Next, import the SVG file into Blender. Clean up and remove any extras from the image. This can be done by using the knife tool to cut away any unwanted parts of the image, and then applying mesh and solidifying each element in the image. This was done by using the mesh modifier to create a 3D mesh for each element in the image. The solidify modifier then added thickness to the mesh, giving it a more realistic appearance.

Depth is added to each segment in the image and extrudes any sharp edges in the model. This was done by extruding the segments along the z-axis. This created a 3D object from the 2D image. Need to identify and clean up any inconsistencies in the model, such as gaps between segments or overlapping vertices. Merge all the similar points and join them into a single group. This helped to simplify the model and make it easier to work with. Then renamed each object present and added color and material to each segment as shown in Fig. 2. This made it easier to identify the different parts of the model and apply different materials to them. Finally, exported the model in FBX (FilmBox) format. The FBX format is a common format for 3D models. Now it can directly import the model into Unity to add functionality (Fig. 2).

4 AR Mapping and Application Building

AR mapping involves the real-time overlay of digital information onto the physical world using computer vision and sensor data, enabling the creation of interactive augmented reality applications that enhance user experiences through spatially-aware content.

4.1 AR Mapping Using Vuforia

To implement augmented reality in Unity using the Vuforia SDK (Software Development Kit), the SDK is downloaded and imported into the application. A database

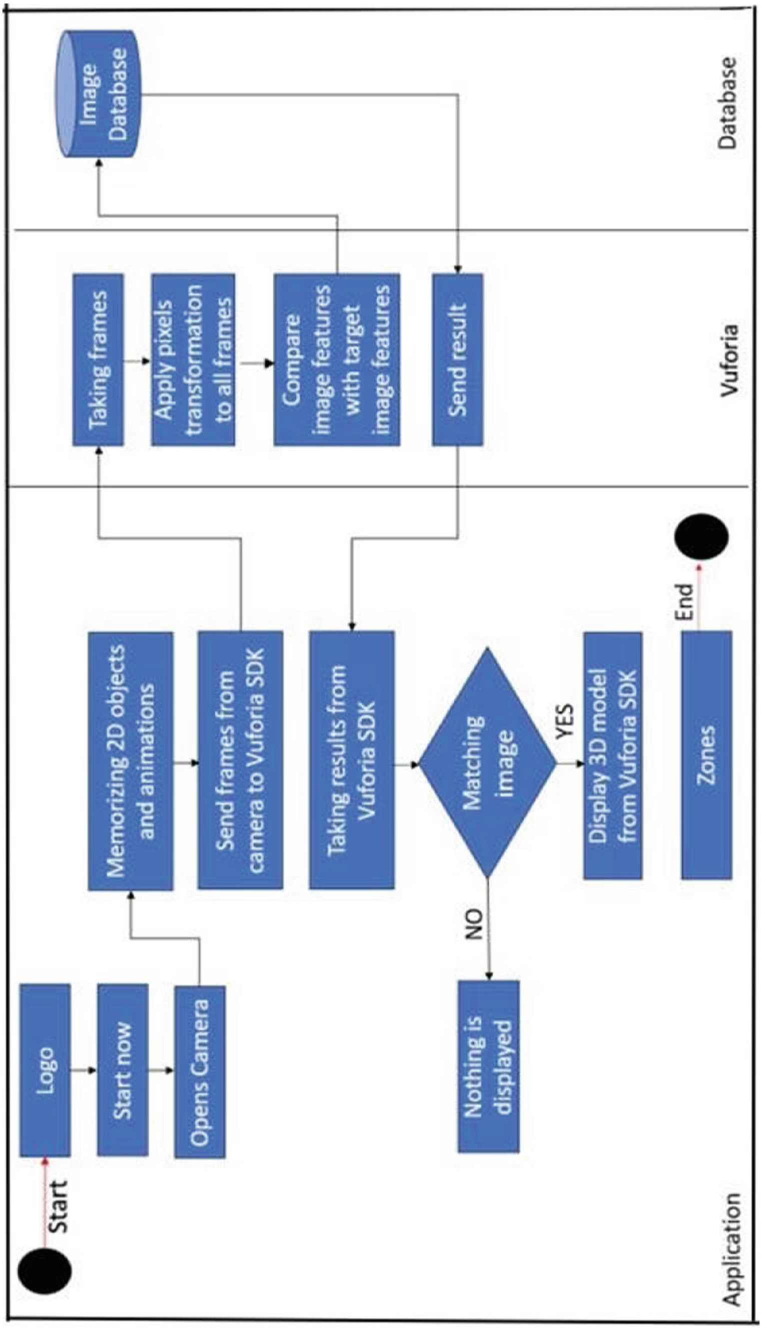


Fig. 1 Flow diagram of the application

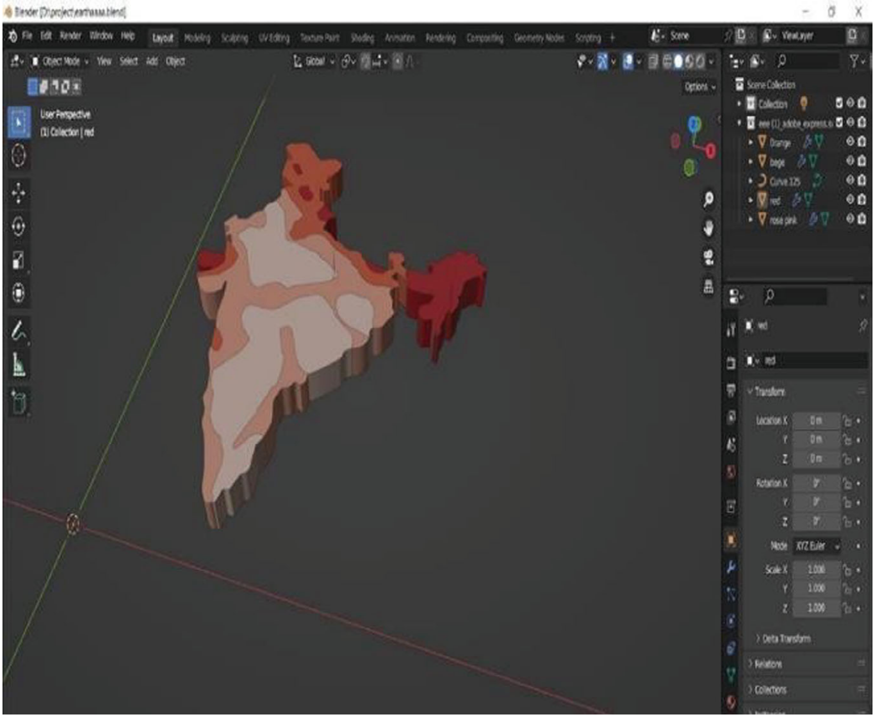


Fig. 2 Earthquake zonal map model in Blender

with a marker image is created. A scene for augmentation is created. The database and model are imported into the hierarchy of the created scene. The model should be a child of the image database. The normal camera is removed from the hierarchy, and an AR camera is added. The AR camera is used to track the marker image and display the model in augmented reality [8], and it is used to track the location of the marker image in the real world.

The Vuforia developer portal is signed into, and a license for the database is created. The license is required to use the Vuforia SDK (Software Development Kit). The Vuforia Configuration App License Key is opened in the AR Camera Inspector. The license key that was created in the Vuforia developer portal is pasted into this field. The database and Image Target behavior target components are changed in the inspector of the Database Image. These components tell the Vuforia SDK how to track the marker image and display the model in augmented reality [9]. Finally, the model is mapped to the Image Target accordingly. This means that the Vuforia SDK is told which model to display when the marker image is detected. The scene is run, game mode is opened, and the target image is shown to the AR Camera to view the results. The model should be displayed in augmented reality over the marker image [10] (Fig. 3).

Fig. 3 Mapping model with target image

4.2 Creation of Zones

Import the specified zone model [11] into the hierarchy, which is created in Blender. This can be done by dragging and dropping the model into the hierarchy view in the Unity editor. Add the most prone states to the zonal model by creating new game objects and parenting them to the zone model. From the Windows menu, add the Animation and Animator windows to the scene. The Animation window will be used to create and edit the animation, and the Animator window will be used to control the animation.

Create a new Animation and name it. This can be done by clicking on the Create button in the Animation window and selecting New Animation. Move the model to recognize the position of the model at the 0th second. By using the Animation tools in the Animation window, Start animating as desired until the first second. This can be done by dragging the playhead in the Animation window or by clicking on the Play button. At the 2nd second, copy the position of the 0th second to animate the model. Run the Animated button to view the animations (Fig. 4).

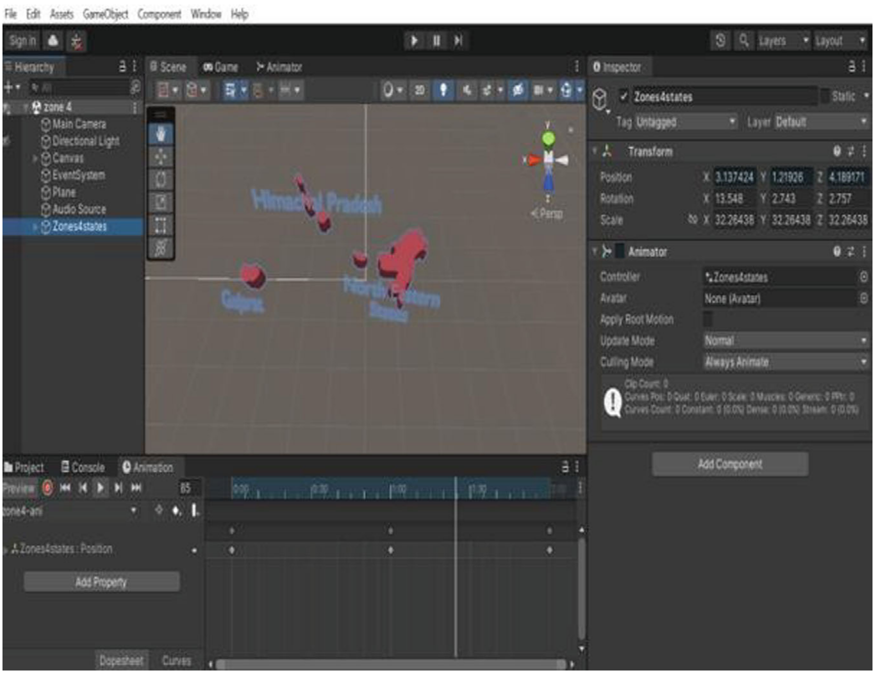


Fig. 4 Animation of the zonal model

4.3 Application Development

The first step is to add the content to the content script. This can be done by writing the content in a text file and then attaching the text file to the text game object. The second step is to create the required buttons. This can be done by creating new game objects and then setting their properties, such as their position, size, and appearance. The third step is to add a scroll bar to the content script. This can be done by creating a new game object and then setting its properties, such as its position, size, and scroll bar style. The fourth step is to make the UI look attractive by adding some colors. This can be done by setting the colors of the text, buttons, and scroll bar. The fifth step is to add an audio source to make the content read out. This can be done by creating a new AudioSource object and then setting its properties, such as its source file and playback speed.

To build an Android application in Unity, you first need to open your project and go to File > Build Settings. In the Platform pane, select Android. Then, in the Player Settings, click on the Android icon and make the necessary changes, such as the target Android version, the minimum SDK version, and the signing settings. Next, add all the scenes that you want to include in your application to the Scenes in Build list. Finally, click on the Build button to build your application [12]. The built application will be saved in the Builds folder of your project. You can then install the APK file on your Android device to view the application. Make sure that your project is set up for Android development. This includes installing the Android Build Support module and setting up the Android SDK.

5 Results

As shown in Fig. 5, the seismic map provides a general picture of the extent of earthquake damage incurred by the various areas of India. This map illustrates India's four distinctive seismic zones in various shades of red. The Department of Disaster Management of the various state governments throughout the nation primarily uses this type of map. They can prepare for a natural calamity like an earthquake thanks to this map. An Indian seismic zoning map makes it easier to determine which parts of India are least, most, and most dangerously vulnerable to earthquakes. Even these maps are examined to determine the degree of seismology in a certain location before building any high-rise structures (Fig. 6).

The nation's several seismic zones, which are clearly visible on the map, are as follows:

Zone-II: This is said to be the least active seismic zone.

Zone-III: It is included in the moderate seismic zone.

Zone-IV: This is considered to be the high seismic zone.

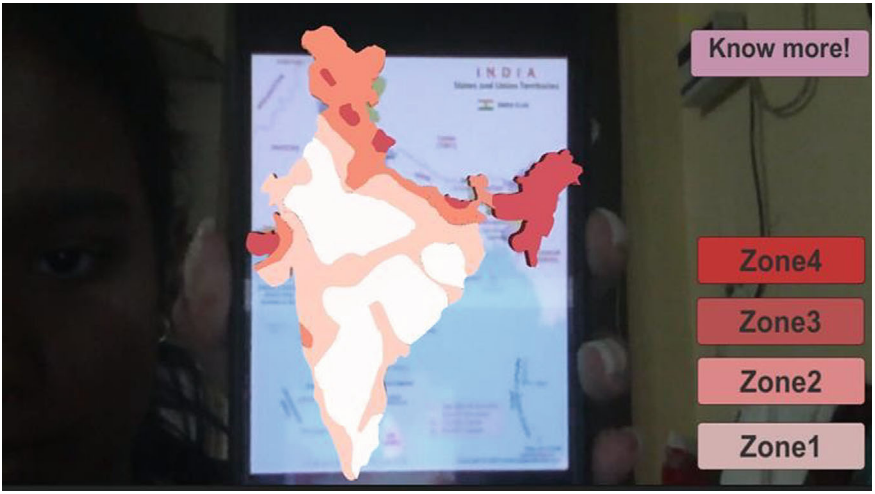


Fig. 5 Seismic zonal map of India

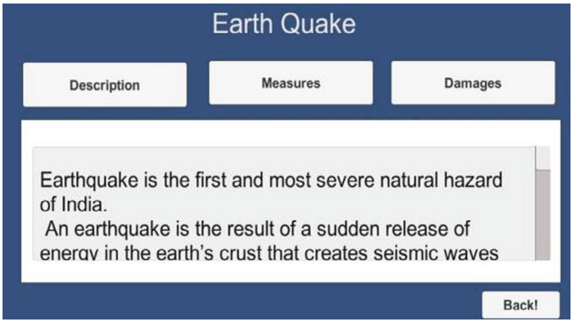


Fig. 6 Earthquake description

Zone-V: It is the highest seismic zone.

6 Conclusion

The project aims to increase awareness about earthquakes and their management strategies among young individuals by utilizing an engaging virtual platform powered by Unity's augmented reality (AR) technology. This innovative approach seeks to combat the monotony of traditional learning methods. The primary objective of the application is to offer children a comprehensive understanding of earthquakes, including their seismic zones, through a user-friendly interface. Provides detailed information about the precautions to be taken during an earthquake and the potential damage it can inflict. Furthermore, the application allows users to explore specific regions and interact with a 3D map of India, highlighting earthquake-prone states, thereby enriching their educational experience.

In addition to the earthquake model, our project will expand its scope to encompass the creation of models for other natural disasters such as cyclones, droughts, and floods, all while focusing on specific measures and management strategies in context to India. These models will be serving as valuable educational tools, providing students with essential information on disaster preparedness and safety measures. To further enhance the learning experience, we will incorporate an interactive quiz component at the end of each module, allowing students to put their newly acquired knowledge to the test. This quiz will enable students to assess their understanding of particular disaster scenarios and the corresponding safety measures. By actively participating in this engaging quiz, students can reinforce their knowledge and hone their disaster management skills in an enjoyable and interactive manner. Furthermore, our long-term vision includes the possibility of extending the application's reach worldwide, addressing disaster education needs on a global scale to ensure that individuals from different regions can benefit from this valuable resource.

References

1. Ye Z, Sitthiworachart J (2021) Curriculum system of preschool education under the background of AR intelligence. In: 2021 2nd International conference on artificial intelligence and education (ICAIED). IEEE
2. Bankoff G et al (2004) Vulnerability as preparedness: bridging the gap between climate change and disaster risk. Natural Hazards 33(2–3):243–269
3. Albala-Bertrand JM (2011) Traditional methods and knowledge for disaster risk reduction. Int J Disaster Risk Reduction 2(1):27–35
4. Milgram P, Kiyokawa K (1994) Augmented reality: a class of displays on the boundary of real and virtual. IEEE Trans Pattern Anal Mach Intell 16(10):753–762
5. Aitsi-Selmi (2012) Disaster risk reduction for children and adolescents: a review of the literature. Int J Environ Res Public Health 9(12):4338–4363
6. Bertel S, Ouellet J (2015) Unity 3D game development cookbook. Packt Publishing Ltd.
7. Koca BA, Çubukçu B, Yüzgeç U (2019) Augmented reality application for preschool children with unity 3d platform. In: 3rd International symposium on multidisciplinary studies and innovative technologies (ISMSIT), IEEE. 978-1-7281-3789-6/19/$31.00.
8. Liu L, Li H, Gruteser M (2020) Object detection in the context of mobile augmented reality. arXiv:2008.06655
9. Batuwanthudawa D, Jayasena HD (2020) Real- time location based augmented reality advertising platform. arXiv:2006.06007
10. Koc IA, Serif T, Gönen S, Ghinea G (2019) Indoor mapping and positioning using augmented reality. In: 2019 7th international conference on future Internet of Things and cloud (FiCloud). IEEE, pp 335–342
11. McGuire RK (2004) Seismic hazard and risk analysis. Earthquake Engineering Research Institute
12. Huang R, Fang C, Qiu K, Cui, L, Dong Z, Zhu S, Tan P (2021) AR mapping: accurate and efficient mapping for augmented reality. arXiv:2103.14846

Enhancing E-Healthcare: Balancing Benefits and Challenges in Securing Patient Data

Yahshaswitha Sirineni, Ashritha Reddy Mulka, and Sangeeta Gupta

Abstract Healthcare has become an indispensable part of the modern world. People have started to focus on healthcare more than ever before. With the advent of technology, healthcare records are now digitally stored, known as electronic health records (EHR). These EHRs contain sensitive and confidential information about a patient's condition and medical history. However, they are vulnerable to data breach attacks. Also, they are highly centralized, hence forming a single point of failure, which may lead to data losses, security, privacy, and trust issues between various stakeholders. The mentioned problems pose a need for a secure, reliable system that maintains privacy among the users. The paper proposes a blockchain-based healthcare management system to solve the problem above. The system is based on Hyperledger Fabric, a distributed ledger platform that helps us achieve security, confidentiality, flexibility, and scalability. The model aims to use smart contracts and digital signatures to achieve modularity, role-based working, and a high level of security while performing all the necessary functions efficiently.

Keywords Blockchain · Hyperledger · Smart contracts · Digital signatures

1 Introduction

In the past, a patient's health data and records were written on paper and stored in file cabinets to ensure safety. However, electronic health records (EHR) have become essential to the modern-day healthcare system. They allow a faster, more convenient

Y. Sirineni (✉) · A. R. Mulka · S. Gupta
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: yahshaswitha.sirineni@gmail.com

A. R. Mulka
e-mail: ashrithareddy.mulka@gmail.com

S. Gupta
e-mail: sangeetagupta_cse@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_59

way to share and store data among various entities. However, they use centralized systems to store data, making it prone to attacks and losses. Centralized systems also act as a bottleneck and cause a significant performance problem. They do not offer trusted sharing among various organizations involved in the healthcare system. As most of these systems store the data in the same form in the database without encryption, accessing and changing the data becomes simpler.

Over recent times, blockchain technology has gained popularity for the various benefits associated with it. It is immutable, making it very difficult to alter or delete the data over the network. Blockchain, being a decentralized network of nodes, solves the problem of having a single point of failure. The decentralized network also ensures that there is faster availability of the data. The transparency in the system helps the users build trust, making them believe that the data provided to the network is safely stored.

Various cryptography techniques can be applied to the blockchain network to encrypt the transactions, making it difficult for any attacker to access them. These cryptographic techniques can also maintain a role-based system using concepts such as public and private keys. For example, digital signatures use these asymmetric cryptographic techniques to verify the authenticity and validity of any transaction.

With its benefits, blockchain can be applied to healthcare systems to ensure security, privacy, modularity, transparency, availability, and performance. In this paper, we have proposed a multi-module-based secure fabric health care architecture. The proposed model tries to cover all basic quality requirements required for a blockchain-based healthcare system, as mentioned in [1].

2 Literature Survey

Marry et al. [2] introduced a model based on decentralized blockchain network built on Hyperledger Fabric. The primary objectives of this system are to bolster data security, foster interaction, empower patients with greater control over their health information, streamline the processes associated with clinical researches, and ultimately enhance healthcare outcomes. Once logged in, the system admin possesses the right to include doctors or hospitals within the system networks. Patients, on the other hand, have the capability to manage access rights to their health records, allowing them to revoke authorization at any point of time. Assuming that the patient's health records are up-to-date and accessible through the blockchain system, this framework enables doctors to conduct more comprehensive assessments of a patient's health. Furthermore, doctors can record any prescribed treatments or medical procedures directly on the blockchain during a patient's visit.

Asha et al. [3] used an immutable ledger to build a healthcare system with three main entities, i.e., doctors, patients, and healthcare providers. The patients can enter their health records and gain access to doctors or healthcare providers. The patient uses a digital signature to grant access rights. This signature must be used during every transaction. Doctors or healthcare providers with the required access can get the patient's corresponding health records. Other forms of data, such as drug details, are given access rights by the admin, who decides who will have the access rights.

A hyperledger and Interplanetary File System (IPFS) model was built by Ramar et al. [4] based on a Social Security Number (SSN). The transactions are stored as blocks, which are secured using cryptography, and every block has a private key, which can be verified with a public key. The patient is the primary owner of his data. He can grant access to data to different hospitals. Digital signatures are used to sign the documents using the private key and transaction data. The transaction's authenticity can be checked by validating the signature and a public key.

Sheeraz et al. [5] proposed a model using Hyperledger Fabric. To start with, organizations are created in the network, which then can create different admin users. Every admin user can have a different role, such as only downloading data or downloading and uploading and modifying data. After connecting with the network, the user can try to upload the data. The data uploading is stored in the IPFS, and the address of the data is hashed and stored in the Hyperledger Fabric-based blockchain network. To download the data, the admin is authenticated using the JWT token and the data from IPFS is presented to the user.

Saranya et al. [6] also build a healthcare system using Hyperledger Fabric. The system contains a chain of different hospitals. Every hospital has a list of patients and doctors associated with it. Every patient has control over their data and can grant different types of access to other doctors to either view or modify the records. Smart contracts define the association between the various entities and maintain access. If any patient moves out of a hospital, his records are removed from that particular hospital and assigned to the new hospital to which he has transferred.

Sinha et al. [7] proposed a blockchain-based healthcare system to manage electronic health records (EHR). The system contains users with the right to authorize hospitals of their health records. Smart contracts are used to develop relationships between various entities. The data is stored as a chain of blocks. Every block in the chain has a hash of the next block, which helps to identify the order of the blocks in the chain. Digital signatures are used to authenticate the data in the blocks and ensure that authorized authors can only access the data. A machine learning module is applied at every node, which trains itself with the data present. The ML helps to check any discrepancy in the prescription and ensures that anything prescribed by the doctor goes in hand with the patient's symptoms. They also help the doctor by predicting the remedies, which allows the doctors to speed up the process.

Haddad et al. [8] proposed a blockchain model for healthcare systems with sharing controls. In the healthcare data management workflow, data is collected from trusted sources like MedLine. This data is then processed using machine learning, particularly convolutional neural networks (CNN), for extracting valuable medical insights while protecting patient privacy. The categorized medical data is securely stored on a blockchain (NEM framework) to ensure data security and integrity. Subsequently, this data is uploaded to a cloud-based system, improving accessibility for authorized healthcare professionals and researchers, with patients retaining control over access permissions.

Ukanah et al. [9] proposed a model which involves secure storage of electronic medical records (EMRs) using blockchain technology. EMRs are stored off-chain, reducing costs, and utilizing the Interplanetary File System (IPFS) for efficient content-addressable storage. Smart contracts within the blockchain govern data ownership, metadata, permissions, and integrity. Patients and healthcare providers have private keys for restricted access. Patients can access their medical data continuously, and the blockchain maintains a secure log of patient-provider relationships.

Mukherjee et al. [10] presented an innovative Web-application that leverages blockchain technology to safeguard patient data. This application empowers patients by allowing them full authority over which medical records they share with healthcare professionals. It generates a list of suggested experts based on real-time patient needs, thereby reducing the need for multiple consultations, and ensuring patients receive the most convenient advice and subsequent treatment.

3 Proposed Framework

3.1 Workflow

The four main entities of the system are patients, doctors, pharmacists, and insurance providers. Whenever a new patient visits the hospital, an account is created for him. The patient can then consult any doctor and provide him with read access to his past prescriptions, which can be studied by the doctor helping in the treatment. If the patient provides write access to the doctor, the doctor can add a new prescription to the patient's data chain. All these transactions are digitally signed to maintain their legitimacy. Further, patients can provide view access to prescriptions to a pharmacist who can digitally view the prescription and issue medicines. This step ensures that the patient is only issued the prescribed drugs. Similarly, the patient can grant access to the insurance providers who can keep track of all the transactions over time, confirm their validity, and automate the insurance claim process. The workflow of the system is depicted in Fig. 1.

3.2 Tech Stack

The proposed model mainly consists of the frontend, backend, and blockchain network. React JS is proposed for the frontend and Django for the backend. The blockchain network consists of Hyperledger Fabric which communicates with the backend using Fabric SDK. The Hyperledger Fabric consists two types of databases: Couch DB and Level DB. In the proposed model, the data is completely stored in JSON format. We do not store data as files because if we do so, the files would have

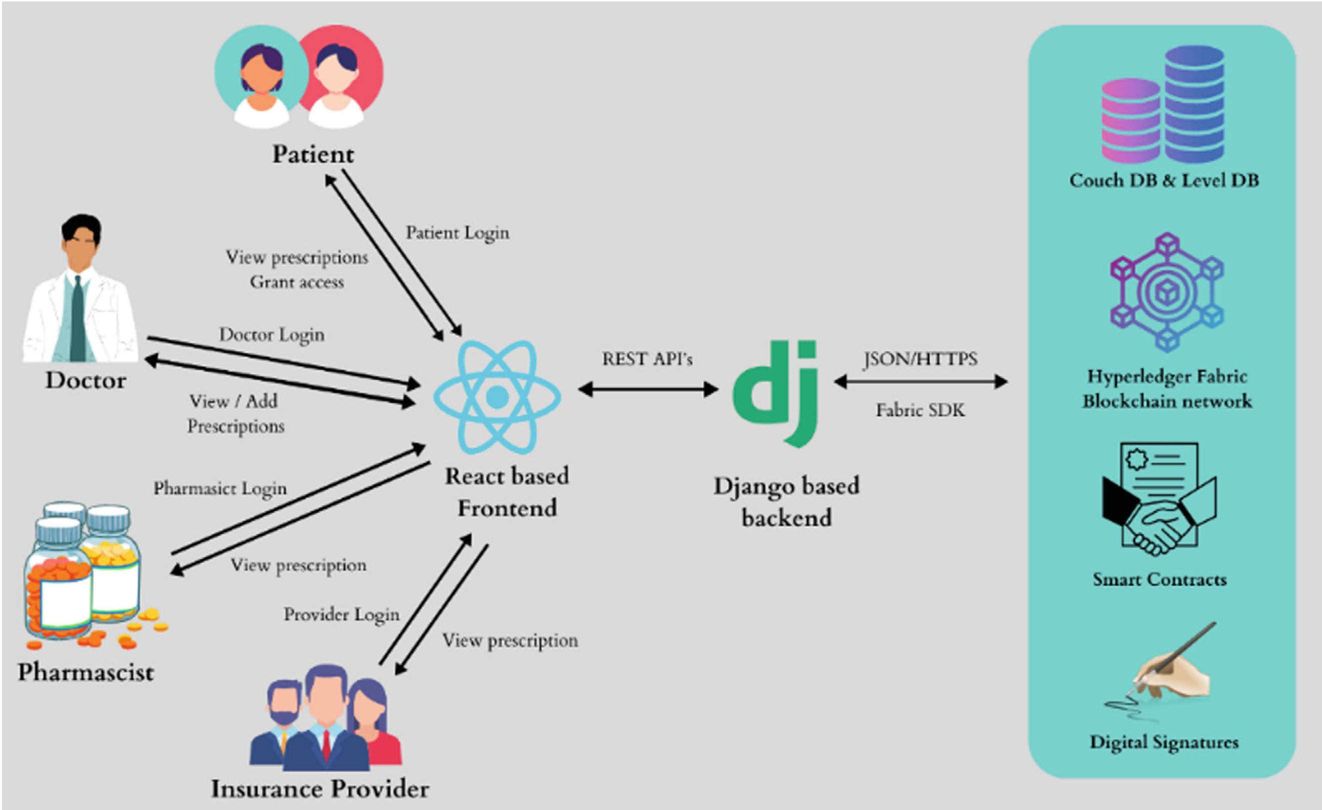


Fig. 1 Multi-module-based secure fabric health care architecture

to be stored in a cloud storage or any other file storage system, making the storage system a single point of failure and removing the essence of a decentralized network.

3.3 Smart Contracts

Smart contracts are self-executing codes that automatically execute and enforce themselves when predefined conditions are met. Two main types of smart contracts are proposed in the system. The first one is to manage patient consent when a patient's data needs to be accessed or shared with doctors, healthcare providers, or pharmacists. The contract should enforce strict access controls and permissions, ensuring only authorized entities can access the data. The second type is to automate the health insurance claims. When a healthcare service is provided, the contract should automatically verify the service, check the policy, and trigger a payment to the healthcare provider if the claim is valid.

3.4 Digital Signatures

Digital signatures are used to verify the legitimacy of the transactions in the system. The Elliptic Curve Digital Signature Algorithm (ECDSA), an asymmetric cryptographic algorithm, is used. The transaction hash is calculated by passing the transaction data through a cipher, and the hash is encrypted using a private key and saved as the digital signature. Every private key has a corresponding public key shared with other transaction stakeholders. Anyone who wants to check if the transaction data is valid and has not been tampered with can pass the transaction data through the same cipher, create a hash and encrypted using the public key. The transaction is valid if this encrypted data is the same as the digital signature. Three different key sizes can be used for this encryption. As the key size increases, the security increases, but the computational time and complexity also increase. The reason that this algorithm is proposed is because it is compatible with the Hyperledger Fabric network and offers a good balance between security and computational requirements.

4 Comparative Analysis

Table 1 gives a comparative analysis between various models built for healthcare system using blockchain technology. The comparison between them reveals intriguing insights into their potential benefits and limitations. Ultimately, the choice between these blockchain models for healthcare systems depends on the specific use

Table 1 Comparison of various models

Proposed/ developed model	Blockchain network	Read–write permissions	Image-based data	Digital signatures
Sirineni, Reddy and Gupta	Hyperledger	Yes	No	Yes
Marry et al. [2]	Hyperledger	No	No	No
Asha et al. [3]	Ethereum	Yes	Yes	Yes
Ramar et al. [4]	Hyperledger	No	Yes	Yes
Saranya et al. [6]	Hyperledger	Yes	No	No

case and regulatory environment. Healthcare organizations must weigh the trade-offs to determine which model aligns best with their objectives and compliance requirements.

5 Challenges and Limitations

One of the significant challenges faced by blockchain-based systems is high maintenance costs. As the complexity and size of the data increase over time, data with each node rises, causing a severe performance and cost challenge. A limitation of the databases supported by Hyperledger Fabric, i.e., Couch DB and Level DB, is that only key-value pairs or JSON-formatted data can be stored in them. Hence, any other form of data, such as visual reports, cannot be held in the system. To store any such data, we would have to store them in a cloud and store the hash of the address in the database. However, the cloud would then become a single point of failure or bottleneck of the system, hence trading-off the main aim to create a decentralized system.

6 Conclusion

The proposed model in the paper overcomes various problems faced by electronic healthcare systems, such as security, privacy, modularity, transparency, availability, and performance. The system is helpful for all the stakeholders involved in a healthcare system as it ensures the safe passage and storage of data without any corruption. Blockchain technology has the potential to improve significantly healthcare by providing secure and transparent data management. It can protect patient information, streamline processes, and enhance trust among healthcare stakeholders. While there are challenges to overcome, the benefits of blockchain in healthcare are promising, paving the way for a more efficient and patient-centered healthcare system in the future.

References

1. Kassab M, DeFranco J, Malas T, Destefanis G, Graciano Neto VV (2019) Investigating quality requirements for blockchain-based healthcare systems. In: 2019 IEEE/ACM 2nd International workshop on emerging trends in software engineering for blockchain (WETSEB), Montreal, QC, Canada, 2019, pp 52–55. https://doi.org/10.1109/WETSEB.2019.00014
2. Marry P, Yenumula K, Katakam A, Bollepally A, Athaluri A (2023) Blockchain based smart healthcare system. In: 2023 International conference on sustainable computing and smart systems (ICSCSS), Coimbatore, India, 2023, pp 1480–1484. https://doi.org/10.1109/ICSCSS57650.2023.10169704.
3. Asha V, Nirmala AP, Saju B, Sk A, Das AK, Sharma A (2023) Role of blockchain technology in healthcare. In: 2023 International conference on sustainable computing and data communication systems (ICSCDS), Erode, India, pp 370–374. https://doi.org/10.1109/ICSCDS56580.2023.10104832.
4. Ramar K, Shanmugasundaram, GPVH, Andraju BP, Baskar S (2022) Digital Healthcare using Blockchain. In: 2022 1st International conference on computational science and technology (ICCST), Chennai, India, 2022, pp 651–655. https://doi.org/10.1109/ICCST55948.2022.10040411
5. Sheeraz MM, Mozumder MAI, Khan MO, Abid MU, Joo M-I, Kim HC (2023) Blockchain system for trustless healthcare data sharing with hyperledger fabric in action. In: 2023 25th International conference on advanced communication technology (ICACT), Pyeongchang, Korea, Republic of 2023, pp 437–440. https://doi.org/10.23919/ICACT56868.2023.10079423
6. Saranya R, Murugan A (2023) A hyperledger fabric-based system framework for healthcare data management. In: 2023 7th International conference on computing methodologies and communication (ICCMC), Erode, India, 2023, pp 552–556. https://doi.org/10.1109/ICCMC56507.2023.10083620
7. Sinha A, Patel A, Jagdish M (2022) Application of blockchain in healthcare. In: 2022 First international conference on artificial intelligence trends and pattern recognition (ICAITPR), Hyderabad, India, 2022, pp 1–4. https://doi.org/10.1109/ICAITPR51569.2022.9844186
8. Haddad A, Habaebi MH, Islam MR Zabidi SA (2021) Blockchain for healthcare medical records management system with sharing control. In: 2021 IEEE 7th international conference on smart instrumentation, measurement and applications (ICSIMA), Bandung, Indonesia, 2021, pp 30–34. https://doi.org/10.1109/ICSIMA50015.2021.9526301
9. Ukanah O, Obimbo C (2020) Blockchain application in healthcare. In: 2020 International conference on computational science and computational intelligence (CSCI), Las Vegas, NV, USA, 2020, pp 1167–1172. https://doi.org/10.1109/CSCI51800.2020.00218
10. Tidke SK, Khedkar V, Banerjee A, Mulik A, Goyal A, Chhabaria Y (2022) An interactive and secure Blockchain web portal for online healthcare services. In: 2022 International conference on decision aid sciences and applications (DASA), Chiangrai, Thailand, 2022, pp 454–459. https://doi.org/10.1109/DASA54658.2022.9764973

Methods and Techniques of Cybersecurity Intrusion Detection: Supervised Machine Learning

Tahseen Jahan, A. Hemalatha Reddy, Jalapala Sinjini, M. Priyadharshini, V. Indumathi, and V. V. Bhavani

Abstract Cybersecurity intrusion detection is crucial for safeguarding digital systems against malicious activities. Supervised machine learning (SML) has emerged as a powerful tool in this domain, offering effective detection capabilities. This paper explores various methods and techniques employed in cybersecurity intrusion detection using SML. It discusses the application of supervised learning algorithms, such as support vector machines, decision trees, and neural networks, in detecting intrusions. Additionally, it examines feature selection, dataset preprocessing, and performance evaluation strategies specific to SML-based intrusion detection systems. The paper also highlights current challenges and future directions in this field.

Keywords Cybersecurity · Intrusion detection · SML · SVM · Decision trees · Neural networks

T. Jahan (✉) · A. H. Reddy · J. Sinjini · M. Priyadharshini · V. Indumathi · V. V. Bhavani
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: tahseen048@cmritonline.ac.in

A. H. Reddy
e-mail: hemareddy@cmritonline.ac.in

J. Sinjini
e-mail: sinjini@cmritonline.ac.in

M. Priyadharshini
e-mail: priyadharshini@cmritonline.ac.in

V. Indumathi
e-mail: indu11mathi@cmritonline.ac.in

V. V. Bhavani
e-mail: bhavaniv@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_60

1 Introduction

Supervised machine learning (SML) plays a pivotal role in enhancing cybersecurity through intrusion detection systems (IDS) [1]. These systems are crucial for identifying malicious activities within computer networks and protecting against cyberthreats. SML techniques empower IDS by enabling them to learn patterns of normal and abnormal behaviour from labelled data, thus facilitating accurate detection of intrusions. Various methods and techniques are employed in SML-based intrusion detection, including but not limited to support vector machines (SVM) [2], random forests [3], decision trees, and neural networks [4]. SVM, for instance, excels in binary classification tasks by identifying an optimal hyperplane that separates different classes of data points. Random forests leverage an ensemble of decision trees to enhance classification accuracy, while neural networks emulate the human brain's learning process, allowing for complex pattern recognition. Data preprocessing, feature selection, and model evaluation are critical stages in deploying SML for intrusion detection. Preprocessing involves cleaning and transforming data to ensure its quality and relevance. Feature selection aims to identify the most informative attributes for training the model, optimizing its performance. Model evaluation assesses the IDS's effectiveness in terms of accuracy, precision, recall, and F1-score.

Cybersecurity Intrusion Detection.

Protecting a company's digital assets requires cybersecurity intrusion detection. System and network intrusion prevention entails keeping an eye out for any suspicious behaviour or policy breaches that can jeopardize the availability, secrecy, or integrity of data and resources. The primary goal of intrusion detection is to identify instances of misuse, illegal access, or unusual activity so that corrective measures can be taken [5]. Intrusion detection makes use of a wide range of methods and tools. One popular approach is signature-based detection, which involves comparing system or network behaviour with a database of recognized attack patterns. Security staff are notified through an alert system in the event that a match is detected. However, anomaly-based detection works by first defining typical behaviour and then marking any changes from this baseline as possible intrusions. Finding zero-day attacks, or threats that have not been seen before, is where this method really shines [6]. Intrusion detection also makes use of behavioural analysis, which is concerned with keeping tabs on how users and other entities operate in order to spot any unusual or suspect activity that could point to a security breach. The best way for security teams to identify and counteract any risks is to examine patterns of behaviour that differ from typical usage patterns. By using rules or algorithms that are based on heuristics, potentially harmful behaviour can be identified. When new or changing threats are detected, this method can help catch them before signature-based detection does.

To make intrusion detection systems more accurate and effective, machine learning and AI approaches are being used more and more. Over time, machine learning algorithms can improve security issue detection by analysing massive amounts of data and discovering patterns that indicate malicious behaviour. Organizational networks are fortified with intrusion detection systems like NIDS and HIDS to keep tabs on system activities and traffic in real time, and security information and event management solutions gather and analyse log data from different sources to identify and handle security incidents.

Combinations of methods and technologies, as well as the ability to continuously monitor, analyse, and respond, are necessary for effective intrusion detection. The ever-changing world of cybersecurity makes it imperative for organizations to maintain a state of constant vigilance and to upgrade their intrusion detection systems on a regular basis (Fig. 1).

Intrusion Detection System

The many stages that can be used as defence mechanisms are illustrated in Fig. 2. Included in these stages are the following: avoiding, identifying, and reacting to breaches. Attacks are averted in advance during the preventative phase.

To monitor host and network data and detect intrusions, analytical tools are created during the detection phase. To address potential intrusions identified by IDS, response tools are utilized. A series of events that go against established security protocols is called an incursion. An intrusion prevention system (IPS) is any kind of defence that may stop attacks in their tracks. Instant prevention systems (IPSs) are intrusion detection systems (IDSs) with the added capacity of blocking identified assaults. But in the current distributed system, it is not feasible to block attacks early on. An intrusion detection system (IDS) is often a combination of hardware and software that keeps tabs on a computer system's activity and finds any suspicious activity. A security countermeasure called an intrusion response system (IRS) is employed to prevent intrusions that have been discovered, based on the alerts from the IDS.

Types of IDS

Intrusion detection systems (IDS) are broadly categorized into two main types based on their detection approach:

In the realm of intrusion detection systems (IDS), there are two primary approaches: signature-based IDS and anomaly-based IDS.

Signature-based IDS, also referred to as knowledge-based or misuse detection systems, function by comparing incoming data packets or events against a database of known attack signatures or patterns. When a match is identified, the system triggers an alert. This method is highly effective at detecting known threats, yet it may struggle when faced with zero-day attacks or novel threats that have not been previously documented [8].

Anomaly-based IDS operate by establishing a baseline of normal network or system behaviour. Subsequently, they continuously monitor current activity and compare it against this established baseline. Any deviations from the expected norm are flagged as potential intrusions. This approach is particularly adept at detecting previously unknown threats and zero-day attacks, as it does not rely on predefined signatures. However, there is a risk of generating false positives if the baseline of normal behaviour is not accurately defined and regularly updated [9].

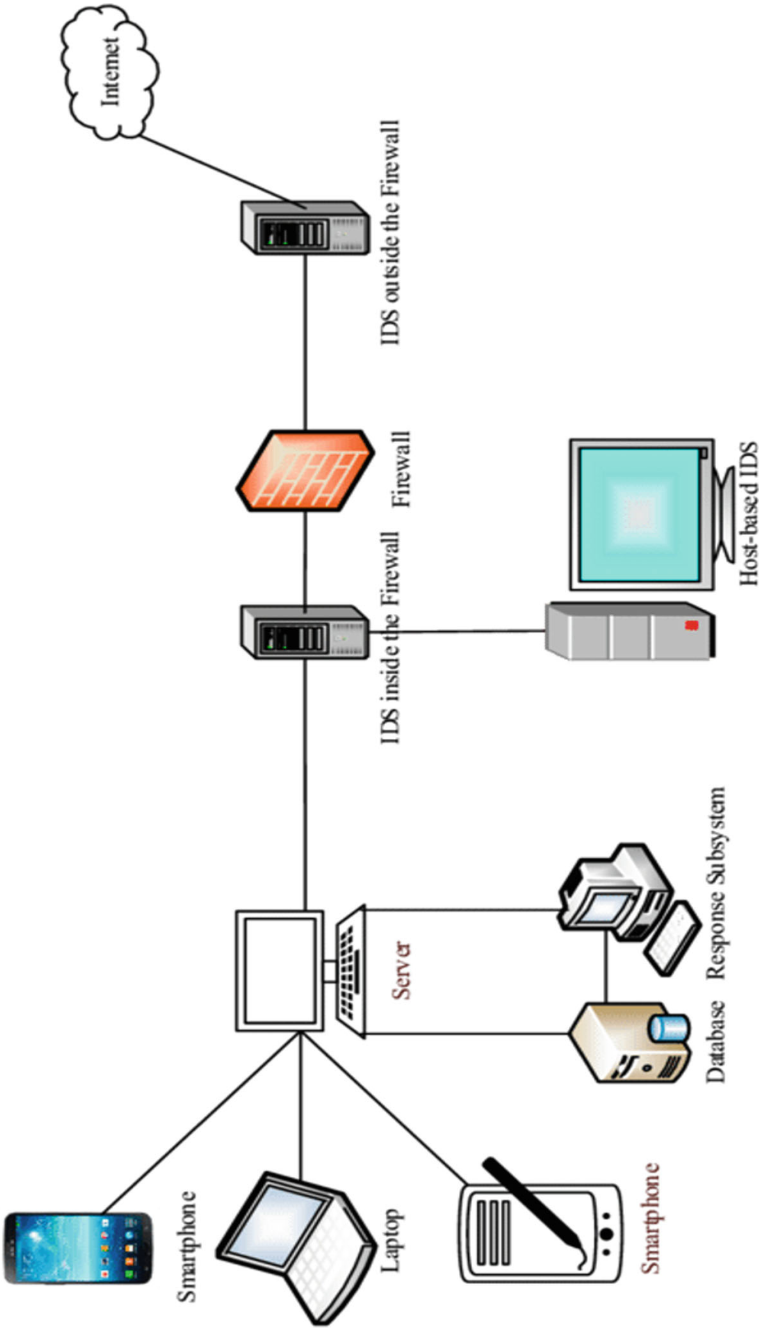


Fig. 1 Intrusion detection system architecture [7]

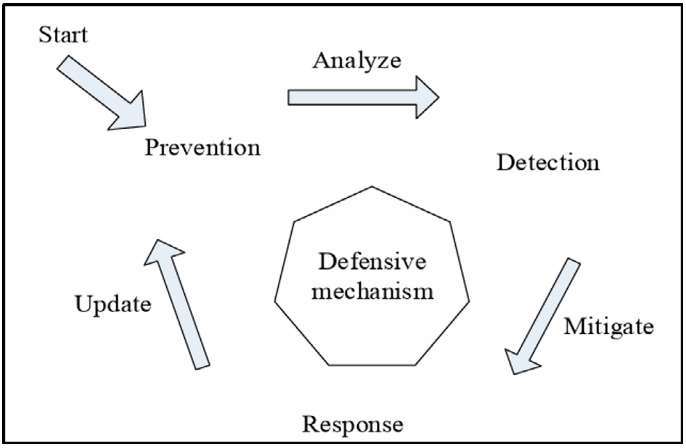


Fig. 2 Defensive life cycle [7]

There are also hybrid IDS that combine elements of both signature-based and anomaly-based detection for improved accuracy and coverage.

Furthermore, IDS can be deployed in various network locations and can operate at different levels of the OSI model:

Network-based IDS (NIDS) operates by monitoring network traffic in real time, scrutinizing packets traversing the network to identify any suspicious activities. Typically, NIDS are strategically positioned within the network, overseeing traffic between internal network segments or between the internal network and external networks, such as the Internet.

- Host-based IDS (HIDS) are designed to reside on individual hosts or servers, actively monitoring activities within the host's operating system and application logs. This type of IDS is particularly adept at detecting insider threats and attacks aimed at specific hosts or applications.
- Network behaviour analysis (NBA) takes a different approach by analysing network traffic patterns and behaviours over time, aiming to pinpoint deviations from normal behaviour. NBA systems are often integrated with traditional IDS to offer a more comprehensive understanding of network activity.
- Protocol-based IDS specializes in analysing specific network protocols like HTTP, FTP, and SMTP. By scrutinizing traffic patterns and anomalies within these protocols, they can detect potential intrusions effectively.
- Application protocol-based IDS operates similarly to protocol-based IDS but focuses exclusively on application-layer protocols such as DNS, HTTP, and SSH. By examining the behaviour of these protocols, they can detect signs of intrusion or misuse.
- Distributed IDS (DIDS) consists of multiple IDS sensors spread across various points within a network, working collaboratively to monitor and analyse network

traffic comprehensively. DIDS offers broader coverage and enhanced scalability compared to single-point IDS deployments.

Each type of IDS has its strengths and weaknesses, and the choice of which type to deploy depends on factors such as the organization's security requirements, network architecture, and available resources.

Intrusion Response System (IRS)

To safeguard the HIS and restore normal system performance after an intrusion, an intrusion response system (IRS) must be in place as part of the self-protecting system. Since the NFA and IRS modules work closely together, it will be more difficult for future attacks with comparable signatures to succeed. Two essential components of an autonomous IRS are the dynamic evaluation of suggested replies and the configuration of appropriate responses. Protection methods known as responses can prevent or lessen the impact of cyberattacks. Different forms of cyberattacks necessitate different configurations and implementations of protection systems to safeguard the system. Nonetheless, distributed denial of service (DDoS) and SQL injection assaults cannot be effectively removed by these answers [10].

2 Background

Security issues have emerged alongside the proliferation of internet use, according [11]. They brought out the fact that, because of the systems' security holes, malicious software may interfere with system operation and compromise data confidentiality. To identify and report intrusions, they pointed out that intrusion detection systems (IDS) had been created. Additionally, they mentioned that IDS systems have increasingly relied on AI-based techniques. Many IDS systems have been developed utilizing the CSE-CIC IDS-2018, UNSW-NB15, ISCX-2012, NSL-KDD, and CIDDS-001 datasets, and this paper thoroughly reviews the literature employing these sets. Beyond that, they normalized the datasets using max–min and classified the data using traditional machine learning methods including K-nearest neighbour (KNN), SVM, and decision tree (DT). They concluded that some of the research mentioned in the literature had achieved more favourable outcomes. They reasoned that the research might be helpful in creating IDS systems based on AI techniques like machine learning.

In their discussion of network security issues, [12] highlighted intrusion detection as a useful tool. A research comparing the efficacy of intrusion detection methods based on supervised probabilistic and predictive machine learning was previously provided. They used decision trees and random forests for prediction, as well as Naive Bayes and Gaussian for probabilistic analysis. Various training datasets were used, all derived from the KDD99 dataset. They tested four different types of attacks—DoS, Probe, R2L, and U2R—and compared the effectiveness of each method. The statistical findings demonstrating how each method reacts to a dataset's attack population were also published. They computed standard deviations with regard to the detection rates of each assault type and compared the strategies' performance; this allowed them to assess the resilience of each methodology.

Cybercriminals are more creative in their attacks on network services, as [13] noted, due to the growing demand for Internet resources. As a result, they realized that network security would soon be an integral aspect of the system. They have also mentioned the requirement of having a strong intrusion detection system (IDS) to identify these types of assaults efficiently and effectively. An intrusion detection system (IDS) is a tool that monitors a network or system and analyses each packet in detail to find malicious activities, as they had mentioned. According to their statement, intrusion detection systems (IDS) primarily function to detect suspicious or undesirable activity and notify the network administrator of it. They went on to say that IDS were crucial for network administrators to protect their systems from known and undiscovered threats that might compromise network resources. They brought up the possibility of using machine learning techniques to create a more effective intrusion detection system (IDS). A total of four machine learning algorithms—RF, DT, MLP, and SVM—were employed for data categorization in their study. According to what they said, all these machine learning models were tested on the NSL-KDD dataset with trained. The purpose of feature choices, as they had previously described, was to filter out superfluous or unneeded characteristics from the dataset. They went on to say that feature selection decreased computational complexity by reducing the dataset's dimensionality. They determined that the suggested model achieved a classification accuracy of over 99.99 percent when tested with three feature subsets taken at random from the NSL-KDD dataset.

In their article, [14] revealed that they have studied intrusion detection with supervised machine learning techniques. Their original statement indicated that classifying supervised machine learning methods and associated intrusion detection systems was the primary objective. Cybersecurity assaults, supervised machine learning methods, and intrusion detection systems were all thoroughly covered in their presentation. After then, they have highlighted pertinent efforts in relation to supervised learning's application to intrusion detection.

In order to determine if network traffic was harmful or not, a new supervised machine learning system was created, according to [15]. They shared that a mix of supervised learning algorithms and feature selection methods were employed to determine the optimal model taking detection success rate into account. In terms of identifying network traffic, the researchers said that the study's findings showed that ANN-based machine learning with wrapper feature selection performed better than the SVM approach. They went on to say that the NSL-KDD dataset was used to categorize network traffic using supervised machine learning techniques like SVM and ANN in order to evaluate the performance. They also mentioned that a research comparing the suggested model to other models found that it outperformed the competition in terms of the success rate of intrusion detection.

According to [16], the difficulty of achieving cybersecurity has been steadily rising in recent years due to the alarming development in the number of computer-related applications and the connection of these devices. A suitable defence mechanism against multiple cyberattacks was also something they had brought up. As a result, they expanded their effort to include creating an intrusion detection system (IDS) that may play a role in cybersecurity and identifying network assaults and inconsistencies. They have detailed how a data-driven intrusion detection system had been created using AI, namely machine learning methods. According to their paper, they used a number of well-known machine learning classification algorithms to identify cybersecurity breaches caused by intelligent service providers. They concluded by mentioning that they had examined the efficacy of several experiments using cybersecurity datasets that included multiple types of cyberattacks and measured the results using accuracy, precision, recall, and F1-score.

Cybersecurity has emerged as a critical field of study and [17] noted that networks are integral to contemporary life. They found out that a crucial cybersecurity tool, an intrusion detection system (IDS), tracked the health of the network's software and hardware. In addition, they mentioned that current IDSs continue to struggle with identifying new assaults, lowering the false alarm rate, and increasing detection accuracy, even after decades of research. Many researchers had concentrated on creating intrusion detection systems (IDSs) that employed machine learning techniques to address the aforementioned issues, as they had described. The authors had claimed that, with little human intervention, machine learning algorithms could reliably identify outliers in datasets. Additionally, they had noted that machine learning approaches may identify unknown threats due to their excellent generalizability. At long last, they admitted that deep learning was a subfield of machine learning; the field's exceptional success had made it a popular site for academic inquiry.

According to [18], while technological advancements have undoubtedly improved people's lives, they have also exposed several security vulnerabilities. Numerous individuals, groups, businesses, etc., were impacted by the evolving nature of assault types, according to their statement. To prevent monetary and psychological harm, they had further stated, intrusion detection systems were created. They had previously said that the CICIDS2017 dataset, which included both benign and the most recent prevalent assaults, was utilized in their research. They professed to have used the Fisher score method to choose the best characteristics. KNN, SVM, and decision tree (DT) algorithms were used to classify real-world data collected from the dataset as either DDoS or benign. The investigation yielded success rates of 0.9997%, 0.5776%, and 0.99%, respectively, as previously stated.

Critical infrastructures (CIs) such as power grids, water distribution systems, and nuclear power plants, rely heavily on supervisory control and data acquisition (SCADA) systems for remote access, monitoring, and control [19]. They had claimed that SCADA systems and CIs were more vulnerable to security threats due to the increasing interconnection, standardized communication protocols, and remote access of contemporary SCADA systems. The SCADA modules and communication networks are strategically important to the functioning of CIs, and any invasive action on them might have terrible effects for nations, according to their addition. The significance of quickly detecting and classifying intrusions into SCADA systems was highlighted as crucial for the operational stability of national CIs. They had detailed how several literature works had suggested a plethora of supervised learning approaches for SCADA intrusion detection and classification (IDC) because of their widely acknowledged and demonstrated efficacy. They boasted that their study offered a critical analysis of current research that modelled SCADA intrusion solutions using supervised learning techniques. The authors had previously said that the paper's goals were to identify important unanswered questions, provide suggestions for future research, and advance the state-of-the-art. Additionally, they had clarified that the goal was to serve as a research resource for scholars investigating the safety of industrial control systems. Lastly, they said that they had conducted a thorough evaluation of several supervised learning methods for SCADA IDC systems, comparing and contrasting them based on methodology, datasets, testbeds, optimization and feature engineering processes, and classification procedures.

According to [20], the Internet has greatly altered economic, political, and social systems, making geographical barriers obsolete in many instances. Because of the Internet's tremendous benefits to commercial transactions and its relative simplicity of use, they had seen, more people were using the Internet, which meant more potential invaders. Security of computer resources should be prioritized, and they emphasized the need of using intrusion detection systems (IDS) with intrusion prevention systems (IPS). Their reasoning was that the conventional rule-based method was inadequate for handling the massive amounts of network traffic that are created every few seconds. As a result, researchers have turned to data mining techniques for intrusion detection, with a focus on improving the accuracy of this method. By carefully selecting relevant features, they have been able to achieve a faster and more accurate detection rate. As a result, they laid out a lightweight intrusion detection system (IDS) that relies on information gain and a multi-layer perceptron neural network in their article. Prior to categorization utilizing neural network, they had mentioned that gain ratio was utilized for picking significant characteristics for attack and regular traffic. Based on their analysis of the UNSW-NB15 intrusion detection dataset, they determined that the lightweight IDS was well-suited for real-time intrusion detection based on empirical results from thirty selected characteristics.

Anomalies detection on various computer and communications systems is facilitated by machine learning techniques, which have enabled the modernization of intrusion detection systems and the assurance of data privacy [21]. In their work, they discussed how they used the UNSW-NB12 dataset to assess the efficacy of several intrusion detection algorithms, including supervised ones like KNN and SVM as well as unsupervised ones like isolation forest and K-means. Their explanation was that the findings demonstrated that the supervised method SVM Gaussian fine was able to accurately categorize normal and abnormal data, with an accuracy of 92%. Regarding the unsupervised algorithms, they had also mentioned that the K-means method appropriately classified the data and allowed for the obvious definition of the required number of groups; however, this dataset was heavily aggregated. Finally, they brought up the fact that isolation forest, although being a strong algorithm for separating out-of-the-ordinary data, had trouble with it. Finally, they came to the conclusion that it should be made explicit that not all distance anomaly detection algorithms were suitable for all datasets.

According to [22], machine learning is finding increasing use across a range of industries, including computer security and self-driving cars. They had seen that machine learning offered a miraculous answer to the difficulties of network traffic analysis and intrusion detection brought on by the massive amounts of network traffic nowadays. Critical industries such as oil and gas, transportation, digital infrastructure, water and sewage treatment, and traffic management might all benefit from incorporating intrusion detection systems into a comprehensive security framework, as they have previously indicated. They claimed to have compared twelve different supervised machine learning techniques in their article. They said that the goal of this comparison research was to show which ML approaches were the most effective at identifying malicious or benign network traffic based on its categorization, particular attack type, and overall health. They found that random forest, Jrip, and J48 performed better in their trials conducted on the CICIDS'2017 dataset (Table 1).

3 Methodology

Machine learning for cybersecurity intrusion detection (Fig. 3).

- Data preprocessing is the initial step in building a robust machine learning model. This phase involves several key tasks. Firstly, the dataset is loaded, which could include various sources such as network traffic logs or system logs. Secondly, the data is cleaned to ensure its quality by removing duplicates and handling missing values appropriately. Categorical variables are encoded if necessary, often using techniques like one-hot encoding. Lastly, the dataset is split into separate training and testing sets to facilitate model evaluation.
- Feature engineering is a crucial aspect of model development where relevant features are extracted from the dataset. This process involves identifying and selecting the most informative features that contribute to the model's predictive performance. Additionally, features may be normalized or scaled if needed to ensure consistency and comparability across different features.
- Model selection involves choosing an appropriate machine learning algorithm based on the nature of the problem and the dataset. Common algorithms include decision trees, random forests, support vector machines (SVM), and neural networks. Multiple models are trained using the training data, and their performance is evaluated using techniques like cross-validation.
- Once the best-performing model is identified based on evaluation metrics, it is trained using the entire training dataset. Model evaluation is conducted using the testing dataset, and performance metrics such as accuracy, precision, recall, F1-score, and ROC curve analysis are calculated. The model's parameters may be fine-tuned if necessary to improve its performance.
- Deployment involves deploying the trained model to monitor real-time network or system activities. Continuous monitoring and updating of the model are essential to adapt to new cyberthreats and changes in the network environment.

Table 1 Comparative analysis

Author	ML technique	Algorithms used	Methodology used	Research area	Conclusion
[11]	Supervised machine learning	Decision tree (DT), (SVM), K-nearest neighbour (KNN)	Literature review, dataset review, classification with machine learning algorithms	Intrusion detection systems (IDS)	More successful results obtained in some studies, IDS systems based on AI with ML approaches
[12]	Supervised Machine learning	Naive Bayes, Gaussian, decision tree, random forest	Study of supervised probabilistic and predictive ML techniques for intrusion detection	Intrusion detection systems (IDS)	Analysis of the performance and robustness of various ML techniques in intrusion detection
[13]	Supervised Machine learning	Random forest (RF), decision tree (DT), multi-layer perceptron (MLP), support vector machine (SVM)	Feature selection, classification with machine learning algorithms	Intrusion detection systems (IDS)	Proposed machine learning model achieved accuracy of classification of more than 99%, highlighting the efficiency of using machine learning for IDS
[14]	Supervised Machine learning	Not specified	Literature review, taxonomy creation based on related works	Intrusion Detection Systems (IDS)	Classification performance of supervised ML algorithms is high and promising, feature selection and data imbalance handling are important aspects for performance enhancement in intrusion detection systems

(continued)

Table 1 (continued)

Author	ML technique	Algorithms used	Methodology used	Research area	Conclusion
[15]	Supervised Machine learning	Support vector machine (SVM), artificial neural network (ANN)	Feature selection, classification with machine learning algorithms	Intrusion detection systems (IDS)	Proposed model using ANN and feature selection outperformed SVM in classifying network traffic, demonstrating the effectiveness of the approach for intrusion detection
[16]	Supervised Machine learning	Bayesian network, random decision forest, Naive Bayes (NB), DT, ANN	Classification with various machine learning algorithms	Intrusion detection systems	Effectiveness of different machine learning algorithms in detecting intrusions in cybersecurity domains, highlighting the importance of employing machine learning for IDS development
[17]	Supervised Machine learning	ML algorithms	Literature review, taxonomy creation based on ML and DL (deep learning)-based IDS literature	Intrusion detection systems (IDS)	Proposed taxonomy framework provides a comprehensive classification system for IDS literature, identifying key issues and solutions using ML and DL techniques for improving detection accuracy and robustness
[18]	Supervised Machine learning	SVM, decision tree (DT), K-nearest neighbour (KNN)	Feature selection, classification with ML algorithms	Intrusion detection systems	Successful classification of network traffic as DDoS or benign using various ML algorithms, demonstrating the feasibility of employing MLfor IDS in real-world scenarios

(continued)

Table 1 (continued)

Author	ML technique	Algorithms used	Methodology used	Research area	Conclusion
[19]	Supervised Machine learning	Machin learning algorithms	Literature review, SCADA intrusion detection and classification, analysis using supervised learning techniques	SCADA intrusion detection and classification	Critical review of supervised learning techniques for SCADA intrusion detection and classification, highlighting key issues and offering insights for future research in industrial control systems security
[20]	Supervised Machine learning	Multi-layer perceptron neural network	Feature selection, classification with neural network	Intrusion detection systems (IDS)	Development of IDS based on feature selection and neural network (NN), demonstrating high accuracy in real-time intrusion detection
[21]	Supervised & Unsupervised Machine learning	K-nearest neighbour (KNN), support vector machine (SVM), isolation forest, K-means	Evaluation of supervised and unsupervised algorithms for intrusion detection	Intrusion detection systems (IDS)	Supervised SVM demonstrated high accuracy in classifying normal and abnormal data, while unsupervised K-means effectively grouped data; Isolation forest faced challenges but exhibited robustness in separating atypical values
[22]	Supervised Machine learning	Random forest, J48, Jrip	Comparative study of twelve supervised machine learning methods for network traffic classification	Network traffic classification	Comparative analysis of machine learning methods for network traffic classification, identifying Jrip, random forest, and J48 as top performers for specific types of traffic classification

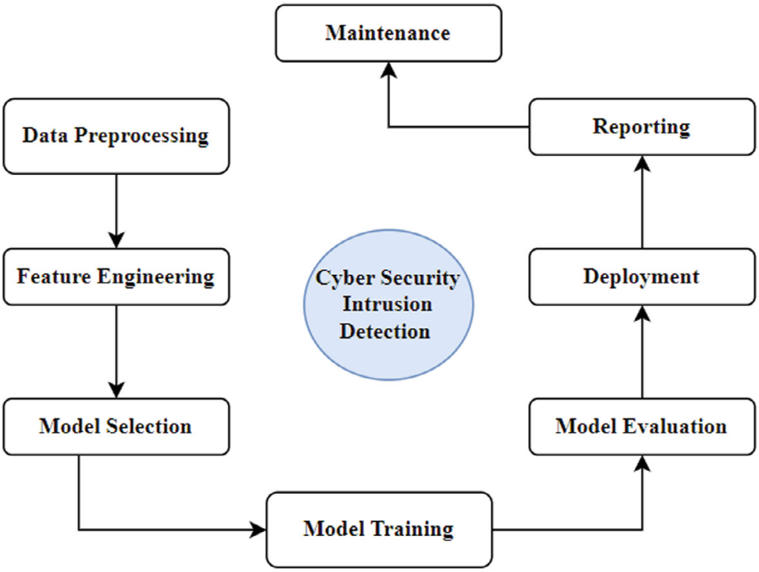


Fig. 3 ML for cybersecurity intrusion detection

- Reporting is crucial for communicating the model's performance and detected anomalies to relevant stakeholders for further action. Reports summarizing the model's performance are generated to facilitate decision-making processes.
- Maintenance involves regular updates of the model with new data to keep it effective against evolving threats. Monitoring the model's performance over time and retraining or updating it as needed ensures its continued effectiveness in detecting and mitigating cyberthreats.

Supervised Machine Learning Techniques for Intrusion Detection

To train a model for intrusion detection (ID) using supervised ML techniques, each occurrence of data is labelled as either benign or malicious. Listed below are some common algorithms. Decision trees are interpretable because they divide the feature space into regions, but they are also vulnerable to overfitting. Support vector machines (SVMs) are delicate to parameter adjustment yet excel at handling complicated data by locating the hyperplane that best divides classes in high-dimensional space.

Logistic regression (LR) models the probability of a binary outcome using a logistic function, offering simplicity and interpretability. Neural networks, particularly deep learning architectures, automatically extract hierarchical features but require substantial computational resources and large amounts of data. Evaluation metrics like accuracy, precision, recall, and F1-score are used to assess model performance, with ensemble methods like random forests often enhancing overall effectiveness by combining multiple classifiers.

Logistic regression [23] is a common supervised ML algorithm [24] used in cybersecurity intrusion detection tasks. It is particularly useful when the outcome to be predicted is categorical, such as whether a network activity is benign or malicious.

In logistic regression, the model predicts the probability that a given input belongs to a certain class. The logistic function (also called the sigmoid function) [25]is used to map the output of the linear equation to a probability value between 0 and 1[25].

The logistic function is defined as:

$$\text{logistic}(z) = \frac{1}{1 + e^{-z}}$$

where z is the linear combination of input features and their corresponding coefficients:

$$z = \beta_0 + \beta_{1\times 1} + \beta_{2\times 2} \times 1 + \cdots + \beta_{n\times n}$$

Here, $x_1, x_2, \ldots, x_n$ are the input features, and $\beta_0, \beta_1, \ldots, \beta_n$ are the coefficients (weights) learned during the training phase.

The predicted probability that the output belongs to class 1 (e.g. malicious) is given by:

$$P(y = 1|x = \frac{1}{1 + e^{-(0+\beta 1\times 1+\beta 2\times 2+\cdots+\beta n\times n)}})$$

And the predicted probability that the output belongs to class 0 (e.g. benign) is:

$$P(y = 0|x) = 1 - P(y = 1|x)$$

where $P(y = 1|x)$ is the probability of the output being in class 1 given input x.

During training, the model's parameters (coefficients) $\beta_0, \beta_1, \ldots, \beta_n$ are learned by minimizing a cost function, typically the cross-entropy loss function, using optimization techniques like gradient descent.

Once trained, the model can make predictions by computing the probabilities using the learned coefficients and applying a threshold (e.g. 0.5) to determine the predicted class.

Support Vector Machines (SVM)

Support vector machines (SVMs) [2] are a popular supervised machine learning algorithm used in cybersecurity for tasks such as intrusion detection[2]. SVMs work by finding the hyperplane that best separates different classes in the feature space. Here is SVMs work.

Linear SVM**:** For a binary classification problem, where we have two classes (+1 and -1), the objective of SVM is to find the hyperplane with the maximum margin that separates the two classes. Mathematically, this hyperplane is represented as:

$$w \cdot x + b = 0$$

where:

weight vector perpendicular to the hyperplane w.input vector x, and bias term b.

The decision function for classifying a new data point $\boldsymbol{xi}$ is given by:

$$f(x_i) = \text{sign}(w \cdot x_i + b)$$

The goal of training an SVM is to find the optimal $\boldsymbol{w}$ and b that maximizes the margin while still correctly classifying the training data.

Nonlinear SVM**:** SVMs can also handle nonlinear decision boundaries by using the kernel trick. The decision function in the nonlinear case becomes:

$$f(X_i) = \text{sign}\left(\sum_{j=1}^{N} y_j \alpha_j K(X_i, X_j) + b\right)$$

where:

- Number of support vectors N.
- y_j is the class label of the j-th support vector.
- α_j are the learned Lagrange multipliers.
- $K(x_i, x_j)$ is the kernel function, which computes the inner product of the

transformed feature vectors $\Phi(x_i)$ and $\Phi(x_j)$ in the higher-dimensional space.

Common kernel functions include the linear kernel ($K(x_i, x_j) = x_i \cdot x_j$, polynomial kernel, Gaussian (RBF) kernel, etc.

Training Objective: In both linear and nonlinear SVMs, the training objective involves solving the optimization problem to find the optimal hyperplane parameters **w** and **b**, or the optimal Lagrange multipliers $\boldsymbol{\alpha_j}$ in the nonlinear case. This is typically done by minimizing the following objective function subject to certain constraints:

$$\min_{\mathbf{w}, b} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N} \xi_i$$

subject to:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

where:

- c is the regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error.
- ξi are slack variables that allow for misclassification.

SVMs are effective in cybersecurity for intrusion detection because they can handle high-dimensional data and are robust against overfitting,

Logistic Regression (LR)

Logistic regression (LR) [26] is utilized in supervised machine learning for cybersecurity intrusion detection due to its effectiveness in binary classification tasks [26]. It models the probability of a binary outcome (intrusion or non-intrusion) based on input features extracted from network traffic or system logs. By training on labelled data, it learns the relationship between features and intrusion events, enabling it to classify new instances. Logistic regression offers interpretability, scalability, and efficiency in analysing large datasets commonly encountered in cybersecurity. Its simplicity and ability to handle high-dimensional data make it a valuable tool for detecting anomalies and identifying potential security threats in network environments.

Neural Networks (NN)

Neural networks, a class of machine learning algorithms, are widely employed in supervised cybersecurity intrusion detection. These models learn patterns from labelled data to classify network activities as normal or malicious [17]. Utilizing various architectures like convolutional neural networks (CNNs) or recurrent neural networks (RNNs), they excel in capturing intricate relationships within data. Through continuous training, neural networks adapt to evolving threats and enhancing the accuracy and efficiency of intrusion detection systems. Their ability to discern subtle anomalies and rapidly process vast amounts of data makes them indispensable in safeguarding networks against cyberthreats.

Decision Trees (DT)

A neural network in supervised machine learning for cybersecurity intrusion detection involves training a model on labelled data to classify network activities as either normal or malicious. By extracting features from network traffic and utilizing a layered network of interconnected nodes, the neural network learns patterns indicative of cyberthreats [27]. Through supervised learning, the model adjusts its parameters iteratively to minimize classification errors and enhance detection accuracy. This approach enables proactive identification of intrusions or suspicious behaviour within computer networks, aiding in the prevention and mitigation of cyberattacks, thus bolstering overall cybersecurity defences.

4 Conclusion and Future Work

This paper has provided a comprehensive overview of the methods and techniques utilized in cybersecurity intrusion detection, specifically focusing on supervised machine learning. Through an extensive analysis of recent literature, we have identified various approaches, including traditional classifiers such as DT and SVM, as well as more advanced techniques like deep learning and ensemble methods. Additionally, we discussed the importance of feature selection, data preprocessing, and evaluation metrics in the design and implementation of intrusion detection systems. By synthesizing these findings, we have highlighted the strengths and limitations of different approaches, emphasizing the need for a tailored solution based on the specific requirements and constraints of the target environment.

Future Work: Moving forward, several avenues for future research in this domain present themselves. Firstly, there is a need for further exploration and experimentation with emerging machine learning algorithms, such as deep neural networks and reinforcement learning, to enhance the accuracy and robustness of intrusion detection systems. Additionally, investigating the applicability of transfer learning and domain adaptation techniques could facilitate the transferability of models across different network environments and attack scenarios. Moreover, the integration of human-in-the-loop approaches, where expert knowledge is combined with machine learning algorithms, holds promise for improving the interpretability and explainability of detection outcomes. Furthermore, given the dynamic nature of cyberthreats, continuous monitoring and adaptation of intrusion detection systems are essential. Therefore, research efforts should be directed towards developing adaptive and self-learning mechanisms capable of detecting novel and evolving attack patterns in real time. Lastly, the evaluation of intrusion detection systems in real-world settings, including large-scale network infrastructures and diverse attack scenarios, remains a crucial area for future investigation to ensure the practical effectiveness and deployment readiness of these solutions. By addressing these research directions, we can contribute to the advancement of cybersecurity intrusion detection and enhance the resilience of digital systems against malicious activities.

References

1. Geetha R, Thilagam T (2021) A review on the effectiveness of machine learning and deep learning algorithms for cyber security. Arch Comput Methods in Eng 28:2861–2879
2. Jha J, Ragha L (2013) Intrusion detection system using support vector machine. Int J Appl Inf Syst (IJAIS) 3:25–30
3. Farnaaz N, Jabbar MA (2016) Random forest modeling for network intrusion detection system. Procedia Comput Sci 89:213–217
4. Bouzida Y, Cuppens F (2006) Neural networks versus decision trees for intrusion detection. In: IEEE/IST workshop on monitoring, attack detection and mitigation (MonAM), vol 28, p 29
5. Ning P, Jajodia S (2003) Intrusion detection techniques. Internet Encyclopedia 2:355–367
6. Majidpour J, Hasanzadeh H (2020) Application of deep learning to enhance the accuracy of intrusion detection in modern computer networks. arXiv:2012.08318
7. Anwar S, Mohamad Zain J, Zolkipli MF, Inayat Z, Khan S, Anthony B, Chang V (2017) From intrusion detection to an intrusion response system: fundamentals, requirements, and future directions. Algorithms 10(2):39
8. Ioulianou P, Vasilakis V, Moscholios I, Logothetis M (2018) A signature-based intrusion detection system for the internet of things. In: Information and communication technology form

9. Jyothsna VVRPV, Prasad R, Prasad KM (2011) A review of anomaly based intrusion detection systems. Int J Comput Appl 28(7):26–35
10. Inayat Z, Gani A, Anuar NB, Khan MK, Anwar S (2016) Intrusion response systems: foundations, design, and challenges. J Netw Comput Appl 62:53–74
11. Kilincer IF, Ertam F, Sengur A (2021) Machine learning methods for cyber security intrusion detection: datasets and comparative study. Comput Netw 188:107840
12. Gharibian F, Ghorbani AA (2007) Comparative study of supervised machine learning techniques for intrusion detection. In: Fifth annual conference on communication networks and services research (CNSR'07). IEEE, pp 350–358
13. Ahanger AS, Khan SM, Masoodi F (2021) An effective intrusion detection system using supervised machine learning techniques. In: 2021 5th International conference on computing methodologies and communication (ICCMC). IEEE, pp 1639–1644
14. Abdallah EE, Otoom AF (2022) Intrusion detection systems using supervised machine learning techniques: a survey. Procedia Comput Sci 201:205–212
15. Taher KA, Jisan BMY, Rahman MM (2019) Network intrusion detection using supervised machine learning technique with feature selection. In: 2019 International conference on robotics, electrical and signal processing techniques (ICREST). IEEE, pp 643–646
16. Alqahtani H, Sarker IH, Kalim A, Minhaz Hossain SM, Ikhlaq S, Hossain S (2020) Cyber intrusion detection using machine learning classification techniques. In: Computing science, communication and security: first international conference, COMS2 2020, Gujarat, India, 26–27 Mar 2020. Revised selected papers 1. Springer, Singapore
17. Liu H, Lang B (2019) Machine learning and deep learning methods for intrusion detection systems: a survey. Appl Sci 9(20):4396
18. Aksu D, Üstebay S, Aydin MA, Atmaca T (2018) Intrusion detection with comparative analysis of supervised learning techniques and fisher score feature selection algorithm. In: Computer and information sciences: 32nd international symposium, ISCIS 2018, held at the 24th IFIP world computer congress, WCC 2018, Poznan, Poland, 20–21 Sept 2018, Proceedings 32. Springer, pp 141–149
19. Alimi OA, Ouahada K, Abu-Mahfouz AM, Rimer S, Alimi KOA (2021) A review of research works on supervised learning algorithms for SCADA intrusion detection and classification. Sustainability 13(17):9597
20. Mebawondu JO, Alowolodu OD, Mebawondu JO, Adetunmbi AO (2020) Network intrusion detection system using supervised learning paradigm. Sci Afr 9:e00497
21. Portela FG, Mendoza FA, Benavides LC (2019) Evaluation of the performance of supervised and unsupervised Machine learning techniques for intrusion detection. In: 2019 IEEE international conference on applied science and advanced technology (iCASAT). IEEE, pp 1–8
22. Ahmim A, Ferrag MA, Maglaras L, Derdour M, Janicke H (2020) A detailed analysis of using supervised machine learning for intrusion detection. In: Strategic innovative marketing and tourism: 8th ICSIMAT, Northern Aegean, Greece, 2019. Springer, pp 629–639
23. Gonaygunta H (2023) Machine learning algorithms for detection of cyber threats using logistic regression. Department of Information Technology, University of the Cumberlands
24. Li Z, Zhang H, Masum M, Shahriar H, Haddad H (2020) Cyber fraud prediction with supervised machine learning techniques. In: Proceedings of the 2020 ACM Southeast conference, pp 176–180
25. Tien Bui D, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016) Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13:361–378
26. Belavagi MC, Muniyal B (2016) Performance evaluation of supervised machine learning algorithms for intrusion detection. Procedia Comput Sci 89:117–123
27. Alom MZ, Taha TM (2017) Network intrusion detection for cyber security using unsupervised deep learning approaches. In: 2017 IEEE national aerospace and electronics conference (NAECON). IEEE, pp 63–69

Analyse Heart Disease Classification Using Convolutional Neural Network

K. Shilpa and T. Adilakshmi

Abstract Heart disease is leading cause of death. Heart disease can be detected by various medical diagnosis mainly by Electrocardiogram (ECG). The measurement of electrical activity is called ECG. Arrhythmia is the name for an unnatural heartbeat. An arrhythmia may cause serious problem. Identifying and classifying normal and abnormal (arrhythmic) heartbeats is the primary goal of the study work. This paper predicts the heart disease at early stage using deep learning approach—convolutional neural network model. The dataset collected from PTB-XL ECG database consists of 21,837 records from 18,885 patients of 10 s length 27 columns. The CNN-proposed model has achieved accuracy of 89.05%, precision has achieved 85.85%, recall has achieved of 68.29%, and loss has 0.24% when compared to other existing model achieved 80.1% and 76.9%. The proposed model has achieved good results though it has taken more datasets.

Keywords Heart disease · Arrhythmia · ECG · Deep learning · PTB-XL Dataset · CNN

1 Introduction

Arrhythmia is a type of cardiovascular heart disease. The arrhythmia may cause heart failure and sudden death. The ECG signal [1–3], which reflects a heartbeat, consists of the P-wave symbolizes "atrial depolarization", "ventricular depolarization"—Q wave, and "ventricular repolarization"—T wave [4]. Arrhythmia is referred

K. Shilpa (✉)
Department of CSE, UCE, OU, Hyderabad, India
e-mail: shilpamtech555@gmail.com

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 Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_61

to as irregular heart rhythm. ECG arrhythmia is identifying the normal and irregular heartbeats in the ECG signal. It has a significant impact on health care.

A deep learning algorithm [5, 6] is called a CNN which are best suited for recognition of image and task processing. The layers of CNN [7] are convolutional, pooling, and FC layers. These are an essential part of CNN filters that are used to extract features from input images. Convolutional layer output has gone through pooling layers, two or three FC layers in the habit of predicting the future or categorizing the image received input from the pooling layers output.

1.1 Objective

To detect arrhythmia early using CNN architecture and evaluate performance using accuracy, precision, and recall.

2 Literature Work

1. Elif Izci et al.—This paper has proposed DL-based innovative 2D-CNN to identify five types of arrhythmia.In this paper first, they were segmented ECG signals into heartbeats, next a heartbeat converted into 2-D greyscale images, and then, these images fed as input to CNN architecture. The proposed system accuracy was 97.42% [4].
2. Saroj Kumar Pandey et al.—In this paper, the author has identified heart disease by using heartbeats of five different types using database of MIT-BIH - Arrhythmia. In this, first they pre-process the data. 12 -layer 1d deep cnn applied to extract features and finally used softmax(cnn) and compared results with ml techniques. Proposed method has achieved 99.40% of accuracy, 98.78%, 98.78%, and 98.74% precision, recall, and F_1 score [8].
3. Sena Yagmur Sen et al.—In this paper, the proposed method used ECG time-series data with a huge volume of data as first step, then spectrograms as fed i/p to a deep CNN. In this, heartbeat was classified as RBBB, normal, and PVC by ECG signals taken from MIT-BIH database of arrhythmia. The first technique time series of ECG signals are used as i/p fed to CNN, and next ECG signals were transform time–frequency domain matrices and sent to CNN. Finally, it improves performance [9].
4. Mohammad Kachuee et al.—In this paper, the authors have proposed a technique which classifies heartbeats based on deep CNN using AAMI EC57 standard for arrhythmia and MI task of classification. The authors used PhysioNet's MIT-BIH [10–12] and PTB diagnostics datasets for evaluation of proposed method. They have achieved 93.4% and 95.9% average accuracy on two arrhythmia and MI classification [13].

5. Taminul Islam et al.—The ECG the heart peaks can be shown to identify heart disease and analyse R peak of arrhythmia disease. Tachycardia and Bradycardia are two types of arrhythmia to detect arrhythmia. The authors are presented various methods like CNN, TQWT, LSTM, and various datasets used to detect arrhythmia. In this work, data pre-processing, feature extraction, and classification were used to classifying ECG signals and to detect arrhythmia [14].

3 Proposed Methodology with Workflow

Methodology

In this research work, there are five steps are involved to classify arrhythmia which consists of ECG signal pre-processing, next segmentation of ECG signals then converting 1-D ECG [15] into 2-D spectrogram images using–short-time Fourier transformation (STFT), and then these images of spectrogram fed as i/p to 2D-CNN, last classification of five types of arrhythmia such as NORM, CD, STTC, HYP, and MI.

3.1 Proposed Workflow

See Figs. 1, 2 and 3 and Table 1.

PTB-XL Dataset

The PTB-XL [16] ECG data consists of 21,799 clinical ECGs 12-lead from 18,869 dataset of patients 10 s length, where male are 52% and female are 48%, and the age ranges from 0 to 95. The SCP-ECG standard has 71 ECG statements that cover diagnostic, form, and rhythm statements. It has form, rhythm and diagnostic, and statements (Fig. 4).

4 Results and Discussion

In this, Python is used to the implementation of the paper. The training accuracy has achieved 89.05%, loss has 0.24%, precision has achieved 85.85%, and recall has achieved 68.29% for 10 epochs (Figs. 5, 6 and 7; Table 2).

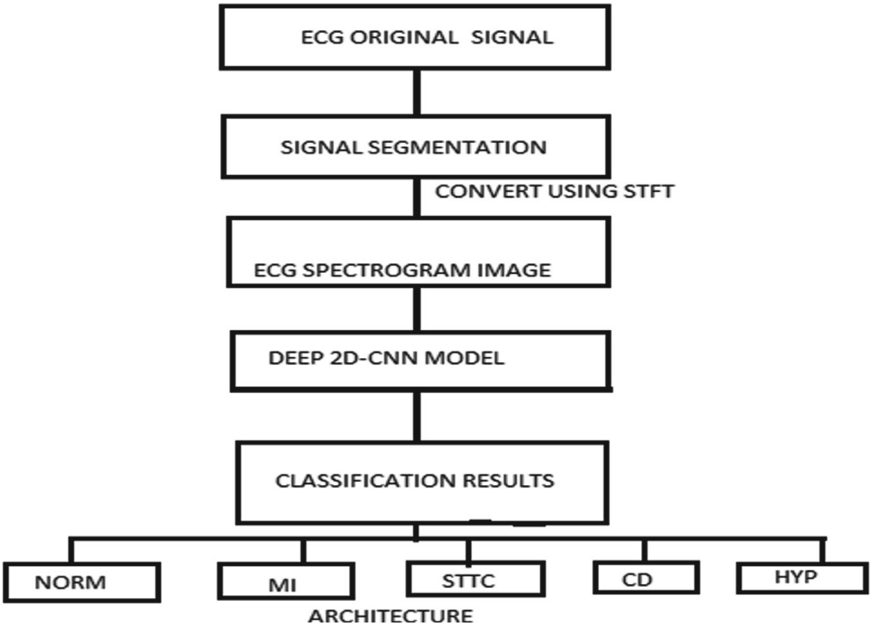


Fig. 1 Proposed workflow of heart disease prediction

5 Conclusion and Future Work

This research paper classify the five different arrhythmia types. The research paper is to identify and classification of normal and abnormal (arrhythmic) heartbeats as primary goal. This paper predicts the heart disease at early stage using deep learning approach—convolutional neural network model. The dataset collected from PTB-XL ECG database consists of 21,837 records from 18,885 patients of 10 s length 27 columns. CNN-proposed model has achieved accuracy of 89.05%, precision has achieved 85.85%, recall has achieved of 68.29%, and loss has 0.24% when compared to other ml existing model achieved 80.1 and 76.9%. The proposed model has achieved good results though it has taken more datasets. In future will work with different CNN architecture to early detection of heart disease.

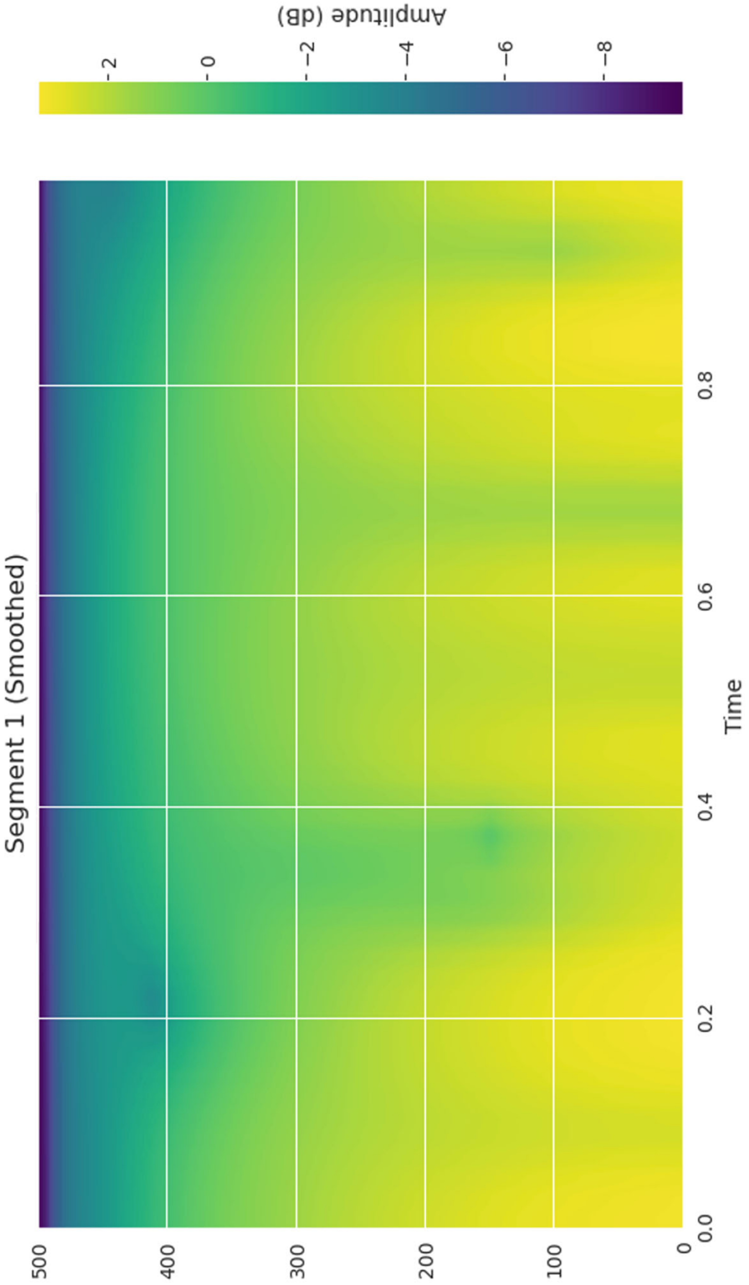


Fig. 2 2D spectrogram image

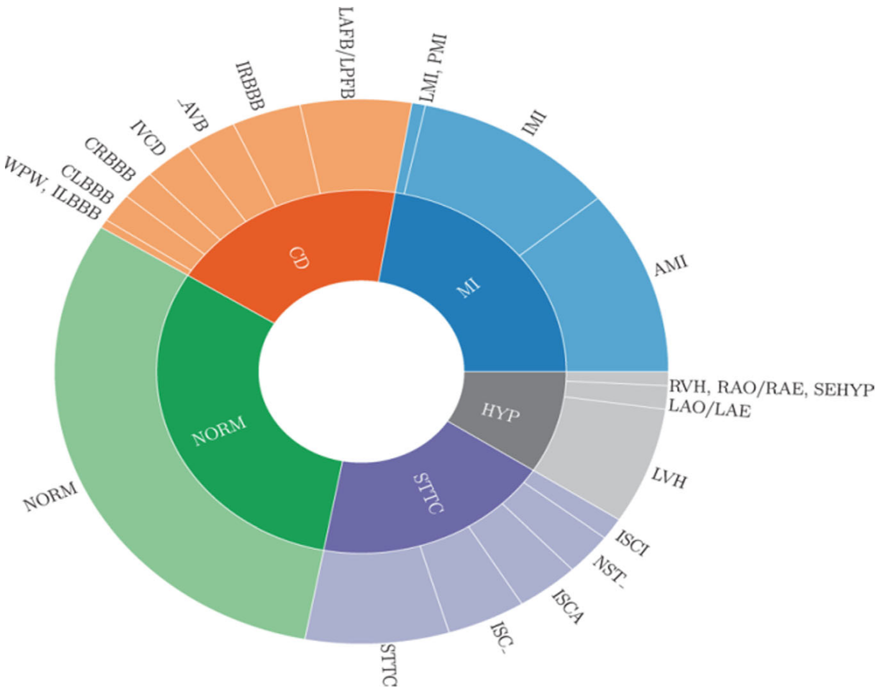
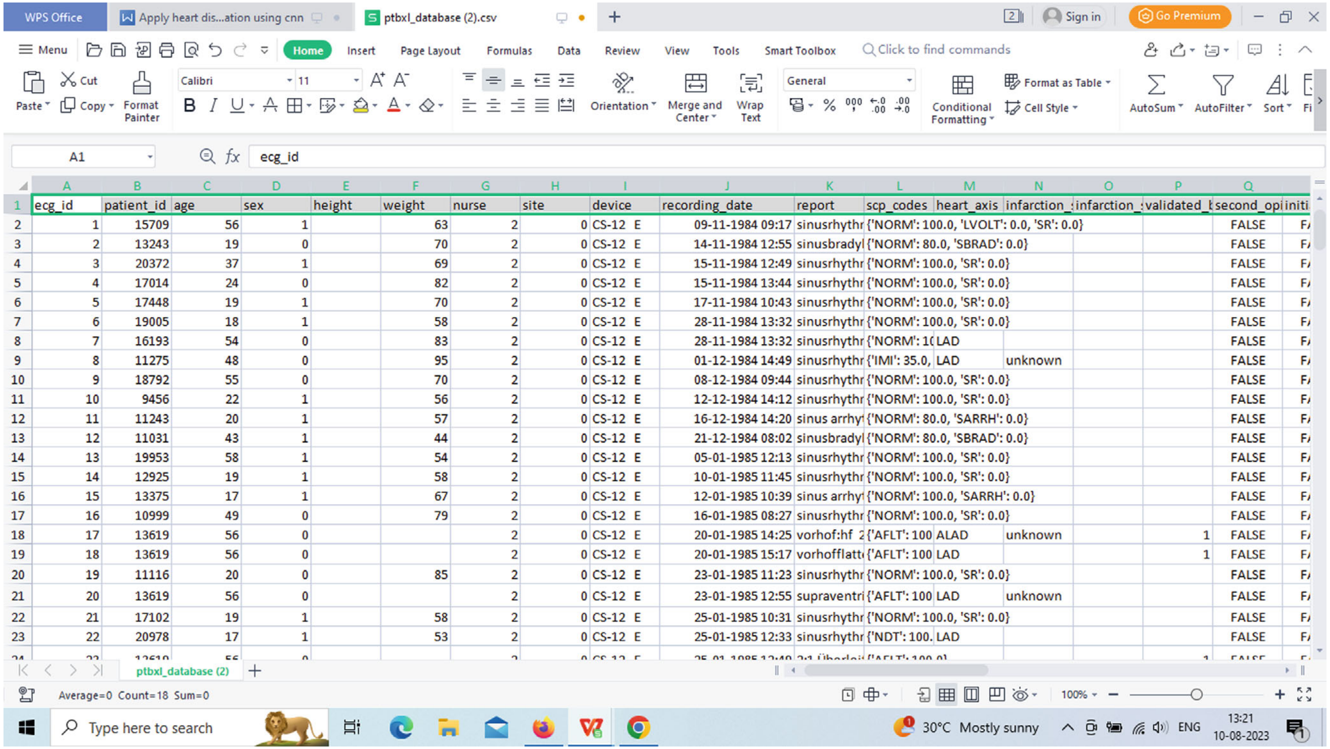


Fig. 3 Arrhythmia types

Table 1 Layers of CNN

Layer (type)	Output shape	Param
X_inputs (InputLayer)	[(None, 7)]	0
X_dense_1 (Dense)	(None, 32)	256
X_drop_1 (Dropout)	(None, 32)	0
X_dense_2 (Dense)	(None, 32)	1056
Y_inputs (InputLayer)	[(None, 1000, 12)]	0
X_drop_2 (Dropout)	(None, 32)	0
Sequential (Sequential)	(None, 1)	4,092,097
Z_concat (Concatenate)	(None, 33)	0
Z_dense_1 (Dense)	(None, 64)	2176
Z_dense_2 (Dense)	(None, 64)	4160
Z_drop_1 (Dropout)	(None, 64)	0
Z_outputs (Dense)	(None, 5)	325
Total params: 4,100,070		
Trainable params: 4,100,070		
Non-trainable params: 0		



	ecg_id	patient_id	age	sex	height	weight	nurse	site	device	recording_date	report	scp_codes	heart_axis	infarction_	infarction_	validated_	second_opi	initi
2	1	15709	56	1		63	2	0	CS-12 E	09-11-1984 09:17	sinusrhythr	{'NORM': 100.0, 'LVOLT': 0.0, 'SR': 0.0}					FALSE	F/
3	2	13243	19	0		70	2	0	CS-12 E	14-11-1984 12:55	sinusbradyl	{'NORM': 80.0, 'SBRAD': 0.0}					FALSE	F/
4	3	20372	37	1		69	2	0	CS-12 E	15-11-1984 12:49	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
5	4	17014	24	0		82	2	0	CS-12 E	15-11-1984 13:44	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
6	5	17448	19	1		70	2	0	CS-12 E	17-11-1984 10:43	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
7	6	19005	18	1		58	2	0	CS-12 E	28-11-1984 13:32	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
8	7	16193	54	0		83	2	0	CS-12 E	28-11-1984 13:32	sinusrhythr	{'NORM': 1(	LAD				FALSE	F/
9	8	11275	48	0		95	2	0	CS-12 E	01-12-1984 14:49	sinusrhythr	{'IMI': 35.0,	LAD	unknown			FALSE	F/
10	9	18792	55	0		70	2	0	CS-12 E	08-12-1984 09:44	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
11	10	9456	22	1		56	2	0	CS-12 E	12-12-1984 14:12	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
12	11	11243	20	1		57	2	0	CS-12 E	16-12-1984 14:20	sinus arrhy	{'NORM': 80.0, 'SARRH': 0.0}					FALSE	F/
13	12	11031	43	1		44	2	0	CS-12 E	21-12-1984 08:02	sinusbradyl	{'NORM': 80.0, 'SBRAD': 0.0}					FALSE	F/
14	13	19953	58	1		54	2	0	CS-12 E	05-01-1985 12:13	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
15	14	12925	19	1		58	2	0	CS-12 E	10-01-1985 11:45	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
16	15	13375	17	1		67	2	0	CS-12 E	12-01-1985 10:39	sinus arrhy	{'NORM': 100.0, 'SARRH': 0.0}					FALSE	F/
17	16	10999	49	0		79	2	0	CS-12 E	16-01-1985 08:27	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
18	17	13619	56	0			2	0	CS-12 E	20-01-1985 14:25	vorhof:hf 2	{'AFLT': 100	ALAD	unknown		1	FALSE	F/
19	18	13619	56	0			2	0	CS-12 E	20-01-1985 15:17	vorhofflatt	{'AFLT': 100	LAD			1	FALSE	F/
20	19	11116	20	0		85	2	0	CS-12 E	23-01-1985 11:23	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
21	20	13619	56	0			2	0	CS-12 E	23-01-1985 12:55	supraventri	{'AFLT': 100	LAD	unknown			FALSE	F/
22	21	17102	19	1		58	2	0	CS-12 E	25-01-1985 10:31	sinusrhythr	{'NORM': 100.0, 'SR': 0.0}					FALSE	F/
23	22	20978	17	1		53	2	0	CS-12 E	25-01-1985 12:33	sinusrhythr	{'NDT': 100.	LAD				FALSE	F/

Fig. 4 PTB-XL dataset

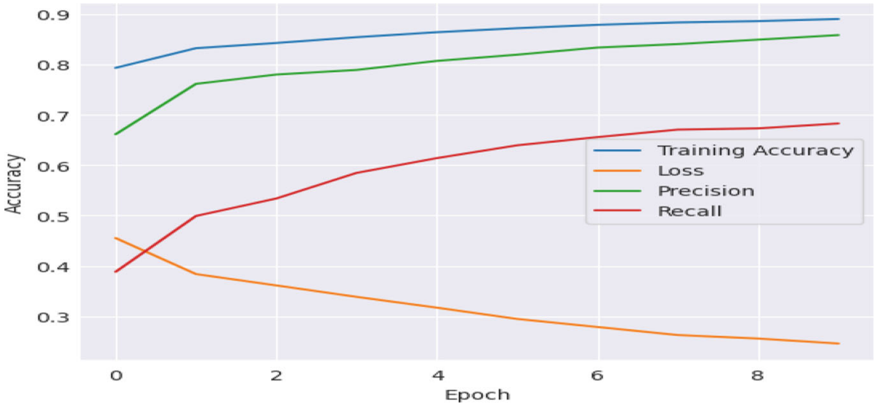


Fig. 5 Performance metrics of proposed method

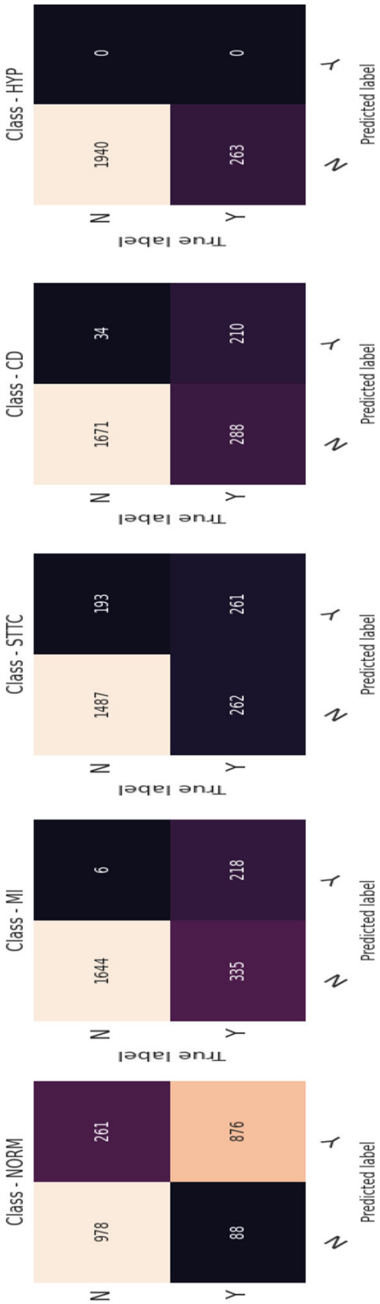


Fig. 6 Confusion matrix of five different arrhythmia NORM, MI, STTC, CD, HYP

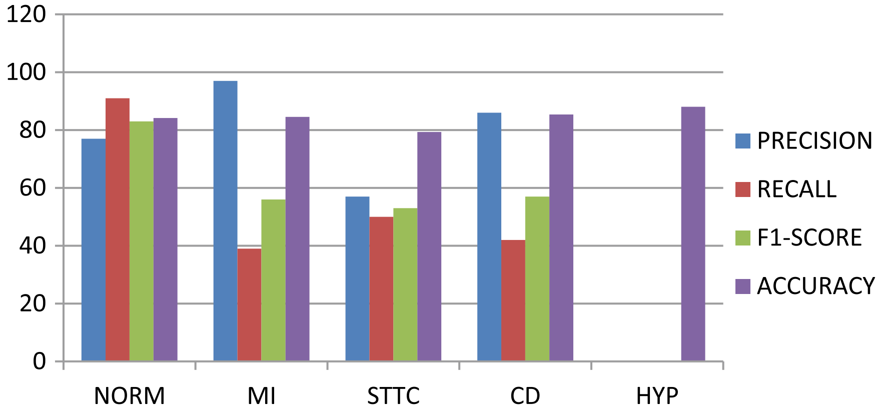


Fig. 7 Graph of performance metrics of proposed method

Table 2 Proposed method performance metrics

	Precision	Recall	F_1-Score	Accuracy
NORM	77	91	83	84.15
MI	97	39	56	84.52
STTC	57	50	53	79.34
CD	86	42	57	85.38
HYP	0	0	0	88.06

References

1. Abdalla FY, Wu L, Ullah H, Ren G, Noor A, Mkindu H et al (2020) Deep convolutional neural network application to classify the ECG arrhythmia. Sig Image Video Process 14(7):1431–1439
2. Acharya UR, Fujita H, Oh SL, Hagiwara Y, Tan JH, Adam M (2017) Application of deep convolutional neural network for auto-mated detection of myocardial infarction using ECG signals. Inf Sci 415:190–198
3. Martis RJ, Acharya UR, Min LC (2013) ECG beat classification using PCA LDA ICA and discrete wavelet transform. Biomed Signal Process Control 8(5):437–448
4. Izci E, Ozdemir MA, Degirmenci M, Akan A (2019) Cardiac arrhythmia detection from 2D ECG images by using deep learning technique
5. Dhillon A, Verma GK (2020) Convolutional neural network: a review of models, methodologies and applications to object detection. Prog Artif Intell 9:85–112
6. Da'u A, Salim N (2020) Recommendation system based on deep learning methods: a systematic review and new directions. Artif Intell Rev 53(4):2709–2748
7. Shaker AM (2020) Generalization of convolutional neural networks for ECG classification using generative adversarial networks. IEEE Access
8. Pandey SK, Shukla A, Bhatia S, Gadekallu TR, Kumar A, Mashat A, Shah MA, Janghel RR (2023) Detection of arrhythmia heartbeats from ECG signal using wavelet transform-based CNN model
9. Şen SY, Özkurt N (2020) ECG arrhythmia classification by using convolutional neural network and spectrogram

10. Zheng Z, Chen Z, Hu F, Zhu J, Tang Q, Liang Y (2020) An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology. Electronics 9(1):121
11. Abdou AD, Ngom NF, Niang O (2020) Arrhythmias prediction using an hybrid model based on convolutional neural network and nonlinear regression. Int J Comput Intell Appl 19(03):2050024
12. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, San Tan R (2017) A deep convolutional neural network model to classify heartbeats. Comput Biol Med 89:389–396
13. Kachuee M, Fazeli S, Sarrafzadeh M (2018) ECG heartbeat classification: a deep transferable representation
14. Islam T, Kundu A, Ahmed T, Khan NI (2022) Analysis of arrhythmia classification on ECG dataset
15. Khan Mamun MMR, Elfouly T (2023) Detection of cardiovascular disease from clinical parameters using a one-dimensional convolutional neural network
16. Strodthoff N, Wagner P, Schaeffter T, Samek W (2021) Deep learning for ECG analysis: benchmarks and insights from PTB-XL. IEEE J Biomed Health Inform 25(5):1519–1528

Collaborative Learning for Personalized Medicine: Federated Approaches to Disease Prediction and Drug Recommendation in Health Care

Sahithi Katoori, Akhil Reddy Vancha, Rachakonda Sai Sathvik, B. Veera Jyothi, and Eliganti Ramalakshmi

Abstract The advancements in machine learning have created a scope for accurate disease prediction and personized drug recommendation. However, the sensitive and private nature of medical data become a challenge in developing robust models for ensuring patient privacy. In this paper, we propose a novel approach that involves federated learning to address the challenges, enabling collaborative model training and knowledge sharing across multiple healthcare entities. The proposed methodology involves establishing a network of healthcare providers, each contributing their local disease prediction and drug recommendation models. Through federated learning, these models are aggregated and refined without sharing raw data, ensuring patient privacy. This project contributes to the advancement of healthcare technology by offering a secure and privacy-preserving solution for disease prediction and drug recommendation. The proposed federated learning approach shows a way for healthcare entities to collaborate and improve models without compromising the data privacy. Through rigorous experimentation and validation, we anticipate that the federated learning framework will prove instrumental in driving the future of data-driven healthcare advancements.

Keywords Disease prediction · Drug recommendation · Federated learning · Privacy-preservation · Collaborative model training

S. Katoori · A. R. Vancha · R. Sai Sathvik · B. Veera Jyothi (✉) · E. Ramalakshmi
Department of Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad 500075, India
e-mail: veerajyothi_it@cbit.ac.in

E. Ramalakshmi
e-mail: eramya2@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_62

1 Introduction

The evolving healthcare sector is driven by advances in machine learning and data-driven technologies [1], and the potential for precise disease prediction and personalized drug recommendations has soared to unprecedented heights [1]. These innovations promise to turn patient care, optimize treatment methods, and ultimately lift healthcare outcomes. However, this journey toward change has come with challenges, particularly concerning the handling of enormous amount of sensitive medical data while upholding the paramount principle of patient confidentiality [2]. In response to these challenges, a groundbreaking research paper introduces a solution that harnesses the power of federated learning to unlock new possibilities in health care [3]. The primary aim is to establish a privacy-preserving and highly efficient for disease prediction and drug recommendation, which is integrated with federated learning techniques [3]. This initiative creative a collaborative network of healthcare providers, each contributing their localized disease prediction and drug recommendation models to a shared network [4]. This innovative approach has federated learning in it, a technique serving as the primary tool for collaborative model development [3]. Federated learning enables the aggregation of the individual models without exposing raw, sensitive data of patients [2].

The soul of this paper lies in its potential to provide a collaborative model which even focuses on the data privacy. By involving federated learning, healthcare entities can collaborate and strengthen the predictive models while maintaining their patients trust. This collaborative effort is poised to bring about a profound transformation in the healthcare sector. The journey of this approach is marked by appropriate experimentation and validation. The federated learning framework is aiming to yield highly accurate and reliable disease prediction models and drug recommendation systems. This research paper tells a significant leap forward in the pursuit of accurate disease prediction and personalized drug recommendations within the healthcare domain. By harnessing the power of federated learning, it not only addresses the inherent privacy challenges associated with healthcare data but also fosters collaborative efforts among healthcare providers. In this data-driven era, it is poised to usher in a new paradigm where healthcare advancements are intrinsically tied to data security, privacy, and, most importantly, the well-being of patients. As this paper forges ahead in the crucible of innovation, it endeavors to pave the way for a future where health care prioritizes both technological advancements and patient-centric ethics.

1.1 Objective

Developing a federated learning and deep learning-based project for disease prediction and personalized drug recommendation: This project aims to create a system that involves both federated learning and deep learning techniques to predict diseases based on individual patient data. Additionally, the system will utilize these models to recommend personalized medicines, accounting the patient's medical history and other relevant factors.

2 Research Methodology

2.1 The Internet of Federated Things

This paper introduces the concept of the Internet of Federated Things (IoFT) as a paradigm shift in the Internet of things (IoT) landscape [5], where decentralized and privacy-preserving model training takes place at the edge devices [2]. It discusses federated learning (FL) data-driven approaches, opportunities, and challenges within three dimensions: global model, personalized model, and meta-learning model [5]. The major findings of this paper are federated learning (FL) data-driven approaches, opportunities, and challenges within three dimensions: global model, personalized model, and meta-learning model [5].

2.2 A Precision Health Service for Chronic Diseases

Advancements and development using wearable devices, machine learning, and deep learning [1]. This paper showcases an integrated precision health services utilizing wearable devices, machine learning, and deep learning for continuous monitoring, early detection, and personalized health promotion in chronic disease management [1]. The study employed wearable devices, open environmental data, smartphone app, and AI-assisted telecare platform for continuous monitoring [1]. Machine learning and deep learning algorithms were used for predictive modeling. Modular chronic disease prediction models achieved an average accuracy of 88.46% for diseases like obesity, panic disorder, and chronic obstructive pulmonary disease. Lifestyle and environmental factors were found to be highly correlated with patient health and improved prediction accuracy compared to using only questionnaire data [1]. The study demonstrated a cost-effective model for prediction that requires only a few features, making it practical for real-world deployment.

2.3 An Ensemble Deep Learning Approach for Disease Prediction Through Metagenomics

The paper presents an ensemble deep learning approach for disease prediction using metagenomic data, addressing the challenges posed by limited samples and high-dimensional features [6]. Unsupervised deep learning methods for feature extraction.

Development of a disease scoring strategy based on deep representations. Ensemble analysis using gradient boosting classifier. The proposed ensemble deep learning framework outperforms existing algorithms in disease prediction using human microbiome data [6]. Deep learning methods, especially convolutional neural networks, show promise in capturing meaningful features from metagenomic data [6]. The study highlights the importance of addressing the high dimensionality and sparsity of microbial data for improved predictive models.

2.4 Deep Reinforcement Learning for Medicine Recommendation

The paper introduces a deep reinforcement learning-based approach to recommend safe and effective medication for patients by integrating electronic health records (EHRs) with both adversarial and synergistic drug–drug interaction (DDI) knowledge graphs [7]. The paper employs deep reinforcement learning, specifically the double deep Q network (DDQN) algorithm, to predict optimal drug combinations for patient therapies. Integration of electronic health records (EHRs) with adversarial and synergistic DDI knowledge graphs to enhance medication recommendations [7]. The integration of synergistic and adversarial DDI knowledge graphs allows for a more comprehensive assessment of drug combinations, leading to safer and more effective medication recommendations. The proposed approach significantly reduces drug side effects and enhances medication safety [7].

The literature survey from these four papers presents an insightful exploration of innovative applications of federated learning and machine learning techniques within the realm of health care [1, 4, 6, 7]. Together, they illuminate the potential of technology driven healthcare advancements [4]. “The Internet of Federated Things” delves into the convergence of the Internet of things (IoT) and federated learning [5], showcasing the possibilities of secure, decentralized data analysis across interconnected devices and platforms. “A Precision Health Service for Chronic Diseases” introduces a precision healthcare service that utilizes wearable devices, machine learning, and deep learning for the monitoring and management of chronic illnesses [1], heralding a new era in personalized health care.

3 Methodology

The methodology for the proposed system involves several steps:

Table 1 Patient dataset

S. No.	Drug	Disease	Gender	Age	Institute
1	ACGEL CL NANO Gel	Acne	Male	23	Apollo
2	Ceriz syrup	Allergy	Female	35	Medicover
3	Debistal GM tablets	Diabates	Male	26	KIMS
4	Dancure shampoo	Fungal infection	Female	20	Virinchi
5	FALRAP KID 150 g	Malaria	Male	23	CARE

3.1 Data Collection and Preparation

3.1.1 Data Gathering

Collaborate with multiple healthcare institutions and entities to collect diverse healthcare data [4]. This includes electronic health records (EHRs), patient demographics, medical histories, genetic information, prescription records, and treatment outcomes.

3.1.2 Data Preprocessing

Standardize data formats and clean the data to remove noise, outliers, and inconsistencies [1]. Anonymize and de-identify patient data to comply with privacy regulations (Table 1).

3.2 Federated Learning Setup

3.2.1 Federated Learning Framework

Choose an appropriate federated learning framework or platform that supports secure and privacy-preserving model training, such as PySyft, TensorFlow Federated, or Open-Minded [3].

3.2.2 Model Selection

Select suitable machine learning or deep learning models for disease prediction and drug recommendation, taking into account the complexity and nature of the healthcare data [6]. There are many models, instead of choosing a predefined frameworks, customized convolutional neural networks will go better and can design the parameters accordingly.

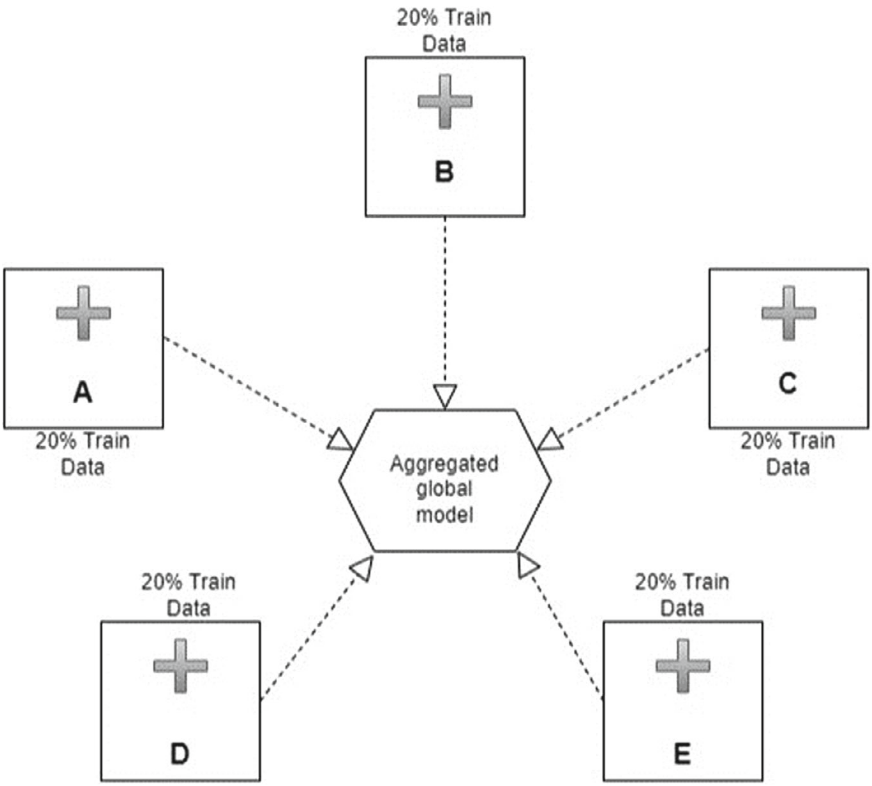


Fig. 1 Training in federated learning

3.2.3 Data Partitioning

Partition the data into multiple subsets, each residing within the infrastructure of a different healthcare institution. Ensure that the data partitioning process maintains data privacy by using techniques like federated data shuffling [3] (Fig. 1).

3.3 Secure Model Training

3.3.1 Local Model Training

Train initial model weights locally within each healthcare institution using the respective partition of data [3]. Employ privacy-preserving techniques like differential privacy to further protect sensitive information during local training.

3.3.2 Gradient Calculation and Aggregation

Calculate model gradients based on local training and share only the gradients (not raw data) among the participating institutions. Aggregate gradients from all institutions using secure aggregation protocols, such as federated averaging, to update the global model.

3.3.3 Iterative Training

Repeat the local training and aggregation steps for multiple rounds (epochs) to allow the global model to converge gradually. Monitor model performance to ensure convergence and accuracy [3].

3.4 Model Development and Analysis

This step involves both the disease prediction and drug recommendation systems.

3.4.1 Disease Prediction System

Evaluate the global federated model's performance using validation datasets or cross-validation techniques. Use appropriate evaluation metrics to assess disease prediction accuracy [6].

3.4.2 Drug Recommendation System

Utilize the federated model to recommend personalized treatment plans for patients [7]. Consider genetic data, medical history, disease predictions, and known drug interactions when making treatment recommendations.

3.5 Validation and Continuous Improvement

3.5.1 Model Validation

Continuously monitor the performance of the federated learning model in real-world healthcare settings [3]. Collect feedback from healthcare providers and patients to validate the effectiveness of disease prediction and drug recommendations [8].

3.5.2 Model Updates

Implement mechanisms for periodic model updates to adapt to evolving healthcare data and research [9]. Use federated learning to collaboratively improve model accuracy and generalizability over time [10].

3.6 Reporting and Knowledge Sharing

3.6.1 Research Findings

Publish research findings and insights gained from disease prediction and drug recommendation using federated learning [3]. Contribute to the broader medical and scientific community's knowledge base.

3.6.2 Collaboration

Collaborate with healthcare institutions, research organizations, and policymakers to share knowledge and promote the adoption of federated learning in health care.

4 Conclusion

This innovative approach addresses major challenges associated with data privacy, security, making it well-suited for the sensitive nature of patients' data [2]. By allowing multiple healthcare institutions to collaborate while keeping patient data localized, federated learning enables the development of robust and accurate predictive models. Through the aggregation of diverse datasets from various sources, federated learning enhances the generalizability and reliability of disease prediction models. As there is a continuous evolution in technology and methodologies, collaborative efforts between researchers, healthcare providers, and policymakers will be essential to unseal the full potential of this approach [2]. With continued advancements, we can anticipate a future where health care becomes more precise, patient-centered, and proactive, ultimately leading to improved population health and well-being.

5 Future Scope

A disease prediction and drug recommendation system using federated learning has the potential to revolutionize health care as it can be expanded to cover more diseases, thereby increasing its utility [3]. With the increasing concern over data privacy, federated learning can provide a solution by training machine learning models locally on devices without any data transfer to the cloud. This technique is already making an impact across the entire care delivery cycle, including collaborative drug discovery. With the rise of the Internet of things (IoT) in health care, federated learning can be integrated with IoT devices for remote patient health monitoring [5].

References

1. Wu CT, Wang SM, Su YE, Hsieh TT, Chen PC, Cheng YC, Tseng TW, Chang WS, Su CS, Kuo LC, Chien JY, Lai F (2022) A precision health service for chronic diseases: development and cohort study using wearable device, machine learning and deep learning
2. Rani S, Kataria A, Kumar S, Tiwari P (2023) Federated learning for secure IoMT-applications in smart healthcare systems: a comprehensive review. Knowl-Based Syst 274. ISSN/ISBN: 0950-7051
3. Qayyum A, Ahmad K, Ahsan MA, Al-Fuqaha A, Qadir J (2022) Collaborative federated learning for healthcare: multi-modal COVID-19 diagnosis at the edge. IEEE Open J Comput Soc 3
4. Wu X, Pei J, Chen C, Zhu Y, Wang J, Qian Q, Zhang J, Sun Q, Guo Y (2023) Federated active learning for multicenter collaborative disease diagnosis. IEEE Trans Med Imaging 42(7)
5. Kontar R, Shi N, Yue X, Chung S, Byon E, Chowdhury M, Jin J, Kontar W, Masoud N, Nouiehed M, Okwudire CE, Raskutti G, Saigal R, Singh K, Sheng Z (2021) The internet of federated things. IEEE Access 9
6. Shen, Zhu J, Deng Z, Lu W, Wang H (2023) EnsDeepDP: an ensemble deep learning approach for disease prediction through metagenomics. IEEE/ACM Trans Comput Biol Bioinform 20(2)
7. Symeonidis P, Chairistanidis S, Zanker M (2022) Deep reinforcement learning for medicine recommendation. In: 2022 IEEE 22nd International conference on bioinformatics and bioengineering (BIBE)
8. Wang Y, Deng G, Zeng N, Song X, Zhuang Y (2019) Drug-disease association prediction based on neighborhood information aggregation in neural networks. IEEE Access 7
9. Guendouzi BS, Ouchani S, Assaad HEL, Zaher MEL (2023) A systematic review of federated learning: challenges, aggregation methods, and development tools. J Netw Comput Appl. ISSN/ISBN: PII: S1084-8045(23)00133-92023
10. Dinakaran S, Anitha P (2022) An efficient drug compound analysis using spectral deep feature classification based compound analyse model for drug recommendation. Neurosci Inform 2(2022). ISSN/ISBN: 2772-5286
11. Chaddad A, Lu Q, Li J, Katib Y, Kateb R, Tanougast C, Bouridane A, Abdulkadir A (2023) Explainable, domain-adaptive, and federated artificial intelligence in medicine. IEEE/CAA J Autom Sin 10
12. Garg S (2021) Drug recommendation system based on sentiment analysis of drug reviews using machine learning. In: 2021 11th International conference on cloud computing, data science & engineering

Fault Diagnosis in Wireless Sensor Networks (WSNs) Using Machine Learning (ML) Approach

Jalapala Sinjini, M. Priyadharshini, V. Indumathi, V. V. Bhavani, Tahseen Jahan, and A. Hemalatha Reddy

Abstract Wireless sensor networks (WSNs) are integral components of modern infrastructure, providing pervasive monitoring and data collection capabilities across diverse domains. However, the reliability of WSNs can be compromised by various faults, including sensor failures, communication errors, and environmental disturbances. Traditional fault diagnosis methods often struggle to cope with the dynamic and resource-constrained nature of WSNs. In response, this paper explores the application of machine learning techniques for fault diagnosis in WSNs. We present the challenges associated with fault diagnosis in WSNs and discuss how machine learning algorithms offer promising solutions by leveraging data-driven approaches. The review encompasses various machine learning paradigms, including supervised, unsupervised, and semi-supervised learning, tailored to the unique characteristics and constraints of WSNs. Additionally, we discuss the key considerations and trade-offs involved in selecting and deploying machine learning models for fault diagnosis in WSNs.

Keywords Fault diagnosis · WSNs · ML · Supervised learning · Unsupervised learning · Semi-supervised learning

J. Sinjini (✉) · M. Priyadharshini · V. Indumathi · V. V. Bhavani · T. Jahan · A. Hemalatha Reddy
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: sinjini@cmritonline.ac.in

M. Priyadharshini
e-mail: priyadharshini@cmritonline.ac.in

V. Indumathi
e-mail: indu11mathi@cmritonline.ac.in

V. V. Bhavani
e-mail: bhavaniv@cmritonline.ac.in

T. Jahan
e-mail: tahseen048@cmritonline.ac.in

A. Hemalatha Reddy
e-mail: hemareddy@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_63

1 Introduction

Wireless sensor networks (WSNs) fault diagnosis is a critical task aimed at ensuring reliable and efficient operation in dynamic and resource-constrained environments. In recent years, machine learning techniques have emerged as powerful tools for addressing this challenge by enabling automated fault detection and classification. This introduction provides a concise overview of the application of machine learning in fault diagnosis within WSNs.

Wireless sensor networks (WSNs) consist of spatially distributed autonomous sensors that monitor physical or environmental conditions and cooperatively transmit data to a central location. Despite their widespread deployment, WSNs are susceptible to various faults such as sensor failures, communication errors, and environmental disturbances. Traditional fault diagnosis methods often rely on predefined thresholds or rules, which may be insufficient for handling complex and dynamic fault scenarios.

Machine learning offers a promising alternative by leveraging data-driven approaches to automatically learn patterns and relationships from sensor readings. Supervised learning algorithms, such as support vector machines (SVMs) [1] and neural networks [2], can be trained on labeled datasets to classify normal and faulty conditions based on sensor data characteristics. Unsupervised learning techniques, including clustering algorithms like K-means, enable the detection of anomalies without requiring labeled data, making them suitable for real-time fault detection in WSNs.

Importance of Fault Diagnosis in WSNs

Fault diagnosis in wireless sensor networks (WSNs) [3] is of paramount importance due to several key reasons:

1. Reliability Assurance: WSNs are often deployed in critical applications such as environmental monitoring, industrial automation, health care, and military surveillance, where reliable operation is essential. Fault diagnosis helps ensure that sensors and communication links are functioning correctly, minimizing the risk of data loss or incorrect information being transmitted.
2. Performance Optimization: Identifying and addressing faults in a timely manner can improve the overall performance of the network. By detecting and isolating faulty nodes or communication channels, the network can adapt and reroute data flows to maintain optimal coverage and connectivity.
3. Cost Reduction: Early fault detection and diagnosis can prevent minor issues from escalating into major failures, reducing maintenance costs and downtime. Proactive fault management strategies, enabled by fault diagnosis, can help extend the lifespan of WSN deployments and minimize the need for costly manual interventions.
4. Energy Efficiency: Faulty nodes in a WSN can consume unnecessary energy due to repeated attempts to transmit data or perform sensing tasks. By identifying and addressing faults promptly, energy consumption can be optimized, prolonging

the overall lifetime of the network and minimizing the need for frequent battery replacements in energy-constrained environments.

5. Data Integrity and Security: Faulty sensors or compromised communication links can lead to inaccurate data being collected or transmitted, potentially jeopardizing the integrity and security of the entire system. Fault diagnosis helps ensure data accuracy and integrity by identifying and mitigating sources of error or tampering.

Fault diagnosis plays a crucial role in ensuring the reliability, performance, cost-effectiveness, energy efficiency, and security of wireless sensor networks across various applications, ultimately contributing to their successful deployment and operation in real-world scenarios.

2 Literature Review

A wireless sensor network (WSN) is defined as a network of autonomous devices dispersed across space and time that use a complex subsystem known as sensors to gather data about their surrounding environment [4]. Environmental factors such as temperature, noise, humidity, wind speed and direction, pollution levels, and so on were monitored by WSN. The presence of sensors in unexpected regions made data defect detection a tough challenge. For applications such as weather forecasting, illness prediction, traffic monitoring, etc., this detection has to be very accurate. Classification accuracy varied from 93 to 100%, according to core analysis performed on different datasets, which is encouraging when compared to the current systems. When data defect detection accuracy was critical, their algorithm's strengths may be put to use [4].

As a result of commercializing innovative technology, the Internet of things (IoT) has established a well-defined infrastructure, as explained in [5]. Due to the proliferation of IoT networks, smart gadgets are now able to collect data about their surroundings and send it to demanding consumers via an IoT gateway. Network constraints caused by the exponential growth of IoT users and sensors significantly reduced the amount of energy that IoT devices could store. An IoT layer based on progressive features, the wireless network was characterized as resilient and experimentally significant. The creation of learning-oriented routing algorithms that minimize energy consumption became paramount in light of the fact that the wireless network's weight distribution is utterly unpredictable and subject to environmental instability. Due to their great degree of precision and flexibility, learning-based routing systems were starting to emerge as possible candidates to accomplish this urgent demand. The time-varying properties of link connections and access status, however, made routing more difficult in dynamic IoT networks. Therefore, it was imperative that cutting-edge learning-based routing systems can respond instantly to changes in the network. This study introduced a reinforcement learning-based intelligent fault-detecting energy-efficient quality-of-service routing algorithm for optimal route discovery with minimum end-to-end delay. But the cluster nodes' leftover energy, which diminished the network's overall existence, was crucial to the cluster head decision. The results were an increase in network resiliency, a reduction in energy consumption during data transmission, and a longer lifespan for the network overall. The experimental findings showed that fault-tolerance solutions, which reduced the likelihood of network failure by including highly trusted computing capabilities, improved network efficiency [5].

In [6], they suggested using a convolutional neural network (CNN) and a Naïve Bayes classifier to detect and fix node problems and enhance convergence performance. In the end, they used real-world datasets to examine CNN, convex hull, and Naïve Bayes algorithms in order to categorize and detect the errors. Using performance criteria, both the simulation and experimental results demonstrated that the CNN approach kept feasibility and efficiency while outperforming the convex hull technique in defect identification [6].

In [7], outlined the ways in which the diverse nature of wireless sensor networks (WSNs) renders them susceptible to many failures that might have catastrophic effects on security, economics, and the dependability of the system as a whole. The variety of deployment and the restrictions in the resources of the sensors made it tough to accurately identify and diagnose malfunctions or defects in WSN. This article explored the possibility of using a supervised machine learning-based approach to analyze sensor data for defect diagnosis and detection. A wide variety of WSN defects were taken into account, including hardover, drift, spike, erratic, data loss, stuck, and random problem. The aforementioned errors were replicated using typical, non-faulty data from a reliable dataset created by UNC researchers and posted online. This dataset included readings from temperature and humidity sensors taken during a multi-hop scenario. In order to simulate real-world WSN conditions, events were created based on fault occurrences. A lightweight ensemble learning-based approach known as extremely randomized trees (extra-trees) was used for the rapid defect detection and diagnosis. There was a considerable decrease in bias and variance errors, and the suggested extra-trees-based detection strategy was resilient to signal noise. Modern machine learning methods including decision trees, neural networks, random forests, and support vector machines were tested against the suggested system to see how well it performed. The accuracy, precision, and F_1-score metrics used to evaluate the proposed scheme's performance all pointed to its efficiency. Furthermore, as compared to state-of-the-art methods, the training time of the suggested scheme was much lower [7].

In [8] attempted to address the issue of error detection during data collecting in WSNs. An effective technological solution was proposed in this research by combining evolutionary computing with machine learning. By using a biological population model to regulate the size of the population and a parallel method for further tuning, they enhanced the traditional particle swarm optimization (PSO). By optimizing the initial weights and biases of a back propagation neural network (BPNN), the suggested RS-PPSO method was able to reduce training time and increase prediction accuracy. The proliferation of WSNs has made them an indispensable backbone for the IoT. Whether the data acquired by the sensor nodes was accurate greatly affected the network's dependability, real-time performance, and energy savings. In this research, we provide an optimized machine learning technology scheme that can successfully detect problem data and guarantee the efficient functioning of WSN [8].

In [9], they looked at the implementation of wireless sensor networks in challenging settings. They were advantageous due to their cheap cost and versatility. However, they may encounter several setbacks, necessitating the need to enhance the precision of the data. When it came to finding and diagnosing errors, many AI approaches have shown remarkable results. Machine learning has recently arisen as a potent AI-based approach to address the issue of WSN failures. This study compares several machines learning methods, including RF, multilayer perceptron, SVM, and probabilistic neural network (PNN), to a multi-fault classification assessed utilizing a deep learning approach based on LSTM classifier. Four metrics were utilized to assess the performance of these described methodologies used for fault detection in WSNs.

Wireless sensor networks (WSNs) were described in [10] as having sensor nodes randomly placed in dangerous areas. It became clear from real-time data that sensor nodes may malfunction. A variety of sensor node defects, brought on by both internal and external factors including poor calibration, low batteries, and the effects of sensor aging, were unavoidable. Identifying and isolating a malfunctioning sensor was crucial for WSNs applications that depended on the accuracy of data given by the sensor nodes. Previous work on defect detection relied heavily on statistical methods, which need expert-level knowledge of the sensor domain as well as data from nearby sensors. Because a malfunctioning sensor reading might resemble data from a working sensor, it may be difficult to identify a sensor issue using a distributed method to data analysis. To far, machine learning algorithms have proven effective in detecting and categorizing different kinds of WSN errors, allowing for their prevention. However, in recent years, there has been a lot of interest from both academics and businesses in using deep learning (DL) techniques. This section describes the application of DL algorithms in WSN fault diagnostics using neural network approaches. Types such as hard, soft, intermittent, and transitory faults were the main emphasis of the fault diagnostic process [10].

To sense (monitor) and collect raw sensor data for later transmission to a base station, wireless sensor networks (WSNs) are usually used in task-related settings with dynamic circumstances, as described in [11]. There are many technological hurdles that need to be overcome before WSNs may be used in actual settings. Conventional methods, crafted for a particular job, made it difficult to respond to changing circumstances that went beyond the original objective. Recent work in WSNs has used machine learning (ML) approaches to successfully learn from dynamic settings, thereby resolving this issue. In particular, WSNs used deep learning (DL) methods, a subset of ML defined by the deployment of deep neural networks, to glean more abstract information from their sensor data. When used to WSNs, ML approaches have many positive effects, such as lowering computing complexity, making it easier to identify optimum solutions, increasing energy efficiency, and so on. But their assessment showed that resource-constrained WSNs have to deal with high energy consumption and lengthy training times in order to achieve acceptable performance. Recent literature reviews have focused on ML approaches and their potential uses in WSNs. On the other hand, evaluations on DL approaches applied to WSNs were few. With a focus on DL approaches, this study showcased the latest advancements in ML for WSNs. Each deep neural network design was discussed in conjunction with the DL approaches that have been developed for use in WSN applications [11].

In [12], they set up in a dispersed fashion to gather different kinds of data. Many problems and obstacles arose in WSN research, including those pertaining to energy efficiency, security, localization, and many more. In order to reduce the amount of inaccurate and noisy data produced by the millions of wireless sensor networks, one of these areas was outlier or anomaly identification. Models for detecting outliers should not skimp on data quality. They had to accurately detect abnormalities in both online and offline modes while improving speed and using minimum network resources. This article provided a comprehensive overview of outlier identification in WSN data.

According to [13], data collection in wireless sensor networks was carried out via a distributed system of inexpensive and energy-efficient sensor nodes. The purpose of a sink node and other coordinating nodes in a network is to facilitate the movement of data or packets from one location to another. One way to think about it is as a network of interconnected devices that relay data gathered from the field of sensors. Wireless connections allowed the data transfer to take place. Due to limitations in power and bandwidth, sensors often have restricted interaction capabilities. Network problems and their solutions were the primary emphasis of this article. They looked at the use of machine learning methods in this network as a potential solution to certain issues with the network. Through the use of training data, a model was created to "learn" from the information it contained; this process is known as machine learning. In this study, we summarized the literature on machine learning algorithms that addressed the complex problems with wireless sensor networks from 2015 to 2020 (Table 1).

3 Methodology

Fault diagnosis in wireless sensor networks (WSNs) using machine learning involves developing mathematical models to detect and classify faults based on data collected from sensor nodes. One common approach is to utilize machine learning algorithms for pattern recognition and classification tasks. Let us outline the mathematical model incorporating equations for this process:

Data Collection and Preprocessing

Data collected from sensor nodes can be represented as a matrix X, where each row corresponds to a set of features from a sensor node, and each column corresponds to a specific feature. Preprocessing involves normalization to ensure that all features have similar scales, reducing the impact of varying magnitudes on the machine learning algorithms. This can be represented as:

Table 1 Background comparison

References	ML algorithms	Fault diagnosis method	Area of research	Techniques used	Sensor network	Methodology	Conclusion
[4]	Enhanced minimum redundancy maximum relevance algorithm	Supervised machine learning-based technique	Wireless sensor network (WSN) fault detection	Filter and wrapper methods using FMID and FMIQ objective functions	WSN	Analysis of various datasets	Proposed algorithm showed promising classification accuracy for fault detection, which can be leveraged in critical situations
[5]	Reinforcement learning	Reinforcement learning	IoT quality-of-service routing and fault detection	Energy-efficient routing, fault-tolerance strategies	IoT	Experimental evaluation	Proposed technique improves network efficiency and robustness, enhancing fault tolerance and prolonging network lifetime
[6]	Naïve Bayes classifier, convolution neural network (CNN)	Supervised machine learning-based technique	Fault detection in WSNs	Convex hull algorithm, Naïve Bayes, CNN	WSN	Simulation and experimental analysis	CNN algorithm showed better performance in fault identification compared to convex hull algorithm
[7]	Ensemble learning-based techniques (e.g., extremely randomized trees)	Supervised machine learning-based technique	WSN fault detection and diagnosis	Extra-trees-based detection scheme, simulation and experimental analysis	WSN	Performance evaluation	Proposed scheme showed robustness toward signal noise and strong reduction of bias and variance error in fault detection

(continued)

Table 1 (continued)

References	ML algorithms	Fault diagnosis method	Area of research	Techniques used	Sensor network	Methodology	Conclusion
[8]	Particle swarm optimization (PSO), back propagation neural network (BPNN)	Evolutionary computing and machine learning	WSN data collection fault detection	RS-PPSO algorithm, improved PSO, BPNN	WSN	Algorithm development and evaluation	Proposed technique optimizes initial weights and biases, enhancing prediction accuracy and training time reduction
[9]	Long short-term memory (LSTM) classifier, support vector machine (SVM), random forest (RF), multilayer perceptron (MLP), probabilistic neural network (PNN)	Deep learning and traditional machine learning	Multi-fault classification in WSNs	LSTM classifier, SVM, RF, MLP, PNN	WSN	Performance comparison	Proposed LSTM classifier showed competitive performance in fault detection compared to traditional machine learning techniques
[10]	Neural network methods	Deep learning	Fault diagnosis in WSNs	Deep learning (DL) algorithms	WSN	Research review and analysis	DL methods offer potential for effective fault diagnosis in WSNs, addressing various fault types
[11]	Deep learning techniques	Deep learning	ML techniques in WSNs	Deep neural network architectures	WSN	Literature review and analysis	DL techniques offer benefits in WSNs but pose challenges such as large training time and energy consumption

(continued)

Table 1 (continued)

References	ML algorithms	Fault diagnosis method	Area of research	Techniques used	Sensor network	Methodology	Conclusion
[12]	Various machine learning techniques	Supervised and unsupervised machine learning	Outlier detection in WSNs	Supervised and unsupervised ML techniques	WSN	Research review and analysis	Various ML techniques show potential in outlier detection in WSNs, offering accuracy and performance improvements
[13]	Various machine learning techniques	Supervised and unsupervised machine learning	Network issues and solutions in WSNs	Supervised and unsupervised ML techniques	WSN	Research review and analysis	ML techniques present potential solutions for addressing network issues in WSNs, offering improvements in various aspects

$$x_i^i = \frac{x_i - \mu}{\sigma}$$

where x_i is the original feature value, μ is the mean of the feature across all samples, σ is the standard deviation of the feature across all samples, and x_i' is the normalized feature value.

Feature Extraction

Feature extraction aims to identify relevant features that can discriminate between normal and faulty operation. This can involve techniques such as principal component analysis (PCA) or wavelet transform to reduce the dimensionality of the data while preserving relevant information.

Machine Learning Model

Let us consider a supervised learning approach using a support vector machine (SVM) classifier, which aims to find the optimal hyperplane that separates data points into different classes.

The decision function of an SVM can be represented as:

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right)$$

where α_i are the Lagrange multipliers, y_i are the class labels, $K(x_i, x)$ is the kernel function, and b is the bias term.

Training the Model

Training the SVM involves finding the optimal values of α_i and b by solving the optimization problem:

$$\underset{\alpha}{\text{minimize}}\ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i x_j) - \sum_{i=1}^{n} \alpha_i$$

Subject to the constraints

$$0 \le \alpha_i \le C, \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$

where C is the regularization parameter.

Fault Diagnosis

Once the SVM is trained, it can be used to predict the class labels of new data points. The decision function output, $f(x)$, indicates whether a sensor node is operating normally or experiencing a fault.

Model Evaluation

The performance of the fault diagnosis model can be evaluated using metrics such as accuracy, precision, recall, and F_1-score, which can be calculated as follows:

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}}$$

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

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$F_1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

This mathematical model integrates machine learning techniques, such as SVM, with data preprocessing, feature extraction, and evaluation metrics to perform fault diagnosis in wireless sensor networks.

Workflow of Fault Diagnosis Process

A fault diagnosis framework in wireless sensor networks (WSNs) using machine learning (ML) involves several key components. First, there are the sensor nodes dispersed throughout the network, each collecting data from its surrounding environment. These nodes transmit their data to a central processing unit or a data aggregator for further analysis. The collected data undergoes preprocessing to clean and prepare it for analysis. This step involves filtering out noise, handling missing values, and scaling the features to make them suitable for ML algorithms. Once the data is preprocessed, relevant features are extracted from it. Feature extraction techniques may include statistical measures, frequency domain analysis, or wavelet transforms. The extracted features are then fed into a machine learning model, which is responsible for classifying faults or anomalies in the sensor data. Various ML algorithms such as support vector machines (SVM), decision trees, or random forests can be used for this purpose. During the training phase, the ML model learns from labeled data, while in the testing phase, its performance is evaluated on unseen data. The output of the fault diagnosis system consists of diagnosed faults or anomalies identified by the ML model. This output can be used to trigger alarms, adjust sensor parameters, or take corrective actions based on the detected faults. Additionally, there is a feedback loop in the framework, where information from the fault diagnosis output is used to improve the performance of the system over time by incorporating new information (Fig. 1).

Machine Learning in Fault Diagnosis

Machine learning (ML) techniques have gained significant traction in fault diagnosis, offering advanced capabilities for automated fault detection, classification, and prediction in various systems, including wireless sensor networks (WSNs). Here

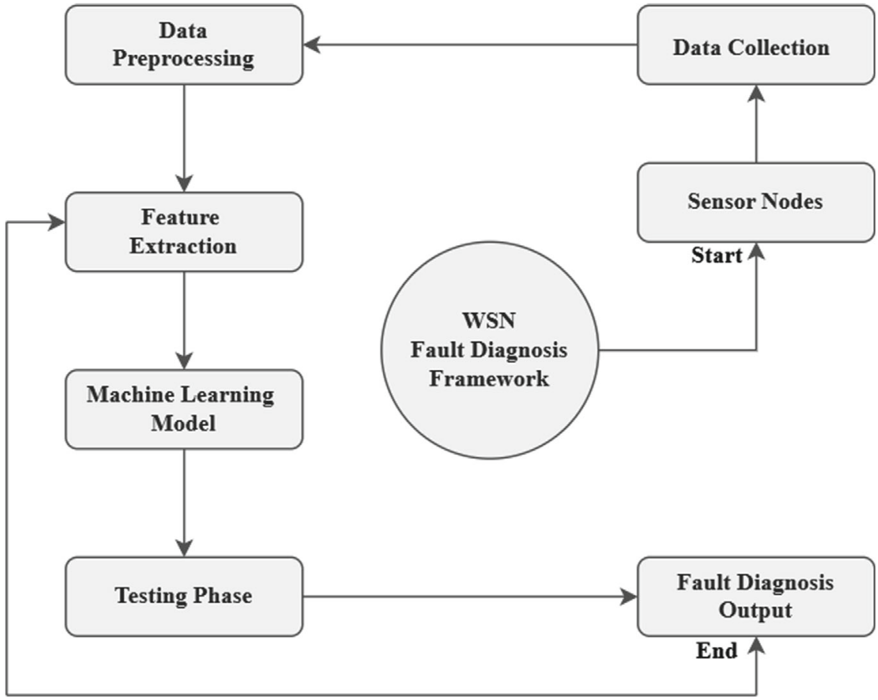


Fig. 1 Workflow of fault diagnosis process

are some key ways in which machine learning is applied in fault diagnosis within WSNs:

1. Anomaly Detection: ML algorithms can learn the normal behavior patterns of sensor data and identify deviations or anomalies indicative of faults. Techniques such as isolation forests, one-class SVM, and autoencoders are commonly used for anomaly detection in WSNs.
2. Classification of Faults: Supervised learning algorithms, such as support vector machines (SVM), decision trees, random forests, and neural networks, are utilized to classify different types of faults based on labeled training data. These algorithms can distinguish between various fault classes, such as sensor failures, communication errors, or environmental disturbances.
3. Regression Analysis: ML regression models are employed to predict the remaining useful life (RUL) of sensors or components within WSNs. By analyzing historical sensor data and failure events, regression models can estimate the time to failure or degradation of components, enabling proactive maintenance and resource allocation.
4. Fault Localization: ML techniques facilitate the localization of faults within the network by analyzing spatial and temporal correlations in sensor data. Clustering

algorithms like K-means or DBSCAN are employed to group sensors exhibiting similar behavior, aiding in pinpointing the location of faults.

5. Online Learning and Adaptation: WSNs operate in dynamic environments where sensor characteristics and fault patterns may change over time. Online learning algorithms, such as online sequential extreme learning machine (OS-ELM) and incremental support vector machine (ISVM), enable continuous adaptation and updating of fault diagnosis models based on incoming data streams.
6. Integration with IoT Platforms: ML-based fault diagnosis systems can be integrated with IoT platforms to enable remote monitoring, real-time analytics, and automated response mechanisms. This integration enhances the scalability and accessibility of fault diagnosis solutions in large-scale WSN deployments.
7. Machine learning plays a pivotal role in enhancing fault diagnosis capabilities within wireless sensor networks, enabling more accurate, efficient, and proactive approaches to fault detection, classification, and mitigation.

4 Conclusion and Future Work

This paper has provided an in-depth analysis of fault diagnosis in wireless sensor networks (WSNs) employing a machine learning (ML) approach. We have explored various ML techniques utilized for fault detection, classification, and localization in WSNs. The effectiveness of ML algorithms such as support vector machines (SVM), decision trees, neural networks, and ensemble methods has been discussed in detail. Additionally, we have highlighted the importance of feature selection and extraction in enhancing the performance of fault diagnosis systems in WSNs. There are several avenues for future research in this domain. Firstly, the integration of advanced ML techniques such as deep learning and reinforcement learning could be explored to improve fault diagnosis accuracy in WSNs further. Additionally, investigating novel feature selection and extraction methods tailored specifically for WSNs could enhance the efficiency of fault detection algorithms. Moreover, the development of distributed fault diagnosis frameworks capable of handling large-scale WSNs with dynamic topologies remains an important area for future exploration. Furthermore, conducting real-world experiments and case studies to validate the performance of ML-based fault diagnosis systems in practical WSN deployments would be valuable. Finally, considering the energy efficiency and resource constraints of WSNs in the design of fault diagnosis algorithms will be crucial for practical implementation in real-world scenarios. Overall, addressing these research directions can contribute significantly to the advancement of fault diagnosis techniques in WSNs, thereby improving the reliability and performance of these networks in various applications.

References

1. Osisanwo FY, Akinsola JET, Awodele O, Hinmikaiye JO, Olakanmi O, Akinjobi J (2017) Supervised machine learning algorithms: classification and comparison. Int J Comput Trends Technol (IJCTT) 48(3):128–138
2. Wang X, Lin X, Dang X (2020) Supervised learning in spiking neural networks: a review of algorithms and evaluations. Neural Netw 125:258–280
3. Zhang Z, Mehmood A, Shu L, Huo Z, Zhang Y, Mukherjee M (2018) A survey on fault diagnosis in wireless sensor networks. IEEE Access 6:11349–11364
4. Priya PI, Muthurajkumar S, Daisy SS (2022) Data fault detection in wireless sensor networks using machine learning techniques. Wirel Pers Commun 1–22
5. Mahmood T, Li J, Pei Y, Akhtar F, Butt SA, Ditta A, Qureshi S (2022) An intelligent fault detection approach based on reinforcement learning system in wireless sensor network. J Supercomput 78(3):3646–3675
6. Regin R, Rajest S, Singh B (2021) Fault detection in wireless sensor network based on deep learning algorithms. EAI Endorsed Trans Scalable Inf Syst 8(32)
7. Saeed U, Jan SU, Lee YD, Koo I (2021) Fault diagnosis based on extremely randomized trees in wireless sensor networks. Reliab Eng Syst Saf 205:107284
8. Fan F, Chu SC, Pan JS, Lin C, Zhao H (2023) An optimized machine learning technology scheme and its application in fault detection in wireless sensor networks. J Appl Stat 50(3):592–609
9. Azzouz I, Boussaid B, Zouinkhi A, Abdelkrim MN (2020) Multi-faults classification in WSN: a deep learning approach. In: 2020 20th International conference on sciences and techniques of automatic control and computer engineering (STA). IEEE, pp 343–348
10. Panda M, Gouda BS, Panigrahi T (2020) Fault diagnosis in wireless sensor networks using a neural network constructed by deep learning technique. Nat Inspired Comput Wirel Sens Netw 77–101
11. Kim T, Vecchietti LF, Choi K, Lee S, Har D (2020) Machine learning for advanced wireless sensor networks: a review. IEEE Sens J 21(11):12379–12397
12. Dwivedi RK, Rai AK, Kumar R (2020) A study on machine learning based anomaly detection approaches in wireless sensor network. In: 2020 10th International conference on cloud computing, data science & engineering (Confluence). IEEE, pp 194–199
13. Kaur R, Sandhu JK, Sapra L (2020) Machine learning technique for wireless sensor networks. In: 2020 Sixth international conference on parallel, distributed and grid computing (PDGC). IEEE, pp 332–335

ODBot—Object-Detecting Robot

Shamish Bramhekar, Sharvari Bodas, Varada Deshmukh, Madhura Birajdar, and Siddharth Bhorge

Abstract The idea of an object-detecting robot built with the Raspberry Pi camera module and employing YOLO image processing algorithms is presented in this paper. Our robot uses YOLO's adaptability and Raspberry Pi's processing capacity to detect and recognize objects in its environment. The hardware arrangement consists of a camera module coupled to a microcontroller as well as a robot chassis with motors and a motor driver board. The camera module starts a live stream and analyzes it using the algorithm. The robot can recognize objects based on specified patterns or attributes by examining the photographs it has taken. The camera module with YOLO integration gives the robot the capacity to observe and engage with its surroundings on its own. Applications in the fields of surveillance, autonomous navigation, object monitoring, and human–robot interaction are made possible by this initiative. The object-detecting robot provides a viable strategy for creating intelligent robotic systems that can perceive and engage with their environment. Through this research, we want to show how advanced robotics technologies may be used effectively, opening the door for the creation of more complex and effective robotic systems.

Keywords Object detecting robot · Image processing · Computer vision · YOLO · Raspberry Pi 4 · Autonomous navigation

S. Bramhekar · S. Bodas (✉) · V. Deshmukh · M. Birajdar · S. Bhorge
Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India
e-mail: sharvari.bodas21@vit.edu

S. Bramhekar
e-mail: Shamish.bramhekar211@vit.edu

V. Deshmukh
e-mail: varada.deshmukh.21@vit.edu

M. Birajdar
e-mail: madhura.birajdar21@vit.edu

S. Bhorge
e-mail: siddharth.bhorge@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_64

1 Introduction

The development of intelligent and interactive robotic systems now has more options because of developments in the field of computer vision technology. One such application is the creation of an object detection robot utilizing YOLO image processing methods and the Raspberry Pi 4.

Blurred images and noise are some of the main problems with object detection. However, object detection using deep learning, mainly using YOLO, is able to learn more complex scenes and improve the average precision [1]. The object-detecting robot uses the ARM7 microcontroller's strength and the YOLO library's capabilities to perceive and engage with its surroundings. The YOLO library offers the required resources for image processing and object identification, while the camera module acts as the visual sensor, taking pictures of the environment. The microcontroller serves as the brain of the robot and is renowned for its great computing capability and networking features. It accepts picture data from the camera module and applies YOLO algorithms to process it. The widely used open-source computer vision library YOLO provides a wide range of tools for manipulating images, extracting features, and identifying objects. The robot makes judgments and takes action depending on the observed items. The motor driver enables the robot to traverse its surroundings on its own, recognize things of interest, and respond to them appropriately.

This offers a wide range of possible applications, such as situations involving human–robot interaction, autonomous robots, item tracking, image segregation, and surveillance systems. We have developed a robotic system that can sense and understand visual information, allowing it to interact intelligently with its environment, by merging hardware and software components.

Robot grasping is also an established and well-researched problem for example text grasping. The text underscores how important it is for robots to grasp objects effectively in real-world tasks. It talks about the difficulties robots face in understanding how to grasp objects correctly, especially in figuring out the best way to position their grippers. To tackle this, it suggests using advanced techniques like deep learning, transfer learning, and domain adaptation. One interesting approach mentioned is taking a CNN that was originally trained to recognize text boxes and adapting it to recognize the best way to grasp different objects. This method has shown great potential, performing well not only in standard tests but also in real-world situations [2].

2 Literature Review

Earlier, MATLAB was used for indoor localization systems which were expensive and time-consuming. So, the authors developed a system that used the YOLO algorithm. There is an increase in the color range. The system uses a camera input to analyze visual data, extract relevant features, estimate location, and provide navigation assistance. The system's components include camera, image processing algorithms, and localization modules. Experimental results show promising localization accuracy and navigation guidance [3].

The paper provides an overview of indoor localization methods, focusing on their applications in navigation, tracking, and context-aware services. Indoor localization is widely used in domains like AR, autonomous robots, and monitoring. The authors analyze each method's strengths, weaknesses, and potential challenges, and showcase real-world implementations and applications [4].

There are four types of object detection methods: machine learning-based processes, OBIA-based processes, knowledge-based processes, and template matching-based processes. It is difficult to obtain high-resolution airborne images and detection of objects in optical remote sensing. The study highlights the significance of scene understanding and recognition, particularly in contexts like land-use analysis and environmental monitoring [5].

Accuracy is a key factor in object detection and recognition. With artificial neural networks (ANNs) for image evaluation we can achieve it. The three key steps of the system are interpretation, ANN processing, and image processing. The form and color of the item are extracted as numerical values during the image processing stage. The ANN stage for object recognition is then fed these values. The ANN's output is translated into human-understandable linguistic phrases [6].

The importance of object detection in various technical fields and their explanations are highlighted where the popular deep learning framework TensorFlow and the computer vision library OpenCV are used. Real-time implementation and detecting objects effectively have become practically possible, which expands its potential for a wide range of applications [7].

This paper highlights the use of convolutional neural network (CNN) as an effective method for image recognition. The You Only Look Once (YOLO) approach to object detection reimagines the task of localizing objects as a regression problem, simplifying the process by predicting both bounding boxes as well as class probabilities directly from entire images in a single evaluation. YOLO achieves this efficiency and optimization by employing a single neural network [8].

This research paper delves into the fascinating world of object detection, particularly honing in on the YOLO algorithm and its advancements when compared to traditional CNNs. It highlights how YOLO keeps getting better and better, continuously evolving to meet new challenges. The paper also breaks down how YOLO actually works, showing how it spots targets and picks out important features. Plus, it explores practical uses of YOLO, like in finance, giving us a glimpse into how these technologies are shaping our world [9].

This paper highlights the importance of object detection and the complexities introduced by real-world images, including problems like noise, blurriness, and rotational variations. YOLO, a real-time object detection algorithm based on CNN, is used for this purpose. The survey primarily revolves around various adaptations made to the YOLO network, all aimed at enhancing the efficiency of object detection [10].

YOLOv3-Tiny adopts a single-shot detection approach, processing the entire image in a single forward pass through the neural network. The YOLO architecture is optimized for real-time applications by reducing the model size and computational complexity. The "Tiny" version typically involves sacrificing some precision for improved speed and efficiency. Modifications to the YOLOv3 architecture are involved, such as adjusting the network architecture, reducing the number of layers, or incorporating lightweight components to achieve a balance between accuracy and inference speed. This makes it suitable for resource-constrained environments like edge devices or real-time systems [11].

This paper discusses object detection in computer vision, with a focus on object detectors which are two-stage as well as single stage. The two-stage detectors use complex architectures with selective region proposals, while single-stage detectors employ simpler architectures to detect objects in one shot. Two-stage detectors generally have better detection accuracy, but single-stage detectors have faster inference times. The accuracy was significantly improved in single-stage detectors, sometimes surpassing two-stage detectors [12].

3 Methodology

Our project is to design and build an autonomous object-detecting robot using a Raspberry Pi 4. The robot is capable of detecting and avoiding obstacles in its environment using computer vision techniques. The robot utilizes a camera module to capture images and process them in real time to detect objects. To set up the Raspberry Pi, we assembled the hardware components on the chassis and connected them to the Raspberry Pi (camera module, motor driver, and ultrasonic sensors) and installed Raspbian OS on the Raspberry Pi. For capturing the image, first initialized the camera module and used Python and the PiCamera library. The captured image will be used to process and to identify objects or obstacles with the help of You Only Look Once (YOLO) algorithm. We have also developed a decision-making algorithm to analyze the object detection data and determine the robot's actions. This algorithm helps the robot decide whether to move forward, backward, left, right, or stop based on the detected objects and obstacle distances. Moving to the hardware part, we implemented a motor control library to control the movements of the robot's wheels. It adjusts the motor control signals according to the decisions made by the algorithm. We have integrated ultrasonic sensors to detect obstacles in the robot's vicinity. We are also using the sensor data to supplement the camera-based object detection and aid in obstacle avoidance. The whole combination of the camera-based object detection, ultrasonic obstacle detection, and decision-making algorithm makes our robot enable autonomous navigation. Regular testing and fine-tuning are essential to achieve reliable performance in real-world situations, and to achieve this, we tested our robot in various environments and scenarios to fine-tune the object detection and navigation algorithms and adjusted parameters as needed for optimal performance.

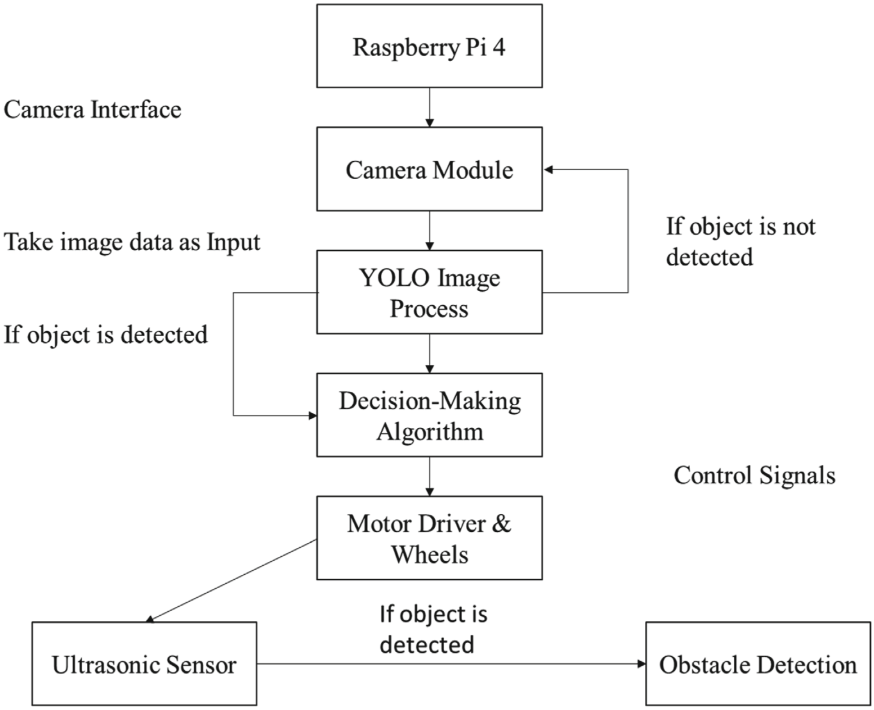


Fig. 1 Block diagram

Figure 1 shows the block diagram for the system. Raspberry Pi 4 is used as a camera interface to collect images of the robot's surroundings. This image data is then forwarded to YOLO image processing. If an object is detected, it takes into account factors such as size, type of the object, and its location to make the decision. After deciding the location, the algorithm sends the signals to the motor driver and with the help of an ultrasonic sensor object is detected.

Algorithm

YOLO is a well-known technique for object detection in computer vision and deep learning applications. It is a single-stage detector that can detect objects in real time. Unlike prior object detection methods, YOLO considers object detection to be a regression problem, learning to predict the bounding boxes as well as class probabilities for all things in an image in a single run.

Image as input

Output: Detected objects with associated bounding boxes and class labels.

1. Load pretrained weights into the neural network.
2. Make a grid of cells out of the image.
3. Using the neural network, forecast a fixed number of bounding boxes and their related class probabilities for each cell.

4. Apply a confidence score criterion to eliminate low-confidence forecasts.
5. Making use of non-max suppression to eliminate any unnecessary bounding boxes.
6. Return the remaining bounding boxes, together with their class labels and confidence scores.

4 Result and Discussion

Efficiency of YOLO algorithm for detection of some household objects:

First of all to test out the YOLO algorithm, we have used the webcam of our laptop for testing purposes. Figures 2 and 3 show the detection of indoor objects such as TV remote, bottle, and sofa.

We have images of different categories like cats, dogs, birds, traffic lights, potted plants, umbrellas, and kites for analyzing the confidence score of the respective images. We have analyzed the confidence scores for each object under normal, dark and bright lighting conditions and mentioned it in Table 1. Under normal lighting conditions, the images have the highest confidence score and also a high accuracy in identifying objects, and the objects are well recognized. Under bright lighting conditions, the accuracy of the images slightly drops, the confidence score of the images is slightly lower than the normal lighting conditions, and the objects are still identified reasonably accurately. Under dark lighting conditions, the images have the lowest confidence score, and so detecting the objects is difficult; for example, the traffic light image is not detected.

We can see in Fig. 4, that the algorithm is also capable of detecting objects in which it can camouflage with the background with impressive accuracy. But with the increase in the level of the camouflage, the algorithm finds it hard to get the accurate result. It can also be inferred for the above figures. There are three levels of intensity

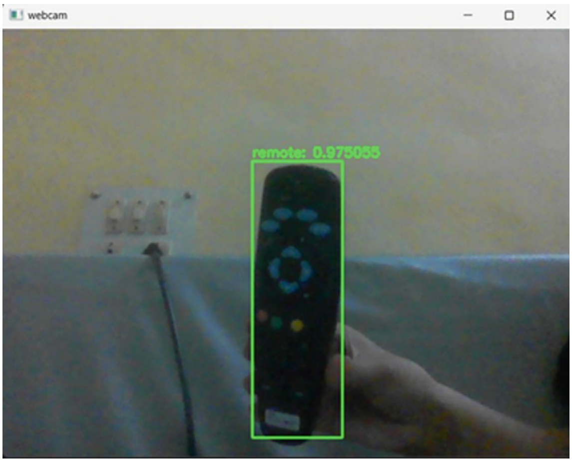


Fig. 2 Detection of a remote

Fig. 3 Detection of a bottle

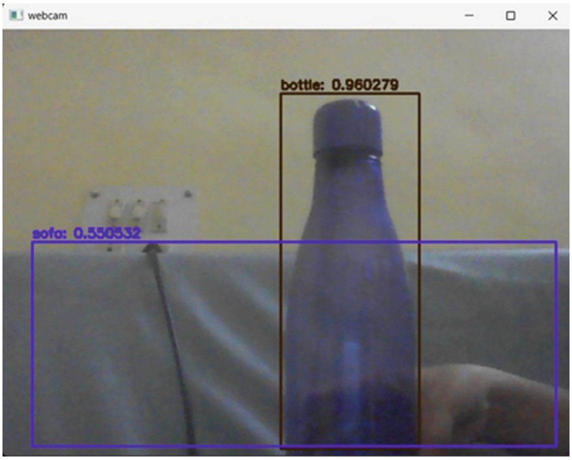


Table 1 Confidence scores under different lighting conditions

Object	Normal	Dark	Bright
Cat	0.9954	0.9724	0.9812
Dog	0.9993	0.9840	0.9992
Bird	0.9955	0.9978	0.9962
Traffic	0.7649	Not detected	Not detected
Potted plant	0.9974	0.9490	0.9962
Umbrella	0.8121	0.7465	0.7751
Kite	0.8876	0.7720	0.6010

of camouflage, and it can detect the bird in the first two cases but is not able to do that in the third case.

Testing the algorithm for COCO dataset

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

where:

TP Positive object class is correctly predicted
TN Negative object class is correctly predicted
FP Positive object class is predicted incorrectly
FN Negative object class is predicted incorrectly

Table 2 Shows the values of accuracy, precision, and recall, respectively, for different objects in the dataset

Fig. 4 Camouflage level 1, 2, 3

Table 2 Accuracy, precision, and recall values for some classes of COCO dataset

Object	Accuracy (%)	Precision (%)	Recall (%)
Cat	94	95	82
Dog	95	95	85
Bird	92	94	78
Traffic	89	91	72
Potted plant	94	94	78

5 Future Scope

Increasing object detection accuracy and the number of recognizable objects would be the first task to improve the system. Object-detecting robots like ODBot could find numerous applications in various technology domains by adding specific features. First application is to operate as a helping hand in the military, in dangerous situations where it is risky to send a soldier or any human, object-detecting robots can easily access the place or the area and can send required information like analyzing the environment and searching for people or specific things like bombs or ammunition. Another application can be in the emergency situations like fire or gas leakage in a building or in mining industries because accessing such places can be risky and sending an actual person to check the place can be dangerous, and robots like these can be deployed to get the required images and sensing of the area. In addition to cameras, we can add a Li-DAR sensor to map the area in 3D so that we can better detect the objects and increase the accuracy.

6 Conclusion

In conclusion, the development of an object-detecting robot is a significant advancement in robotics and artificial intelligence. The robot's ability to accurately detect and recognize objects in its environment has the potential to revolutionize various industries, such as manufacturing, logistics, health care, and home automation. With advancements in computer vision, sensor technologies, and machine learning algorithms, object-detecting robots have become more efficient, precise, and versatile.

However, there are still challenges that need to be addressed in the development of object-detecting robots, such as improving their robustness in different environments, addressing ethical considerations related to privacy and security, and ensuring that they are cost-effective and accessible to many different user types.

References

1. Liu C, Tao Y, Liang J, Li K, Chen Y (2018) Object detection based on YOLO network. In: 2018 IEEE 4th information technology and mechatronics engineering conference (ITOEC). IEEE, pp 799–803
2. Karaoguz H, Jensfelt P (2019) Object detection approach for robot grasp detection. In: 2019 International conference on robotics and automation (ICRA). IEEE, pp 4953–4959
3. Heya TA, Arefin SE, Chakrabarty A, Alam M (2018) Image processing based indoor localization system for assisting visually impaired people. In: 2018 Ubiquitous positioning, indoor navigation and location-based services (UPINLBS). IEEE, pp 1–7
4. Morar A, Moldoveanu A, Mocanu I, Moldoveanu F, Radoi IE, Asavei V, Gradinaru A, Butean A (2020) A comprehensive survey of indoor localization methods based on computer vision. Sensors 20(9):2641
5. Fatima SA, Kumar A, Pratap A, Raoof SS (2020) Object recognition and detection in remote sensing images: a comparative study. In: 2020 International conference on artificial intelligence and signal processing (AISP). IEEE, pp 1–5
6. Cruz JPN, Dimaala ML, Francisco LGL, Franco EJS, Bandala AA, Dadios EP (2013) Object recognition and detection by shape and color pattern recognition utilizing artificial neural networks. In: 2013 International conference of information and communication technology (ICoICT). IEEE, pp 140–144
7. Talele A, Patil A, Barse B (2019) Detection of real time objects using TensorFlow and OpenCV. Asian J Converg Technol (AJCT) ISSN-2350-1146
8. Mittal N, Vaidya A, Kapoor S (2019) Object detection and classification using Yolo. Int J Sci Res Eng Trends 5(2)
9. Jiang P, Ergu D, Liu F, Cai Y, Ma B (2022) A review of Yolo algorithm developments. Procedia Comput Sci 199 (2022):1066–1073

10. Rekha BS, Marium A, Srinivasan GN, Shetty SA (2020) Literature survey on object detection using YOLO. Int Res J Eng Technol (IRJET) 7(06)
11. Adarsh P, Rathi P, Kumar M (2020) YOLO v3-Tiny: object detection and recognition using one stage improved model. In: 2020 6th International conference on advanced computing and communication systems (ICACCS). IEEE, pp 687–694
12. Diwan T, Anirudh G, Tembhurne JV (2023) Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimed Tools Appl 82(6):9243–9275

Industry Automation Using IOT

Puja Chavan, Soham Borkar, Prasad Rathod, Swapnil Gawali, Shivprasad Kavathe, and Amit Dolas

Abstract In this paper, we propose a practical IoT-based industry automation system utilizing the ESP32 microcontroller and various sensors, such as PIR, smoke, fire, and humidity/temperature sensors. The system aimed to overcome challenges faced by the previous manual monitoring system and enhance operational efficiency and safety measures in industrial environments. The previous system's limitations included manual monitoring, delayed responses, and increased human error. Real-time data collection and analysis were absent, preventing proactive measures and prompt response to hazardous situations. To address these issues, we proposed an IoT-based automation system that streamlined operations, improved safety, and enabled real-time monitoring and control. Integration of sensors enhanced the system's capabilities. The IoT-based automation system offered several advantages. Real-time data collection and analysis enabled proactive monitoring and timely. Automated processes increased operational efficiency and productivity by eliminating manual intervention interventions, minimizing accident risks. Sensor integration improved safety by providing early detection and response to fire and gas incidents. These enhancements resulted in cost savings, reduced downtime, and increased overall effectiveness.

P. Chavan (✉) · S. Borkar · P. Rathod · S. Gawali · S. Kavathe · A. Dolas
Vishwakarma Institute of Technology, Pune, MH 411037, India
e-mail: puja.cholke@vit.edu

S. Borkar
e-mail: soham.borkar22@vit.edu

P. Rathod
e-mail: prasad.rathod22@vit.edu

S. Gawali
e-mail: swapnil.gawali22@vit.edu

S. Kavathe
e-mail: shivprasad.kavathe22@vit.edu

A. Dolas
e-mail: amit.dolas22@vit.edu

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_65

Keywords Industry · IOT · Automation · Sensors · Safety · Industrial safety · Industry automation · Proactive monitoring

1 Introduction

The Internet of Things (IoT) has revolutionized various industries by connecting devices and sensors to the Internet, enabling them to communicate and share data. One of the key areas where IoT has made significant strides is in industrial hazard detection, its management, and automation. The use of linked devices and sensors to automate industrial processes and tasks is known as IoT-enabled industrial automation [1]. This system has the ability to greatly raise industrial environments' productivity, effectiveness, and safety.

Proposed system focuses on practical way to the integration of sensors, actuators, and cloud infrastructure to create intelligent systems capable of detecting and responding to hazards in industrial environments. Fire detection is addressed by incorporating fire sensors that activate fans to mitigate fire spread. Smoke detection involves integrating smoke detectors with exhaust fans to eliminate smoke and reduce risks [2]. Motion sensors are implemented to automatically control lighting based on detected motion, optimizing energy consumption, and improving safety. Temperature monitoring, facilitated by temperature sensors, enables remote control of fan and cooling systems, ensuring efficient regulation of the industrial environment, and preventing overheating or underutilization of cooling resources [3]. By leveraging these technologies, proposed system aims to enhance safety, energy efficiency, and productivity in industrial settings in a practical manner. Building upon these previous studies, proposed system aims to integrate multiple sensors and actuators into a cohesive IoT system capable of hazard detection and industrial automation [4]. By combining fire, smoke, motion, and temperature sensors, we seek to create an intelligent system that responds in real-time to various hazards, optimizing safety, energy efficiency, and productivity in industrial environments. The benefits of this system include increased productivity, efficiency, and safety, as well as cost savings [5]. However, there are also potential challenges such as data security, integration with existing systems, and the need for specialized skills and knowledge. Proposed system stands out due to its practical implementation and its comprehensive approach which sets it apart from other papers in the field of IoT-based industrial automation. While many studies describe theoretical ideas, our research goes one step further and shows how sensors—such as motion, temperature, smoke, and fire—are really implemented in an actual industrial setting [6]. Our unique selling point is proposed system focus on usability and real-time monitoring through the use of the interactive and user-friendly Blynk mobile application. Systems focus is on safety and, demonstrating how IoT may be used in real industrial practice to improve industrial safety and efficiency. In the field, proposed system is unique due to its practical application and emphasis on real-time safety and monitoring.

Through a literature review and case studies, this paper aims to provide insights into the benefits, challenges, and future potential of IoT-enabled industrial automation. The rest of the paper is organized as follows: The relevant work in this subject is outlined in Sect. 2 of this document. The methodology of the study is outlined in Sect. 3, while Sect. 4 explores the advantages and difficulties associated with IoT-driven industrial automation. This section presents case studies to illustrate the application of this technology in various industries. Last section concludes with some last remarks and explores potential future study topics.

2 Literature Review

IoT-enabled industrial automation has received considerable focus in recent times. The benefits of this technology have been widely acknowledged, including improved efficiency, productivity, and safety. This section provides a review of relevant literature published in the past 5 years [7]. Numerous researches have been centered on the application of IoT-enabled industrial automation in specific industries. For example, Chen et al. explored the use of IoT in smart agriculture, where sensors and actuators are used to automate irrigation, pest control, and other farming activities. The authors found that the use of IoT-enabled automation can improve crop yields and reduce labor costs [8].

In the manufacturing industry, IoT-enabled automation has been used to optimize production processes and improve product quality. A framework for the use of IoT in intelligent manufacturing was put forth by them, where sensors and other IoT devices are used to collect and analyze data to optimize manufacturing processes. The authors found that the use of IoT-enabled automation can lead to significant improvements in productivity and quality [9].

Another area of research has focused on integrating IoT-driven automation with other technologies, such as machine learning (ML) and artificial intelligence (AI). For example, Zhou et al. proposed an AI-based approach for the optimization of industrial automation systems. The authors used a combination of IoT-enabled sensors, ML algorithms, and cloud computing to optimize industrial processes and reduce energy consumption [10].

In addition to the benefits of IoT-enabled industrial automation, there are also challenges that must be addressed. Several studies have explored these challenges and proposed solutions. Jafri et al. discussed the challenges of data security in IoT-enabled industrial automation and proposed a framework for secure data transmission. The authors found that by implementing appropriate security measures, the risk of cyberattacks can be mitigated [11].

Another challenge is the integration of IoT-enabled automation with existing systems. Several studies have proposed solutions to address this challenge. For example, Zhang et al. 2018 suggested a middleware-based approach for the integration of IoT-enabled automation with legacy systems. The authors found that this approach can enable seamless integration of new and existing systems [12].

In summary, research on IoT-enabled industrial automation has increased significantly in the previous 5 years. Numerous studies have explored the interests of this technology in various industries, including agriculture and manufacturing. In addition, research has focused on addressing the challenges of data security and integration with existing systems. The integration of IoT-enabled automation with other technologies, such as AI and ML, has also been explored. Overall, the literature supports the potential of IoT-enabled industrial automation in improving efficiency, productivity, and safety in industrial environments.

3 Methodology

3.1 System Design

Figure 1 showcases the digital connections between various components used in the industry automation system. It demonstrates how the mentioned components are interconnected to form a functional system. Proposed system involves the integration of sensors, devices, and actuators with the Internet to enable communication and control of various industrial processes remotely. The proposed methodology for implementing IoT-based industrial automation can be outlined as follows:

Identifying the Industry Process. The first step is to identify the industrial process that needs to be automated. It can be a manufacturing process, supply chain management, or any other process that requires monitoring and control.

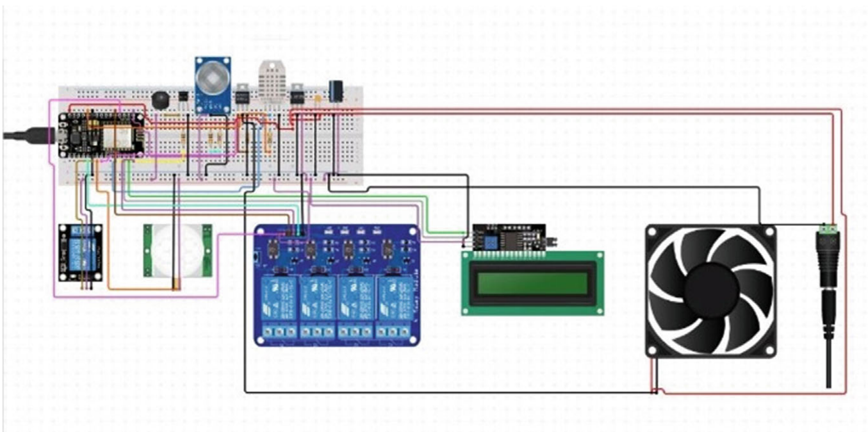


Fig. 1 Connection diagram for industry automation

Identifying the Relevant Data. Once the process is identified, the next step is to identify the relevant data that needs to be collected to automate the process. This can include temperature, humidity, pressure, and other sensor data.

Choosing the Right Sensors and Devices. After identifying the relevant data, the next step is to choose the right sensors and devices that can collect and transmit this data to the cloud.

Setting Up the Cloud Infrastructure. The data gathered from the sensors and devices must be stored and processed using the cloud infrastructure. Its configuration enables it to process enormous amounts of data and offer real-time insights.

Creating Hardware Connections. The connection of sensors, actuators, and motors, with microcontroller with the help of breadboard is the next step in the development of this IOT system.

Integrating the Hardware and Software. The next step is to integrate the hardware and software components to enable communication between the sensors and devices, cloud infrastructure, and the software application.

Testing and Validating the System. The final step is to test and validate the system to ensure that it is working as intended. This can involve testing the sensors and devices, cloud infrastructure, and the software application.

Enabling Remote Access and Control. System have remote access capabilities, allowing authorized personnel to monitor and control the industrial process from any location enabling flexibility and timely response to critical situations.

Overall, the literature supports the potential of IoT-enabled industrial automation in improving efficiency, productivity, and safety in industrial environments.

3.2 Proposed Methodology

Figure 2 explains working flow of the system. As if fire detected, smoke detected, humidity or temperature increases, and PIR sensor senses movement, then data is sent to the Blynk server to show and control it remotely. We have set up an industrial IoT system for monitoring and controlling motion, temperature, humidity, and gas detection, integrated with the Blynk server for remote access.

Sensor Setup. Selected appropriate sensors for motion, temperature, humidity, and smoke detection. Installed a PIR motion sensor for accurate motion detection. This system enhances industrial process control and safety while providing remote access for monitoring.

Blynk Server Integration. Created an account on the Blynk server and obtained necessary authentication tokens. Established remote control and visualization through the Blynk app. Connected the hardware to the server using Wi-Fi or Ethernet.

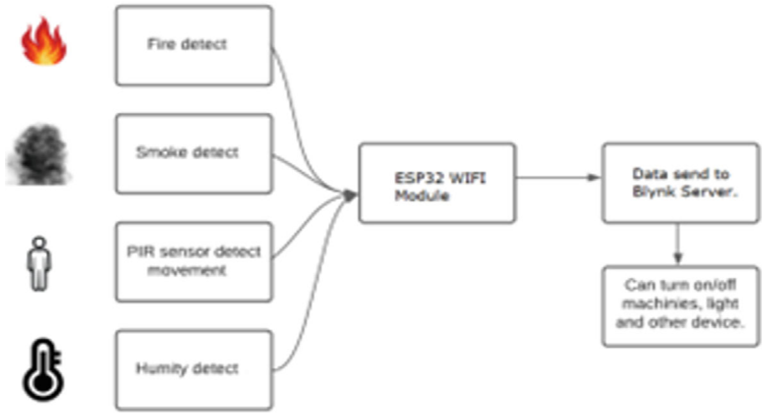


Fig. 2 Working flow for industry automation

Motion Detection and Alert System. Configured a motion sensor to detect motion. Programmed the microcontroller to send motion data to the Blynk server.

Set up email notifications for the administrator upon motion detection.

Temperature and Humidity Regulation. Calibrated sensors according to system requirements. Continuously monitored data and activated the exhaust fan when predefined thresholds, like a temperature above 32°, were reached. Sent temperature and humidity data to the Blynk server for remote monitoring.

Gas Detection and Safety Measures. Configured a sensor to detect hazardous gases or smoke. Integrated the sensor with a microcontroller to send data to a Blynk server. Defined safety protocols for triggering alarms and emergency shutdown systems when gas is detected.

Remote Lighting Control. Connected the lighting system to a microcontroller using relay modules. Created a Blynk app button for remote lighting control.

Testing and Validation. Conducted thorough testing of the system, including sensor accuracy and automation functionality. This system enhances safety and convenience by responding to gas hazards and offering remote lighting control. Regular maintenance is important for system reliability.

By implementing methodology above, we successfully implemented industry automation using IoT and sensors as described in the methodology. The system offers improved efficiency, safety, and remote monitoring capabilities. Future enhancements may include integrating additional sensors, implementing predictive analytics, or expanding the control functionalities.

4 Result and Discussion

In Fig. 3, user interface for Blynk application is shown. Data sensed from sensors in the industry is shown in the above figure. We have also given option to turn on and turn off the light and motor.

Blynk Mobile Application Platform. The Blynk mobile application is used to visualize and monitor the system's results. The application displays real-time values of temperature and humidity, providing insights into the environmental conditions within the industrial setting. The status of the smoke and fire sensing is indicated by values of 0 and 1, allowing for immediate detection and response to potential hazards. The application also includes two switch controls for motor and light control, enabling convenient remote operation for turning them on and off.

Temperature and Humidity Monitoring. On the Blynk mobile application, the system accurately measures and shows the temperature and humidity data. Users can easily monitor and track the environmental conditions in the industrial conditions in

Fig. 3 Online remote control

real-time. This feature ensures that the temperature remains within acceptable limits, while maintaining optimal humidity levels. We have also implemented automatic turning on and off fan along with remote control.

Smoke and Fire Sensing. The smoke and fire detection sensors accurately detect the presence of hazardous gases or smoke. The Blynk mobile application reflects the status of smoke sensing and fire sensing as values of 0 and 1, respectively. This enables quick identification and response to potential fire hazards, ensuring the safety of the industrial environment and personnel. If smoke is detected, then exhaust is automatically turned on.

Motor and Light Control. The Blynk mobile application includes two switch controls for motor and light control. Users can remotely turn on and off the motor and lighting systems with a simple toggle on the application. Along with remote control, it turns on and off automatically as motion detected. This feature provides convenient control over industrial processes, allowing for efficient operation and energy conservation.

Overall System Success. The successful practical integration of sensors, Blynk server, and the mobile application demonstrates the effectiveness of the industry automation system. Real-time monitoring, control, and visualization capabilities contribute to improved efficiency, safety, and convenience within the industrial environment.

5 Conclusion

In an industrial setting, proposed system used sensors and actuators to detect changes in temperature, motion, smoke, and fire. In order to stop the spread of the fire, the fire sensor was essential in initiating fan activation. By turning on an exhaust fan, the smoke detector decreased the risk to people and property. Motion sensors improved productivity, optimized energy use, and increased safety by detecting motion and adjusting lights. Temperature sensors allowed for the remote management of cooling and fan systems, allowing for environmental regulation, increased operational efficiency, and lower energy costs. Proposed system showed off real-time hazard identification and industrial automation capabilities, successfully integrating sensors and actuators. As a USP, we described an industrial automation and hazard detection and its resolution in practical way. Predictive analytics and machine learning for more advanced hazard detection in the industry can be considered as some improvements in the system. Integration of the system with existing infrastructure will add to safety in the environment of the industry. By providing the practical and physical world application of IOT-based industrial hazard detection and automation, proposed system differentiates itself from other theoretical papers.

References

1. Wang X, Li J, Wang F, Li Y, Xue F (2019) An IoT-enabled intelligent manufacturing system for product quality improvement. Int J Adv Manuf Technol 105(5–6):2201–2216
2. Uttam BK, Baspusaheb GA, Shivaji JP, Pisal RS (2017) Industrial automation using IoT
3. Ramaiah NS and Kiran RC, Dharshan DS (2019) IoT based industrial automation. In: Proceedings of international conference on sustainable computing in science, technology and management (SUSCOM), Amity University Rajasthan, Jaipur, India, 26–28 Feb 2019
4. Bello LL, Steiner W (2019) A perspective on IEEE time-sensitive networking for industrial communication and automation systems. Proc IEEE 107(6):1094–1120. https://doi.org/10.1109/JPROC.2019.2905334
5. Sundaram GS, Patibandala B, Santhanam H (2013) Bluetooth communication using a touchscreen interface with the Raspberry Pi. IEEE. 978-1-4799-0053-4/13/31.00 2013 IEEE
6. Piyare R, Lee SR (2013) Smart home control and monitoring system using smart phone ICCA, ASTL 24:83–86
7. Breivold HP, Sandström K (2015) Internet of things for industrial automation—challenges and technical solutions. In: 2015 IEEE international conference on data science and data intensive systems, Sydney, NSW, Australia, pp 532–539. https://doi.org/10.1109/DSDIS.2015.11
8. Chen Z, Jiang P, Wang Y, Zhang W (2021) A smart agriculture system based on the internet of things. IEEE Internet Things J 8(4):2457–2465. Author F, Author S (2016) Title of a proceedings paper. In: Editor F, Editor S (eds) Conference 2016, LNCS, vol 9999. Springer, Heidelberg, pp 1–13
9. Masetti G, Marazzi F, Di Cecilia L, Rovati L (2018) IOT-based measurement system for wine industry. 2018 Workshop on metrology for industry 4.0 and IoT, Brescia, Italy, pp 163–168. https://doi.org/10.1109/METROI4.2018.8428308
10. Zhou Y, Wang Z, Cheng M (2020) An AI-based optimization approach for industrial automation systems in the era of the industrial internet of things. IEEE Trans Industr Inf 16(4):2284–2294
11. Jafri MM, Saleem K, Sher M, Arshad R (2020) Cybersecurity in industrial internet of things (IIoT): Challenges and solutions. IEEE Internet Things J 7(11):10421–10433
12. Zhang L, Liu S, Yang Y (2018) An IoT middleware for integrating legacy industrial systems. IEEE Internet Things J 5(3):1693–1703

Natural Language Processing for Sentiment Analysis: Neural Network Models

Balusupati Anil Kumar, Pirangi Vijay Kumar, Talari Swapna, Rajendhar Reddy Gaddam, S. Dhanalakshmi, and B. Pradeep

Abstract Sentiment analysis, a subfield of natural language processing (NLP), has garnered significant attention in recent years due to its wide range of applications in various domains, such as social media monitoring, customer feedback analysis, and market research. In this review paper, we focus on the application of neural network models for sentiment analysis tasks. Neural network models have shown remarkable performance in capturing intricate patterns and nuances in textual data, making them particularly suitable for sentiment analysis tasks. We provide a comprehensive overview of neural network architectures employed in sentiment analysis, including recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformer-based models. Additionally, we discuss various preprocessing techniques, feature representations, and training strategies commonly used in conjunction with neural network models for sentiment analysis. Furthermore, we present a comparative analysis of different neural network architectures in terms of their effectiveness, efficiency, and scalability for sentiment analysis tasks. Finally, we outline current challenges and future directions in leveraging neural network models for sentiment analysis, aiming to provide insights and guidance for researchers and practitioners in the field of NLP.

Keywords NLP · Sentiment analysis · Neural network · Machine learning · Customer feedback

B. Anil Kumar (✉) · P. Vijay Kumar · T. Swapna · R. R. Gaddam · S. Dhanalakshmi · B. Pradeep
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: anilkumarb@cmritonline.ac.in

P. Vijay Kumar
e-mail: vijaykumar@cmritonline.ac.in

T. Swapna
e-mail: swapnat@cmritonline.ac.in

R. R. Gaddam
e-mail: rajendarreddy.gaddam@cmritonline.ac.in

B. Pradeep
e-mail: pradeepb26@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_66

1 Introduction

The rise of social media, e-commerce, and online reviews has generated an unprecedented amount of unstructured text data, providing important insights into human thoughts, attitudes, and emotions. Businesses, researchers, and policymakers can use sentiment analysis, a subfield of natural language processing (NLP) [1], to automatically analyze and interpret textual data to gain actionable insights into public opinion, customer feedback, and market trends. Because they can learn complex patterns and representations from massive text data, neural network models have revolutionized sentiment analysis. These models, from simple feedforward neural networks to complex architectures like RNNs, CNNs, and transformer-based models, have captured human language's nuanced semantics and context dependence.

1.1 Natural Language Processing (NLP) in Sentiment Analysis

Sentiment analysis is a subfield of computational linguistics that relies heavily on natural language processing (NLP) to glean subjective information from text data. Automatically identifying the positive, negative, or neutral emotion or opinion conveyed in a text is the goal of sentiment analysis. Thanks to natural language processing (NLP) methods, computers can comprehend and interpret human speech, opening the door to massive sentiment analysis across many media types (e.g., social media, reviews, news articles, etc.).

Natural language processing facilitates sentiment analysis through various methods, such as tokenization, syntactic parsing [2], part-of-speech tagging [3], and semantic analysis. Tokenization breaks down text into smaller units (tokens), such as words or phrases, while part-of-speech tagging assigns grammatical categories to each token, aiding in understanding the syntactic structure of sentences. Syntactic parsing involves analyzing the grammatical structure of sentences, which helps in understanding the relationships between words. Semantic analysis delves deeper into the meaning of words and sentences, enabling systems to infer sentiment from context and linguistic cues. Natural language processing-powered sentiment analysis systems employ machine learning algorithms, including neural network models, to classify text into sentiment categories. These models learn from labeled data to recognize patterns and features indicative of sentiment, allowing them to accurately classify text as positive, negative, or neutral.

2 Literature Review

2.1 Traditional Sentiment Analysis (SA) Techniques

The evolution of sentiment analysis, especially with the advent of deep learning, has been a focal point in recent research. Kansara and Sawant [4] conducted an experimental comparison of classification algorithms, showcasing the superiority of deep learning techniques like LSTM and CNN over traditional methods like Naïve Bayes and logistic regression. Dhola and Saradva [5] highlighted the significance of sentiment analysis in understanding consumer preferences from Internet and social media data. Lagrari and Elkettani [6] emphasized the importance of sentiment analysis in comprehending user-generated content, especially with recent advancements in deep learning algorithms. Singh and Karur [7] categorized sentiment analysis into different types and discussed their respective significance, challenges, and applications.

Yildirim [8] evaluated the performance of RNN-based architecture in sentiment analysis compared to traditional bag-of-words methods, showcasing the advantages of neural network approaches. Kanakaraj and Guddeti [9] proposed strategies to enhance sentiment classification by incorporating semantics into feature vectors, resulting in improved prediction accuracy compared to traditional methods. Hasan et al. [10] developed a framework for sentiment analysis of Twitter data, achieving high accuracy by combining natural language processing techniques with models like Bag of Words and TF-IDF. Chong et al. [11] outlined their approach to sentiment analysis in tweets, focusing on subjectivity, semantic connection, and polarity analysis to better capture sentiment in short text formats.

2.2 Neural Network Models in Sentiment Analysis (SA)

In research, Ouyang et al. [12] demonstrated the superior performance of deep LSTM architectures in speech recognition tasks, highlighting the advantages of deep hierarchical models. Kumar and Rani [13] proposed a self-adaptive probabilistic neural network (PNN) for sentiment analysis of tweets, achieving improved accuracy by optimizing the smoothing value. Severyn and Moschitti [14] introduced a novel methodology for initializing convolutional neural networks (CNNs) for tweet sentiment analysis, achieving competitive results in sentiment analysis tasks. Lin et al. [15] presented a comparison-enhanced bi-LSTM model with multi-head attention (CE-B-MHA) for text sentiment analysis, outperforming several existing models on sentiment analysis datasets. Chen et al. [16] proposed a CNN-RLSTM deep neural network model for target-based sentiment analysis, achieving superior performance compared to SVM and other neural network models. Borele and Borikar [17] summarized the current state of sentiment analysis methods, emphasizing machine learning techniques, such as Naïve Bayes, maximum entropy, and support vector machine. Paliwal et al. [18] utilized artificial neural networks (ANN) for sentiment analysis of tweets, demonstrating the accuracy and efficiency of ANN in predicting sentiment polarity.

2.3 Neural Network Architectures for Sentiment Analysis

In research, Trofimovich [19] participated in the SentiRuEval-2016 Twitter sentiment analysis assignment, focusing on tweets mentioning financial institutions and telecom providers. They compared three methods based on neural networks: GRU (Gated Recurrent Unit) neural network (NN), convolutional neural network (CNN), and SVM classifier with domain adaptation. The GRU-based solution proved most effective, achieving state-of-the-art results in sentiment classification.

Stojanovski et al. [20] developed a deep learning method for analyzing Twitter conversations and identifying emotions. They tested various classification techniques, including a convolutional neural network (CNN), and evaluated performance using datasets from the SemEval competition. Their architecture outperformed existing methods in sentiment analysis and emotion detection tasks. Camacho-Collados and Pilehvar [21] explored the impact of text preprocessing on neural text classifiers. They evaluated different preprocessing techniques, such as tokenization, lemmatization, and lowercasing, highlighting the importance of this stage in achieving optimal performance. Çano and Morisio [22] focused on building large datasets for sentiment analysis, particularly in the domain of musical emotions. They examined various neural network architectures and found that parallel convolutions with filter lengths up to 3, along with adjusted max pooling area sizes, yielded effective results in sentiment analysis models. Jebbara and Cimiano [23] addressed sentiment analysis as a relation extraction problem, breaking it down into identifying aspect and opinion terms, labeling opinion terms with sentiment, and extracting relations between them. They proposed a neural network-based architecture that outperformed existing methods in aspect-opinion relation extraction tasks. Shuang et al. [24] introduced SICENN (Sentiment Information Collector-Extractor architecture) for sentiment analysis, comprising SIC (Sentiment Information Collector) and a Sentiment Information Extractor (SIE). The SIC gathered sentiment information from phrases using Bidirectional Long Short-Term Memory (Bi-LSTM), while the SIE accurately extracted sentiment information. Their ensemble technique for merging outputs from multiple sub-extractors resulted in improved performance compared to existing approaches.

These studies highlight advancements in sentiment analysis using neural network architectures, preprocessing techniques, and innovative approaches to tackle sentiment analysis tasks.

3 Mathematical Model

Input Representation: Let us denote the input text data as X, where each x_i represents a word or token in the input sentence. We need to convert this text data into a numerical representation suitable for feeding into a neural network. This can be done using techniques like word embeddings (e.g., Word2Vec, GloVe) or character-level embeddings.

For example, if we use word embeddings, we can represent each word x_i as a d-dimensional vector w_i. Thus, the input sentence X can be represented as a sequence of word vectors: $X = [w_1, w_2, \ldots, w_n]$, where n is the length of the input sentence.

Neural Network Architecture: Neural network architecture for sentiment analysis, such as a feedforward neural network or RNN (Recurrent Neural Network). Let us denote the parameters of the neural network as Θ.

Forward Propagation: Given the input sentence X and the parameter Θ, the neural network computes the output scores for each sentiment category. Let us denote the output scores as S. The forward propagation process involves passing the NN layers data to compute the output scores.

Mathematically, the forward propagation can be represented as:

$S = f_\Theta(X)$ where $f(\Theta)$ represents the function computed by the neural network with parameters Θ.

Activation Function: Typically, a softmax activation function is used in the output layer of the neural network to compute the probabilities of each sentiment category. Let us denote the softmax function as $\mathrm{softmax}(\cdot)$.

The output scores S are passed through the softmax function to obtain the probability distribution over sentiment categories: $P = \mathrm{softmax}(S)$ where P is a vector of probabilities corresponding to each sentiment category.

Loss Function: During training, the model's predictions are compared to the ground truth labels to compute the loss. Let us denote the ground truth sentiment label for the input X as Y. We will use a suitable loss function, such as cross-entropy loss L_{CE}.

The loss function is defined as:

$$L_{\mathrm{CE}}(P, Y) = -\sum_i Y_i \log(P_i).$$

where Y_i is the ground truth probability for sentiment category i, and P_i is the predicted probability for sentiment category i.

Training Objective: The objective during training is to minimize the loss function L_{CE} with respect to the parameters Θ. This is typically done using optimization algorithms like Adam or SGD (stochastic gradient descent).

Mathematically, the training objective can be formulated as: $\min_\Theta L_{\mathrm{CE}}(P, Y)$.

Backpropagation: The gradients of the loss function with respect to the parameters θ are computed using backpropagation, and the parameters are updated iteratively using the chosen optimization algorithm.

Mathematically, the parameter updates can be represented as: $\theta \leftarrow \theta - \alpha \nabla_{\theta} L_{\mathrm{CE}}$ where α is the learning rate, and $\nabla_{\Theta} L_{\mathrm{CE}}$ represents the gradient of the loss function with respect to the parameters θ.

This mathematical model outlines the key components and equations involved in building a neural network model for sentiment analysis. By training the model on labeled datasets and optimizing the parameters using techniques like backpropagation, the model can learn to classify text data into different sentiment categories effectively.

4 Neural Network Architectures for Sentiment Analysis

4.1 Recurrent Neural Networks (RNNs)

The directed cycles' internal memory is used by RNN to process sequential information. RNN may recall and reuse information from earlier computations on the next input element, unlike typical neural networks. Long short-term memory (LSTM) RNNs can use long memory for hidden layer activation functions (Fig. 1).

4.2 Convolutional Neural Networks (CNNs)

Figure 2 illustrates various filters with different region sizes to get multiple 1-max pooling results. Concatenating 1-max pooling data from feature maps creates a CNN feature. CNN features are, intuitively, feature map maximum values. NN layers synthesize high-level features from CNN features to connect these values.

4.3 BERT

BERT (Bidirectional Encoder Representations from Transformers) has significantly advanced sentiment analysis tasks by capturing bidirectional context information from text data. Unlike traditional models that process text in a unidirectional manner, BERT employs a transformer architecture, enabling it to consider both preceding and succeeding words when encoding representations of words or sentences. This bidirectional approach allows BERT to capture richer contextual information, making it particularly effective for sentiment analysis [27] tasks where understanding the relationships between words is crucial for accurate sentiment classification.

BERT is pre-trained on large corpora of text data using unsupervised learning tasks, such as masked language modeling [28]. This pre-training process enables BERT [29] to learn general language representations [30], which can be fine-tuned

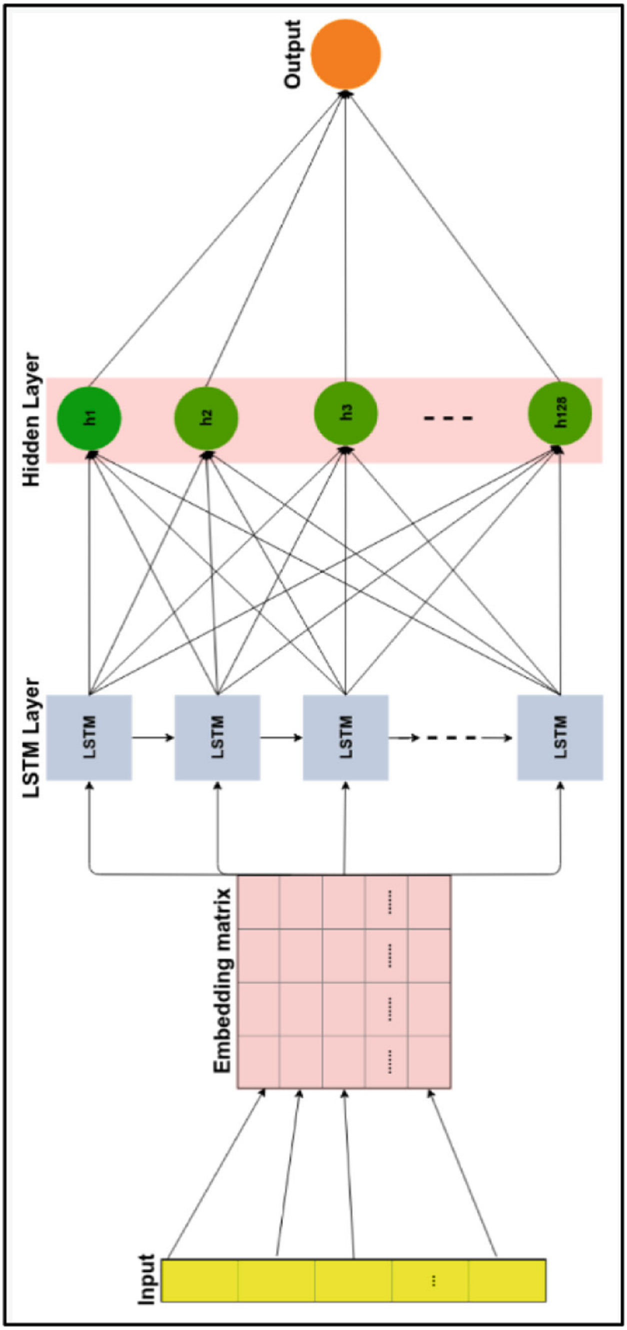


Fig. 1 RNN-LSTM [25]

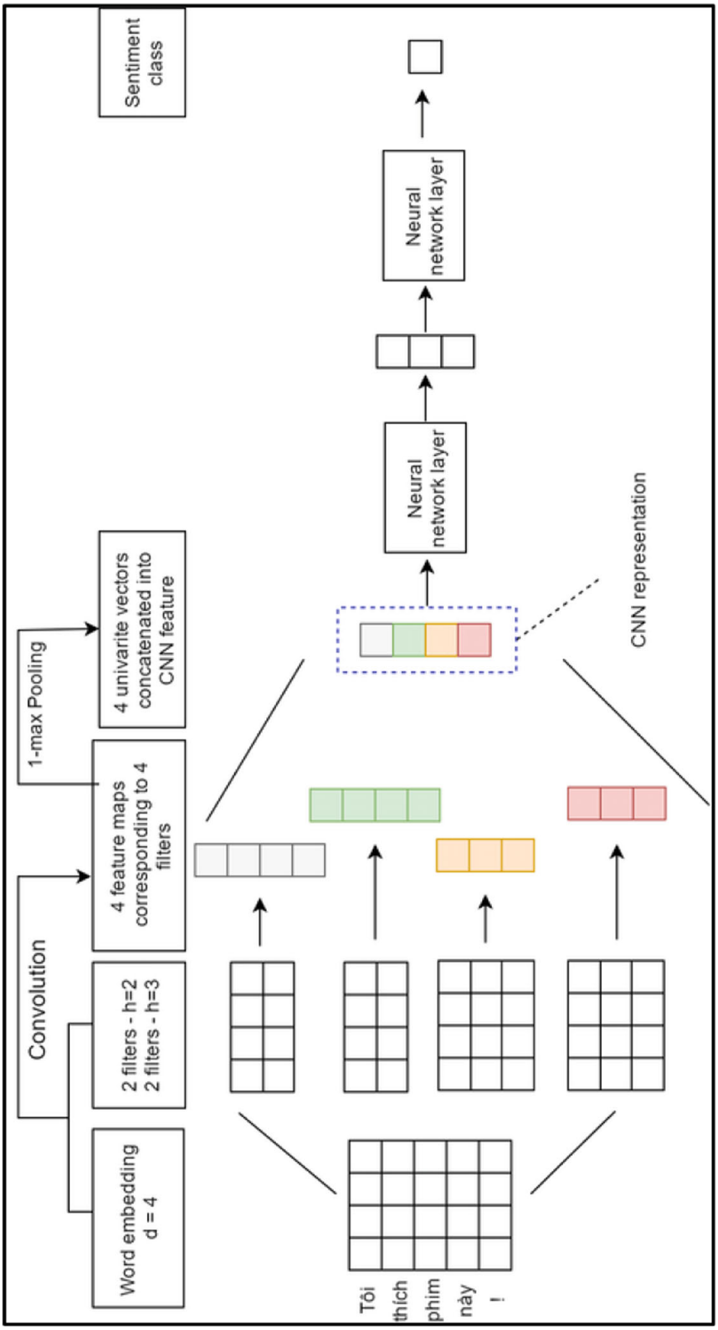


Fig. 2 CNN model for sentiment analysis [26]

on downstream tasks like sentiment analysis with relatively small amounts of labeled data. Fine-tuning BERT involves updating its parameters on task-specific datasets, allowing it to adapt to the nuances and characteristics of the sentiment analysis task at hand.

By leveraging the contextual embeddings learned by BERT, sentiment analysis models can better capture the nuanced semantics and dependencies within text data, resulting in improved performance on sentiment classification tasks. BERT-based models have achieved state-of-the-art results in various sentiment analysis benchmarks, demonstrating their effectiveness in capturing fine-grained sentiment information across different domains and languages.

BERT has emerged as a powerful tool for sentiment analysis, offering enhanced capabilities in capturing contextual information and achieving superior performance on sentiment classification tasks compared to traditional models. Continued research and advancements in transformer-based architectures promise to further enhance the capabilities of BERT and its variants for sentiment analysis and other natural language processing tasks.

5 Conclusion and Future Work

This review paper has examined the advancements in natural language processing (NLP) for sentiment analysis, specifically focusing on Neural Network Models. Through an extensive analysis of recent literature, it is evident that neural network models have significantly improved the correctness and robustness of sentiment analysis (SA) tasks across various areas and languages. The utilization of deep learning architectures, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformer-based models like BERT, has demonstrated remarkable performance in capturing semantic nuances and contextual dependencies within text data, thereby enhancing sentiment classification outcomes. Several challenges and opportunities for future research have been identified. Addressing issues related to data scarcity, domain adaptation, model interpretability, and bias mitigation remains crucial for advancing the field. Furthermore, exploring innovative techniques for leveraging multimodal data and incorporating external knowledge sources could further enhance the sophistication and applicability of sentiment analysis systems.

References

1. Egger R, Gokce E (2022) Natural language processing (NLP): an introduction: making sense of textual data. In: Applied data science in tourism: interdisciplinary approaches, methodologies, and applications. Springer International Publishing, Cham, pp 307-334

2. Pei S, Wang L, Shen T, Ning Z (2019) Da-bert: enhancing part-of-speech tagging of aspect sentiment analysis using bert. In: Advanced parallel processing technologies: 13th international symposium, APPT 2019, Tianjin, China, Aug 15–16 2019, Proceedings 13. Springer International Publishing, pp 86–95
3. Gómez-Rodríguez C, Alonso-Alonso I, Vilares D (2019) How important is syntactic parsing accuracy? An empirical evaluation on rule-based sentiment analysis. Artif Intell Rev 52:2081–2097
4. Kansara D, Sawant V (2020) Comparison of traditional machine learning and deep learning approaches for sentiment analysis. In: Advanced computing technologies and applications: proceedings of 2nd International conference on advanced computing technologies and applications—ICACTA 2020. Springer, Singapore, pp 365–377
5. Dhola K, Saradva M (2021) A comparative evaluation of traditional machine learning and deep learning classification techniques for sentiment analysis. In: 2021 11th International conference on cloud computing, data science & engineering (confluence). IEEE, pp 932–936
6. Lagrari FE, Elkettani Y (2021) Traditional and deep learning approaches for sentiment analysis: a survey. Adv Sci Technol Eng Syst J 6(4):1–7
7. Singh S, Kaur H (2021) Comparative sentiment analysis through traditional and machine learning-based approach. In: Cognitive behavior and human computer interaction based on machine learning algorithm, pp 315–338
8. Yildirim S (2020) Comparing deep neural networks to traditional models for sentiment analysis in Turkish language. In: Deep learning-based approaches for sentiment analysis, pp 311–319
9. Kanakaraj M, Guddeti RMR (2015) NLP based sentiment analysis on Twitter data using ensemble classifiers. In: 2015 3rd International conference on signal processing, communication and networking (ICSCN). IEEE, pp 1–5
10. Hasan MR, Maliha M, Arifuzzaman M (2019) Sentiment analysis with NLP on Twitter data. In: 2019 International conference on computer, communication, chemical, materials and electronic engineering (IC4ME2). IEEE, pp 1–4
11. Chong WY, Selvaretnam B, Soon LK (2014) Natural language processing for sentiment analysis: an exploratory analysis on tweets. In: 2014 4th International conference on artificial intelligence with applications in engineering and technology. IEEE, pp 212–217
12. Ouyang X, Zhou P, Li CH, Liu L (2015) Sentiment analysis using convolutional neural network. In: 2015 IEEE International conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing. IEEE, pp 2359–2364
13. Kumar A, Rani R (2016) Sentiment analysis using neural network. In: 2016 2nd International conference on next generation computing technologies (NGCT). IEEE, pp 262–267
14. Severyn A, Moschitti A (2015) Twitter sentiment analysis with deep convolutional neural networks. In: Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval, pp 959–962
15. Lin Y, Li J, Yang L, Xu K, Lin H (2020) Sentiment analysis with comparison enhanced deep neural network. IEEE Access 8:78378–78384
16. Chen S, Peng C, Cai L, Guo L (2018) A deep neural network model for target-based sentiment analysis. In: 2018 International joint conference on neural networks (IJCNN). IEEE, pp 1–7
17. Borele P, Borikar DA (2016) An approach to sentiment analysis using artificial neural network with comparative analysis of different techniques. IOSR J Comput Eng (IOSR-JCE) 18(2):64–69
18. Paliwal S, Khatri SK, Sharma M (2019) Sentiment analysis and prediction using neural networks. In: Advanced informatics for computing research: 2nd international conference, ICAICR 2018, Shimla, India, July 14–15 2018, Revised Selected Papers, Part I 2. Springer, Singapore, pp 458–470
19. Trofimovich J (2016) Comparison of neural network architectures for sentiment analysis of Russian tweets. In: Computational linguistics and intellectual technologies: proceedings of the international conference dialogue, pp 50–59

20. Stojanovski D, Strezoski G, Madjarov G, Dimitrovski I, Chorbev I (2018) Deep neural network architecture for sentiment analysis and emotion identification of Twitter messages. Multimedia Tools Appl 77:32213–32242
21. Camacho-Collados J, Pilehvar MT (2017) On the role of text preprocessing in neural network architectures: an evaluation study on text categorization and sentiment analysis. arXiv preprint arXiv:1707.01780
22. Çano E, Morisio M (2019) A data-driven neural network architecture for sentiment analysis. Data Technol Appl 53(1):2–19
23. Jebbara S, Cimiano P (2017) Aspect-based relational sentiment analysis using a stacked neural network architecture. arXiv preprint arXiv:1709.06309
24. Shuang K, Zhang Z, Guo H, Loo J (2018) A sentiment information collector–extractor architecture based neural network for sentiment analysi. Inf Sci 467:549–558
25. Dang NC, Moreno-García MN, De la Prieta F (2020) Sentiment analysis based on deep learning: a comparative study. Electronics 9(3):483
26. Vo QH, Nguyen HT, Le B, Nguyen ML (2017) Multi-channel LSTM-CNN model for Vietnamese sentiment analysis. In: 2017 9th International conference on knowledge and systems engineering (KSE). IEEE, pp 24–29
27. Sayeed MS, Mohan V, Muthu KS (2023) BERT: a review of applications in sentiment analysis. HighTech Innov J 4(2):453–462
28. Chung YA, Zhang Y, Han W, Chiu CC, Qin J, Pang R, Wu Y (2021) W2v-bert: combining contrastive learning and masked language modeling for self-supervised speech pre-training. In: 2021 IEEE automatic speech recognition and understanding workshop (ASRU). IEEE, pp 244–250
29. Pota M, Ventura M, Fujita H, Esposito M (2021) Multilingual evaluation of pre-processing for BERT-based sentiment analysis of tweets. Expert Syst Appl 181:115119
30. Deepa MD (2021) Bidirectional encoder representations from transformers (BERT) language model for sentiment analysis task. Turk J Comput Math Educ (TURCOMAT) 12(7):1708–1721

A Machine Learning Approaches for Elevating Content Quality and Authenticity Assessment on Quora

Ch. Ravikumar, Mulagundla Sridevi, Macha Sarada, M. Radha, and Vadapally Praveen Kumar

Abstract In the rapidly evolving landscape of digital technology, the Internet has become an indispensable tool for communication and information sharing. Platforms like Quora play a pivotal role in this digital age, enabling users to ask and answer questions on a wide range of topics. However, ensuring the quality and authenticity of content on Quora remains a challenging task. This research paper delves into the critical task of classifying messages on Quora, distinguishing between useful and irrelevant content, and assessing the authenticity of user-generated questions. The current system employed on Quora falls short of achieving optimal accuracy in these tasks, prompting the development of an enhanced system leveraging machine learning techniques. This study utilizes Naïve Bayes, logistic regression, and support vector machine algorithms to create a robust model. Remarkably, the proposed system achieves an impressive accuracy rate of up to 99.87%, effectively addressing the limitations of the existing system. This research highlights the potential of machine learning in improving content quality and authenticity assessment on question-and-answer platforms contributing to a more reliable and informative online environment.

Keywords Naïve Bayes · Support vector machine · Spam · Insincere questions · Regression · Web application

---

Ch. Ravikumar (✉)
Department of Artificial Intelligence & Data Science, Chaitanya Bharathi Institute of Technology, Hyderabad, India
e-mail: chrk5814@gmail.com

M. Sridevi · M. Radha
Department of CSE, CVR College of Engineering, Hyderabad, India

M. Sarada
Department of CSE, Priyadarshini Institute of Science & Technology for Women, Khammam, India

V. Praveen Kumar
Department of CSE (Data Science), CVR College of Engineering, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_67

1 Introduction

In today's rapidly evolving technological landscape, ubiquitous Android applications like WhatsApp, Telegram, Instagram, and Twitter have revolutionized global communication, enabling instant message exchange. While this convenience enhances connectivity, it introduces persistent threats, such as spam messages and malicious content, posing risks to data integrity and user privacy. To safeguard digital communication, there is an urgent need for effective message evaluation [1]. Simultaneously, platforms like Quora face challenges in classifying user-generated content as honest or deceptive, necessitating innovative solutions involving machine learning, natural language processing, and data analytics [2].

A significant contribution to email spam filtering comes from Alsadoon et al.'s comprehensive review [1], highlighting the evolution of spam filtering techniques. Another notable effort by Alurkar et al. introduces an enhanced technique using support vector machine (SVM) algorithms for accurate email spam detection [2]. In the social media domain, Wang et al.'s research on "Detecting Inappropriate Content in Social Media with Network Embedding" emphasizes the role of machine learning and network embedding techniques in monitoring content quality and authenticity [3]. These efforts collectively strive to ensure the reliability of online platforms amidst the challenges of our increasingly digital world.

Drawing inspiration from the insights provided by influential researchers in the field, particularly Fig. 1's questions designed to foster genuine and meaningful conversations and Fig. 2's questions often lacking sincerity, our research addresses the intricate challenges of spam detection, content authenticity assessment, and user experience enhancement in contemporary communication platforms and Q&A forums.

Fig. 1 Key indicates sincere questions list

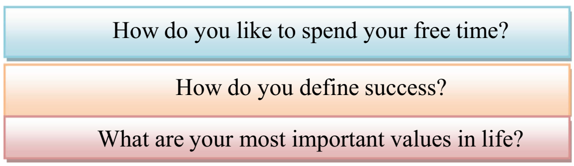
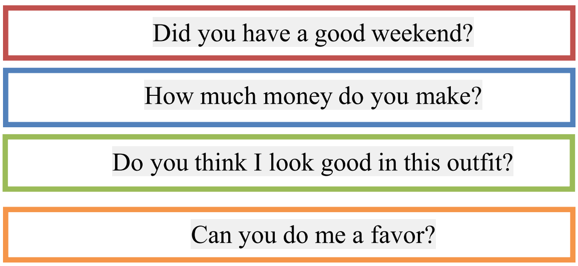


Fig. 2 Key indicates insincere questions list



To achieve this, our proposed system employs regression algorithms for efficient classification of SMS messages and Quora queries. Additionally, we prioritize user satisfaction by incorporating a user-friendly graphical interface (GUI) developed using the Python framework Flask. Looking forward, our ambition is to further enhance our system to fortify user security against evolving spamming technologies.

In the subsequent sections of this research paper, we will elaborate on the architectural framework and methodologies that form the basis of our proposed system. This builds upon the invaluable contributions of the researchers mentioned earlier, aligning our work with the principles of fostering genuine communication and meaningful interactions.

2 Literature Review

Quora, a widely used platform for knowledge sharing, faces challenges, such as deceptive questions and spam due to its open nature. This literature review explores research on improving message classification and question authenticity assessment on Quora, with a focus on leveraging machine learning techniques.

Muhana et al. introduced a method using K-means clustering, K-nearest neighbor (KNN), and decision trees for spam differentiation, but our proposed approach surpasses it in time efficiency and accuracy for both emails and Quora inquiries [4]. Aakash Atul et al. presented a system distinguishing spam from legitimate messages, emphasizing the superiority of their machine learning-based approach over manual configuration. Their system aims to classify incoming emails based on spam criteria, leveraging publicly available Internet domains [5]. Hendri et al. categorized Quora queries as sincere or insincere, addressing data imbalance issues and recommending further research on their approach [6]. Suresh Babu et al. faced challenges in classifying Quora queries, whereas our system efficiently handles both email and Quora classifications using regression algorithms [7]. Moses et al. introduced a comprehensive approach for email spam evaluation, achieving high accuracy with support vector machine versatility for both Quora inquiries and emails [8]. Karthika Renuka et al. presented a diverse set of techniques for email spam detection and classification [9].

2.1 Contributions

In summary, the literature review emphasizes the importance of using machine learning techniques to enhance message classification and question authenticity assessment on platforms like Quora. Our study contributes significantly by employing Naïve Bayes, logistic regression, and support vector machines, achieving a remarkable accuracy rate of up to 99.87%. This addresses shortcomings in the current system, providing users with more reliable message classification and authenticity assessment.

3 Methodology

In the dynamic tech landscape, challenges arise from deceptive messages, spam, and dishonest questions on platforms like Quora. To combat this, our team developed a system employing diverse machine learning techniques, particularly text classification methods [10]. Notably, our proposed system excels in efficiency and accuracy, addressing challenges in existing systems related to accuracy and time complexity. These are the steps:

(a) Gathering the Dataset Information,
(b) Preprocessing,
(c) Training the Model,
(d) Model Deployment on Web Pages.

The proposed system operates on a three-tier structure, depicted in the architectural diagram. Input data is captured and processed, and then passed to a trained model for precise results. The trained model, constructed using the SMS Spam Collection dataset, employs machine learning algorithms, including Naïve Bayes, non-probabilistic algorithms, and regression algorithms like logistic regression, to accurately categorize messages and queries (Fig. 3).

Proposed System Architecture

Result presentation utilizes scripting languages, such as HTML and CSS, to create a visually appealing, user-friendly interface. Python framework Flask is employed to develop a web application, hosted locally or non-locally, providing accessibility for users. This architecture supports multiple users, ensuring the delivery of precise results. Cascading style sheets enhance the website's aesthetic appeal.

3.1 Message Classification

Within the system, SMS messages are categorized as either useful (ham) or non-useful (junk). This classification relies on specific words or parameters distinguishing between these categories [11, 12].

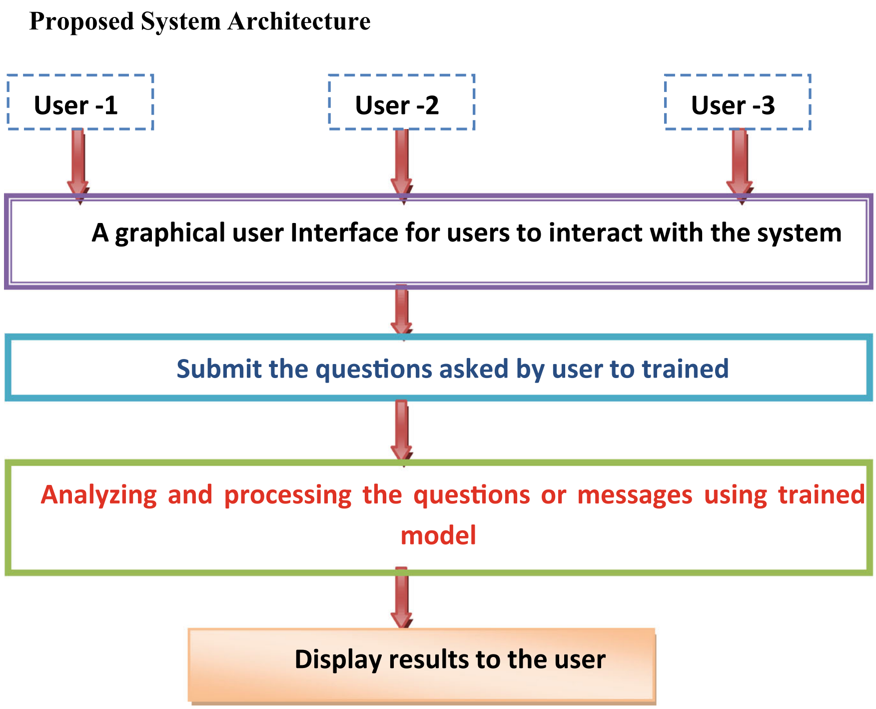


Fig. 3 Architecture of proposed system

4 Implementation and Results

The system implementation requires expertise in machine learning algorithms (logistic regression, support vector machines, Naïve Bayes) using Python, Scikit-Learn, and TensorFlow. This integration enhances user experience on Quora, ensuring content reliability. Logistic regression is applied to Quora queries and SMS messages, involving data retrieval, vectorization, splitting, classifier selection, model fitting, prediction, accuracy evaluation, confusion matrix construction, and data visualization. Support vector machines and Naïve Bayes classifiers are similarly implemented, with results analyzed for accuracy, time complexity, and throughput. The Naïve Bayes classifier excels in accuracy despite assuming attribute independence, while support vector machines demonstrate effectiveness in categorizing new text data (Table 1).

Table 1 Machine learning attributes for proposed system

S. No.	Methods	Accuracy (%)	Precision	Time (s)	Recall
1	Naïve Bayesian	96	0.95	1.2	0.87
2	Support vector machine	99.87	0.98	0.12	0.96
3	Logistic regression	98	0.97	0.20	0.88

5 Conclusion

In this study, we utilized diverse machine learning algorithms, including non-probabilistic, regression, and support vector machine, to categorize spam messages and insincere questions on Quora. Our system demonstrated remarkable accuracy with minimal processing time, achieving an impressive 99.87% accuracy rate. Future efforts will focus on integrating cutting-edge technologies like Blockchain to enhance content classification, ensuring unparalleled accuracy, data integrity, and trust in the dynamic landscape of digital communication and knowledge-sharing platforms.

References

1. Alsadoon A et al (2013) Email spam filtering: a review. In: 2013 International conference on control, decision and information technologies (CoDIT). IEEE
2. Alurkar AA et al (2017) An enhanced technique for spam email detection using SVM algorithm. In: 2017 International conference on trends in electronics and informatics (ICEI). IEEE
3. Wang Z et al (2020) Detecting inappropriate content in social media with network embedding. IEEE Trans Comput Soc Syst 7(2):371–382
4. Raza M, Jayasinghe D, Muslam MMA (2021) A comprehensive review on email spam classification using machine learning algorithms. In: 2021 International conference on information networking (ICOIN), pp 327–332
5. Alurkar AA et al (2017) A proposed data science approach for email spam classification using machine learning techniques. In: 2017 Internet of Things business models, users, and networks, pp 1–5
6. Octaviani NL et al (2020) Comparison of multinomial Naïve bayes classifier, support vector machine, and recurrent neural network to classify email spams. In: 2020 International seminar on application for technology of information and communication (iSemantic), pp 17–21
7. Anitha PU, Rao CVG, Babu S (2017) Email spam classification using neighbor probability-based Naïve Bayes algorithm. In: 2017 7th International conference on communication systems and network technologies (CSNT), pp 350–355
8. Nagdeve AS, Ambekar MM (2020) Spam detection by designing machine learning approach in twitter stream. In: 2020 International conference on smart innovations in design, environment, management, planning and computing (ICSIDEMPC), pp 126–130
9. Vinitha VS, Renuka DK (2019) Performance analysis of email spam classification using different machine learning techniques. In: 2019 International conference on advances in computing and communication engineering (ICACCE), pp 1–5
10. Priyambowo H, Adriani M (2019) Insincere question classification on question answering forum. In: 2019 International conference on electrical engineering and informatics (ICEEI), pp 390–394

11. Chen J, Siu SW (2020) Machine learning approaches for quality assessment of protein structures. Biomolecules 10(4):626. https://doi.org/10.3390/biom10040626
12. Parisot O, Bruneau P, Hitzelberger P (2023) Astronomical images quality assessment with automated machine learning. In: Proceedings of the 12th International conference on data science, technology and applications. https://doi.org/10.5220/0012073800003541

Enhancing Forest Fire Detection Using SOS Framework by Using DLBT

Deepsikha Adidam, Lalasa Reddy, and Sangeeta Gupta

Abstract In this groundbreaking forest fire detection and SOS alert system, cutting-edge technology and blockchain seamlessly converge to transform the landscape of wildfire monitoring and response. The process initiates with the continuous collection of CCTV footage across forested regions, undergoing comprehensive video preprocessing to enhance image quality. A sophisticated fire detection algorithm meticulously scrutinizes video frames, incorporating a confirmation mechanism to minimize false alarms. Upon confirming the presence of a fire, the system autonomously triggers the generation of SOS alerts, delivering crucial details for immediate action. At the heart of this innovative system lies its integration with the Polygon blockchain. Here, smart contracts take charge of managing SOS alert operations, ensuring transactions are secure and authenticated through authorized keys. Polygon's smart contracts emit events that promptly notify the fire detection system, maintaining its vigilant state. Prior to deployment in fire-prone areas, the system undergoes rigorous testing, and continuous monitoring and maintenance are implemented to ensure long-term reliability and effectiveness. To further enhance safety and response efficiency, comprehensive public awareness campaigns disseminate essential information regarding the system's utilization during forest fire emergencies. This multifaceted approach revolutionizes forest fire management, significantly improving environmental preservation and the well-being of communities at risk.

Keywords CCTV cameras · Forest fires · Polygon · SOS alert system · Smart contracts

D. Adidam · L. Reddy (✉) · S. Gupta
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: lalasareddy2904@gmail.com

S. Gupta
e-mail: Sangeetagupta_cse@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_68

1 Introduction

1.1 Improved Forest Fire Detection

Forest fires pose a significant threat to our environment, wildlife, and human communities, requiring swift detection and response. Traditional methods of detection rely on human surveillance, leaving room for delays and inaccuracies. To address these challenges, this research project introduces an innovative system for early forest fire detection using strategically placed closed-circuit television (CCTV) cameras in forested areas. Powered by state-of-the-art deep learning (DL) algorithms, this system continuously analyzes camera data to improve accuracy, reduces false alarms, and enhances overall effectiveness in identifying wildfires.

1.2 Challenges of Previous Paper

The foundation of this system is built upon the work of Wang et al., who introduced novel wildfire image classification and region detection algorithms based on DL. However, their reliance on the FLAME dataset and potential real-time applicability issues limited their research. This current research seeks to address these limitations and expand upon them, offering a more comprehensive solution for early forest fire detection.

1.3 Reliable Alerts with Blockchain

Upon detecting a potential forest fire, the system activates an SOS alert mechanism that leverages blockchain technology for rapid notification of nearby communities and relevant authorities. The integration of blockchain technology ensures data integrity and security, making the system's findings reliable and tamper-proof. Collaborative partnerships with local forestry departments further enhance the system's situational awareness by mining historical data, enabling well-coordinated and swift responses through the use of smart contracts. In essence, this research bridges critical gaps in forest fire detection and response by combining cutting-edge computer vision technology, blockchain security, and intelligent alert systems to better protect our forests and communities from wildfires.

2 Literature Survey

In a study conducted by Zhang et al. [1], the focus was on forest wildfire detection within the context of a Machine Vision course. The authors introduced two key algorithms: one for wildfire image classification based on Reduce-VGGnet and another for wildfire region detection optimizing CNN with spatial and temporal features. These algorithms were incorporated into an experiment that required expertise in digital image processing, machine learning, and deep learning. The research was motivated by the increasing frequency and intensity of forest fires, necessitating automatic wildfire detection. It is worth noting that their proposed methods were tested exclusively on the FLAME dataset, potentially limiting their applicability to other datasets. Additionally, the wildfire region detection algorithm exhibited high computational intensity, making it potentially unsuitable for real-time applications [1].

The proposed methodology [2] involves using deep learning with RCNN and CNN with transfer learning models for accurate and efficient detection of forest fires. The fast region-based CNN is utilized, which includes the fast region-based-CNN and RPN to share the CNN model after removing the fully connected layer. Additionally, the YOLO v3 model, with 53 convolutional layers, is employed for faster and smoother detection of forest fires. The application includes online and offline modes for reporting fires and utilizes deep learning and transfer learning techniques for efficient and cost-effective fire prevention and reporting.

This paper [3] conducted research on the impact of forest fire particles and ash on transmission line conductor-plane gap breakdown characteristics. Their simulation experiments demonstrated that forest fires generate significant particles and ash, which can alter gap dielectric properties and reduce AC breakdown voltage. While the study offered recommendations to mitigate the effects of forest fires on transmission lines, it should be noted that it relied on simulations and focused solely on conductor-plane gaps, necessitating further real-world validation and comprehensive transmission line protection strategies.

This paper [4] proposed a dynamic task assignment scheme for forest firefighting aimed at reducing task completion time and ensuring uniform task completion times. This scheme was based on a fire-spreading model that predicted future fire behavior and reduced the search domain. The paper also explored optimal static task assignment and compared the proposed dynamic task assignment scheme with an auction algorithm through simulation examples to showcase its performance.

This paper [4] conducted an analysis of the effects of flame temperature, conductivity, and particles on gap breakdown characteristics of overhead transmission lines under forest fire conditions. They combined theoretical analysis with experimental data to propose different discharge mechanisms based on various factors. However, the study had limitations, such as not systematically considering the impact of vegetation-burning flames on transmission line gap insulation strength and not providing a clear understanding of the discharge mechanism of gaps under forest fire conditions. They also noted the importance of considering the polarity effect of DC voltage discharge in such scenarios.

This paper [5] proposed the methodology that involves the implementation of a blockchain-based open emergency alert system called EmergencyBlock. The system architecture includes a data aggregation layer, processing layer, and target area. The blockchain model in EmergencyBlock uses Merkle Tree to make data unchanging and irreversible. The workflow involves devices with processing power processing emergency data and putting it into a blockchain transaction. A proof of concept was implemented in Python version 3, and mining tests were conducted on different platforms to validate the system's performance. The methodology aims to collect surveillance data from various devices, store them securely, and initiate emergency warnings if necessary.

This paper [6] evaluated the accuracy of four satellite-derived fire products in a challenging environment, specifically the cloudy and mountainous area over subtropical China. They employed two methods for accuracy assessment: a spatiotemporal clustering algorithm based on historical fire records and a density-based random sampling and estimation method using Landsat 8 fire scenes. Their findings indicated that the VNP14DL product performed the best in detecting fires.

This paper [7] proposed an onboard fuzzy logic approach for active fire detection in the Brazilian Amazon Forest using satellite remote sensing. This method aimed to identify and detect active fire spots in the forest, minimizing the negative impact on the affected area. Fuzzy logic was chosen for its ability to mitigate image noise and recognize patterns in data, making it suitable for this pure send's location, and the ESP-32 cam takes pictures of the surroundings.

This paper [8] methodology involves using a Raspberry Pi, GSM, GPS, and ESP-32 cam modules to create a safety device. When the switch is turned on, the system is activated. The GSM module sends message and call alerts, the GPS module system is designed to be wearable and independent of mobile phones, providing a self-sufficient solution for women's safety.

This paper [9] proposed methodology involves the development of the NEMO smartphone SOS alert app using Kotlin, Java, and C++. The app utilizes sensors to track motion, orientation, and environmental factors in Android devices. It also employs the Android sensor framework to access raw sensor data and Google Play services for location APIs to provide emergency SMS alert services. Additionally, the methodology includes the use of Gradle for building automation, Unwired Labs API for device positioning, Postman API tool for HTTP request testing, and Firebase Authentication for user authentication. The aim is to address technical limits of telecommunications networks and societal concerns about monitoring and surveillance in emergency situations.

This paper [10] focuses on optimizing data collection for forest fire monitoring using a fixed wing unmanned aerial vehicle (UAV), treating it as a Dubins traveling salesman problem with neighborhood (DTSPN). The proposed Bi-Level Hybridization-based Metaheuristic Algorithm (BLHMA) employs differential evolution for continuous optimization and a self-organized multi-agent competition heuristic for discrete collection sequence determination. A secondary local search refines positions and headings. Simulation experiments, based on a real-world forest fire monitoring scenario, demonstrate BLHMA's superiority over three state-of-the-art algorithms, ensuring efficient UAV data collection and energy conservation in forest fire monitoring missions [10].

Finally, Cao et al. [11] proposed an Attention Enhanced Bidirectional Long Short-Term Memory Network (ABi-LSTM) for video-based forest fire smoke recognition. Addressing the challenge of detecting slow-moving, less salient smoke in long-distance wildfires, the model integrates spatial feature extraction, Bidirectional LSTM, and a temporal attention subnetwork. This enables the capture of discriminative spatiotemporal features in image sequences while assigning varying attention levels to different patches. Experimental results demonstrate superior accuracy (97.8%) and reduced false alarms across diverse scenarios, showcasing a 4.4% improvement over conventional image-based deep learning models.

3 Proposed Framework

3.1 Workflow

In this meticulously designed workflow, the forest fire detection and SOS alert system seamlessly fuse cutting-edge technology with the security and transparency of blockchain to revolutionize fire monitoring and emergency response in forested regions.

The process begins with the continuous and vigilant collection of critical data through a network of strategically placed CCTV cameras in forested areas. These cameras serve as the system's vigilant sentinels, capturing real-time video footage of the environment. To ensure the highest accuracy and reliability, the captured video data then undergoes a meticulous phase of video preprocessing. This phase leverages advanced techniques, such as those offered by OpenCV, to enhance image quality, stabilize frames, and reduce noise. By optimizing the quality of the incoming data, the system is better equipped to analyze it effectively.

As the video preprocessing step refines the raw data, a sophisticated fire detection algorithm takes center stage. This algorithm, often powered by state-of-the-art deep learning models like YOLOv3, tirelessly scrutinizes each video frame for potential signs of wildfires. It has been fine-tuned to minimize false alarms, ensuring that alerts are only triggered when there is a genuine threat. To further enhance the reliability of the system, a confirmation mechanism is in place, acting as a safeguard against false positives. This mechanism carefully validates fire detections, adding an extra layer of confidence to the alert generation process.

When a fire threat is confirmed, the system springs into action, generating SOS alerts with crucial details. These alerts are not just ordinary notifications; they are blockchain-based SOS alerts. The heart of the system lies in its integration with the Polygon blockchain, a robust and efficient layer-2 scaling solution for Ethereum.

Smart contracts within Polygon govern the management of SOS alerts, enforcing strict protocols for their creation, validation, and secure storage. This blockchain integration ensures data integrity, tamper-proof records, and authentication through authorized keys. As soon as a fire is detected and an SOS alert is triggered, the system initiates the creation of blockchain-based SOS alerts, signing transactions with authorized keys to guarantee their authenticity. Polygon smart contracts play a pivotal role by emitting events that signal the fire detection system, which remains vigilant and ready to respond. Rigorous testing, including simulations and real-world validation, precedes the deployment of the system in fire-prone areas. Continuous monitoring and maintenance are imperative to uphold long-term reliability and responsiveness to evolving conditions and technologies.

Moreover, the impact of this system extends beyond its technical capabilities. Public awareness campaigns are an integral part of the workflow, actively educating the community on how to utilize the system effectively during forest fire emergencies. This collective effort strengthens safety and response effectiveness, ultimately contributing to safer forested environments and rapid emergency responses when they are most needed. The holistic workflow of the system, as depicted in Fig. 1, symbolizes a transformative approach to forest fire management that prioritizes both environmental preservation and human well-being.

3.2 Tech Stack

The forest fire detection and SOS alert system leverage a sophisticated tech stack to ensure its effectiveness. CCTV cameras are strategically deployed to capture continuous footage in forested areas, serving as the system's eyes on the ground. This raw video data is then subjected to meticulous preprocessing using OpenCV, enhancing image quality, stabilizing frames, and reducing noise to optimize the data for analysis. The core of the fire detection process relies on YOLOv3, a state-of-the-art deep learning model, which works tirelessly to identify fires in the preprocessed video streams. Its advanced algorithms are fine-tuned to minimize false alarms, ensuring that alerts are only triggered when there is a genuine threat of a forest fire.

The integration of Polygon blockchain technology further enhances the system's capabilities. Polygon nodes and smart contracts play a pivotal role in building, validating, and securely storing SOS alerts. These blockchain-based alerts not only ensure data integrity but also provide a tamper-proof record of critical information, contributing to a more reliable and secure emergency response system.

In essence, this tech stack seamlessly combines video preprocessing, innovative fire detection with YOLOv3, and the robust security of Polygon blockchain integration to create a comprehensive forest fire detection and SOS alert system, safeguarding both the environment and human well-being.

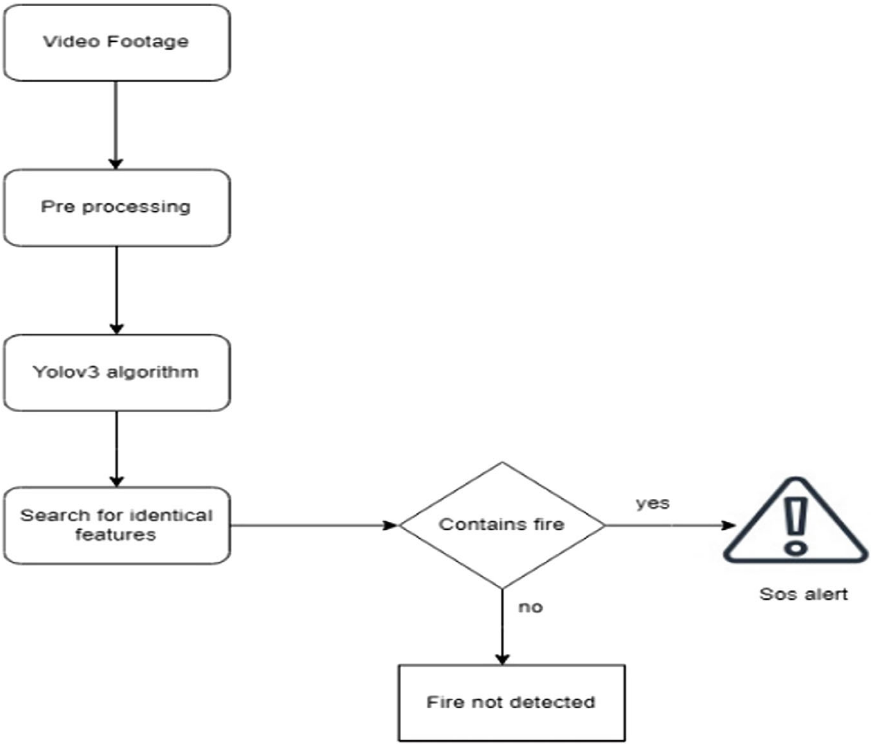


Fig. 1 Proposed workflow of forest fire detection system using SOS framework

3.3 Smart Contracts

In the proposed forest fire detection and SOS alert system, smart contracts play a pivotal role in ensuring the efficient and secure management of SOS alerts within the Polygon blockchain. These smart contracts encompass a range of essential functions, each contributing to the overall integrity and reliability of the system. The "CreateSOSAlert" smart contract is at the forefront of SOS alert generation, allowing authorized users, particularly the fire detection system, to initiate the creation of SOS alerts. By limiting this capability to designated entities, the system ensures that only valid and verified alerts can be generated, reducing the risk of false or unauthorized alerts that could lead to unnecessary emergency responses. Simultaneously, the "ValidateSOSAlert" contract takes on the crucial task of verifying the authenticity of SOS alerts. It employs various validation mechanisms, including digital signatures, to confirm the legitimacy of these alerts. This step is essential to maintain the system's trustworthiness, as it prevents fraudulent or malicious alerts from triggering unwarranted responses. Only authorized validators are granted the privilege to confirm the validity of these alerts, adding an extra layer of security to the process.

To further safeguard the integrity and accessibility of SOS alerts, a dedicated storage and data handling contract is employed. This contract oversees the decentralized storage and retrieval of SOS alerts within the Polygon blockchain. By utilizing blockchain technology for data storage, the system ensures that SOS alerts are tamper-proof and resistant to unauthorized modifications. This approach provides a transparent and verifiable record of each emergency event, which can be crucial for post-incident analysis, reporting, and coordination of emergency responses.

Lastly, an event emission and monitoring contract completes the ecosystem by emitting events upon the creation and validation of SOS alerts. These events serve as notifications that can be monitored by external systems, enabling them to respond promptly to critical events. This real-time communication ensures that relevant parties are alerted swiftly, enhancing the overall effectiveness of the forest fire detection and emergency response system.

4 Experimental Evaluations

See Table 1.

5 Challenges and Limitations

Implementing the forest fire detection and SOS alert system, as outlined in the proposed workflow, presents several significant challenges. Firstly, ensuring the clarity and dependability of the CCTV footage is crucial. Forested areas often face adverse weather conditions like rain, fog, and smoke, which can obscure visibility, making it essential to maintain high-quality and uninterrupted video feeds for accurate fire detection.

Reducing false alarms is another critical challenge. False alerts can lead to unnecessary panic and resource allocation, requiring fine-tuning of the fire detection algorithms to minimize false positives and ensure alerts only trigger during genuine fire threats.

Managing transaction costs in the blockchain is complex. Blockchain transactions often incur fees, and as the system generates a substantial volume of SOS alerts, controlling these costs while maintaining affordability is vital.

Encouraging user adoption is a significant challenge. Public awareness campaigns and education initiatives are necessary to inform communities about the system's capabilities and how to utilize it during forest fire emergencies.

Ensuring compliance with rules and regulations, including privacy and data protection laws, is of utmost importance while safeguarding sensitive data.

Lastly, ongoing monitoring and maintenance are crucial for the system's long-term effectiveness. As technology evolves and environmental conditions change, the system must adapt to remain reliable and responsive to emerging challenges.

Table 1 Observations of various models

S. No.	Paper title	Methodology	Observations
1	A Deep learning-based experiment on forest wildfire detection in machine vision course	Reduce-VGGnet and CNN for wildfire detection	Tested on FLAME dataset; high computational intensity
2	Faster and 45 smoother detection of forest fire using YOLO v3	Deep Learning with RCNN and YOLOv3 for forest fire detection	Online and offline reporting modes; efficient and cost-effective
3	Influence of forest fire particles on the breakdown characteristics of air gap	Simulation experiments on transmission line breakdown	Forest fires alter dielectric properties; needs real-world validation
4	Multi-agent dynamic task assignment based on forest fire point model	Dynamic task assignment for forest firefighting	Based on fire-spreading model; outperformed auction algorithm
5	Research on the mechanism of gap breakdown characteristics of overhead transmission line under the condition of forest fire flame	Analysis of flame temperature and particles on transmission lines	Limited consideration of vegetation-burning flames; polarity effect not clear
6	A Blockchain-based public emergency alert system	Blockchain-based open emergency alert system	EmergencyBlock implementation; performance validated through mining tests
7	Evaluation of four satellite-derived fire products in the fire-prone, cloudy, and mountainous area over subtropical China	Evaluation of satellite-derived fire products	VNP14DL performed best in cloudy and mountainous areas
8	Onboard fuzzy logic approach to active fire detection in Brazilian amazon forest	Fuzzy logic approach for active fire detection	Onboard approach for the Brazilian Amazon Forest
9	Emergency alert system for women safety using Raspberry Pi	Raspberry Pi, GSM, GPS, and ESP-32 cam for safety device	Wearable, independent safety solution for women
10	NEMO-SOS—a future trends in smart accident alert application	NEMO smartphone SOS alert app using Kotlin and Java	Utilizes sensors for motion tracking and location APIs for emergency SMS alerts

These challenges collectively underscore the complexity of implementing a forest fire detection and SOS alert system, emphasizing the need for careful planning and continuous improvement to overcome them successfully.

6 Conclusion

In summary, the proposed workflow for the forest fire detection and SOS alert system represents a powerful fusion of innovative technology and blockchain, with the goal of revolutionizing fire monitoring and emergency response in forested regions. This innovative system seamlessly combines several crucial components to enhance its over- all effectiveness. Firstly, the integration of CCTV data preprocessing ensures that the incoming video footage is optimized for analysis. This step enhances the system's ability to accurately detect and respond to potential forest fires by improving the quality of the input data. Secondly, the utilization of a precise fire detection algorithm, such as YOLOv3, enhances the system's capability to identify wildfires swiftly and accurately. This advanced deep learning approach minimizes false alarms, making the system more reliable and efficient. Including Polygon blockchain technology is a significant step toward ensuring the system's reliability and security. By leveraging blockchain, the system maintains data integrity and creates a tamper-proof environment. It achieves this by using smart contracts to govern the management of SOS alerts, which are vital for timely notification of nearby communities and relevant authorities. However, it is important to acknowledge that the successful implementation of this system comes with its set of challenges. These challenges include addressing issues related to data quality, minimizing false alarms, ensuring scalability to cover large, forested areas, and gaining user adoption and acceptance. To overcome these challenges and ensure the long-term reliability and effectiveness of the system, rigorous testing is essential. This testing should encompass various scenarios and conditions, simulating real-world forest fire situations. Additionally, continuous monitoring and maintenance are crucial to keeping the system up-to-date and responsive to evolving conditions and technologies. Furthermore, community awareness campaigns play a pivotal role in educating the public about the system's capabilities and how to utilize it during forest fire emergencies. This collective effort contributes to creating safer forested environments and enables rapid emergency responses when needed.

References

1. Wang L, Zhang H, Zhang Y, Hu K, An K (2023) A deep learning-based experiment on forest wildfire detection in machine vision course. IEEE Access 11:32671–32681. https://doi.org/10.1109/ACCESS.2023.3262701
2. Panda S, Padhi A, Rath BS, Rout B, Mishra KS (2022) Faster and smoother detection of forest fire using YOLO v3. In: 2022 International conference on machine learning, computer systems and security (MLCSS), Bhubaneswar, India, pp 100–104
3. Li P et al (2016) Influence of forest fire particles on the breakdown characteristics of air gap. IEEE Trans Dielectr Electr Insul 23(4):1974–1984. https://doi.org/10.1109/TDEI.2016.7556469

4. Chen Y, Guo Z, Qiu B, Xin Q-S, Jia WG (2022) Multi agent dynamic task assignment based on forest fire point model. IEEE Trans Autom Sci Eng 19(2):833–849. https://doi.org/10.1109/TASE.2021.3061757
5. Dündar Z, Kocatürk O, Seçinti G (2022) EmergencyBlock: a blockchain-based public emergency alert system. In: 2022 IEEE 19th annual consumer communications & networking conference (CCNC), Las Vegas, NV, USA, pp 963–964. https://doi.org/10.1109/CCNC49033.2022.9700574
6. Jiao M, Quan X, Yao J (2022) Evaluation of four satellite-derived fire products in the fire-prone, cloudy, and mountainous area over subtropical China. IEEE Geosci Remote Sens Lett 19(1–5). Art no. 6513405. https://doi.org/10.1109/LGRS.2022.3188259
7. Leal BEZ, Hirakawa AR, Pereira TD (2016) Onboard fuzzy logic approach to active fire detection in Brazilian amazon forest. IEEE Trans Aerosp Electron Syst 52(2):883–890. https://doi.org/10.1109/TAES.2015.140766
8. Vineeeth Reddy C, Lohitt Venkata Saai N, Konguvel E, Sumathi G, Sujatha R (2022) Emergency alert system for women safety using Raspberry Pi. In: 2022 2nd international conference on next generation intelligent systems (ICNGIS), Kottayam, India, pp 1–4. https://doi.org/10.1109/ICNGIS54955.2022.10079823
9. Behera SK, Navin SK, Saran VS, Vasundhara L, Saisudha G (2022) NEMO-SOS—a future trends in smart accident alert application. In: 2022 international conference on power, energy, control and transmission systems (ICPECTS), Chennai, India, pp 1–5. https://doi.org/10.1109/ICPECTS56089.2022.10047341
10. Zhang H, Dou L, Xin B, Chen J, Gan M, Ding Y (2021) Data collection task planning of a fixed-wing unmanned aerial vehicle in forest fire monitoring. IEEE Access 9:109847–109864. https://doi.org/10.1109/ACCESS.2021.3102317
11. Cao Y, Yang F, Tang Q, Lu X (2019) An attention enhanced bidirectional LSTM for early forest fire smoke recognition. IEEE Access 7:154732–154742. https://doi.org/10.1109/ACCESS.2019.2946712

A Study and Analysis of Autism Spectrum Disorder for Different Classes of Human Lives Using Multi-Machine Learning Techniques

Sravanthi Kadiri, A. Prakash, L. Arokia Jesu Prabhu, and Vijender Solanki

Abstract Autism Spectrum Disorder (ASD) is a neurodevelopment condition marked by challenges in social interaction, communication, and repetitive behaviors. Treatment typically involves tailored therapies and interventions to address individual needs. It explores the use of various machine learning techniques, such as AdaBoost, random forest, decision tree, logistic regression, support vector machine, linear discriminate analysis, and a voting classifier, to predict and analyze autism spectrum disorder across different age-groups—toddlers, children, adolescents, and adults. It aims to achieve accurate predictions and early detection of autism spectrum disorder. The results show that the highest accuracies are achieved at 96.6% accuracy and 99.5% precession for the Voting Classifier Toddler subset, 99.5% accuracy for the Voting Classifier Children subset, 99.8% precision and F_1-score for the Voting Classifier Adult subset, and 95.2% accuracy and recall for the Voting Classifier-based Adolescents subset by scaling data differently for different age-groups. The proposed framework shows promising results for early autism spectrum disorder detection compared to existing approaches.

Keywords Autism spectrum disorder · Machine learning · Classification · Feature scaling · Feature selection technique

S. Kadiri · A. Prakash · L. Arokia Jesu Prabhu (✉)
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, Telangana, India
e-mail: arokiajeruprabhu@gmail.com

S. Kadiri
e-mail: Sravanthi.kadiri@gmail.com

A. Prakash
e-mail: csehod@cmritonline.ac.in

V. Solanki
Computer Science and Engineering Department, Stanley College of Engineering and Technology for Women (Autonomous), Hyderabad, India
e-mail: spesinfo@yahoo.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_69

1 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopment disorder characterized by challenges in social interaction, communication, and repetitive behaviors [1, 2]. It encompasses a wide range of symptoms [3, 4]. A autism spectrum disorder manifests differently in each person, leading to a spectrum of abilities and challenges. Diagnosis of autism spectrum disorder is complex and often occurs around the age of two, although it can be identified later depending on its severity [5–7]. In this work, it proposes an effective framework using machine learning (ML) techniques for early autism spectrum disorder detection. Autism spectrum disorder datasets (Babies, Kids, Youths, and Adults) [13, 14] and preprocess them, including handling missing data and encoding. Four scaling methods are applied to prepare the datasets, followed by evaluation using eight classification approaches (quantile transformer, random forest, decision tree, logistic regression, support vector machine, linear discriminate analysis, and a voting classifier) [15]. Feature scaling methods' significance is explored, and four feature selection techniques are employed to identify important autism spectrum disorder risk factors across different age-groups [8–10]. The main goal is to develop an effective prediction model using ML methods to detect autism in individuals of various ages, and the proposed project aims to assist healthcare practitioners in screening autism spectrum disorder cases by considering essential features [11, 12]. However, a limitation of the research is insufficient data to create a generalized model applicable to all age-groups.

2 Literature Review

2.1 Efficient Machine Learning Models for Early Stage Detection of Autism Spectrum Disorder

Autism spectrum disorder (ASD) is a neurodevelopment disorder that severely impairs an individual's cognitive, object recognition, communication, and social abilities although early detection of ASD can assist in diagnosing and taking proper steps to mitigate its effect. This study aimed to propose a machine learning model that investigates ASD data of different age levels and to identify ASD more accurately. In this work, we gathered ASD datasets of toddlers, children, adolescents, and adults and used several feature selection techniques. Then, different classifiers were applied into these datasets, and we assessed their performance with evaluation metrics including predictive accuracy, the F_1-measure, and we found that support vector machine (SVM) performed better than other classifiers where we gained 97.82% accuracy for the RIPPER-based toddler subset; 99.61% accuracy for the correlation-based feature selection (CFS) and Boruta CFS intersect (BIC) method-based child subset; 95.87% accuracy for the Boruta-based adolescent subset; and 96.82% accuracy for the CFS-based adult subset.

2.2 A Deep Learning Approach to Predict Autism Spectrum Disorder Using Multisite Resting-State fMRI

Autism spectrum disorder (ASD) is a complex and degenerative neurodevelopment disorder. Most of the existing methods utilize functional magnetic resonance imaging (fMRI) to detect ASD with a very limited dataset which provides high accuracy but results in poor generalization. In this paper, we propose an ASD detection model using functional connectivity features of resting-state fMRI data. Simulation results indicate that the proposed model outperforms state-of-the-art methods in terms of accuracy. The mean accuracy of the proposed model was 88%, whereas the mean accuracy of the state-of-the-art methods ranged from 67 to 85%. The sensitivity, F_1-score, and area under receiver operating characteristic curve (AUC) scores of the proposed model were 90%, 87%, and 96%, respectively.

2.3 A New Machine Learning Model Based on Induction of Rules for Autism Detection

Autism spectrum disorder is a developmental disorder that describes certain challenges associated with communication, social skills, and repetitive behaviors. Typically, autism spectrum disorder is diagnosed in a clinical environment by licensed specialists using procedures other pervasive developmental disorders. One promising direction to improve the accuracy and efficiency of autism spectrum disorder detection is to build classification systems using intelligent technologies, such as machine learning. Machine learning offers advanced techniques by users and clinicians to significantly improve sensitivity, specificity, accuracy, and efficiency in diagnostic discovery. This article proposes a new machine learning method called Rules-Machine Learning. Empirical results on three datasets related to children, adolescents, and adults show that Rules-Machine Learning offers classifiers with higher predictive accuracy, sensitivity, harmonic mean, and specificity.

2.4 Machine Learning in Autistic Spectrum Disorder Behavioral Research: A Review and Ways Forward

Autistic spectrum disorder (ASD) is a mental disorder that retards the acquisition of linguistic, communication, and social skills and abilities. Despite being diagnosed with ASD, some individuals in the last few years, ASD has been investigated by social and computational intelligence scientists utilizing advanced technologies, such as machine learning to improve diagnostic timing, precision, and quality. Machine learning is a multidisciplinary research topic which are utilized in prediction to improve decision-making. Machine learning techniques, such as support vector machines, decision trees, logistic regressions, and others, have been applied to datasets related to autism to construct predictive models. This article critically analyzes these recent investigative studies on autism, not only articulating the aforementioned issues. Future studies concerning machine learning in autism research will greatly benefit from such proposals.

3 Methodology

In previous research, they evaluated the ASD features using rule-based ML (RML) techniques and discovered that RML helps classification models improve classification accuracy. Another study used the random forest (RF) and Iterative Dichotomiser 3 (ID3) algorithms to create predictive models for children, adolescents, and adults. Another study demonstrated a feature-to-class and feature-to-feature correlation value using cognitive computing and applied support vector machines (SVM), decision trees (DT), and logistic regressions (LR) as ASD diagnostic and prognosis classifiers (Fig. 1).

- Dataset: This module is responsible for loading data into the system. It involves examining the structure and content of the dataset to gain insights into its features and characteristics.

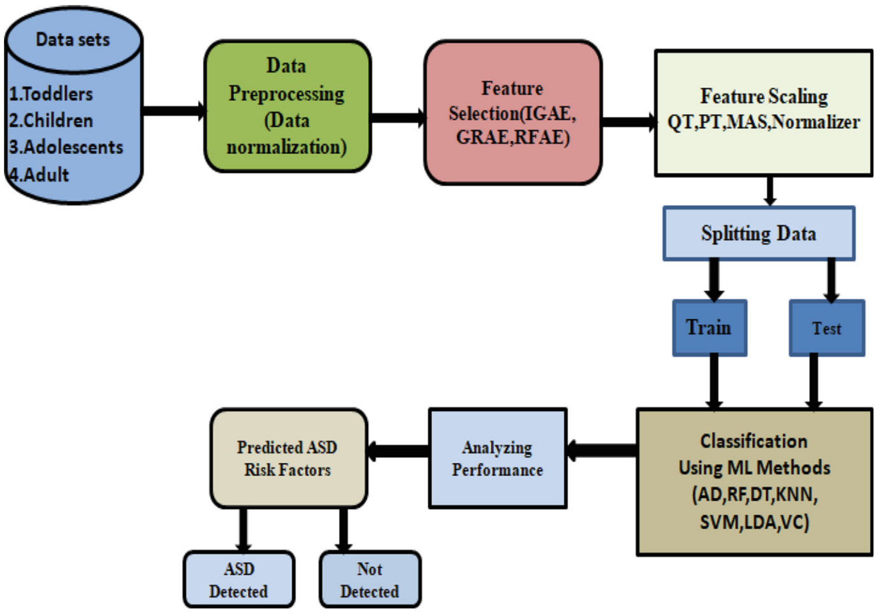


Fig. 1 System architecture

- Preprocessing: In this module, the data is read and prepared for further analysis. This may include handling missing values, encoding categorical variables, and normalization to ensure the data is suitable for modeling.
- Splitting Data into Train and Test: This module divides the dataset into training and testing sets. The training set is used to train the machine learning models, while the testing set is used to evaluate their performance.
- Model Generation: This module involves building machine learning models using different algorithms and techniques. Specifically, four different feature scaling methods are applied.

3.1 Features in a Dataset Include

Quantile Transformer: This method transforms the features to follow a uniform or a normal distribution. Therefore, for a given feature, this transformation tends to spread out the most frequent values. It also reduces the impact of (marginal) outliers:

Power Transformer: Power transforms are a family of parametric, monotonic transformations that are applied to make data more Gaussian-like. This is useful for modeling issues related to heteroscedasticity (non-constant variance), or other situations where normality is desired.

MaxAbsScalar: Maximum absolute scaling scales the data to its maximum value; that is, it divides every observation by the maximum value of the variable. The result of the preceding transformation is a distribution in which the values vary approximately within the range of -1 to 1.

Normalizer: The normalizer of S in G is the set of elements of G that satisfy the weaker condition of leaving the set. Fixed under conjugation. The centralizer and normalizer of S are subgroups of G. Many techniques in group theory are based on studying the centralizers and normalizers of suitable subsets of S.

3.2 Classification Algorithms

AdaBoost: An ensemble method used with decision trees (often decision stumps) to improve classification accuracy.

$$H(x) = \mathrm{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right). \tag{1}$$

Random Forest: A popular ensemble learning algorithm that combines the outputs of multiple decision trees to reach a single result.

Mathematically, the prediction of a random forest for classification can be represented as:

$$\mathrm{K}^{\Lambda} = \mathrm{mode}(k_1, k_2, \ldots k_n). \tag{2}$$

where k^{Λ} is the predicted class, and $k_1, k_2, \ldots k_n$ are the predicted classes by individual decision trees.

Decision Tree: A nonparametric supervised learning algorithm used for both classification and regression tasks, characterized by its hierarchical tree structure.

Given an instance x, we can predict its output y using the decision tree as follows:

$$Y = \sum_{m=1}^{M} (I(x \,\text{follows path to leaf}\, L_m) \times \text{valu of leaf}\, L_m). \tag{3}$$

where $I(x\,\text{follows the path to leaf}\, L_m)$ is an indicator function that returns 1 if the instance x follows the path to leaf L_m, and 0 otherwise.

Support vector machine (SVM): A powerful supervised algorithm suitable for classification tasks, particularly effective on smaller datasets and complex problems.

$$f(x) = w.x + b. \tag{4}$$

where f is decision function, w is the weight, vector is the bias term, and x is the input feature vector (Fig. 2).

Linear Discriminate Analysis (LDA): A supervised learning algorithm used for classification tasks, aiming to find a linear combination of features that best separates classes.

$$g_k(x) = x^T \cdot \sum_K^{-1} \cdot\mu_k - \frac{1}{2} \cdot \mu_{\mathrm{k}}^{\mathrm{T}} \cdot \sum_K^{-1} \cdot\mu_k + \mathrm{In}(\pi_k). \tag{5}$$

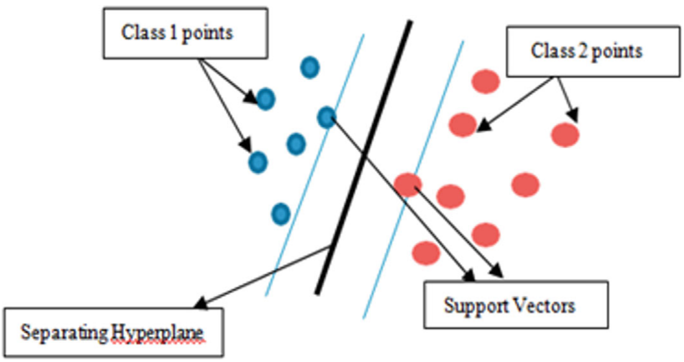


Fig. 2 SVM classifier [1]

Fig. 3 Voting classifier

Voting Classifier (AB + RF): A machine learning estimator that aggregates predictions from multiple base models, combining the decisions of each estimator (Fig. 3).

$$f(x) = \text{sign}(F\text{AdaBoost}(x)F\text{RandomForest}(x)). \tag{6}$$

4 Implementation

In this work, various algorithms and techniques are used in a machine learning project for predicting autism spectrum disorder (ASD) and aim to predict autism spectrum disorder (ASD) using the dataset "autism_screening.csv".

- Data loading and Exploration: It imports necessary libraries, such as NumPy and Pandas. It then lists all files under the input directory to verify the location of the dataset. The dataset "autism_screening.csv" is loaded into a Pandas Data Frame called DF. The first few rows of the Data Frame are displayed using the head () method. The shape and information of the Data Frame are printed to understand its size and data types. The counts of unique values in the "ethnicity" column are displayed.
- Data Preprocessing: Some preprocessing steps are performed on the Data Frame Replacing values in the "ethnicity" column to standardize them, mapping categorical values in columns like "jaundice", "autism", "used app before", and "gender" to numerical values and visualizing the distribution of "ethnicity" and "gender" using counterplots. In feature engineering, label encoding is applied to certain columns using a custom function label encoder (). Correlation between features is visualized using a heat map.
- Feature Transformation Techniques: Quantile Transformer: Transforms features to follow a uniform or normal distribution, robust to outliers. Power Transformer: Applies parametric, monotonic transformations to make data more Gaussian-like, useful for addressing issues related to heteroscedasticity. MaxAbsScalar: Scales data to its maximum value, resulting in values varying approximately within the range of -1 to 1. Normalizer: A technique that normalizes data, ensuring that every observation is fixed under conjugation.

- Model Training: The dataset is split into features (X) and targets (y). Train-test split is performed using train_test_split() from scikit-learn. Three classification models are trained: logistic regression, XGBoost (XGBClassifier), and support vector machine (SVC).

5 Experimental Results

Accuracy: Measures the ability of a model to differentiate between positive and negative cases correctly, calculated as the proportion of true positives and true negatives among all evaluated cases.

$$\begin{aligned} \text{Accuracy} = {} & \text{True Positive} + \text{True Negative}/\text{True Positive} \\ & + \text{True Negative} + \text{False Positive} + \text{False Negative}. \end{aligned} \tag{7}$$

The performance of machine learning approaches in terms of accuracy is listed in Table 1. After performing detailed analysis, it has been observed that voting classifier can be best fit for predicting autism spectrum disorder in toddlers. We have obtained 98.4% accuracy, 99.6% precision, 95.4% recall, and 98.2% F_1-score.

According to Fig. 4, the performance of all the algorithms is included in the ML model; the voting classifier (AB + RF) got effective results in accuracy, precision, recall, and F_1-score in the Toddler Dataset.

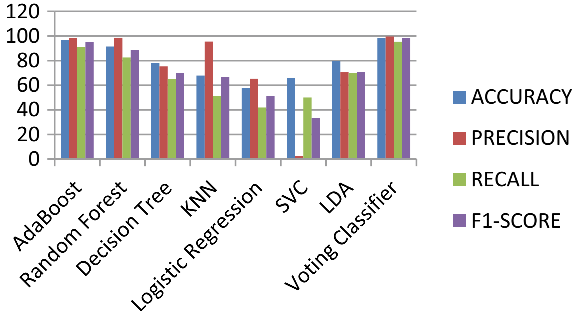
Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

$$\begin{aligned} \text{Precision} &= \text{True Positives}/(\text{True Positives} + \text{False Positives}) \\ &= \text{True Positive}/(\text{True Positive} + \text{False Positive}). \end{aligned} \tag{8}$$

Table 1 Toddler performance table

ML model	Accuracy	Precision	Recall	F_1-score
AdaBoost	96.6	98.5	90.9	95.2
Random forest	91.5	98.6	82.6	88.4
Decision tree	78.2	75.3	65.2	69.8
KNN	67.8	95.5	51.4	66.7
Logistic regression	57.6	65.3	41.9	51.3
SVC	66.1	2.6	50.0	33.3
LDA	79.7	70.5	70.0	70.7
Voting classifier	98.4	99.6	95.4	98.2

Fig. 4 Toddler performance graph



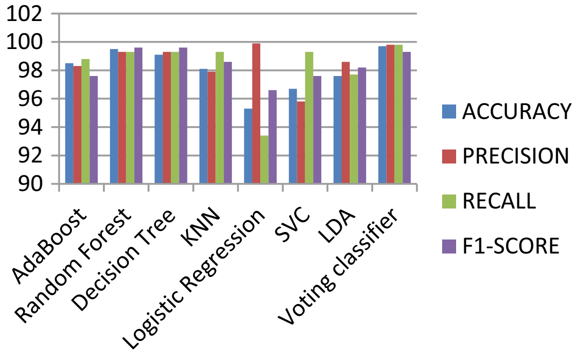
The performance of machine learning approaches in terms of accuracy is listed in Table 2. After performing detailed analysis, it has been observed that voting classifier can be best fit for predicting autism spectrum disorder in Children. We have obtained 99.7% accuracy, 99.8% pecision, 99.8% recall, and 99.3% F_1-score.

According to Fig. 5, the performance of all the algorithms is included in the ML model; the voting classifier (AB + RF) got effective results in accuracy, precision, recall, and F_1-score in the Children Dataset.

Table 2 Children performance table

ML model	Accuracy	Precision	Recall	F_1-score
AdaBoost	98.5	98.3	98.8	97.6
Random forest	99.5	99.3	99.3	99.6
Decision tree	99.1	99.3	99.3	99.6
KNN	98.1	97.9	99.3	98.6
Logistic regression	95.3	99.9	93.4	96.6
SVC	96.7	95.8	99.3	97.6
LDA	97.6	98.6	97.7	98.2
Voting classifier	99.7	99.8	99.8	99.3

Fig. 5 Children's performance graph



Recall: Measures a model's ability to identify all relevant instances of a particular class, calculated as the ratio of correctly predicted positive observations to the total actual positives.

$$\text{Recall} = \text{True Positive}/\text{True Positive} + \text{False Negative}. \quad (9)$$

The performance of machine learning approaches in terms of accuracy is listed in Table 3. After performing detailed analysis, it has been observed that voting classifier can be best fit for predicting autism spectrum disorder in Toddlers. We have obtained 99.3% accuracy, 99.8% precision, 99.7% recall, and 99.8% F_1-score.

According to Fig. 6, the performance of all the algorithms is included in the ML model; the voting classifier (AB + RF) got effective results in accuracy, precision, recall, and F_1-score in the Adults Dataset.

F_1-score: The F_1-score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

Table 3 Adults performance table

ML model	Accuracy	Precision	Recall	F_1-score
AdaBoost	98.8	98.7	99.9	99.8
Random forest	99.5	99.6	99.8	99.8
Decision tree	98.7	99.8	99.5	99.9
KNN	94.3	94.7	90.4	92.2
Logistic regression	98.3	98.6	99.5	95.4
SVC	97.3	99.3	98.5	98.3
LDA	93.6	88.7	93.6	90.7
Voting classifier	99.3	99.8	99.7	99.8

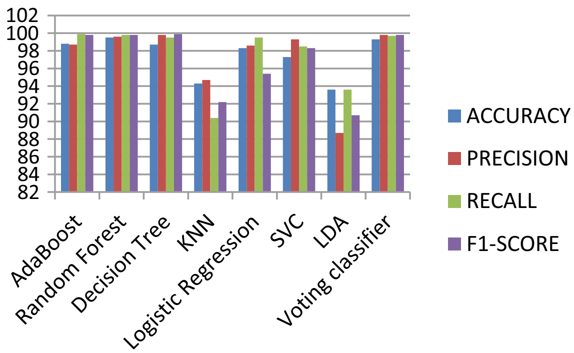


Fig. 6 Adults performance graph

Table 4 Adolescents performance table

ML model	Accuracy	Precision	Recall	F_1-score
AdaBoost	81.5	76.9	90.9	83.3
Random forest	85.7	84.6	91.7	88.7
Decision tree	81.7	76.9	90.9	83.7
KNN	85.7	99.9	81.2	89.7
Logistic regression	99.8	99.8	99.6	99.7
SVC	99.5	99.6	99.8	99.8
LDA	85.7	84.6	91.7	88
Voting classifier	99.2	99.3	99.5	99.3

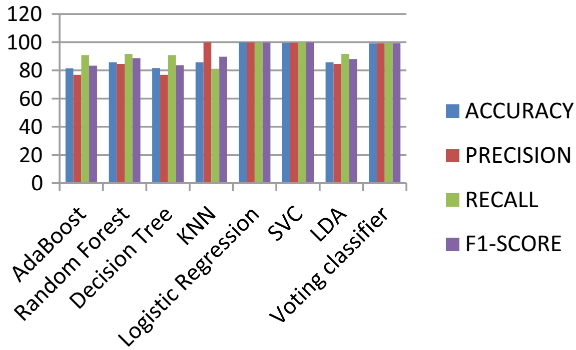


Fig. 7 Adolescents performance graph

$$F_1 \text{ - Score} = 2(\text{Precision})(\text{Recall})/\text{Precision} + \text{Recall}. \tag{10}$$

The performance of machine learning approaches in terms of accuracy is listed in Table 4. After performing detailed analysis, it has been observed that voting classifier can be best fit for predicting autism spectrum disorder in Toddlers. We have obtained 99.2% accuracy, 99.3% precision, 99.5% recall, and 99.3% F_1-score.

According to Fig. 7, the performance of all the algorithms is included in the ML model; the voting classifier (AB + RF) got effective results in accuracy, precision, recall, and F_1-score in the Adolescents Dataset.

6 Conclusion

In this work, the comprehensive machine learning framework is used for autism spectrum disorder (ASD) detection across various age-groups, through extensive experimentation and analysis, to predict models based on machine learning techniques in accurately identifying ASD cases. Preprocess ASD datasets and apply four different types of feature scaling techniques to ensure optimal data representation for classification and use eight different machine learning classifiers to classify ASD cases based on the scaled datasets. Various statistical evaluation measures are utilized to assess the classification performance of each model and calculate feature importance values using four different feature selection techniques to identify the most significant features for ASD prediction. The performance and prediction decision is assesed. It is a comprehensive pipeline for ASD prediction using machine learning techniques.

7 Future Scope

Explore deep learning models for enhanced pattern recognition. Extend the framework to incorporate diverse data sources like genetics and imaging. Develop a system for continuous ASD risk assessment and intervention. Investigate ASD progression using longitudinal datasets for stage-specific insights. Collaborate with healthcare professionals to implement the framework in real-world practice, ensuring user-friendly interfaces and interpretability.

References

1. Bala M, Ali MH, Satu MS, Hasan KF, Moni MA (2022) Efficient machine learning models for early stage detection of autism spectrum disorder. Algorithms (15)5:166
2. Petrucci D, Teofani A, Milanesi M, Fosso B, Putignani L, Messina F, Pesole F, Desideri A, Chillemi G (2022) Machine learning data analysis highlights the role of parasutterella and alloprevotella in autism spectrum disorders. Biomedicines (10)8:2028
3. Sreedasyam R, Rao A, Sachidanandan N, Sampath N, Vasudevan SK (2017) Aarya—a kinesthetic companion for children with autism spectrum disorder. J Intell Fuzzy Syst 32(4):2971–2976
4. Amudha J, Nandakumar H (2018) A fuzzy based eye gaze point estimation approach to study the task behavior in autism spectrum disorder. J Intell Fuzzy Syst 35(2):1459–1469
5. Omar KS, Mondal P, Khan NS, Rizvi MRK, Islam MN (2019) A machine learning approach to predict autism spectrum disorder. In: Proceeding international conference electrical, computer and communication engineering (ECCE), pp 1–6
6. Thabtah F (2018) Machine learning in autistic spectrum disorder behavioral research: a review and ways forward. Inform Health Soc Care 44(3):278–297
7. Tyagi B, Mishra R, Bajpai N (2019) Machine learning techniques to predict autism spectrum disorder. In: Proceeding IEEE Punecon, pp 1–5
8. Kollias K-F, Syriopoulou-Delli CK, Sarigiannidis P, Fragulis GF (2021) The contribution of machine learning and eye-tracking technology in autism spectrum disorder research: a systematic review. Electronics 10(23):2982
9. Ahmed IA, Senan EM, Rassem TH, Ali MAH, Shatnawi HSA, Alwazer SM, Alshahrani M (2022) Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques. Electronics 11(4):530
10. Sukumaran P, Govardhanan K (2021) Towards voice based prediction and analysis of emotions in ASD children. J Intell Fuzzy Syst 41(5):5317–5326
11. Hossain MD, Kabir MA, Anwar A, Islam MZ (2021) Detecting autism spectrum disorder using machine learning techniques. Health Inf Sci Syst 9(1):1–13

12. Thabtah F, Kamalov F, Rajab K (2018) A new computational intelligence approach to detect autistic features for autism screening. Int J Med Inform 117:112–124
13. Dritsas E, Trigka M (2022) Stroke risk prediction with machine learning techniques. Sensors 22(13):4670
14. Chang V, Bailey J, Xu QA, Sun Z (2022) Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. Neural Comput Appl Early Access 1–17
15. Crippa A, Salvatore C, Perego P, Forti S, Nobile M, Molteni M, Castiglioni I (2015) Use of machine learning to identify children with autism and their motor abnormalities. J Autism Develop Disord 45(7):2146–2156

Faculty and Student Perceptions of Library Resource and Service Utilization: A Clustering Algorithm Approach

P. B. Dhore, Uma Patil, and P. S. Shinde

Abstract In the digital age, academic libraries play a crucial role in providing resources and services to support the educational and research needs of faculty and students. This study aims to explore faculty and student perceptions using a clustering algorithm approach. By categorizing participants into distinct clusters based on their perceptions, the study seeks to uncover patterns in how different groups interact with library resources and services. The results of this research could provide insights into tailoring library offerings to better meet the diverse needs of faculty and students. This study employs a novel approach by integrating clustering algorithms to analyze and categorize the perceptions of these stakeholders. Through surveys administered to faculty members and students, data on resource usage, preferences, challenges, and satisfaction are collected. The study contributes to enhancing library services by tailoring them to specific user segments and addressing their distinct needs. Ultimately, the research underscores the critical role of understanding user perceptions in optimizing library resources and services for the evolving digital age.

Keywords Clusters · Surveys · Analyze · Distinct · Patterns

1 Introduction

Academic libraries have evolved beyond traditional repositories of physical books to become digital hubs offering a wide array of resources and services. The success of a library is often measured by its ability to align its offerings with the needs and preferences of its users. Faculty members and students are two primary user groups with unique requirements and expectations. This study aims to gain a deeper understanding of how these groups perceive and engage with library resources and services. In the contemporary educational landscape, academic libraries have evolved into dynamic hubs of information, catering to the diverse needs of faculty members and students.

P. B. Dhore (✉) · U. Patil · P. S. Shinde
Nutan College of Engineering and Research, Talegoan Dabhade, Pune, India
e-mail: dhoreprasad89@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_70

With the proliferation of digital resources and services, understanding how these user groups perceive and interact with library offerings has become paramount. This study delves into the nuanced realm of faculty and student perceptions of library resource and service utilization, employing an innovative approach through clustering algorithms.

2 Literature Review

The literature review will delve into studies related to library usage patterns, user perceptions, and clustering algorithms. Previous research on user preferences, satisfaction levels, and barriers to library resource utilization will be explored. Additionally, the application of clustering algorithms in educational research and user segmentation will be discussed, highlighting their potential to identify distinct groups within a larger user population.

Evolving Role of Academic Libraries Academic libraries have undergone a profound transformation in recent years, transitioning from repositories of physical books to dynamic centers of digital information [1, 2]. Scholars like Bates and Lewis emphasize the need for libraries to adapt their services to the digital age, catering to diverse user preferences and learning styles. This evolution underscores the importance of understanding how faculty and students engage with the new library landscape [3, 4].

User Perceptions and Engagement Patterns The role of user perceptions in shaping library engagement has been widely explored. Studies by Rundle and Johnson have highlighted that user satisfaction and perceived utility strongly influence library resource utilization. Additionally, Wang and Wu (2015) found that positive perceptions of library services are linked to increased academic success [5]. These findings underscore the significance of studying user perceptions to enhance service offerings [6].

Segmentation and User Profiles Clustering algorithms have gained prominence in segmenting users based on behavior and preferences. In the context of libraries, Segal and Demczynski (2018) applied clustering algorithms to identify user groups with distinct borrowing patterns. Similar applications have been observed in e-learning environments (Liu et al. 2017), illustrating the potential of clustering techniques in educational contexts [6].

Application of Clustering Algorithms in Education Clustering algorithms, such as K-means and hierarchical clustering, have been widely used in education research. Kim and Williams (2019) utilized clustering algorithms to categorize students based on their learning preferences and performance. Such approaches enable educators to tailor interventions to specific student groups, enhancing overall learning outcomes [7, 8].

Personalization of Library Services Personalization is a growing trend in various industries, including education and information services. Chen and Huang (2016) emphasize the value of personalized library services in improving user engagement and satisfaction. The application of clustering algorithms aligns with this trend, allowing libraries to create targeted interventions for different user segments [9, 10].

Gaps in the Current Literature While user perceptions and clustering algorithms have been studied in separate contexts, there is a gap in the literature regarding their combined application in the domain of academic libraries. This research seeks to bridge this gap by investigating how clustering algorithms can uncover distinct faculty and student perceptions, informing the customization of library resources and services.

3 Methodology

In clustering algorithm, the collected data will be subjected to a clustering algorithm, possibly K-means or hierarchical clustering, to group participants based on similarities in their responses. The algorithm will uncover clusters of individuals who share similar perceptions and behaviors. Using a clustering algorithm means, we going to give the algorithm a lot of input data with no labels and let it find any groupings in the data it can. Those groupings are called clusters. A cluster is a group of data points that are similar to each other based on their relation to surrounding data points. Clustering is used for things like feature engineering or pattern discovery. The execution of clustering algorithms can be computationally intensive, especially for large datasets. Therefore, considerations should be given to the computational resources available and the scalability of the chosen algorithm. Additionally, careful interpretation and validation of the results are necessary to ensure that the clustering captures meaningful patterns and insights from the data (Fig. 1).

Clustering algorithms are a type of machine learning algorithm that aims to group similar data points together based on certain criteria or features [11, 12].

The execution of a clustering algorithm involves several steps:

3.1 Data Preparation

When applying clustering algorithms to analyze library perception data, data preparation is a crucial step to ensure accurate and meaningful results. Here are some considerations for data preparation in the context of library perception analysis. When I define the variables that capture the library perception aspects you want to analyze. These variables could include factors, such as library facilities, staff expertise, resources, accessibility, user satisfaction, and any other relevant aspects of library perception. Data collection is collecting data from library users, such as

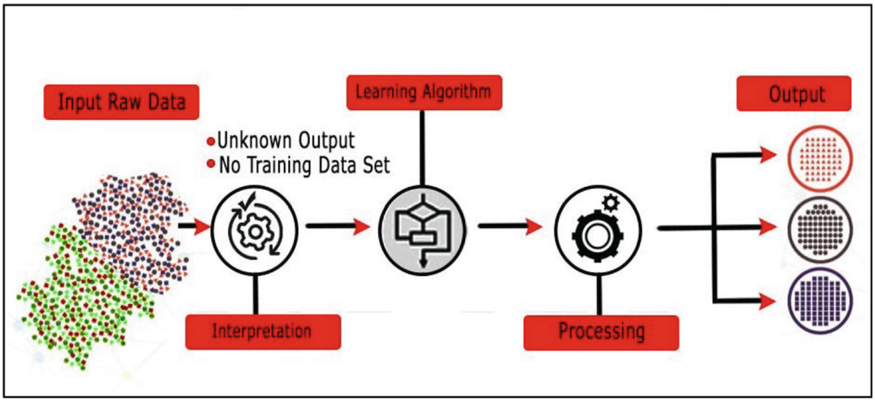


Fig. 1 Clustering execution process

through surveys, feedback forms, or interviews. The data can be quantitative (rating scales, Likert scales) or qualitative (open-ended responses), depending on the nature of the perception data you want to capture.

Data visualization is, before applying clustering algorithms, visualizing the data to gain insights and detect any patterns or relationships. Visualization techniques like scatter plots, bar charts, or heatmaps can help identify potential clusters or groupings in the data.

3.2 Algorithm Selection

When selecting a clustering algorithm for library perception analysis, it is important to consider the nature of the data, the objectives of the analysis, and the specific requirements of the library perception study. Clustering algorithm depends on various factors, such as the nature of the library perception data, the desired interpretation of the results, the number and shape of expected clusters, and the algorithm's assumptions and limitations. It is often beneficial to experiment with different algorithms and assess their performance using evaluation metrics or visual inspection to select the most suitable algorithm for the library perception analysis.

Gaussian Mixture Models (GMM) assumes that the data points are generated from a mixture of Gaussian distributions. It models the data using a weighted combination of Gaussian components. GMM can identify clusters with different shapes and sizes and provides probabilistic cluster assignments. It is suitable for datasets with overlapping clusters. OPTICS follow Ordering Points to Identify the Clustering Structure (OPTICS) is another density-based clustering algorithm. It extends DBSCAN by providing a density-based ordering of points. OPTICS allows for the detection of clusters with varying densities and can handle datasets with noise and outliers. It also provides a visual representation of the cluster structure.

3.3 Parameter Selection

Parameter selection is an important aspect of using clustering algorithms for library perception analysis. The specific parameters to consider depend on the chosen clustering algorithm. Here are some common parameters to consider and approaches to parameter selection.

In number of clusters (k) for algorithms that require specifying the number of clusters, such as K-means or GMM, selecting an appropriate value for k is crucial. There are several methods to determine the optimal number of clusters, including:

3.4 Algorithm Execution

Algorithm execution refers to the process of implementing and running a specific algorithm to solve a given question which assigns to students and teaching faculty. First, we understand the problem gain: a clear understanding of the problem you want to solve and the specific goals or objectives you aim to achieve. Then choose algorithm selection: an appropriate algorithm that is suitable for addressing the problem at hand. Consider factors, such as the problem domain, available data, computational requirements, and the desired output.

In algorithm implementation, it implement the chosen algorithm in a programming language of your choice. This involves translating the algorithm's steps and logic into code that can be executed by a computer. Then prepare the input data for the algorithm. This may involve data cleaning, preprocessing, transformation, or feature engineering, depending on the requirements of the algorithm and the nature of the data. Configure the parameter configuration. If the algorithm has configurable parameters, set the appropriate parameter values based on the problem and data characteristics. These parameter values can significantly impact the algorithm's behavior and performance.

3.5 Evaluation of Algorithm

Clustering algorithm evaluation is an essential step to assess the quality and effectiveness of the clustering results. It helps determine the reliability of the clusters, the coherence of the grouping, and the overall performance of the algorithm (Fig. 2).

Here are some common evaluation methods for clustering algorithms:

Internal Evaluation Metrics

- Silhouette Score: Measures the compactness and separation of clusters. It assigns a score to each data point, indicating how well it fits within its assigned cluster

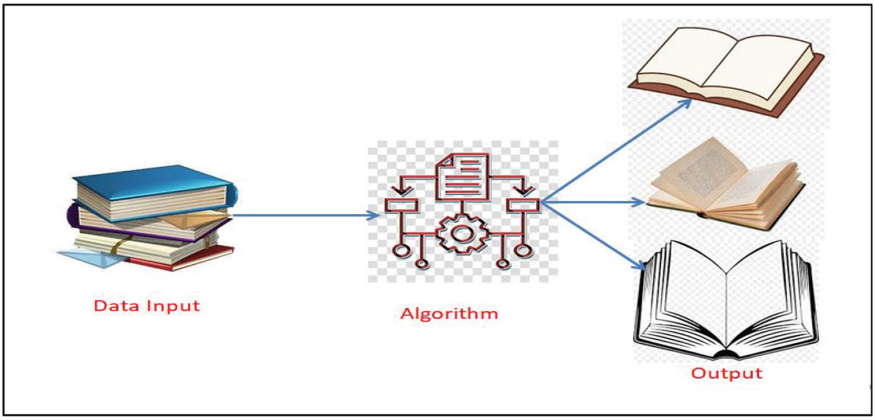


Fig. 2 Clustering process

compared to other clusters. Higher silhouette scores indicate better-defined clusters.
- Davies–Bouldin Index: Quantifies the average similarity between clusters while considering their separation. Lower values indicate better clustering results.
- Calinski–Harabasz Index: Computes the ratio of between-cluster dispersion to within-cluster dispersion. Higher values suggest better-defined clusters.

External Evaluation Metrics

- Rand Index: Compares the clustering results against a reference clustering or ground truth labels. It measures the similarity between the predicted clusters and the true clusters. A higher Rand index indicates better agreement between the predicted and reference clusters.
- Adjusted Rand Index: Similar to the Rand index but accounts for the chance agreement between clusters. It adjusts for the expected random agreement and provides a more accurate measure of clustering performance.

Visual Inspection

- Scatter Plots: Visualize the data points in a scatter plot, with each point represented by its features or attributes. Assess the separation, compactness, and cohesion of the clusters. Visual inspection can provide intuitive insights into the clustering quality.
- Heatmaps: Use heatmaps to visualize the similarity or dissimilarity matrix of the data points. Clusters should exhibit high similarity within the clusters and low similarity between the clusters.

Stability Evaluation

- Cluster Stability: Assess the stability of the clustering results by performing multiple runs of the algorithm with slightly perturbed data or random initializations. Compare the resulting clusters and measure their similarity or overlap. Higher stability indicates more reliable clustering results.

Domain-Specific Evaluation

- Domain Expert Validation: Involve domain experts or stakeholders who possess expertise in the problem domain to evaluate the clustering results. Their knowledge and judgment can provide insights into the meaningfulness and interpretability of the clusters.

3.6 Iteration and Refinement

Iteration and refinement are important steps in the clustering process to improve the quality, accuracy, and meaningfulness of the clustering results. In initial clustering, it start by applying the chosen clustering algorithm to the data and obtain the initial clustering results. This serves as the baseline for further refinement. In evaluation and analysis, it evaluate the initial clustering results using appropriate evaluation metrics and visualization techniques. Assess the quality of the clusters, such as their cohesion, separation, and interpretability. Analyze the clusters' characteristics and patterns to gain insights into the underlying data structure. In feedback and domain expertise, it seeks feedback and insights from domain experts or stakeholders who have knowledge of the problem domain and the data.

Their input can help identify any issues or areas that need improvement in the initial clustering results. In iterative clustering, it applies the adjusted parameters and refined data preprocessing steps to perform the clustering algorithm again. Obtain updated clustering results and compare them to the initial results. In evaluation and comparison, it evaluates the refined clustering results using the same evaluation metrics and visualization techniques as before. Compares the refined results with the initial results to assess the improvements achieved through the iterative process. If the clustering results still need improvement, iterate by going back to step and repeating the process. Iterate until the clustering results meet the desired quality, validity, and interpretability criteria or until convergence is achieved.

4 Results and Discussion

The results section will present the identified clusters, describing the distinct characteristics and perceptions of each group. This might include clusters of highly engaged users, occasional users, users with specific preferences, etc. Visualizations like dendrogram plots or cluster profiles will aid in illustrating the findings.

5 Experimental Result

The experimental results of students and teaching faculty perception in a library can vary depending on the specific experiment conducted and the factors being tested. Here are some potential outcomes that may arise from experimental studies on perception in a library:

Impact of library layout in experimental results may indicate that a well-designed and organized library layout positively influences the perception of students and teaching faculty. Participants may express higher levels of satisfaction, report easier navigation, and perceive the library as a more welcoming and conducive learning environment. The experiment could reveal that maintaining appropriate noise levels in designated library areas has a significant impact on users' perception. Students and teaching faculty may show preference for quiet areas for focused study and express higher satisfaction when noise levels are controlled.

The experimental findings might suggest that adequate lighting positively affects users' perception of the library. Students and teaching faculty may report higher comfort levels, increased concentration, and improved overall satisfaction when lighting conditions are optimized. The experimental results may indicate that the provision of well-designed collaborative spaces positively influences the perception of students and teaching faculty. Participants may express higher levels of engagement, increased opportunities for collaboration, and greater satisfaction with group work areas.

The experiment could reveal that integrating technology-based services in the library positively impacts users' perception. Students and teaching faculty may appreciate the convenience of self-checkout kiosks, mobile apps, or online resources, leading to higher levels of satisfaction and improved perception of library services.

The experimental outcomes might suggest that information literacy programs or workshops have a positive impact on users' perception. Participants may demonstrate improved research skills, increased confidence in finding and evaluating information, and expressed higher levels of satisfaction with library resources and services.

The experimental results may indicate that diversifying the library's collection by introducing new types of resources positively affects users' perception. Students and teaching faculty may show increased usage of resources, express satisfaction with resource availability, and perceive the library as more comprehensive and aligned with their needs. A library is a collection of items, such as books or television programs, that may be used rather than just seen [5, 3].

5.1 *Activity of Library Resources*

Users' frequent and infrequent activities reveal that the most common tasks they perform in the library include reading individual readers (83.9%), newspapers (67.1%), conducting research (65.8%), and completing class assignments (65.3%).

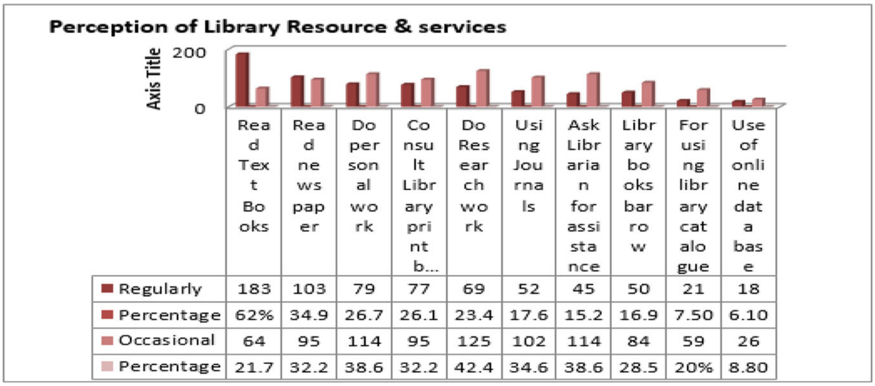


Fig. 3 Activity of library resource and services

More than half of the respondents (53.8%) had referred library pattern books, used print and online periodicals, and gotten advice from librarians. 45.4 percent of all people have checked out books from the library. Online databases (14.9%) and library catalogues (27.5%) are not frequently used as shown in Fig. 3.

5.2 Major Issues of Library Resource and Services

The lack of computers and/or Internet (75.6%), a lack of present resources (60.7%), a warm situation (46.1%), a lack of reading space and chairs (38.3%), and poor lighting (36.6%) are the main issues that students have with using the College Library effectively. The least significant deterrent to using the library was noise (18.3%) as shown in Fig. 4.

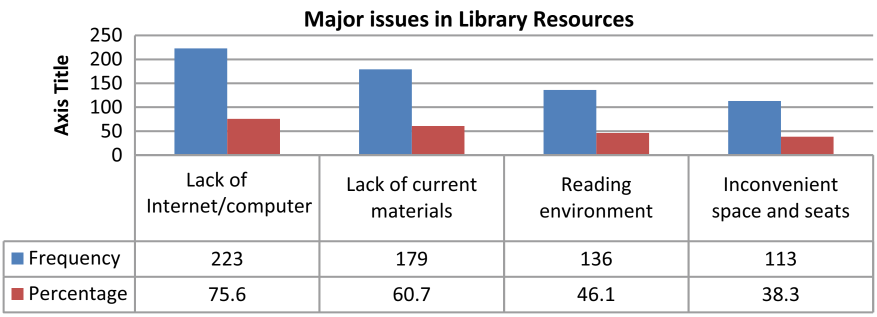
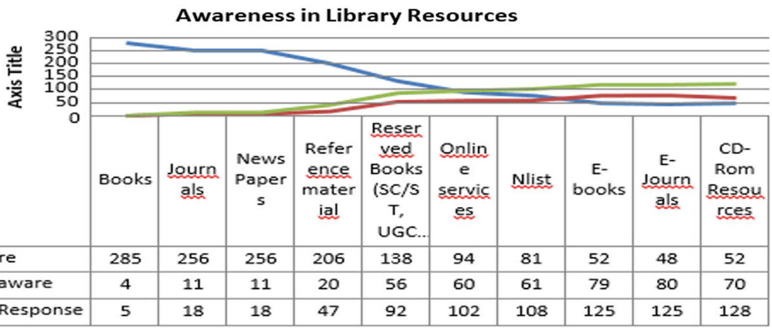


Fig. 4 Major issues in library resources



	Books	Journals	News Papers	Reference material	Reserved Books (SC/ST, UGC..	Online services	Nlist	E-books	E-Journals	CD-Rom Resources
Aware	285	256	256	206	138	94	81	52	48	52
Not aware	4	11	11	20	56	60	61	79	80	70
Not Response	5	18	18	47	92	102	108	125	125	128

Fig. 5 Awareness in library resources

5.3 Faculty Awareness of Library Resources and Services

Awareness of library resources refers to the understanding and knowledge that library users have regarding the available resources and services offered by the library. The following table shows that people are quite aware of print resources. 96.6 percent of students are aware that the library has books available. Journals, newspapers, and reference materials are all known to a large percentage of people (86.8%, 86.8%, and 69.8%, respectively). However, there is little knowledge of the library's e-Resources' existence. Over a third of respondents (36.6–44.4%) do not indicate that the public library needs all of the mentioned e-Resources, and over 20% of respondents are unaware that they are available there. This translates to an average of 60% of respondents not knowing about or responding that the library has e-Resources as shown in Fig. 5.

5.4 Students Awareness of Library Resources and Services

See Fig. 6.

5.5 Average Time Spend in Library

See Fig. 7.

5.6 Information Sources in Library

See Fig. 8.

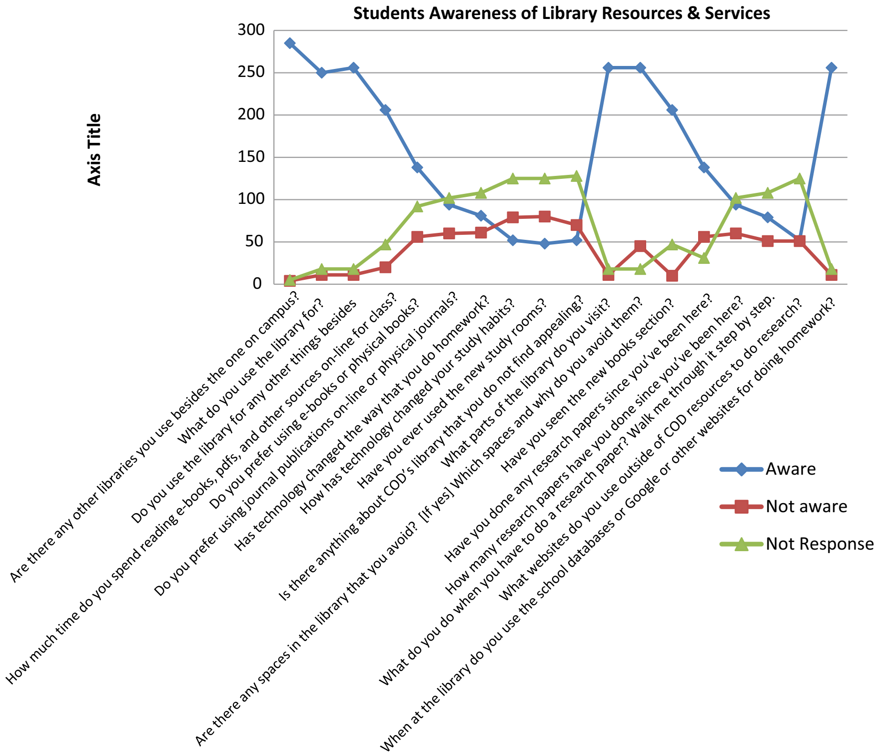
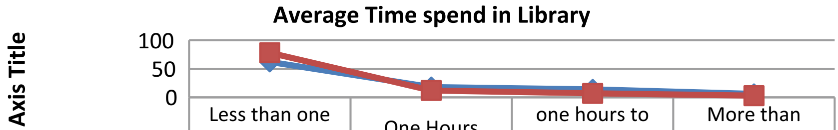
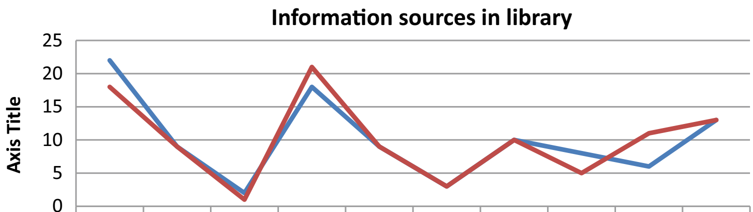


Fig. 6 Students awareness of library resources and services



	Less than one Hour	One Hours	one hours to two hours	More than Three Hours
Students	62	18	14	6
Teaching Faculty	78	12	7	3

Fig. 7 Average time spend in library



	Books	Journals and Periodicals	Databases	E-books and E-journals	Reference Materials	Audiovisual Materials	Special Collections	Government Publications	Theses and Dissertations	Online Resources
Students	22	9	2	18	9	3	10	8	6	13
Teaching Faculty	18	9	1	21	9	3	10	5	11	13

Fig. 8 Information sources in library

6 Conclusion

The conclusion will summarize the study's findings and emphasize the significance of understanding user perceptions in optimizing library resources and services. This study delved into the perceptions of faculty members and students regarding library resource and service utilization, employing a pioneering approach through clustering algorithms. In conclusion, this research sheds light on the dynamic relationship between user perceptions and library resource utilization. By employing clustering algorithms, libraries can tap into a new dimension of user understanding, redefining engagement strategies and fostering an environment where resources and services align seamlessly with user expectations.

References

1. Shieber SM (1996) Predicting individual book use for off-site storage using decision trees. Libr Quart 66(3):266–293
2. Shi Y, Zhu Y (2015) Book loan quantity prediction using time series data mining. In: Proceeding international conference electromechanical control technology and transportation
3. Evans GE (1970) Book selection and book collection usage in academic libraries. Libr Quart 40(3):297–308
4. Mohamed A-R, Dahl GE, Hinton G (2012) Acoustic modeling using deep belief networks. IEEE Trans Audio Speech Lang Process 20(1):14–22
5. Baba K, Minami T, Nakatoh T (2016) Predicting book use in university libraries by synchronous obsolescence. Proc Comput Sci 96:395–402
6. Zhou L, Pan S, Wang J, Vasilakos AV (2017) Machine learning on big data: opportunities and challenges. Neurocomput 237:350–361
7. Gao T, Li X, Chai Y, Tang Y (2016) Deep learning with consumer preferences for recommender system. In: Proceeding IEEE international conference information and automation (ICIA), Aug 2016, pp 1556–1561

8. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
9. Pritchard SJ (1980) Purchase and use of monographs originally requested on interlibrary loan in a medical school library. Libr Acquisitions: Pract Theor 4(2):135–139
10. Kambatla K, Kollias G, Kumar V, Grama A (2014) Trends in big data analytics. J Parallel Distrib Comput 74(7):2561–2573
11. Maharjan S, Arevalo J, Montes M, González FA, Solorio T (2017) A multi-task approach to predict likability of books. In: Proceeding 15th conference European chapter association computational linguistics, Long papers, vol 1, Apr 2017, pp 1217–1227
12. Ntakaris A, Magris M, Kanniainen J, Gabbouj M, Iosifidis A (2018) Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods. J Forecasting 37(8):852–866

Performance Improvement of Software Through Genetic Algorithm by Using Test Case Generation

Thonta Ashwini and Katta Subba Rao

Abstract The phase of software development that is most time-consuming in software testing. It is desired to minimize the amount of effort expended while maximizing the number of faults detected. Therefore, the generation of test cases may be regarded as an optimization problem. One of the primary challenges in software testing is the automatic generation of test data that meets a specified adequacy criterion. The automatic generation of test cases will result in a significant reduction in costs and effort.

Keywords Software testing · Test case generation · Automatic test case generation · Genetic algorithm · Boundary value analysis

1 Introduction

Software program checking out is a critical technique in which both new and current test cases are applied to software to affirm its functionality throughout improvement and before release. The length of the checking out phase is a key element, relying at the scope and complexity of the software program. It has been located that checking out money owed for greater than 50% of software program development time because it covers a range of sports. These activities include producing check instances, developing check suites, executing take a look at suites, analysing execution outcomes, and developing complete checking out hobby reviews.

Software testing is a crucial factor of software program exceptional assurance and serves as the ultimate assessment of specification, design, and coding. Its number one purpose is to come across defects and estimate reliability. The objective of software

T. Ashwini · K. S. Rao (✉)
Department of Computer Science Engineering, B V Raju Institute of Technology Vishnupur, Narsapur, Telangana, India
e-mail: subbarao.k@bvrit.ac.in

T. Ashwini
e-mail: 21211D5803@bvrit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_71

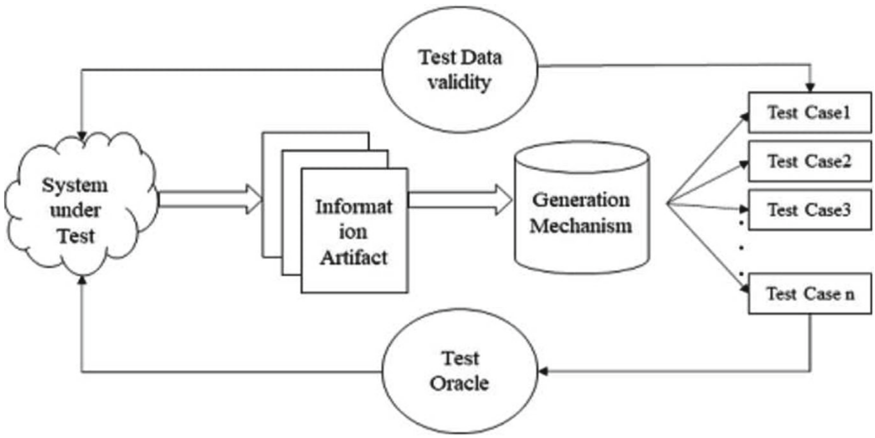


Fig. 1 Basic test case generation process

program checking out is to expand a minimal variety of test cases and check facts that can disclose as many faults as possible. Take a look at statistics era includes figuring out a set of software input information that meets a given testing criterion. Each software program undergoes testing in numerous, strategically planned environments. Testing is a systematic process aimed at identifying errors within a software system. The primary objective of testing is to minimize the disparity between the actual output and the anticipated output. While testing endeavours to produce accurate results based on specific functions, it is not always possible to identify all defects. Instead, testing provides a comparative analysis of the product's principles or mechanisms, which enables users to recognize potential issues. Typically, a test case comprises data that serves as input for software testing.

The primary objective of test case generation is to produce test suites that may correctly become aware of software program bugs. These test cases and suites are designed to affirm that the software program gadget meets the specified standards, reliability, and patron specs. The take a look at case generation manner comprises four key activities, specifically statistics artifacts, era mechanism, test case validity, and formation of test oracle. Those sports are fundamental to the essential check case era procedure, as illustrated in Fig. 1.

2 Literature Survey

"Automatic test data generation by applying the genetic algorithm". In this paper, they defined test data, input data as the population in genetic algorithm. Each chromosome or bit string acts like a test data in initial population. The complete set of chromosomes or the bit strings is being used to generate test cases for all possible paths [1]. The whole process of generating the test data and test cases by applying genetic algorithms is coded and programmed using numerical computing environment MATLAB. Then, the primary population is randomly generated, after that the individual chromosomes based on their fitness value are evaluated. After those genetic algorithm, operations such as evaluation, mutation, crossover, and selection are applied to bring out the next level of generation. This process is continued until the genetic algorithm finds best possible results [2].

"The application of the genetic algorithm in the field of software testing". The authors have utilized this algorithm to explore a vast search space and identify the most optimal solution from a range of frequently occurring solutions [3]. The use of the genetic algorithm is the study's main goal to minimize the number of test cases, reduce costs, save time, and minimize the effort required for the development of high-quality software. The paper highlights that the genetic algorithm can adapt the software to any situation and identify the best possible solution [4]. In the realm of software testing, the genetic algorithm has made significant strides over an extended period of experimentation, resulting in the identification of ideal solutions. This research paper has successfully resolved a critical numerical problem through the application of the genetic algorithm in software testing [5].

"Automatic test data and test case generation using evolutionary genetic algorithm". The reduction of equivalent inputs or input data occurs as a result of the production of ideal test data and test cases using a genetic algorithm [6]. This process takes into account various factors, such as the influence of branch conditions, coverage concepts, and the level of competition between individuals. In order to evaluate the qualities and merits of individual species, a fitness function is designed while considering these factors [7].

"Genetic algorithm technique to find the most critical paths for improving software testing efficiency". The authors present a method for optimizing software testing efficiency by identifying the most critical path clusters in a program [8]. The improved adaptive genetic algorithm used in this paper presents a unique method for the automated production of test cases while maintaining population diversity. The method enhances search efficiency by adjusting crossover rates and mutation rates to maintain population diversity, which is determined by the difference between fitness values and individual similarity [9].

To create test cases for software systems, many strategies are used. In order to choose a collection of test cases, random approaches rely on assumptions regarding fault distribution. On the other side, goal-oriented approaches look for test cases that cover a certain objective, such as a statement or branch, independent of the direction taken [10]. Conversely, test cases for specification-based methodologies are created using explicit requirement descriptions. In complicated software systems, creating test cases from non-formal requirements might lead to inaccurate implementations, and hence, it is crucial to test for conformity to the specifications [8]. A genetic algorithm has been developed with the capability to "automatically generate test cases for the purpose of testing a specific path". This algorithm is designed to pursue a predetermined path as its target and employs a series of iterative operator engagements to develop test cases. The resulting test cases are able to direct program execution towards the intended target path [11].

3 Problem Statement

To reduce the cost and effort of software population by using genetic algorithm with test cases.

4 Objectives

- In order to enhance the current test data and test cases utilized in software testing, it is recommended to employ appropriate genetic algorithm operators to either improve upon the existing ones or generate new ones.
- The examination of genetic algorithm for enhancing and creating test data and test cases. The current test data and test cases utilized for automated software testing.

5. Existing Work

Effective testing is a process that must be carried out. Due to the availability of sources, exhaustive testing is not feasible. It has been noted in the past that test cases belong to various classes. Additionally, in particular categories of test cases, programmers make greater errors. Boundary values are the most crucial class of these since they present the greatest risk of program failure. Therefore, as they are more important for software, these test case classes must be the focus of autonomous test case generation. The input bounds of the program's variables are referred to as the program's boundary values (Fig. 2).

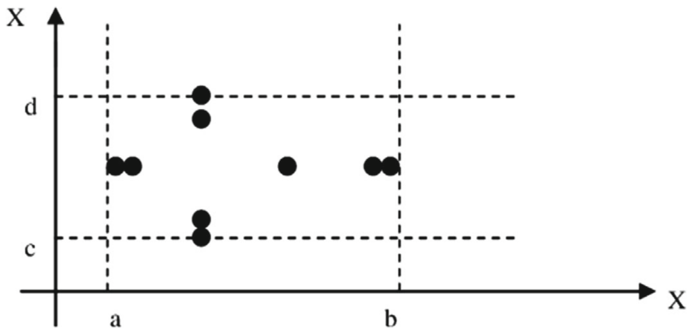


Fig. 2 Examples of boundary value evaluation via functions

5 Proposed Work

For the development of test cases to analyse boundary values, a proposed genetic algorithm is described. The study used a genetic algorithm and random testing to automatically identify these limits, and then, they compared the outcomes of both methods. Both genetic algorithms and random testing begin with a random beginning population, but GA advances towards the optimums using an individual's fitness; in contrast, random testing makes decisions at random throughout the run. The fitness of each chromosome in this experiment will be determined by the distance between the borders.

6 Methods

To create test cases, a technology called the genetic algorithm is used.

Representation The testing process necessitates efficient execution, yet it is impractical to carry out comprehensive testing owing to resource constraints. Empirical evidence has demonstrated that test cases can be categorized into discrete groups and that developers are prone to committing more errors in certain categories of test cases.

Initial Population As previously mentioned in the representation section, the first population is generated at random. To be more precise, a total of pop size vectors is produced at random, each with a length of c_size. In this case, the population size is denoted by pop size, and the number of variables is indicated by c_size.

Fitness Function The fitness of every chromosome is decided by way of its deviation from the variable barriers. The closer a variable is to the bounds, the better its degree of health is asserted.

Selection Upon completion of the fitness computation for each test case within the present population, the algorithm proceeds to designate certain members of said population as effective parents for the forthcoming generation.

Crossover At the individual level, the operation is finished. In the crossover system, two parental chromosomes exchange genetic material or sub-string statistics at a randomly chosen location on the chromosome, leading to the creation of two new offspring strings (Fig. 3).

7 Result

All inputs are sourced from the user, thereby facilitating ease of testing with varying parameters. The user interface during execution in MATLAB is presented as follows:

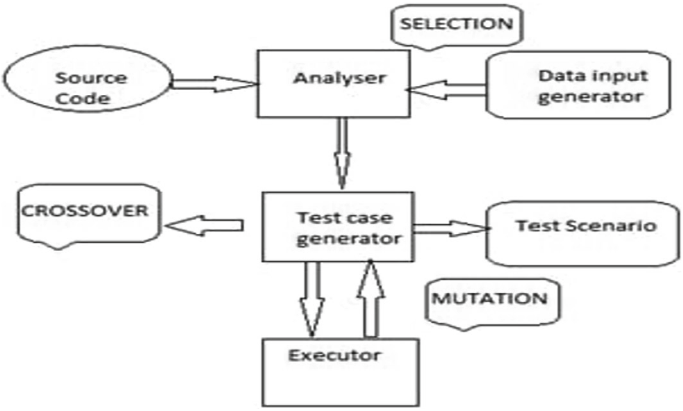


Fig. 3 Using a genetic algorithm, create test cases for software testing

INPUT: (Table 1)

Number of individuals in population: 20
Number of variables: 2
Number of generations: 200
Limits of first variable: Lower limit: 6 and Upper limit: 16
Limits of second variable: Lower limit: 10 and Upper limit: 20 (Figs. 4, 5 and 6).

OUTPUT:

Table 1 Genetic algorithm against random testing

	Genetic algorithm		Random testing	
S. No.	Variable 1	Variable 2	Variable1	Variable 2
1	14.15	12.6	14.14	14.83
2	7.78	10.85	10.15	12.14
3	15.01	15.14	10.34	15.24
4	13.13	11.98	11.77	16.86
5	7.50	10.01	9.11	12.62

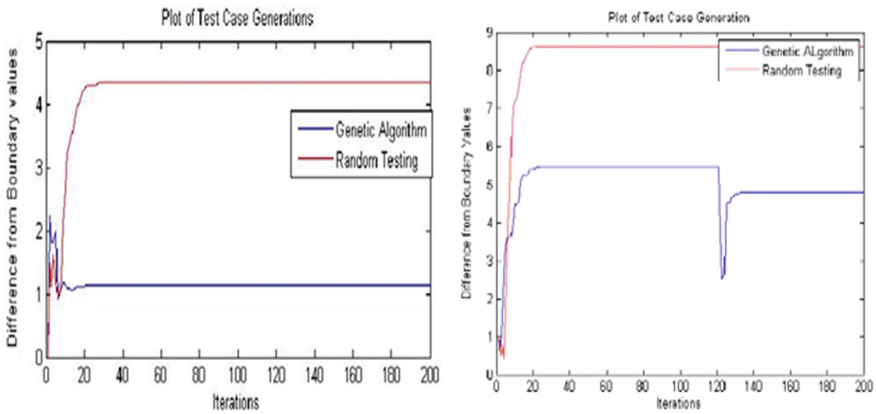


Fig. 4 Difference between first access boundary and second access boundary values from Table 4.1 for 200 generations

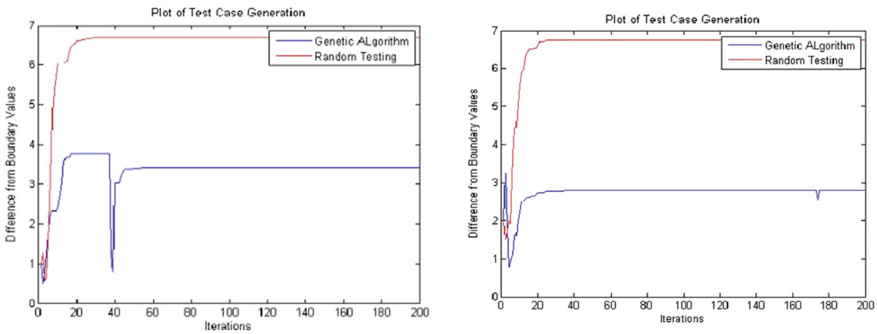


Fig. 5 Difference between third access boundary and Fourth access values from Table 4.1 for 200 generations

8 Discussion

To be able to beautify the efficiency and pleasant of the trying out method, it is far imperative to utilize an automated take a look at case generator. This method has been validated to noticeably enhance the overall pleasant of trying out while compared to random testing. The effects exhibit a promising technique for automated take a look at case generation utilizing the boundary fee analysis idea. Shifting ahead, the test case generator may be in addition advanced with the aid of incorporating different types and parameters of crossover (inclusive of mixture crossover, single point, uniform, and so on.), mutation (together with basic bit, uniform, and so forth.), and choice (such as rank and so on.). While boundary price evaluation is currently applied for automatic check case generation, other extra fields consisting of manipulate glide graphs, route testing, regression testing, etc., also can be hired for test case technology.

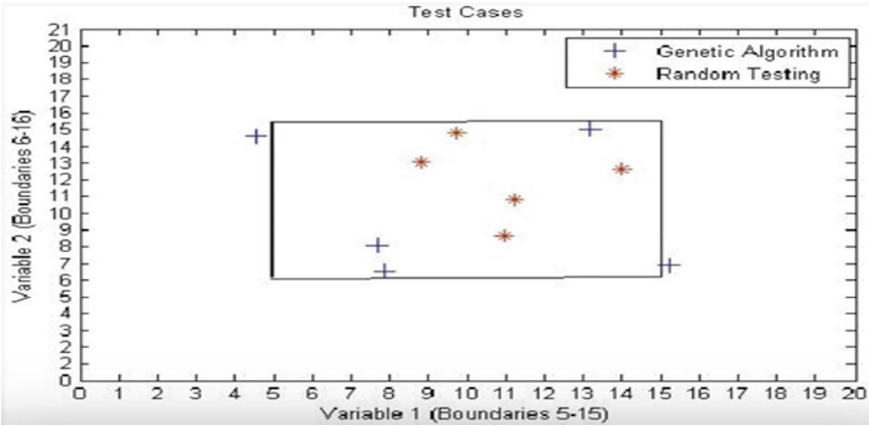


Fig. 6 Results displaying analysis of the most recent test instances

9 Conclusion

This essay provides analysis of the utility of evolutionary techniques, mainly genetic algorithms, inside the context of software checking out. The study demonstrates the efficiency of genetic algorithms in facilitating automatic check case technology for software program checking out. The use of genetic algorithms in software program trying out has been found to decorate the performance and method time of the testing system, even as also presenting a method of generating test instances routinely.

References

1. Verma AS, Choudhary A, Tiwari S (2023) Software test case generation tools and techniques. Int J Math Eng Manage Sci 8(2):293–315
2. Chakraborty S, Gujar VB, Choudhury T, Dewangan BK (2022) Improving software performance by automatic test cases through genetic algorithm. Int J Comput Appl Technol 68(3):10
3. Ajayi W, Bukola O, Ahmed J (2022) Improving test case automation using genetic algorithm. Int J Adv Res Comput Sci 13(2)
4. Esnaashari M, Damia AH (2021) Automation of software test data generation using genetic algorithm and reinforcement learning. Expert Syst Appl 183:30
5. Swathi B, Dr. Tiwari H (2020) Genetic algorithm approach to optimize test cases. Int J Comput Trends Technol 68(10)
6. Md Hussain K, Dr. Nandan Prasad K (2019) different techniques used in automatic test case generation. In: International conference ICGTES-19, Oct 2019
7. Chauhan A, Singhal RS, Yadav PK (2018) Test case generation techniques. Int J Comput Trends Technol 57(2)
8. Alrawashed TA, Almomani A, Althunibat A, Tamimi A (2019) An automated approach to generate test cases from use case description model. CMES—Comput Model Eng Sci 119(3):409–425

9. Abdallah SA, Moawad R, Fawzy EE (2018) An optimization approach for automated unit test generation tools using multi-objective evolutionary algorithms. Future Comput Inf J 3(2):178–190
10. Khari M (2020) Empirical evaluation of automated test suite generation and optimization. Arab J Sci Eng 45(4):2407–2423
11. Khalifa EM, Jawawi DNA, Jamil HA, bin Deris S (2019) An efficient method to generate test cases from UML—USE CASE DIAGRAM. Int J Eng Res Technol. 12(7):1138–1145. ISSN 0974-3154

Asymmetrically Clipped Optical Orthogonal Frequency Division Multiplexing (ACO-OFDM) in Underwater Optical Communication

Y. Pavan Kumar Reddy, Y. Sunanda, T. Uday Kiran, M. Venkata Sadvika, and K. Shankar

Abstract In recent development in underwater communication, OFDM system plays a vital role. One among of those OFDM systems are ACO-OFDM technique. Underwater communication systems have become increasingly important in recent years, with a focus on wireless networks in response to the growing demand for ocean data mining. This has recently emerged for both high and moderate turbulence as a viable wireless carrier choice for signal transmission. To reduced inter symbol interference (ISI) and severe signal attenuation, advanced modulation techniques such as OFDM's are used. This project will investigate two optical wireless orthogonal frequency techniques, ACO-OFDM and DCO-OFDM. The performance of ACO-OFDM and DCO-OFDM with and without turbulences will be compared. The examination of BER vs SNR of ACO-OFDM and DCO-OFDM for the condition of 4-QAM modulation will be used to compare these two techniques. In this work, the Monte-Carlo ray tracking system will be used for simulating the free impulse response and route loss using the simulation environment provided by the MATLAB technical computing language.

Keywords ACO-OFDM · DCO-OFDM · Underwater visible light communication · BER · SNR

Y. Pavan Kumar Reddy (✉) · Y. Sunanda · T. Uday Kiran · M. Venkata Sadvika · K. Shankar
Department of Electronics and Communication Engineering, Annamacharya University, Rajampet, India
e-mail: ratnasena.reddy@gmail.com

Y. Sunanda
e-mail: sunanda.bujji@gmail.com

T. Uday Kiran
e-mail: atuday9@gmail.com

M. Venkata Sadvika
e-mail: sadvimanyam@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_72

1 Introduction

As mysterious water sources, the seas must be investigated using a variety of signal processing techniques and applications, including undersea monitoring, watching marine life, and tsunami early warning, among others. In these circumstances, UWC has emerged as a crucial thing. Some of the advantages of UWC are.

- It avoids data spoofing and privacy leakage.
- It is advantageous in pollution monitoring.

There are many several advantages of these UW communications. Acoustics, RF, and optics are the options for implementing UW communications. Recent studies advocate adopting UWC technology to study the maritime environment, such as electromagnetic waves (EM) in terms of RF, acoustic, and optical waves' propagation. To work better in underwater mediums, UWRF signaling needs a large-sized antenna [1]. Similar to this, radio frequency waves are also used at ultra-low frequency bands in order to go forward with little attenuation, but at the expense of a huge budget as well as slow data rate it offers modest data rates. In addition to UWRF, acoustic waves are also receiving a lot of attention for their ability to transmit signals over great distances in bandwidth-constrained underwater media, whereas traditional acoustic wireless carriers have a slow propagation rate over long distances of about 1500 m/s and a low data throughput [2]. In general, channel deficiencies have an impact on how quickly acoustic waves propagate [2]. Contrarily, changes in the water's temperature and salinity result in a significant rise in the speed of acoustic waves. Acoustic signals are distorted by influencing elements such scattering, reflection, and absorption by suspended particles [2].

For underwater communications, underwater visible light communication (UVLC) also has major advantages, also it is thought that UVLC is a useful addition to traditional acoustic communication, providing a dependable and efficient system connecting surface vehicles, submerged equipment, divers, and bottom infrastructure [3]. Along with these advantages, creating trustworthy underwater visible light communication is extremely challenging due to physical barriers, scattering, absorption, turbulence, and misalignment [4]. UVLC has highest transmission data rate and highest band width. Lowest cost implementation is also one of the advantages. So it is also used in applications' real-time data transmissions. In these under visible light communications, several modulation techniques are used, but OFDM systems are more advantageous in terms of several factors. OFDM systems has again various types. In this paper, ACO-OFDM and DCO-OFDM will be investigated.

These OFDM techniques are needed in UVLC system to reduce Inter symbol interference (ISI). In an UVLC system, there will be signal attenuation, so as a part of recent trends these both OFDM techniques are needed to reduce signal attenuation to some extent. Because of orthogonality, OFDM techniques are excellently resistant to ISI and ICI. Especially ACO-OFDM modulates only remaining half of the subcarriers, which makes it a significant one in UVLC applications. Even under small turbulence, it constantly degrades. The SC system known as orthogonal frequency division multiplexing (OFDM) has an orthogonal subcarrier spacing. The high spectral efficiency, ISI resistance, and OFDM is frequently employed in the VLC system, along with frequency-selective fading. Even while OFDM beats remaining modulation techniques in a band-limited channel, the channel's bandwidth also affects how well it operates.

Even and odd subcarriers are modulated in DCO-OFDM. Negative peaks could not be eliminated without a very high bias. However, minor bias is utilized in DCO-OFDM. All subcarriers were impacted by the clipping noise that arises with moderate bias. Only the odd carriers were utilized and modulated in ACO-OFDM. The non-negativity requirement is satisfied by zeroing even subcarriers, which creates a bias signal. Whole subcarriers transmit symbol data in DCO-OFDM, whereas sporadic subcarriers transmit symbol data in ACO-OFDM [5]. ACO-transmitter OFDM's and DCO-transmitter OFDM's are comparable, but their receivers are the same. Although there are factors that affect the ACO-OFDM signal. Even subcarriers only incur interference as a result of zero clipping [5]. Full form of BER is bit error rate. It is defined as no of bit errors per unit time. BER is key parameter for measuring the characteristics of channel, i.e., the performance of wireless or wired data channel. Generally BER is calculated by comparing the transmitted bits to received bits and then count the number of errors. Full form of SNR is signal-to-noise ratio. Signal-to-noise ratio is term used in engineering applications that gives comparison of the desired signal-to-background noise.

$$\mathrm{SNR} = P_{\mathrm{signal}}/P_{\mathrm{noise}} \tag{1}$$

SNR specifications are found in many products that includes audio, including speakers and many more wireless communications. If the SNR value is less, than the product or that device is more efficient. If the incoming signal is weak, it will be closer to noise floor. BER versus SNR analysis is used for comparison of ACO-OFDM versus DCO-OFDM in this paper.

2 Literature Survey

In [2], the author has discussed that, a potential wireless carrier contender for signal transmission in extremely risky, unexplored, and stressed water mediums like oceans has just developed. This technology is known as underwater visible light communication. Unfortunately, the oceans have not received nearly enough attention throughout the history of oceanography. Significant signal attenuation and extremely turbulent channel conditions are present in UVLC systems. This study provides a complete summary of recent improvements in UVLC implementations due to the difficulties inherent with optical signal propagation. In this regard, a comprehensive description of current project plans, channel constraints, different modulation approaches, underwater sensor networks including energy harvesting techniques, hybrid communication opportunities, and the evolution of the Internet of undersea things was offered.

A detailed summary of the project's present plans, channel impairments, different optical signal modulation methods, networks of underwater sensors with renewable energy approaches, hybrid communication possibilities, and the development of the Internet of beneath the water things were provided in this regard.

In [3], the author has suggested the effectiveness turbulence channel of spectrally effective DC-biased optical orthogonal frequency division multiplexing (DCO-OFDM) coupled to transceiver spatial diversity. Absorption, dispersion, and turbulence all negatively impact the UVLC system's ability to communicate underwater. The veracity of analysis is confirmed by analyzing the simulation results with the analytical forecast. It has been demonstrated that spatial variety can significantly lessen channel fading brought on by turbulence. The results can be used to construct, forecast, and assess the DCO-OFDM UVLC system in a weak ocean environment.

In visible light communication (VLC), the visual modulation method orthogonal frequency division multiplexing is utilized, according to [5]. The bit error rate (BER) of numerous 16 QAM and 64 QAM hybrid photonic OFDM systems is proposed to be compared. SEAHACO-OFDM, AAO-OFDM, ADO-OFDM, and LACO-OFDM are among these approaches. Each modulation approach's core concept is presented, and the bit error rate is restricted. According to simulation data, SEA-HACO-OFDM has the best BER efficiency in both scenarios when compared to alternative modulations.

In [6], the author has suggested the intensity modulated/direct-detection (IM/DD) optical systems using the two OFDM techniques which are compared in this research. These are ACO-OFDM and DCO-OFDM. Analysis of the BER versus SNR of ACO-OFDM and DCO-OFDM for intensity-modulated direct-detection systems is used to compare. For lower SNR values, DCO-OFDM is less used than ACO-OFDM because of the issue with optical power. However, it is power-efficient at higher SNR values.

In [7] he author discussed that the current frequency allocation techniques cannot satisfy the demands of growing higher data rate devices. One of the promising methods to alleviate the spectrum shortage and other issues with WiFi networks is Visible Light Communication (VLC). The various modulation techniques appropriate for VLC networks are reviewed in this paper. Additionally, performance measurement of various schemes is reviewed in terms of normalized power need, dimming factor, SNR, bit error rate (BER) and spectral efficiency, VLC applications are shown in the end.

In [8], the author mentioned that, due to its adaptability when transmission includes multipath distortion, orthogonal frequency division multiplexing has acquired widespread usage in wired as well as wireless communication systems. Recently, unipolar waveform has been implemented using new OFDM techniques such as DC-biased OFDM, ACO OFDM, and SSB OFDM. Three different OFDM methods for direct-detection single mode fiber systems are examined along with a linear system model. Bit error rate, normalized OSNR vs. dispersion indices, and OSNR for different fiber lengths are reviewed.

3 Methodology

3.1 Existing Method

3.1.1 DCO-OFDM:

DC-biased OFDM is one of the most used OFDM technique. Since only positive message signals can be delivered, the DC character of the bipolar OFDM signals is added to the DCO-OFDM signal [5]. Hermitian symmetry is used to generate authentic transmitter signals, zero pickings cannot be avoided by adding a DC bias, while this is happening, adding the DC could skew peak power cuts. Before the DC bias is removed at the recipient, M-QAM data cannot be retrieved and demodulated [5]. In Fig. 1, DCO-OFDM system is displayed. All subcarriers in DCO-OFDM system transmits only data symbols. The IFFT is used with complex data signal as input (IFFT). Hermitian symmetry must exist in the input signal [6]. The Hermitian symmetry of input results in a real rather than a complex output signal for the IFFT, then the signal is next transformed from parallel to serial (P/S), a cyclic prefix is added, and output is then converted from digital to analog converter and low pass filter to produce low frequency signal [6].

The resulting signal can be considered as a Gaussian variable that is random for large subcarriers. After adding a suitable DC bias, the remaining negative peak values are removed [6]. Because OFDM signals have an extraordinarily high peak-to-average power ratio, a large bias is required to totally remove negative peaks.

Generally for a DCO-OFDM systems, data symbols are transmitted on odd and even subcarriers, along with that all subcarriers are highly impacted by clipping noise. Any residual negative peak after adding the level of DC bias is eliminated at zero, the optical modulator then receives the clipped signal as input, the intensity of the optical signal output in this case is straight proportional with electrical input current because a perfect optical modulator is being used [6].

A flat underwater turbulence channel is used to transmit the generated signal. In the electrical realm, short noise++ that has an impact on the signal is introduced as AWGN-additive white Gaussian noise, at the receiver, photodiodes are used to convert the obtained signal from an optical to an electrical [6]. After the photodiode's output is filtered, the signal is converted from analog to digital, and the received signal is then changed from serial to parallel. Thereafter, everything is the same as a regular OFDM receiver, once the cyclic prefix is removed. FFT receives the signal as its input, the data stream is then deciphered, transformed from parallel to serial, and finally retained [6].

DC bias IDC is shown below:

$$\mathrm{IDC} = g\sqrt{Es} \tag{2}$$

With help of RC and EGC, the obtained electrical signal can be expressed as:

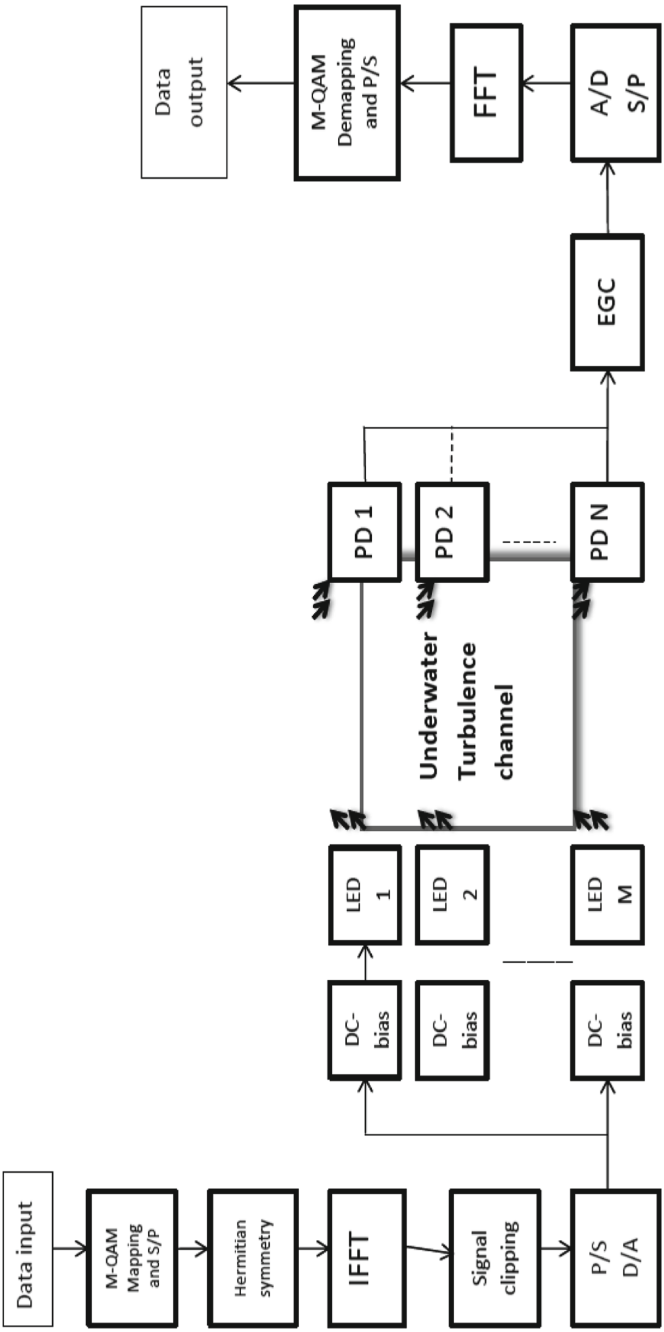


Fig. 1 Block diagram of DCO-OFDM system

$$y(t) = (\eta kPs/MN)\,(x(t)) \sum \sum \alpha_{mn} h_{mn} + n(t). \tag{3}$$

Limitations:

- Even and odd subcarriers are modulated in DCO-OFDM.
- Negative peaks would need to be eliminated with a very strong bias. However, minor bias is utilized in DCO-OFDM.
- Moderate bias results in clipping noise, which affects all sub carriers.

Because of the above limitations, ACO-OFDM is further proposed.

3.2 *Proposed Method*

ACO-OFDM

As mentioned earlier, In ACO-OFDM half subcarriers gets modulated. Only the odd sub-carriers helps in transmitting data symbols in ACO-OFDM, on the other side even subcarriers help in formation of bias signal, so that non-negativity requirement is satisfied [6] (Fig. 2).

The IFFT's input signal only contains odd components, like this, also the vector's components are also required to have Hermitian symmetry [6]. The outcome is a real-time domain signal with the anti-symmetry feature. The ACO-OFDM transmitter's front and end is almost same to a DCO-OFDM transmitter in that a cyclic prefix is added after it has been serialized.

An ideal LPF is used to D/A convert the signal and send it. Signal is clipped at zero since In an IM/DD system, negative samples cannot be communicated, resulting in ACO-OFDM Signal, mainly because of the anti-symmetry of the signal, clipping does not cause any information to be lost, the ideal optical modulator then receives the ACO-OFDM signal as input, and the output signal is broadcast across an AWGN channel [6]. The signal is produced after the D/A conversion and perfect LPF transmission. In an IM/DD system, negative samples cannot be broadcast, and hence the signal is cut at zero, producing the ACO-OFDM signal. The anti-symmetry ensures that there is no information loss.

Due to ACO-OFDM clipping, the signal on the odd subcarriers of the discrete frequency transform vector retains the transmitted data, while the even subcarriers experience interference that resembles noise. Then sent into a perfect optical modulator, with the output signal being sent over a flat AWGN channel [6]. With the exception of the fact that the odd subcarriers in ACO-OFDM gets demodulated because they carry the data, The receiver's processing is similar with that found in an DCO-OFDM receiver. The procedure is identical as for a standard OFDM receiver: the photodiode output undergoes filtering, the signal is transformed from analog to digital, then the resulting signal undergoes a conversion from sequential to parallel format after the prefix cyclic is removed, and the signal is then input into FFT, which

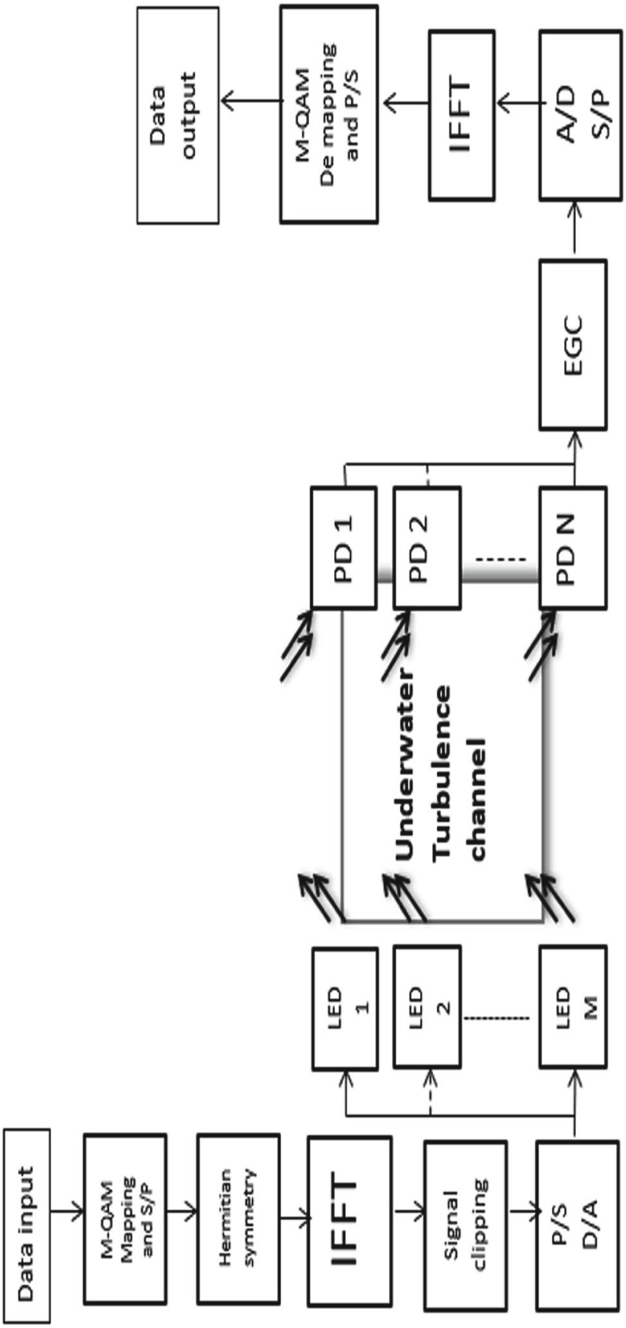


Fig. 2 Block diagram of ACO-OFDM system

subsequently decodes the data stream, converts it from parallel to serial, and then retains it [6].

The electrical signal-to-noise ratio (SNR) of ECG can be derived as:

$$\gamma_{\mathrm{EGC}} = (\eta k Ps/MN\sigma)^2 E[x^2(t)] \left(\sum\sum amnhmn\right)^2 \quad (4)$$

Attenuation loss can be estimated using Beer—Lambert's law as shown below:

$$\mathrm{PL}_{\mathrm{att.}} = e - c(\lambda)d \quad (5)$$

BER performance of QAM OFDM is estimated as:

$$P_{\mathrm{e}} = \left(4\sqrt{L} - 1/\sqrt{L}\log_2(L)\right)Q\left(\sqrt{(3/L-1)}\gamma\mathrm{ff}\right). \quad (6)$$

4 Results and Analysis

In the above plot, BER versus SNR analysis without turbulence for an ACO-OFDM system is displayed. For different values of normalized biases, the BER value is estimated for an ACO-OFDM system. The theoretical and simulated results are represented with different colors for proper comparison. For $g = 1$, the BER value is between 10–1 and 10–2. It is known factor that the BER should be as less as possible for a system. However, in DCO-OFDM, the BER value is even more high. The simulated and analysis BER value for $g = 1$ and $g = 2$ is almost same. Even though $g = 4$ and $g = 5$ values have low BER values at high SNR, but as the normalized values increases the electrical power efficiency decreases for an ACO-OFDM System. As the SNR value increases, the electrical power efficiency also degrades for a system. So for a perfect conditions and efficiency, $g = 3$ is considered (Figs. 3 and 4).

As stated earlier, $g = 3$ is considered as efficient normalized bias value. So keeping it as a constant value, different order values are considered. Different order values like 1×1 2×2 and 4×4 are considered. It is known that as oceanic turbulence strength increases the BER of a system gets worse, but with different spatial orders, the turbulence strength have less impact on BER of system. In the above plot, oceanic turbulence is considered as turbulence condition. For a diversity order of 1×1, the BER value is approximately between 10–1 and 100 for a minimum SNR value. And the BER value is very minimum for a high SNR value of 15 dB. But, when the diversity orders increases, the low BER value is achieved at less SNR values. The previous statement is justified by considering the spatial order 4×4, and it is clearly displayed that for lowest SNR value which is around 7 Db, the lowest BER value 10–4 is achieved. So it is proved that with larger diversity orders the BER performance can be increased.

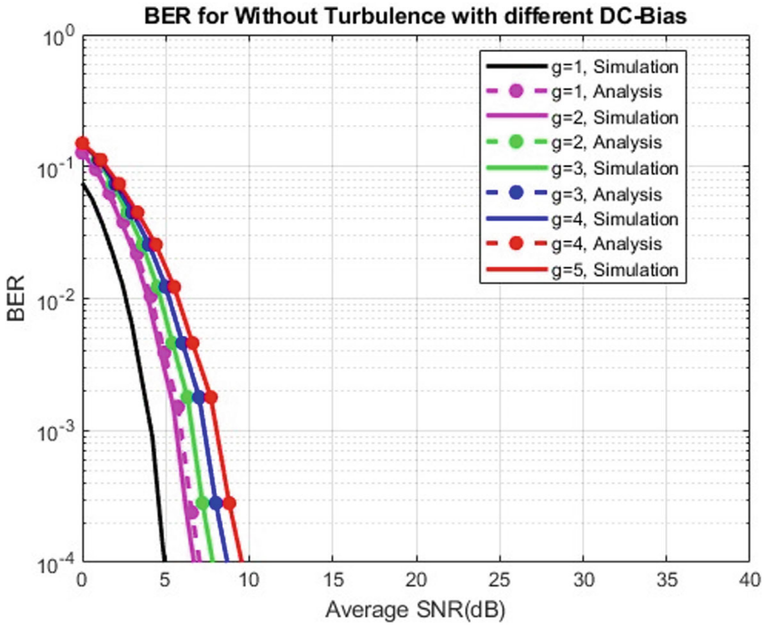


Fig. 3 BER versus SNR analysis without turbulence for an ACO-OFDM system

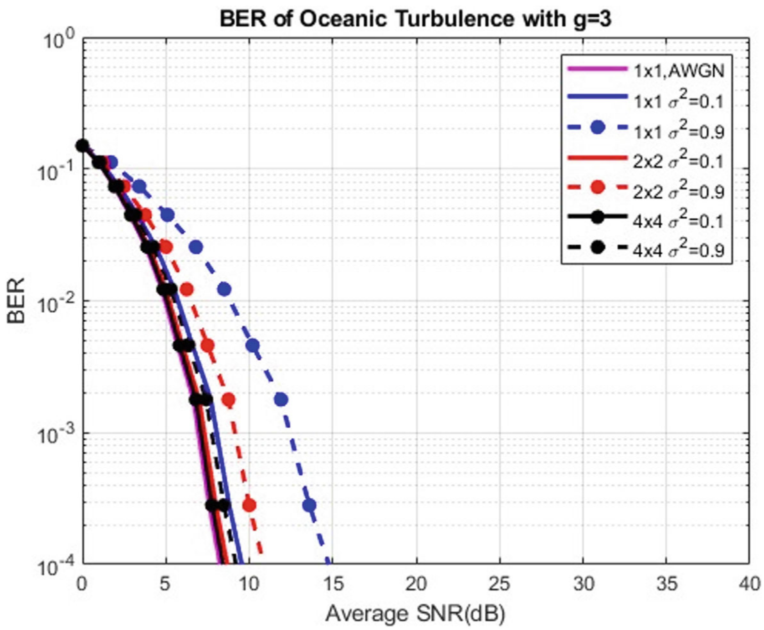


Fig. 4 BER versus SNR analysis for an ACO-OFDM system in oceanic turbulence channel

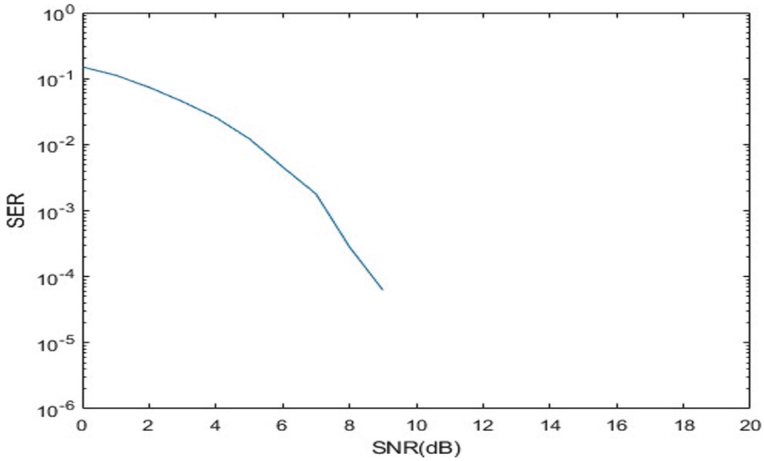


Fig. 5 SER versus SNR analysis of ACO-OFDM system

SER versus SNR analysis of ACO-OFDM system is displayed in Fig. 5. Larger SER values generally indicates low quality communication. So as per the above plot, the SER value is high which is around 10–1 for a low SNR values. The lower SNR values also signify that the system is weak. So the less SER value 10–4 is achieved SNR value around 9 Db. To reduce the SER value for ACO-OFDM system, the SNR of a system is also needed to increase.

Comparison between ACO-OFDM and DCO-OFDM:

Table 1 gives a brief comparison of BER versus SNR analysis of both ACO-OFDM and DCO-OFDM system without turbulence. As displayed in the above table, low BER values are achieved at higher SNR values in DCO-OFDM, whereas in the ACO-OFDM the lower BER values are achieved at low SNR values compared to DCO-OFDM. But with the increase in the normalized value g, the optical power of a system is decreased greatly. For a DCO-OFDM system, even though higher SNR values are achieved, with increased normalized value makes it an inefficient system if we consider optical power. In the case of ACO-OFDM, there is no concern with the optical power because only half of the carriers gets modulated in ACO-OFDM system.

If an oceanic turbulence is considered, the comparison between different two systems for different spatial diverse orders is provided in Table 2. As the diversity order increases, the BER performance of ACO-OFDM system also increases for strong turbulence conditions which makes an ACO-OFDM system efficient one.

Table 1 BER versus SNR (without turbulence)

ACO-OFDM			DCO-OFDM	
	BER	SNR(dB)	BER	SNR(dB)
$g = 1$	0.1	0	0.1	0
	0.0001	5	0.01	19
$g = 2$	0.01	4	0.01	16
	0.0001	6	0.0001	19
$g = 3$	0.001	6.5	0.1	10
	0.0001	7	0.0001	24
$g = 4$	0.1	2	0.1	10
	0.0001	8	0.0001	24

Table 2 BER versus SNR (with oceanic turbulence condition)

ACO-OFDM			DCO-OFDM	
Order	BER	SNR (dB)	BER	SNR (dB)
1×1	0.1	0	0.1	4
	0.0001	8	0.0001	38
2×2	0.01	7	0.01	22
	0.0001	11	0.0001	30
4×4	0.001	7	0.1	0
	0.0001	9	0.0001	24

5 Conclusion

In this chapter, the BER versus SNR analysis is compared for both ACO-OFDM system and DCO-OFDM system under different conditions like without turbulence and with oceanic turbulence for diverse spatial orders with normalized values (g). From the analysis, it is clear that DCO-OFDM is inefficient in terms of optical power when compared to ACO-OFDM. Even though SNR values are high in DCO-OFDM, optical power also need to be efficient for underwater communications. For underwater communication (UWC), optical power is also need to be considered as major requirement. So ACO-OFDM is superior to DCO-OFDM for low SNR values and with low normalized bias value. When there is weak turbulence condition, ACO-OFDM is more efficient than DCO-OFDM in underwater communications.

References

1. Zhu S, Chen X, Xiaoyan L, Zhang G, Tian P (2020) Recent progress in and perspectives of underwater wireless optical communication. Prog Quantum Electron 73:100274
2. Ali MF, Jayakody DNK, Li Y (2022) Recent trends in underwater visible light communication (UVLC) systems. IEEE Access 10:22169–22225. https://doi.org/10.1109/ACCESS.2022.3150093
3. Jiang H, Qiu H, He N, Popoola W, Ahmad Z, Rajbhandari S (2020) Performance of spatial diversity DCO-OFDM in a weak turbulence underwater visible light communication channel. In: J Lightwave Technol 38(8):2271–2277. https://doi.org/10.1109/JLT.2019.2963752
4. Wang Z, Mao T, Wang Q (2017) Optical OFDM for visible light communications. In: 2017 13th International wireless communications and mobile computing conference (IWCMC), pp 1190–1194. https://doi.org/10.1109/IWCMC.2017.7986454.
5. Pall SC, Tanim MMH, Jabiullah M, Ahsan EE, Hasan MG (2021) BER performance analysis of different O-OFDM techniques in VLC. In: International conference on information and communication technology for sustainable development (ICICT4SD), pp 461–465. https://doi.org/10.1109/ICICT4SD50815.2021.9396866
6. Saju SC, Mr. George AJ (2015) Comparison of ACO-OFDM and DCO-OFDM in IM/DD systems. Int J Eng Res Technol (IJERT). 4(04). ISSN 2278-0181
7. Idris S, Aibinu AM, Koyunlu G, Sanusi J (2019) A survey of modulation schemes in visible light communications. In: 3rd international conference on trends in electronics and informatics (ICOEI), pp 1–7. https://doi.org/10.1109/ICOEI.2019.8862538
8. Sangeetha A, Sardar S, Rao S, Kumar KS (2017) Performance analysis of different OFDM techniques in DD-SMF systems. In: 2017 International conference on technical advancements in computers and communications (ICTACC), pp 77–82. https://doi.org/10.1109/ICTACC.2017.29.4

Speckle Reduction for the SAR Images Using an Improved Truncated Variational Method—Based Array

M. Tejaswi, S. Fahimuddin, P. Hari Obulesu, Shaik Karimullah, and Syed Javeed Basha

Abstract The primary challenge for coherent imaging systems is speckle reduction. We present an essay shortened non-convex non-smooth method for silencing speckle in SAR images in this paper. It includes a non-convex regularisation term that has been shortened as well as an I-divergence fidelity term. This shortened, standard (0 p 1) area keep is also used to improve restoration of fresh corners while ignoring contrast reducing artefact. The I-divergence fidelity descriptor effectively suppresses multiplicative noise. It is suggested that an efficient technique based on variable splitting as well as the multiplier's replacement direction method can be used to solve the model (ADMM). Exhaustive study results on scope of SAR pictures show which the suggested technique outpaces state-of-the-art image active noise cancellation methods alike qualitative and quantitatively.

Keywords Synthetic aperture radar is involved in noise removal · Remove noise · As well as directional synchronous edge detection

1 Introduction

Weapons programmes, entire world maps, and environmental control all require high-resolution broad-area imaging. As it is frequently required to capture of above pictures at dark even during inclement weather, synthetic aperture radar (SAR) offers this capability. SAR devices provide greater images by utilising lengthy sound flow characteristics of radar signals along with the advances in data processing capacity of today's digital electronics. SAR is a useful tool for photographers and other photodetectors imaging techniques because it is not limited by time of day or wind conditions, as well as specific reactions of terrain and openness to change to Doppler frequencies.

M. Tejaswi · S. Fahimuddin · P. Hari Obulesu · S. J. Basha
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India

S. Karimullah (✉)
Smartsoc Solutions Pvt. Ltd, Bengaluru, India
e-mail: Munnu483@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_73

Mostly in late 1970s, image pixels adaptive threshold techniques including Lee, Kuan, and Freeze–thaw filters have been commonly then used achieve SAR despeckling [2]. These filters are easy for using and require a poor amount of computational complexity, but they struggle to maintain picture sharpness or reduce speckle [2]. Since the advancement of despeckling techniques, plenty of transform domain approaches, such as wavelet wastage principal components analysis as well as model structure, have been widely used to remove speckle from SAR images[2]. Such reshape context filter techniques can outperform image pixel filtering techniques in terms of edge preservation, however they can yield certain aircrafts [2]. Non-local reduction prior has finally confirmed its overpowering gain in photo or video quantisation [3]. Zhang et al. [3, 4] tried to include the non-linear reduced previous into Matched filter despeckling. They did, however, limit their use of the proposed reduced concept to log-linear SAR images [3].

2 Literature Survey

Speckle reduction is a primary issue in coherent imaging techniques, according to [1]. In this paper, we propose a new shortened non-convex non-smooth model to reduce speckle in SAR images. It includes a non-convex similarity that was clipped together with I-divergence occur in real time. That shortened, standard (0 p 1) regularisation can restore tidy corners while trying to prevent contrast reduction artefacts. I-divergence quality structure is utilised to effectively suppress multiplicative noise. We also provide an efficient technique for solving the model that relies on variable splitting and or the interchanging direction technique of multipliers (ADMM). Experimental findings on a wide range of SAR images show that the recommended method is qualitatively and quantitatively superior to state-of-the-art image noise suppression methods.

In [2], this study proposed two newer models for tackling picture the deblurring problem in the presence of impulsive noise. High-order corrected TVL1 model integrates high-order total variation (TV) regularised term in corrected total variation L1 (CTVL1) model (HOCTVL1). This latest design not only eliminates faults of staircase effect, while also optimises image restoration quality. In most occasions, regularisation parameter in model is fixed value, which could also influence processing results. In order to tackle this problem, the HOCTVL1 model incorporates a regionally suited regularisation parameter selection scheme, as well as spatially adapted HOCTVL1 (SAHOCTVL1) model is developed. When dealing with corrupted images, regularisation parameter in SAHOCTVL1 model can be changed automatically. Many numerical experiments are undertaken.

In [3], spatiotemporal association pattern mining can reveal remarkable interrelatedness among various geographical sources of data. However, majority of spatiotemporal association pattern mining algorithms interpret spatial occurrences as simple spatiotemporal point events. As a result, they cannot be used to predict sophisticated geographic phenomena like storms and air pollution, which continually change their attributes, looks, or locations. Most visible characteristic of such complicated geographic occurrences is the geographic dynamic. This research presents unique extremely complicated occurrence spatiotemporal association pattern mining framework to entirely expose dynamic features of complex geographic phenomena but instead find their associated variables. First, complicated geographic event was modelled hierarchically and represented by a new data structure called directed spatiotemporal pathways. Sequence mining technique was then used to discover the spatiotemporal relationships.

In [4], the simplification of an electronic elevation model (DEM) is used to retain basic aspects of such flat surfaces while restricting its details at various scales. This is a critical component of cartography and landscape analysis. This strategy typically includes data filtration and structural analysis. While gentle structure of a flat surfaces is sometimes ignored, the systemic analysis technique is useful in identifying terrain structural boundaries. Besides that a noise-reducing device may over smooth the structural boundaries of the terrain. As an outcome, we promote combining statistical segmentation plus systemic analytical techniques to conserve both smooth surface characteristics along with terrain structure boundaries. The structural analysis method, in particular, is used to establish all critical relations that are useful for data analysis.

In [5], refinement of digital elevation models (DEM) is employed in mapmaking along with landscape analysis to retain fundamental characteristics of flat surfaces while restricting its specifics at multiple visuals. Quantitative filtering as well as structural analytical techniques are commonly used for this operation. Although the structural analysis method is useful in determining terrain structural boundaries, it ignores the smooth shape of a terrain surface. Furthermore, the filter designed to reduce surface noise may over-smooth its terrain's systemic edges. As a result, in order to preserve both terrain spatial boundaries and smooth morphology, we propose combining quantitative filtering and spatial analysis approaches.

In [6], one of the fundamental topics in maps generalisation is road network generalisation, which focuses on maintaining patterns of road systems across map scales. Previous methodologies defined road network patterns mainly from both topological and geometric perspectives. The concept "digital economy" refers to electronic purchase of goods and services. To be more specific, road segments with a close proximity relationship in the traffic flow structure must be retained together towards small-scale maps to maintain the completeness of the driving route. This study, in this context, offers an improved method for generalising road networks which full absorption geometry, structure, and road traffic patterns. This same phrase "digital economy" refers to electronic purchase of goods and services. Then there are the connections.

In [7], speckle noise has a significant impact on coherent imaging systems, making visual analysis and feature extraction difficult. People display image noise rejection method that utilises weighted nuclear new standard reducing (WNNM) as well as grey hypothesis in this study [8]. First, people apply a logarithmic transformation to the image data, converting speckle noise to additive noise. Second, using grey theory to match local blocks, we will obtain interpolating reduced matrices clustered by related frames of the source patches. The wavelet is then applied to the noisy images to predict the noise variance. Finally, people techniques the picture using WNNM approach [9]. The results show that our method not only improves but also maintains the artistic impression of applied input image [10].

3 Methodology

Existing Methods

SARBM3D

Shading (SB) edition of SAR block-matching 3D (BM3D) filtration system named SB-SARBM3D. SARBM3D is arguably among the most promising as well as accurate despeckling techniques, offering a good balance of noise removal and detail preservation. People adjust the original algorithm to extract the previous available information on the replicated scene, taking scattering concepts into account. The new algorithm is evaluated objectively using a number of synthetic parameters in a variety of distinct and complementary simulated scenarios. Furthermore, for subjective evaluation, some real SAR images, alike inherent to natural and urbanised areas, are compared to various state-of-the-art despeckling algorithms.

DCA

The method of the invention for SAR picture speckle suppression uses a DCA algorithm and includes the following steps: creating a novel second-order operator; creating a non-convex variational model for SAR image speckle suppression with a self-adaptive regular term based on the novel second-order operator; creating a first-order term and a data item; transforming the non-convex variational model; solving the non-convex variational model by using a DCA algorithm; and entering a SAR image with the desired speckle suppression. With self-adaptive regular term based on DCA algorithm, SAR image speckle suppression method can provide speckle suppression outcomes with strong feature retention and smooth area recovery. It also has good adaptability, efficacy, stability, and expansibility.

A sensor positioned on a movable platform, such as an aeroplane, satellite, or the like, detects a target actively generating microwave signals, and a technical method is used to receive an echo signal to create the SAR image. Since the SAR system can continuously transmit microwave signals, a variety of environmental factors, such as night time, bad weather, and the like, may be avoided. As a result, the SAR image has a wide range of application scenarios, particularly for those activities requiring long-term and time-sensitive timing observation.

As an illustration, consider surface scientific studies, rescue and relief efforts, and military reconnaissance. The generated SAR image's spatial resolution has significantly increased as a result of advancements in the SAR system's imaging technology, and the associated application sceneries have grown in size. The resulting

SAR image, however, is invariably accompanied by substantial speckle noise due to the SAR system's intrinsic coherent imaging properties.

Speckle noise, which manifests on an image as extreme light and dark fluctuations between pixels, is typically thought of as multiplicative noise as opposed to the additive Gaussian noise produced by optical sensors. Some extremely important information on the generated SAR image is severely lost as a result of this strong speckle noise such as boundary or texture information. However, such information is extremely important for subsequent interpretation work of SAR images and related applications. Therefore, it is valuable to perform speckle noise removal on SRA images.

The present approach to speckle noise suppression is mostly based on non-local filtering, variational methods, and data driving techniques. Because non-local filtering-based approaches recognise detail information so well, they typically produce excellent suppression outcomes. However, because it also attempts to detect features where none exist, irksome artefacts like brush (PPBit) and ghost (SARBM3D) appear. A lot of work is done using the variation approach based on the good edge protection capabilities and the effective recovery of the uniform region.

Nevertheless, it also inexorably introduces artefacts like the step phenomenon (TV), fuzzy edge (TGV), etc. There are many data-driven techniques that can effectively reduce speckle noise, but the outcomes greatly depend on the training data set chosen. The superior speckle noise reduction techniques listed above all result in artificial effects including ghosting, brush, and boundary blurring, as observed earlier in manuscript.

TGV

In synthetic aperture radar, the issue of noise removal is critical. Unattractive staircasing artefacts are introduced by one often used method, which is founded in total variation (TV) regularisation. This method can produce noticeably sharp edges. In essence, even in areas with clean lines, television procedures tend to produce images based on linear pictures. Such a work proposes a unique approach for noise removal that used a total fundamental variation (TGV) consequence. Given that TGV-based model, which takes into account higher-order smoothing, can lessen the staircasing artefacts of TV, this makes sense. The TGV-based optimisation problem is also resolved using the Nesterov's analysis tool numerical technique. Monte Carlo simulations demonstrate that the suggested technique produces cutting-edge outcomes in terms of both efficiency and speed. The term 'electronic commerce' refers to the sale of goods and services over the internet.

With synthetic aperture radar (SAR) images, speckle noise typically makes it harder to discern meaningful information (textures, corners, curves, and so on). As a result, speckle reduction is critical for other applications that involve instantaneous visual categorisation and segmentation. Until recently, the primary despeckling techniques, with the exception of a dual methodology, could be roughly categorised into four groups: spatial, curvelet, non-local (NL) filtering, variational, along with structural.

Inside the spatial domain, despeckling tactics can be easily divided into two categories. The very first category commonly gets converted multiplicative quality issue into an additional one, but this isn't always adequate. The Lee filter, Frost filter, Gamma optimum a posteriori (MAP), and their different versions fall under the same category [2–9]. These methods work within original image space, but if the periodic review of the SAR image is flawed, they may lose information. Duplex filtering has really been extended to SAR despeckling in order to keep more bounds throughout softening.

The bulk of wavelet-based techniques employ the quantitative wavelet shrinkage in combination with the Mapping criterion. Wavelet-based algorithms generally offer greater overall performance and stronger signal resolution preservation as compared to spatial-domain methods. However, the unusual types of defects seen in curvelet techniques appear as separated models in plain areas of the photo or going to ring all around border lines that may be disturbing to the eye.

SAR despeckling has been effectively improved with a few NL filtering tweaks. A first possibility to directly insert the noise data into the filter's load computation is provided by the so-called Bernoulli NL median filtered [24]. The strong expectation that the photo patch alone provides a reliable approximation of the underlying asset makes this approach biased, especially when there is high speckle noise.

Proposed Method: TRTV

In this paper, people show total variation (TV) regularised approach for SAR image despeckling. The dynamic total variation (ATV) regularised technique, which would be based on a dual formulation, is used to solve the TV regularisation problem. The TV regularisation's parameters are modified based just on sound intensity determined using wavelets. The corner SAR image simply the lack technique utilised in the TV smoothing-based image enhancement model will perform well at preserving image resolution and corners while reducing noises. The optical photos with synthesised speckles were used first in the experiments, accompanied either by SAR images. An evaluation index based on the proportion of the degrees of separation of two neighbourhood social areas of an image with different sizes is planned to evaluate the impact of edge retain despeckling on SAR images. Using both qualitative inspection and objective image quality evaluation indices, the recommended approach basically represses SAR image deficiencies while maintaining image representation curve intensity.

Algorithm

Step 1: Because image noise throughout SAR images is typically regarded as multiplicative noise, the input should be initialised first.

Step 2: Find the sub-problem for the given input.

Step 3: It undergoes many iterations using the Lagrange multipliers method.

Step 4: To effectively solve the non-linear objective model, variable splitting and ADMM are used as a solution method.

Step 5: Hence, using same the strategy we reformulate the problem.

Step 6: When one of the halting conditions is met, then iteration process comes to an end.

4 Results and Analysis

Individuals exhibit and contrast the results for an actual SAR image to unique edges and numerous uniform regions. However, as it is observed, inside of circumstance all tested strategies can cancel out noise. However, DCA as well as TGV confuse the issue corners and homogeneous areas to varying degrees; see Fig. 1 for related zoomed views. SARBM3D also generates ghost artefacts in some areas of the image. For retaining structures inside the corresponding ratio images, our method Triviid produces the best results. The truncated method yields higher ENL and EPI values, which improves recovery.

We present and compare results for a real SAR image with sharp edges and diverse homogeneous regions. As can be observed, all of the tested approaches can remove noise in this case. DCA and TGV, on the other hand, blur borders and homogeneous regions to variable degrees; see Fig. 2 for zoomed images. Furthermore, SARBM3D generates ghost artefacts in some homogeneous zones. In the comparable ratio images, our method TRTVpIdiv produces the greatest results for retaining structures; see Fig. 3, second row. Again, with greater ENL and EPI values, our abbreviated technique achieves superior recovery.

People compare the despeckling outcomes generated by the various methods. To get the best despeckling outcomes, we painstakingly modify the settings of each test technique. We can notice that every attempted approach could indeed successfully reduce speckle. DCA and TGV, especially TGV for its derivative property, obscure and attenuate features, but, to a greater or lesser extent. Structures may be seen in the equivalent ratio photos of DCA and TGV, which shows that not much detail has been preserved. PP Bit experiences photoshop brushes in some homogeneous areas, and SARBM3D encounters ghost effects. One important finding that when angle protection, uniformly region softening, and contrast rebound are being used, TRTVpIdiv produces more appealing results than the other approaches. Variational methods are faster than non-local techniques, and our technique TRTVpIdiv is the quickest of these benefits vary. In addition to being more effective at speckle suppression, the proposed solution required this same least amount of computational complexity.

People demonstrate also compare result for one real SAR image to sharp corners as well as various homogeneous regions. As can be seen, each of the existing method are capable of removing noise in this scenario. DCA and TGV, on other hand, blur corners and uniformly regions to various degrees; zoomed views are shown in Fig. 3. Moreover, SARBM3D produces ghost artworks in some small regions. The term "electronic commerce" refers to the sale of goods and services over the internet. The shortened technique provides better recovery with higher ENL and EPI values.

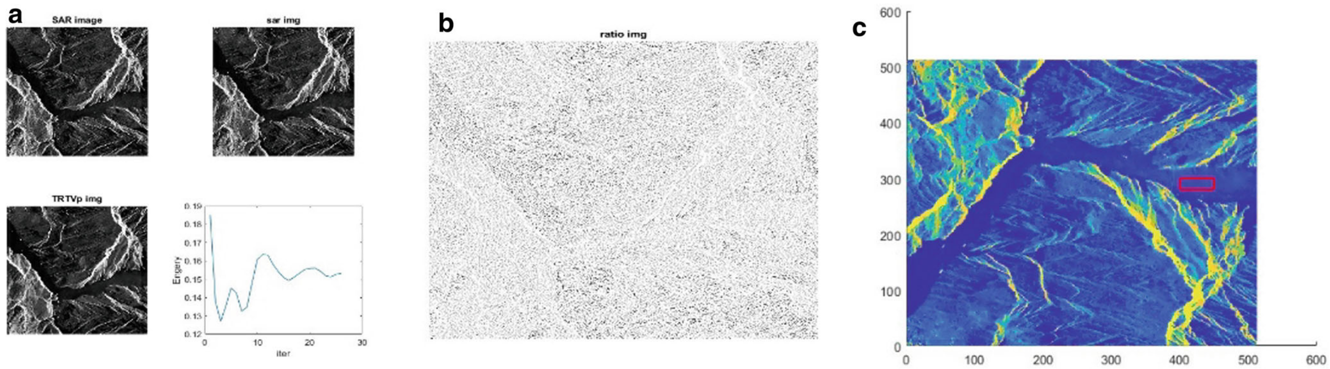


Fig. 1 Noisy input SAR1 with its energy, **b** zoomed view, **c** despeckling results of SAR1

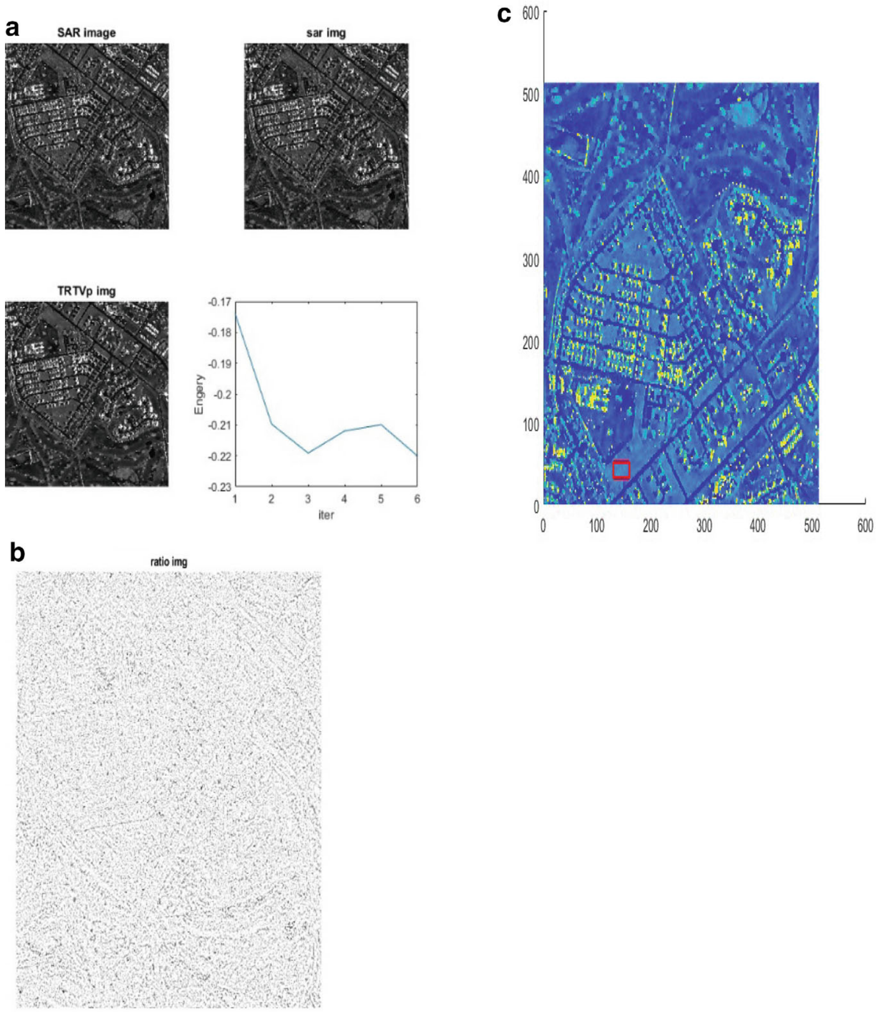


Fig. 2 **a** Noisy input SAR2 with its energy level, **b** zoomed view, **c** despeckling results of SAR2

SAR1 and SAR2 Numerical Review Outcomes

From Table 1 numerical evaluation, it is observed that the SAR1 can be executed with the time limit of nearly 12 s with the parameters ENL and EPI of its values 184.1254 and 0.6641, whereas SAR2 executed with the time limit of nearly 4 s with the parameters ENL and EPI of its values 224.7530 and 0.7521. Here it is observed that the time taken for an image which is not SAR is too high when compared to SAR images, but have less EPI value.

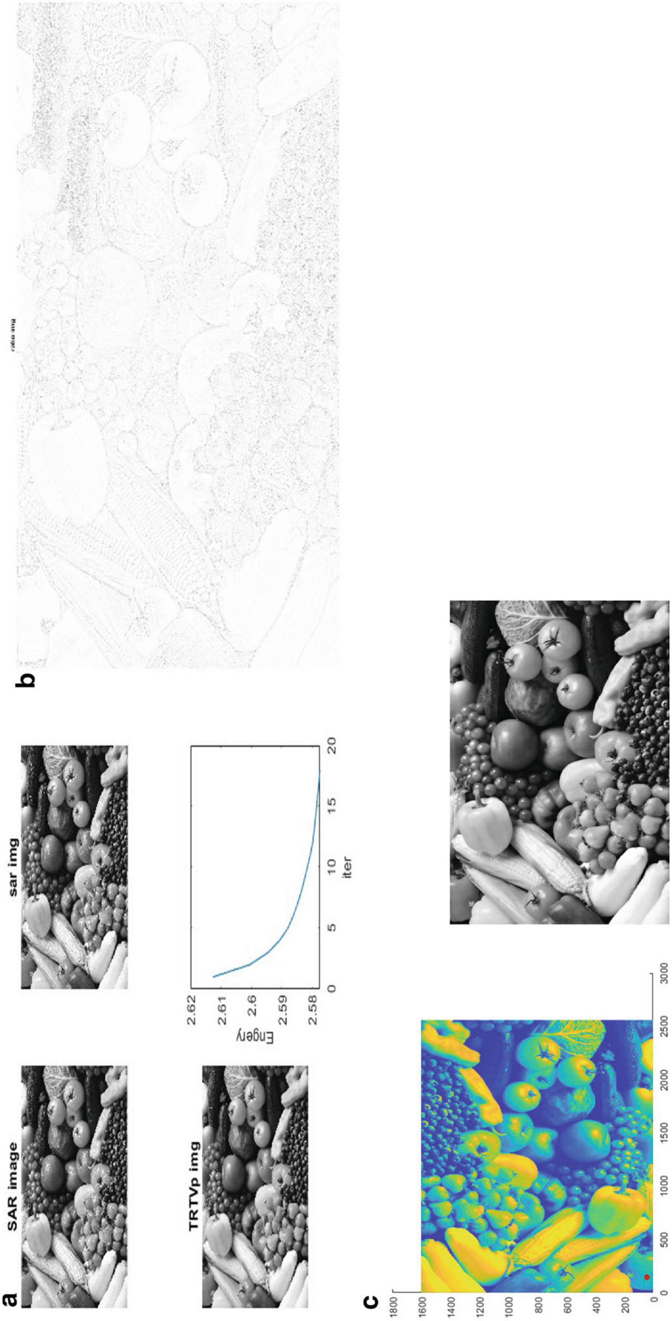


Fig. 3 **a** Noisy input with its energy level, b zoomed view, despeckling results of input image

Table 1 Numerical analysis among ENL, EPI and time

Image	Parameters		
	ENL	EPI	Time (s)
SAR1	184.1254	0.6641	11.979494
SAR2	224.7530	0.7521	3.916971
General image (not SAR)	2.2539e + 04	0.5083	388.645397

5 Conclusion

People display a short story shortened non-convex finite difference theory for noise removal in images acquired in this paper. The current proposal is much more reliable compared to the traditional convex and softening model for repairing clean corners and preserving SAR picture contrast. The method can always be rectified and can use an efficient system utilising variable splitting as well as ADMM. Numerous numerical tests show that the proposed speckle suppression strategy works. We intend to broaden our research beyond the despeckling application to include edge identification, appearance, and structure photo degradation, as well as SAR image reconstruction.

References

1. Guo M, Han C, Wang W, Zhong S, Lv R, Liu Z (2021) A novel truncated nonconvex nonsmooth variational method for SAR imagedespeckling. Remote Sens Lett 12(2):174–183. https://doi.org/10.1080/2150704X.2020.1846820
2. Yu W, Zhang Y, Tinghua A, Chen Z (2020a) An integrated method for dem simplification with terrain structural features and smooth morphology preserved. Int J Geogr Inf Sci 1–23. https://doi.org/10.1080/13658816.2020.1772479.
3. Yu W, Zhang Y, Tinghua A, Guan Q, Chen Z, Haixia L (2020) Road network generalization considering traffic flow patterns. Int J Geogr Inf Sci 34(1):119–149. https://doi.org/10.1080/13658816.2019.1650936
4. Guan D, Xiang D, Tang X, Kuang G (2018) SAR image despeckling based on nonlocal low-rank regularization. IEEE Trans Geosci Remote Sens 57(6):3472–3489. https://doi.org/10.1109/TGRS.2018.2885089
5. Anu Priya D, Vishnuvardhan D, Mamatha G, Shaik F, Karimullah S (2023) VIP development of SPI controller for open-power processor-based fabless SOC. In: Kumar A, Ghinea G, Merugu S (eds) Proceedings of the 2nd international conference on cognitive and intelligent computing. ICCIC 2022. Cognitive science and technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_2
6. Karimullah S, Vishnuvardhan D, Bhaskar V (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Pers Commun 127:3041–3059. https://doi.org/10.1007/s11277-022-09909-2
7. Shaik F, Sharma AK, Ahmed SM (2016) Detection and analysis of Diabetic Myonecrosis using an improved hybrid image processing model. In: 2016 2nd International conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB), Chennai, India, pp 314–317. https://doi.org/10.1109/AEEICB.2016.7538298

8. Karimullah S, Vishnu Vardhan D, Basha SJ (2020) Floorplanning for placement of modules in VLSI physical design using harmony search technique. In: Kumar A, Paprzycki M, Gunjan V (eds) ICDSMLA 2019. Lecture notes in electrical engineering, vol 601. Springer, Singapore. https://doi.org/10.1007/978-981-15-1420-3_197
9. Shaik F, Giri Prasad MN, Rao J, Abdul Rahim B, Soma Sekhar A (2010) Medical image analysis of electron micrographs in diabetic patients using contrast enhancement. In: 2010 International conference on mechanical and electrical technology, Singapore, pp 482–485. https://doi.org/10.1109/ICMET.2010.5598408
10. Karimullah S, Sai Sumanth Goud E, Lava Kumar Reddy K (2023) Spectral efficiency for multi-bit and blind medium estimation of DCO-OFDM used vehicular visible light communication. In: Kumar A, Senatore S, Gunjan VK (eds) ICDSMLA 2021. Lecture notes in electrical engineering, vol 947. Springer, Singapore. https://doi.org/10.1007/978-981-19-5936-3_83

Detection of Skin Disease Using Convolution Neural Network

M. Ravi Kishore, D. Suresh, G. Obulesu, Syed Javeed Basha, and D. Vishnuvardhan

Abstract The expense of faster and more precise diagnostic skin conditions is still prohibitive and high. Therefore, image processing methods aid in the beginning development of an automated dermatological screening system. The classification of skin disorders relies heavily on the extraction of characteristics. In several methods, computer vision plays a part in the identification of skin conditions. Skin infections are prevalent in Saudi Arabia as a result of the deserts and the hot climate. This paper contributes to the understanding of diseases of the skin detection. Based on image processing, we proposed a method for diagnosing skin problems. Using image analysis, a digital image of the diseased skin region is used to diagnose the type of illness. The only expensive pieces of equipment required for our simple, rapid procedure are a camera and a computer. The approach is based on the inputs of a color image. Resize the image after that using a convolution neural network that has been trained to extract features. Utilizing Multiclass SVM, the feature was subsequently categorized. On dermatological macro-images, a modified sigmoid transform based on EfficientNet may be used to improve the contrast between lesion and background areas. The border between the lesion and background areas of the pixel values may be fixed using the modified sigmoid transform. Aids in increasing the accuracy of skin lesion segmentation. Application in practice as a preliminary processing stage in automated tools for skin cancer detection from dermatological macro-images. On dermatological macro-images, skin lesions may not initially contrast or vary in intensity enough from the surrounding tissue. The contrast is further diminished by

M. Ravi Kishore · G. Obulesu
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India

D. Suresh
RNS Institute of Technology, Bengaluru, India

S. J. Basha (✉)
Design Verification Engineer, Smartsoc Solutions Pvt. Ltd., Banglore, India
e-mail: fahimaits@gmail.com

D. Vishnuvardhan
Department of ECE, JNTUACE, Ananthapuramu, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_74

the improper exposure of the light at the time the photograph was taken. Segmentation is negatively impacted by low contrast between the lesion and background areas.

Keywords Image enhancement · Convolution neural network · Macro-images · Skin-cancer sigmoid transform

1 Introduction

One of the most crucial diagnostic techniques for melanoma and other pigmented skin diseases is dermoscopy. Compared to traditional clinical pictures, underlying features are more easily apparent with this less-annoying skin imaging technology, which uses optical magnification and optics that eliminate surface reflection. As a result, there are fewer screening errors and a better ability to distinguish between challenging lesions like Spitz nevi with pigment and tiny, clinically murky lesions [1]. Dermoscopy has been shown to reduce diagnostic accuracy when used by inexperienced dermatologists. Therefore, computerized image analysis approaches are greatly desired to reduce diagnostic mistakes caused by the challenge, and Pictorial analysis is made informally. An automated study of dermoscopy pictures frequently begins with automatic boundary identification, which is significant for majorly two reasons. Because several clinical criteria, including imbalance, border irregularity, and sharp boundary cut-off, are immediately assessed from the boundary, its boundary structure first and foremost gives crucial information for appropriate diagnosis. Second, the precision of boundary identification is required for the abstraction of other relevant clinical structures, such as aberrant melanin structures, blue and white areas along with globules [2]. Automated boundary recognition has a difficult undertaking due to dermoscopy pictures of the lesion and adjacent skin frequently having poor contrast. Additionally, various photos or even the same image was taken in a variety of lighting situations can result in varying image colors, which might compromise segmentation performance [3]. Using a color normalization approach to pre-process the photos, which simultaneously eliminates color fluctuations and improves the images' contrast, this work attempts to solve these concerns. After processing, two distinct segmentation methods are used to segment the pictures. Whereas the second uses a cooperative neural networking technique for lesion edge identification, the first iteratively analyses the picturesque backdrop and produces the best threshold for segmentation. Results scheduled a sizable collection of the lesion photos show that our method provides better segmentation performance, as determined by manual boundaries provided by three knowledgeable dermatologists. Our findings also show that to achieve this precise segmentation, the color normalization process is essential [4].

2 Lesion Segmentation

For lesion separation, applied the following two methods

1. An iterative segmentation scheme.
2. Edge detection with co-operative NN.

The above methods were operated based on corresponding RGB channels of dermoscopy and gray scale channels.

2.1 Iterative Segmentation

This segmentation technique is designed with two requirements in mind: (1) an precise search for the ideal lesion boundaries may be performed by analyzing the entire picture, as a result, and real lesion boundaries were also recovered; approximation of the lesion boundaries inside the modified areas and also curve fitting methodologies were not chosen; and (2) input constraints were generally image dependent and has to be modified depending on corresponding characteristics class of image used in segmentation. Firstly, we use the straightforward noisy reduction technique to lessen artifacts, such as hair, that are frequently visible in dermoscopy pictures. To achieve this, we first remove the background median and then smooth the data using recursive Gaussian methods.

3 Literature Review

The following describes research that has been presented by many authors that relate to the diagnosis of skin lesions and cancer using various methodologies connected to the categorization of skin disorders. Vipin Venugopal and Justin Joseph et al.: used a modified sigmoid transform for Improving Dermatological Macro-Images of Melanoma and Nevi Skin Lesions [5]. Masood A, Al-Jumaily AA et al.: provided statistics and findings from the most significant implementation to date were provided [6]. They analyzed the results and contrasted the effectiveness of several classifiers created especially for diagnosing skin lesions [7]. When information is available, it is documented how numerous factors influence how well an approach works. They examine the outcomes based on various models and provide a methodology for comparing assessments of skin cancer diagnostic models. The shortcomings of several of the available studies are emphasized, and recommendations for more studies are given. Vijayalakshmi MM demonstrated a fully automated approach for identifying dermatological diseases from lesion photos [8], a machine intervention as opposed to traditional detection methods that rely on the medical staff. Our model is developed in three stages: collecting and enhancing compromised data, creating the model, and lastly forecasting. For a better structure and greater accuracy of 85%, we combined many AI techniques, including Convolution Neural Networks [9] and Support Vector Machines, with image processing technologies [10] suggested a method for computer-aided detection and classification of skin lesions for diagnostic assistance that combines medical knowledge with several cutting-edge technologies [11]. Image processing, pattern classification, statistical learning, and mimicking techniques of model-based classifiers, also proposed an approach demonstrated to produce great results [12] offering an accurate classification of up to 98% of cases.

4 Methodology

4.1 By Applying CNN

Transfer learning was used to fine-tune the CNN model's ability to differentiate between malignant and benign tumors. With pictures taken from the ISIC 2019 and PH2 datasets, we retrained by using GoogleNet, ResNet-101, and NasNet-Large. The methods were created using the deep learning toolset models such as Google Net, ResNet-101, and NasNet-Large networks available in MATLAB. For the above-specified networks, load the pre-trained networks using a deep learning toolbox model. The network design and other information about the network layers were then seen using the MATLAB function "analyze Network" [13]. The last classification layer and a last learnable layer both employ characteristics that were retrieved by the convolutional layers to categorize the pictures. The information on how to integrate data into loss value, class probabilities, and projected labels is found in the two last layers of all three networks. The final two layers of the pre-trained networks were swapped out with new layers that better matched our photographs to retrain them to be able to identify fresh images. The layer names are then located and replaced using the MATLAB function "find Layers To Replace." A new, fully linked layer with two outputs is added instead of the last layer that contains learnable weights. The need to automatically configure an output class (common nevus and also melanoma).Using a "training Options" function in MATLAB, we can modify several training settings. The "Initial Learn Rate," "Mini batch Size," and "Max Epochs" parameters are configured. The initial learning rate was changed to a lower value to delay training in the unfrozen layers [14] (Fig. 1).

4.2 Proposed Method

The transformation of the HSV color space is carried out using a modified sigmoid technique. This adjustment is crucial for the successful completion of the subsequent steps in the procedure. Subsequently, the modified sigmoid transform is employed

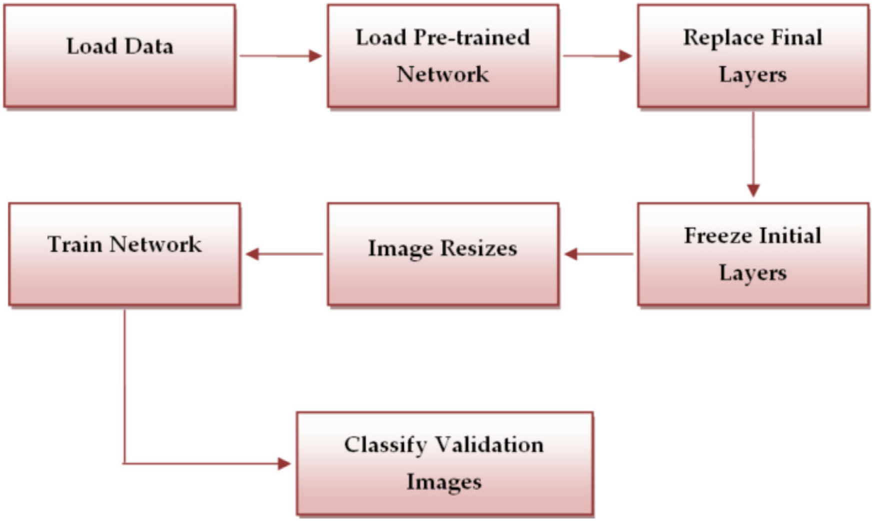


Fig. 1 CNN-based system

to predict a pivotal point utilizing a modified EfficientNet regressor. This crucial juncture serves as an obstacle, effectively dividing the image into areas of abnormality and parts of the background [15]. The objective is to eliminate subjectivity and manual work from the segmentation process.

The typical EfficientNet's classifier layer is substituted with a regression layer in order to tailor the EfficientNet design for this particular objective, yielding the modified EfficientNet regressor. The purpose of this modification is to enhance the model's performance for regression tasks, particularly in predicting the crossover point within the sigmoid-transformed HSV color space.

Transfer learning is employed to enhance the training of the updated EfficientNet regressor. Utilizing a pre-trained model and leveraging its knowledge on the current task results in a reduction in both training time and dataset size. This is a common technique used to enhance the efficiency of deep learning models and utilize knowledge acquired from one task to enhance performance on another. The training data for the modified EfficientNet regressor consists of crucial elements extracted from the macroscopic representation of the training dataset in the HSV color space. These elements are crucial for instructing the model to differentiate patterns linked to lesions and background in the specific color space.

The technique aims to identify suitable crossing locations by utilizing the dice similarity coefficient (DSC) values. These borders represent the areas where the segmented output photos, acquired using Otsu's thresholding, are most accurately matched with the ground truth images. By maximizing the dice similarity coefficient (DSC) values, the segmentation process ensures that the intended regions of interest are precisely captured. This makes the system a dependable tool for automated image

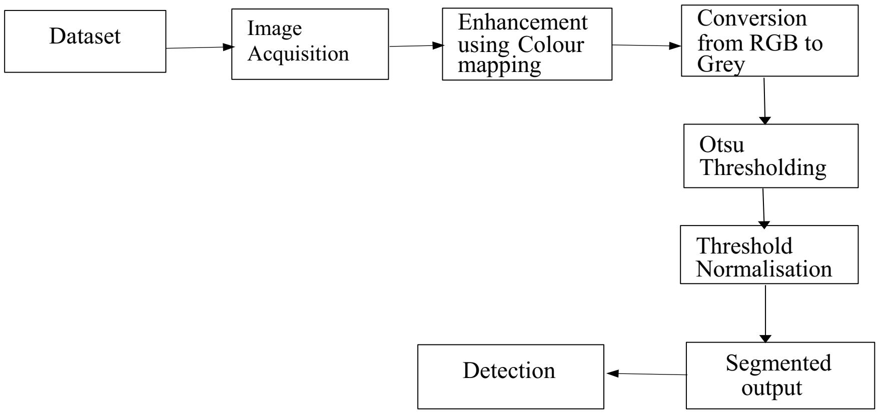


Fig. 2 Proposed process diagram

analysis, particularly in medical imaging and other fields where precise segmentation is crucial (Fig. 2).

5 Results

By applying the above-proposed method for different input images, the threshold value and DSI values are compared which are given in Table 1. The result of the proposed method is shown below for different input images (Figs. 3 and 4).

Table 1 Comparison of DSI value for various input images

Input image	Threshold value	DSI_OTSU
Image 1	214	0.7673
Image 2	116.5000	0.1922
Image 3	83.5000	0.1544
Image 4	126	0.4740
Image 5	119	0.1168
Image 6	139	0.0699
Image 7	66	0.3415
Image 8	107.5000	0.2430
Image 9	69	0.3354
Image 10	101.5000	0.4974

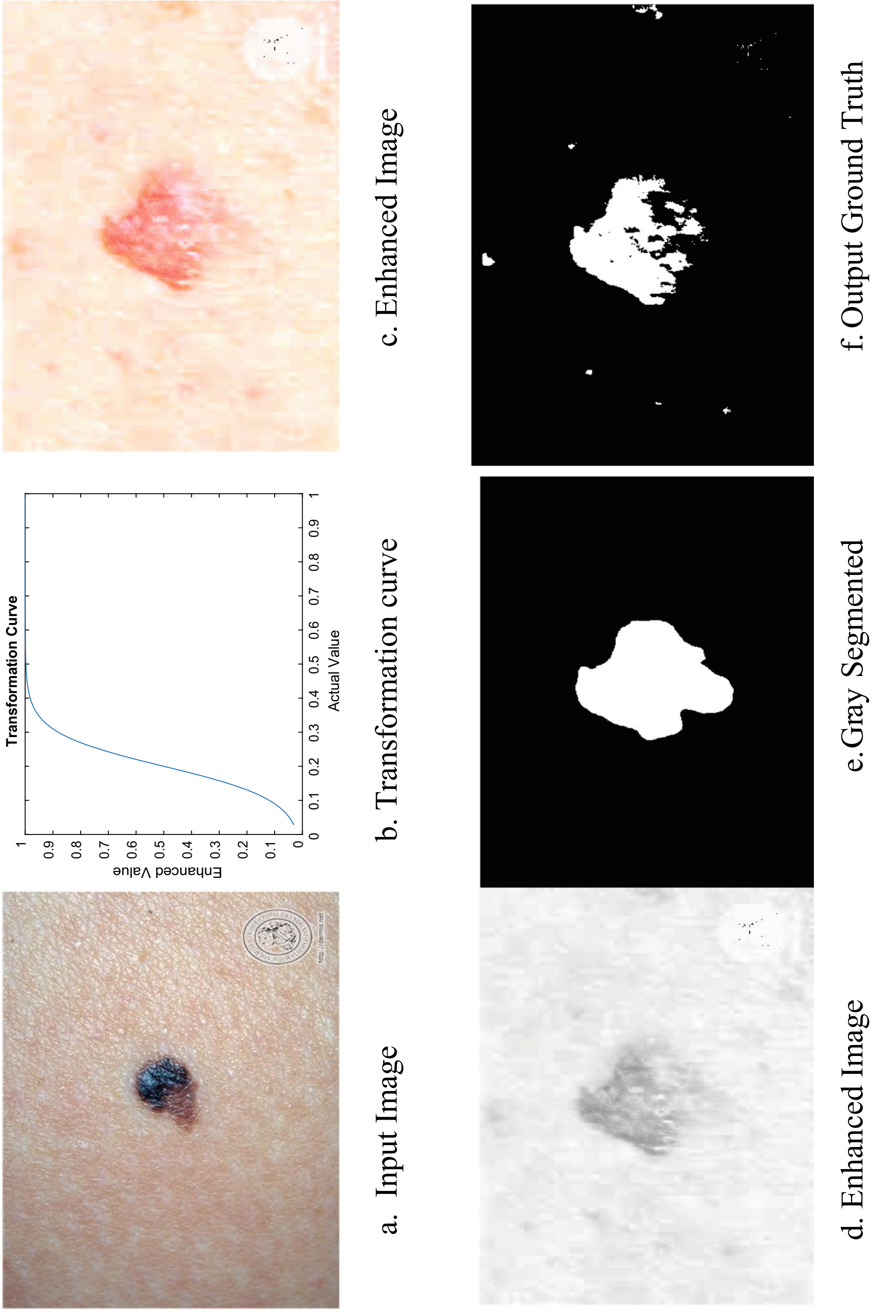


Fig. 3 Result of proposed method for sample 1

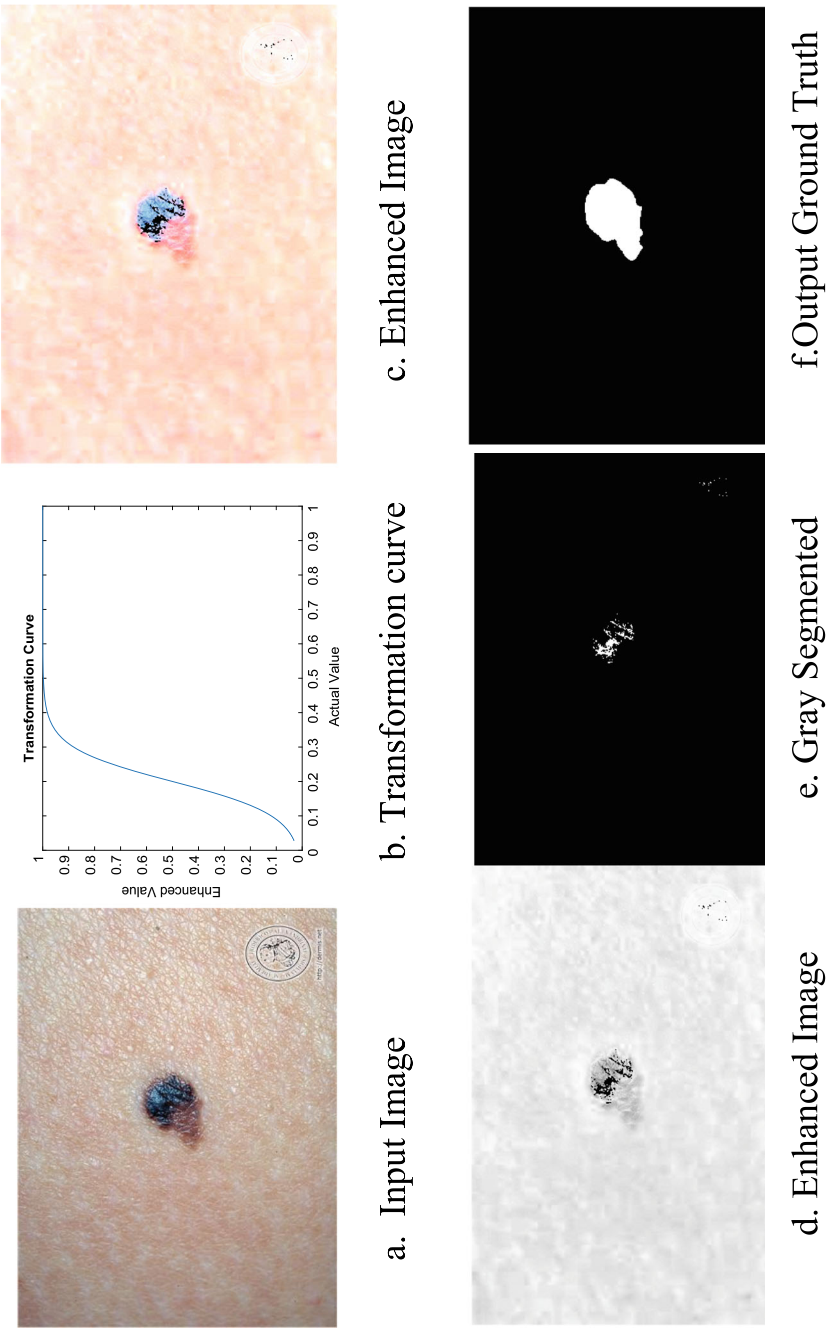


Fig. 4 Result of proposed method for sample 2

6 Conclusions

The proposed approach classifies the type of skin cancer for different images which would consistently increase the difference in background and lesion contrast and supports its use in ADC of dermatological macro-images**.** It also gives increasing accuracy of skin lesion segmentation. The innovative aspect of the current methodology is that it should do the detection quickly, assisting the technicians in honing their diagnostic abilities.

References

1. Venugopal V, Joseph J, Vipin Das M, Nath MK (2022) An efficient net-based modified sigmoid transform for enhancing dermatological macro-images of melanoma and nevi skin lesions. Comput Methods Programs Biomed 222:106935, ISSN 0169-2607. https://doi.org/10.1016/j.cmpb.2022.106935
2. Anu Priya D, Vishnuvardhan D, Mamatha G, Shaik F, Karimullah S (2023) VIP development of SPI controller for open-power processor-based fabless SOC. In: Kumar A, Ghinea G, Merugu S (eds) Proceedings of the 2nd international conference on cognitive and intelligent computing. ICCIC 2022. Cognitive science and technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_2
3. Vijayalakshmi MM (2019) Melanoma skin cancer detection using image processing and machine learning. Int J Trend Sci Res Dev (itsRD) 3(4):780–784, ISSN: 2456-6470. https://www.ijtsrd.com/papers/ijtsrd23936.pdf
4. Karimullah S, Lalitha M, Mounika, Mahesh Reddy S, Mahammad Waseem S, Irfan SM (2023) Performance enhancement of wireless systems using hybrid RIS technique. In: Proceedings of the 4th international conference on data science, machine learning and applications. ICDSMLA 2022. Lecture Notes in Electrical Engineering, vol 1038. Springer, Singapore. https://doi.org/10.1007/978-981-99-2058-7_17
5. Prasad PS, Pathak R, Gunjan VK, Ramana Rao HV (2020) Deep learning based representation for face recognition. In: Kumar A, Mozar S (eds) ICCCE 2019. Lecture Notes in Electrical Engineering, vol 570. Springer, Singapore. https://doi.org/10.1007/978-981-13-8715-9_50
6. Karimullah S, Vishnuvardhan D, Bhaskar V (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Pers Commun 127:3041–3059. https://doi.org/10.1007/s11277-022-09909-2
7. Rajab M, Wolfson M, Morgan S (2004) Application of region-based segmentation and neural network edge detection to skin lesions. Comput Med Imaging Graph 28(1–2):61–68
8. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. Springer Plus 5:507. https://doi.org/10.1186/s40064-016-2152-2
9. Yasir R, Rahman A, Ahmed N (2014) Dermatological disease detection using image processing and artificial neural network. In: Dhaka: International conference on electrical and computer engineering
10. Santy A, Joseph R (2015) Segmentation methods for computer-aided melanoma detection. In: Global conference on communication technologies
11. Kishore MR, Suresh D (2022) Novel method for the segmentation of brain images using the FCM clustering approach as well as rough set. In: Gunjan VK, Zurada JM (eds) Modern approaches in machine learning & cognitive science: a walkthrough. Studies in Computational Intelligence, vol 1027. Springer, Cham. https://doi.org/10.1007/978-3-030-96634-8_42

12. Karimullah S, Vishnuvardhan D (2020) Iterative analysis of optimization algorithms for placement and routing in ASIC design. In: Kumar A, Paprzycki M, Gunjan V (eds) ICDSMLA 2019. Lecture Notes in Electrical Engineering, vol 601. Springer, Singapore. https://doi.org/10.1007/978-981-15-1420-3_199
13. Suganya R (2016) An automated computer aided diagnosis of skin lesions detection and classification for dermoscopy images. In: International conference on recent trends in information technology
14. Karimullah S, Sai Sumanth Goud E, Lava Kumar Reddy K (2023). Spectral efficiency for multibit and blind medium estimation of DCO-OFDM used vehicular visible light communication. In: Kumar A, Senatore S, Gunjan VK (eds) ICDSMLA 2021. Lecture Notes in Electrical Engineering, vol 947. Springer, Singapore. https://doi.org/10.1007/978-981-19-5936-3_83
15. Mendes DB, da Silva NC (2018) Skin lesions classification using convolutional neural networks, arXiv arXiv:1812.02316

Railway Interlocking System Using AI Technique

N. Gomathi, Devaki Sai Sanuth, G. L. Smitha, Jaslin, K. Saravanan, and Obili Setty Sai Naga Tejesh

Abstract The railway interlocking system (RIS) is very important part in the railway signaling system. It enhances the train movement safety. With the development of railway signaling system, it has experienced three stages: Mechanical interlocking system, Relay interlocking system, and Computer-based interlocking system. This paper aims in the development of interlocking systems and discusses the future development of the railway interlocking system.

Keywords Interlocking system · Railway interlocking system · National Rail and Transportation Institute · Machine learning · Reinforcement learning

1 Introduction

One of the main service providers in the nation's transportation industry is the Indian Railways. Its 170-year existence is already complete. On April 16, 1853, a train was initially operated between Thane and Boribandar (Mumbai V.T.). The Indian Railways has successfully adjusted to the nation's shifting transportation and updated its equipment to meet the expanding demands of passenger and high traffic travel needs. It has also incorporated a number of developments in the field of railway technology and continuously.

N. Gomathi (✉) · D. S. Sanuth · O. S. S. N. Tejesh
Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, India
e-mail: Gomathin@veltech.edu.in

G. L. Smitha
SIMATS Engineering, Chennai, India

Jaslin
JJ College of Engineering, Trichy, India

K. Saravanan
RMD Engineering College, Chennai, India
e-mail: saravanan.it@rmd.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_75

2 Interlocking System

1. **Definition**

The Interlocking System (IS) was introduced on 1875 by J. M. Toucey and William Buchanan. It is mainly used for changing the tracks by the trains. The interlocking system, which is a component of the railway signaling system, guarantees that the equipment will work according to the schedule in order to guarantee the safe movement of the train.

2. **Category**

Interlocking system has undergone a number of significant changes. The research presented here classifies the advances into the following three stages based on the device used for implementing interlocking mechanisms.

a. Mechanical components are used to achieve the interlocking interaction. It includes mechanical interlocking system and mechanical integrated with electrical interlocking system were both stages that it went through.
b. Interlocking of relays. Relays and circuits built on the basis of wire logic are used to create the interlocking relationship.
c. Interlocking with a computer. The manifestation of interlocking relationships using processor or computer software. Relay-based computer interlocking system, fully electronic computer interlocking system, and fully digital computer interlocking system were its three stages.

3. **Mechanical Interlocking System**

Charles Hutton Gregory created and developed the lever-based signal system that was put in London's Bricklayers Arms Junction in 1843. It is done with help of the centralized control of switches and signals was made possible by this control paradigm. Centralized control has evolved over the last few years into the fundamental tenet of interlocking systems. The activities are the sequence of the equipment that will be pre-applied to prevent the arrangement of a conflicting path, which is another important component of the interlocking system.

John Saxby was first person to create mechanical interlocking system on year 1856, and Bricklayers Arms also started to adopt it as shown in Fig. 1.

4. **Relay Interlocking System**

The mechanical interlocking system was earlier replaced with a relay interlocking system as shown in Fig. 2 that is based on circuitry because the operating bar was replaced by a tiny button.

5. **Computer-Based Interlocking System** [1–8]

The bulky relay and convoluted relay circuit are slowly being replaced by electronic devices such as computer technology, electronic technology, and network technology advance. The software- and hardware-based interlocking of computers entered the historical stage.

Fig. 1 Mechanical interlocking images

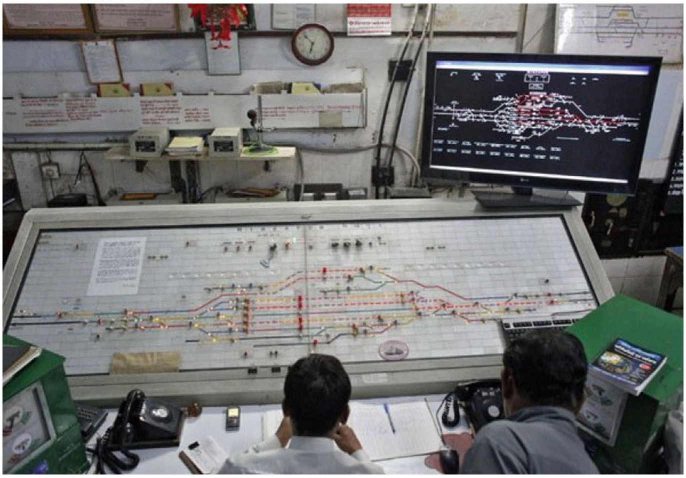


Fig. 2 Relay interlocking system image

Computer-based interlocking systems have the following benefits over relay interlocking systems: high safety, high reliability, high integration, high efficiency, simple maintenance, and so on as shown in Fig. 3. Relay interlocking systems have the following disadvantages over computer-based interlocking systems: very large in size, expensive, great failure rate, difficult in fault diagnosis, and low automation efficiency.

The computer-based interlocking system was implemented in India in 1993. At the Gujarati railway station of Baroda, the first such system was put in place. The National Rail and Transportation Institute (RRTI) and the Indian Railways worked together to create the system. Compared to the conventional relay-based interlocking system, the computer-based interlocking system is more sophisticated and dependable. The

Fig. 3 Computer-based interlocking system images

signals and switches on the railway tracks are managed by computers, removing the chance of human error.

6. **Grouping of Routes for Entry and Exit of Station**

Interlocking systems [1–12] are created to put the constraints in the control table into practice. The control table contains details about the points' locations and the aspect of the signal that will be shown, as well as information about when a route can be defined.

- For Left to Right Movement
  (1) Track G1 can move to right {Route 1, Route 2, Route 3} only if {Route 1, Route 2, Route 14, Route 12, Route 4, Route 3} are Empty
  (2) Track G3 can move to right {Route 1, Route 6, Route 7}only if {Route 1, Route 2, Route 8, Route 10, Route 11, Route 7, Route 13, Route 3} are Empty
  - For Right to Left Movement
  (1) Track G4 can move to left {Route 8, Route 9, Route 10} only if {Route 3, Route 4, Route 12, Route 14, Route 6, Route 10, Route 8, Route 2, Route 1} are Empty
  (2) Track G6 can move to left {Route 8, Route 13, Route 14} only if {Route 3, Route 4, Route 12, Route 14, Route 15, Route 5} are Empty.

Combination of Routes that Does Not Conflicts
The possible combinations of routes that does not conflicts are as follows:

- Route 1 cannot have a conflict with routes Route 2, Route 3, Route 4, Route 5, Route 6, Route 11, Route 13
- Route 2 cannot have a conflict with routes Route 1, Route 3, Route 5, Route 7, Route 13
- Route 3 cannot have a conflict with routes Route 1, Route 2, Route 4, Route 6, Route 12, Route 14
- Route 4 cannot have a conflict with routes Route 1, Route 3, Route 5, Route 7, Route 8, Route9, Route 11

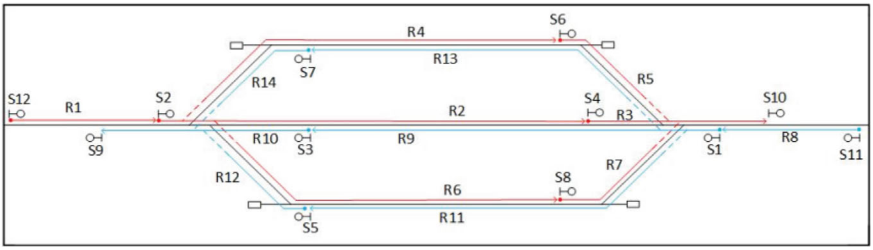


Fig. 4 Procedure that is followed after a train is found in the system

- Route 5 cannot have a conflict with routes Route 1, Route 2, Route 4, Route 6, Route 10, Route 12.

Combination of Routes that Conflicts

The possible combinations of routes that conflicts are as follows as in Fig. 4.

- Route 1 can have a conflict with Route 10, Route 12, Route 14
- Route 2 can have a conflict with routes Route 4, Route 6, Route 9, Route 10, Route 12, Route 14
- Route 3 can have a conflict with routes Route 5, Route 7, Route 8, Route 9, Route 11, Route 13
- Route 4 can have a conflict with routes Route 2, Route 6, Route 10, Route 12, Route 14
- Route 5 can have a conflict with routes Route 3, Route 7, Route 8, Route 9, Route 11, Route 13.

Train Qualities:

- Train ID: Each train is identified in a specific way. ID Alphanumeric is an option
- Train Priority: Determined by the ID that the using side equipment, the system will access the database and obtain information on the train must be aware of train information and its status.
- Stopping station: Following the retrieval of the data, the system will make a decision regarding the train's timetable whether or not the train stops at the station. If Yes, then it ought to be forwarded to either platform as if the train has not been mentioned in the database and Stop, then direct traffic to the station's exit from the available routes.

(3) Recognizing the actual situation: The system checks to see if the optimized route is available before assigning the train's path.
(4) Route assignment and conflict resolution: If the optimized route is free, the route is set. If the optimized route is already taken, the system looks for an other route and assigns it.
(5) Locking of Points: In a railway yard, points should not be simultaneously put to the normal and reverse positions.

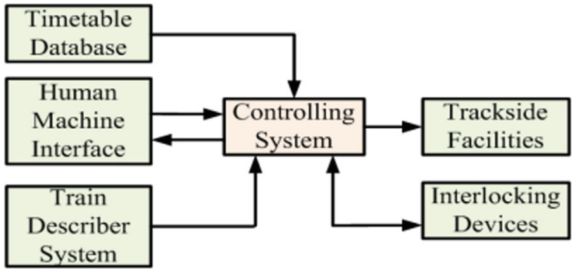


Fig. 5 Flowchart

The system locks or sets the other points in direction once the path has been determined, avoiding them if possible. This feature prevents train collisions.

7. Interlocking System Using AI

To prevent human error, which causes many mishaps, such as the Odisha Accident, we are intending to combine the interlocking System with Artificial Intelligence. We're planning to create an interlocking system that runs automatically with AI's assistance Fig. 5. The locking system won't be changed to the same track if only one train is utilizing it. It contributes to the reduction of railway crashes.

Rule-Based System

Rule-based system can be used for this project as it incorporates precise guide lines and requirements for movement depending on certain weight. We can create if–then rules as required.

Machine Learning Models (ML)

The Following Machine Learning Models that are used for this system are Linear Regression, Neural Networks, and Reinforcement Learning.

Linear Regression: We can also use linear regression to forecast weight if a linear relationship is obtained between data obtained from the sensors and the weight.

Neural Networks: Neural Networks can also play a major role to forecast weight and regulate movement accordingly. It also incorporates with more intricate patterns / algorithms and patterns the Data.

Reinforcement Learning: Reinforcement Learning is used because it is used when a system needs to learn from feedback and modify its movement.

The below graph in Fig. 6 is depicting the average used Routes of the reinforcement learning for the different Routes of wall following with different action selection algorithm (x-axis = trains, y-axis = average usage). As it shows, the proposed approach has given a superior result compared to that of the output of the below graph.

The proposed action selection algorithm is more rewarding than others, with Route 8–Route 11 being more used tracks at initial days but maintaining the same average usage for remaining routes. The proposed selector is as rewarding as Route 4–7 and Route 12–14 are used almost same times, but its average used route rises more after

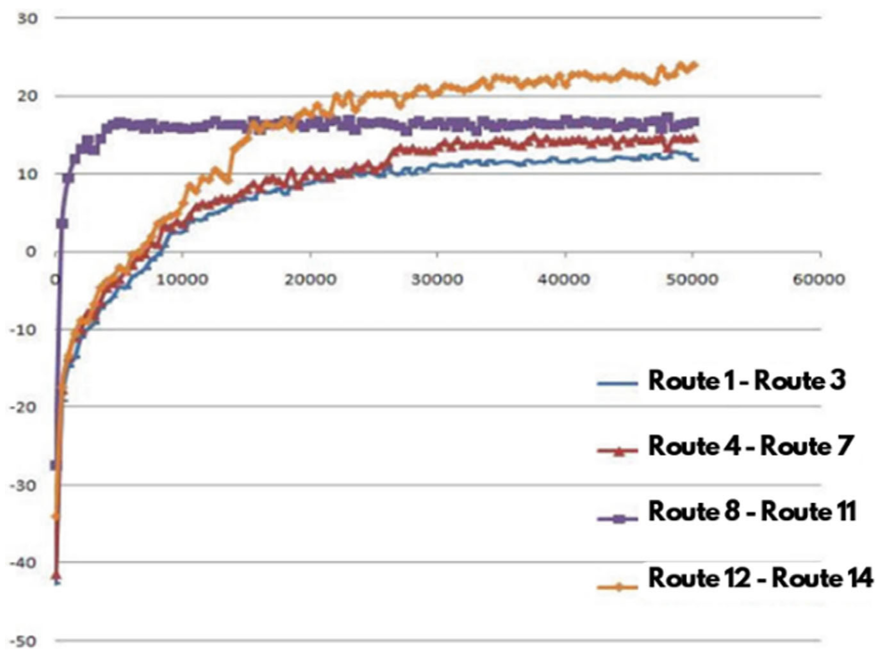


Fig. 6 Reinforcement Learning

that. The suggested method has a built-in knowledge of which route is used more times, unlike the Route 4–7, which is not used more in the exploration section.

Risk Models:

It takes a lot of information, time, and resources to create and maintain a risk model for the railway industry. Checking the availability of historical data in the railway sector is required for this. Big-data modeling of risk also requires operational data, asset data, and safety records. A Machine Learning (ML) approach to analyze all the historical data and find new event patterns and correlations that could forecast the decline in safety performance. The distinction between a proactive and a predictive approach must be made clear. The ability to foresee an event based on routine operating data would allow the operators to take action before any safety is compromised.

3 Conclusion

Finally, it should be noted that the Indian Railway interlocking system is an essential part of the nation's rail infrastructure, ensuring the efficient and safe movement of trains throughout its vast network. This study examined the system's complex features, emphasizing how important it is for averting collisions, derailments, and other potential mishaps. It became clear through a thorough analysis of its technical components, operational difficulties, and safety procedures that the interlocking system is crucial to improving railway safety and operating effectiveness.

Modernizing the Indian railway system by incorporating cutting-edge technologies into the interlocking system, such as computer-based interlocking and centralized control, has made it possible to more precisely handle complex train operations. Remote switch, signal, and route monitoring and control through computerized.

References

1. Zhixi Z (1999) Computer interlocking system technology. China Railway Press, Beijing
2. TieXin (2018) German railway digital interlocking technology and its development. Modern Urban Transit 2018(10):76–77
3. Lishixiang, Lupeiling, Tangshijun (2016) Research on system design of train control system and interlocking system integration. Railway Signlling Commun 2016(3):1–3
4. Huang L (2020) The application of data mining in fault location of computer interlocking system in railway signling. In: IEEE 5th International conference on cloud computing and Big Data analytics (ICCCBDA), Chengdu, China, 2020, pp 331–335
5. Baviskar Y, Suryawanshi U, Sheikh A (2018) Modelling of track layout for intelligent railway signalling system: a machine learning application. In: 2018 IEEE 4th Southern power electronics conference (SPEC)
6. Handbook on basic concepts of railway signalling (India)
7. Huang L (2020) The past, present and future of railway interlocking system. In: 2020 IEEE 5th International conference on intelligent transportation engineering (ICITE)
8. Yoshinori Kon. The current status of signal control systems, and research and development. JR EAST Tech Rev 3
9. Signals at Greenford. http://www.roscalen.com/signals/Greenford/index.htm.2010.8
10. Alstom Signaling Incorporated (2011) A Centennial: History of Alstom Signaling Inc. Alstom. 12
11. General Railway Signal Company. The NX system of electric interlocking. OCLC, Rochester, New York, 184909207
12. He W (2007) 6502 electrical centralized circuit. China Railway Press, Beijing

Image-Based Emotion Recognition: Understanding Facial Expressions

Komuravelly Sudheer Kumar, Pillalamarri Vyshnavi, Ega Deepak, Chinthapatla Suchithra, and Kusuma Naveen

Abstract In this project, we explore the field of emotion recognition using deep learning techniques. We conducted an analysis of a selected dataset consisting of 748 images. Our focus was on evaluating the accuracy and efficiency of algorithms, namely MobileNet, DenseNet, and VGG in recognizing emotions. The DenseNet algorithm achieved an accuracy of 80%, while MobileNet V2 performed lower at 79.33%. The VGG 16 and VGG 19 models demonstrated accuracies of 77.67% highlighting the varying performance of models when it comes to recognizing emotions. Notably, MobileNet emerged as the leading performer, boasting an impressive accuracy rate of 81.33%. Developed a user-friendly interface where users can simply upload an image. The predicted emotion from the image will be displayed as the label, whether it is happiness, sadness, disgust, angry, or surprise. Furthermore, our research highlights the vital role of context-aware datasets, showcasing the nuanced understanding achieved through the integration of custom images.

Keywords DenseNet · Facial emotion recognition · MobileNet · MobileNet V2 · VGG 16 · VGG 19

1 Introduction

Facial emotion recognition, also known as analysis of facial expression, falls under artificial intelligence. Its aim is to determine the distinct human emotions like happy, sad, angry, disgust, and surprise. This involves utilizing algorithms and machine learning methods to analyze the features and patterns in order to understand an emotional mood of a person. The core of facial emotion classification lies in understanding subtle facial expressions and translating them into distinct emotional states.

K. Sudheer Kumar
School of Computer Science & Artificial Intelligence, SR University, Warangal, Telangana, India

P. Vyshnavi (✉) · E. Deepak · C. Suchithra · K. Naveen
Computer Science & Artificial Intelligence, SR University, Warangal, Telangana, India
e-mail: pillalamarrivyshnavi@gmail.com

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_76

The goal is to develop AI systems that can accurately recognize and categorize these expressions, providing insights into emotions.

Some of the models that are popular are VGG, ResNet, DenseNet, and MobileNet. This approach leverages existing knowledge to push the boundaries of emotion classification. In our research project, we investigated the performance of three learning models: DenseNet, MobileNet, and VGG 16/VGG 19 for emotion recognition. Among the models assessed, MobileNet stood out for its exceptional accuracy in recognizing emotions, making it the top performer in our evaluations.

Facial emotion classification has applications across sectors, such as improving the interaction between humans and digital assistants, enhancing gaming experiences, and helping with online education. Moreover, it plays a role in healthcare by assisting in the diagnosis and monitoring of some types of medical conditions. In marketing field, understanding customer emotion is vital for developing strategies and designing products that resonate with customers. Furthermore, facial emotion recognition also contributes to enhancing security measures in access control systems and surveillance systems.

However, the rapid progress in facial emotion classification technology raises ethical concerns, particularly regarding data protection and social bias.

2 Literature Survey

Reference [1] "Recognizing Facial Expressions Using Deep Learning" Within this scholarly outline, they have adapted the applications of deep learning methods, with a firstly pivot on CNNs, to address the intricate task of recognizing and categorizing fundamental human facial expressions. The research findings are based on two datasets taken from Kaggle. To tackle this challenge, the researchers employ ensemble learning techniques, which involve combining the predictions of multiple models, to achieve an initial accuracy rate of 67.2%. However, the paper doesn't stop there. It also explores the potential of transfer learning. Through transfer learning, the model leverages pre-learned knowledge from a different task or dataset to boost its performance on the specific task at hand. The research reports a notable accuracy rate of 78.3% after employing transfer learning techniques.

Reference [2] "An efficient deep learning technique for facial emotion recognition" In this research, the researchers focused on efficiency as their key factor. They focus their methodology on employing a highly effective CNN model technique. The model has a dual functionality: it excels not only in accurately classifying emotions from facial images but also in proficiently identifying both age and gender through human facial emotions. The researchers also conducted a sequence of thorough experiments.

Reference [3] "Extended deep neural network for facial emotion recognition" In the context of this scholarly paper, the authors have introduced an innovative DNN tailored for recognition of facial emotions. To comprehensively assess its effectiveness, the researchers utilized two well-known datasets sourced from Kaggle. Notably, they combined fully connected networks (FCN) with deep residual blocks, leading to substantial enhancements in performance. This study marks a significant advancement in the field of recognizing facial emotions, demonstrating the effectiveness of their model through inventive architectural design and successful utilization of diverse datasets.

Reference [4] "Improving Facial Emotion Recognition with Image Processing and Deep Learning" Within this academic framework, the researchers present a methodology aimed at enhancing the accuracy of emotion recognition. Their methodology incorporates two key components: unsharp masking and histogram equalization. To comprehensively evaluate the efficacy of their approach, the researchers conducted testing on the FER2013 dataset, a recognized benchmark in the domain of recognizing facial emotions. Initially, their CNN achieved a success rate of 69.46%, representing a respectable baseline. However, the paper doesn't stop there. It explores the potential of transfer learning, a potent technique in deep learning by employing pre-trained models like FaceNet, SENet-50, ResNet50, and VGG 16. These pre-trained models, fine-tuned through transfer learning, significantly enhance recognition accuracy. In fact, the ensemble of seven models propelled the accuracy to an impressive 76.01%.

Reference [5] "Human facial emotion detection using deep learning" In this study, researchers delve into the realm of detecting facial emotions using DNN. Their focus is specifically on CNNs and the influential ResNet50 architecture. Their innovative approach introduces a design that harnesses the capabilities of ResNet50 for classifying human facial emotions. To train and evaluate their model the researchers utilize the FER2013 dataset, a recognized dataset specifically designed for recognizing emotions. The results of this research are highly significant as they demonstrate advancements compared to existing methods. These improvements include enhanced precision, robustness, and overall performance, showcasing the capability of this approach to influence the sphere of emotion detection.

Reference [6] "Deep learning-based facial emotion recognition for human–computer interaction applications" This research focuses on transfer learning for emotion recognition. The study included popular pre-trained networks Resnet50, VGG 19, Inception V3, and MobileNet. The CK+ database identified emotions with 96% accuracy in experiments. Model accuracy: VGG 19 (96%), Resnet50 (97.7%), Inception V3 (98.5%), MobileNet (94.2%). MobileNet surprised everyone with its precision.

Reference [7] "Facial emotion recognition based on Textural pattern and Convolutional Neural Network" The paper describes ways to recognize basic fac expressions in textural pictures using a CNN, a deep learning framework. This unique method generates textured facial images using Local Binary Pattern (LBP). Next, they trained the CNN model on LBP-generated images. They tested the proposed face expression detection method on a modified CK+ dataset. They also compared the accuracy of face emotion recognition using the standard CNN model on the original CK+ dataset with the suggested technique on the changed dataset.

Reference [8] "GA-SVM-Based Facial Emotion Recognition Using Facial Geometric Features" This study uses vectorized landmark and landmark curvature to recognize facial emotions. This method uses a GA for parameter and feature selection and an SVM for classification. CNN outperformed SVM on the MUG dataset (96.56 vs. 99.62), whereas SVM exhibited slightly higher accuracy on CK+ (95.43% vs. 95.85%).

Reference [9] "Facial Expression Recognition via Deep Learning" In this paper, they tackled the intriguing and challenging subject of automated facial expression identification in computer vision. Facial expression recognition using machine learning techniques is difficult. Deep learning, a novel machine learning method, categorizes images. The model's predictions were 96% accurate.

Reference [10] "Facial emotion recognition using convolutional neural networks (FERC)" This study introduced face emotion identification using convolutional neural networks (FERC). A two-layer CNN underpins the FERC. Part 1 entails cropping the shot to remove the background. Second, determine the image's sentiment. An expressional vector (EV) extracts face expressions in the FERC model. The software predicted emotions 96% of the time.

Reference [11] "Facial Expression Recognition System Using Machine Learning" This paper proposes a dual-purpose facial emotion recognition system using photo features. The major topics are facial recognition and expression detection. Face detection uses Haar-like features to reset the region of interest to limit appearance changes. The suggested method correctly recognized a person's facial expression with an F1 score of 0.8759 in testing.

Reference [12] "Deep learning and machine learning-based facial emotion detection using CNN" No clear templates for identifying emotions made facial emotion recognition problematic. Nonverbal facial expressions reveal a lot about the feelings. In this paper, they identified emotions using machine learning, deep learning, and neural networks. This study provides a CNN-based method for facial emotion recognition that accurately detects anger, contempt, fear, sadness, calm, surprise, and delight.

Reference [13] "Real-Time Facial Expression Recognition using Deep Learning" This paper shows facial expression recognition using convolutional neural networks. This notion helps analyze how people feel while viewing movies or listening to lectures. Scientists trained a CNN model on the FER2013 dataset and reached 0.8978 test accuracy experimentally.

Reference [14] "Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges" This study compared classic ML- and DL-based facial emotion recognition (FER) methods. Traditional face detection, feature extraction, and emotion classification techniques included random forest, AdaBoost, KNN, and SVM. DL-based techniques reduced face physics model dependence by speeding preprocessing and learning end-to-end.

Reference [15] "Human emotion recognition based on facial expressions via deep learning on high-resolution images." This study offered a face-sensitive convolutional neural network (FS-CNN) for computer vision challenges to simplify human emotion perception. Finding faces in huge pictures and considering locations helps it predict emotion recognition. The FS-CNN has two parts: patch cropping for face detection, landmark analysis for expression prediction, and pyramid pictures for scale invariance. On the UMD Faces dataset, the FS-CNN had a mean average accuracy of almost 95%.

Reference [16] "Emotion recognition through facial expression analysis based on a neurofuzzy network" The article presents an emotion identification system based on psychological research on emotion representation and facial expressions. They evaluated system performance using experimental data. They characterized facial expressions in a continuous two-dimensional emotion space using a rule-based neurofuzzy framework. This approach had great data classification and quadrant clustering accuracy.

3 Methodology

In this project, we have prepared our own dataset consisting of 5 different emotions classified into 5 different classes. Each class in the dataset is transformed into numerical labels ranging from 0 to 4 so that the operations can be done easily by the model. These images stored in the dataset are manually resized and stored in array format to maintain consistency and for systematic processing of the data. We have mainly focused on 5 evaluation models based on our project constraints. The models that we have used for evaluation are DenseNet, MobileNet, MobileNet V2, VGG 16, VGG 19. MobileNet demonstrated exceptional accuracy, achieving 81.33%, making it the top performer. DenseNet achieved an accuracy of 80%, while VGG 16 and VGG 19 attained 77.67%. MobileNet version 2 showed a respectable accuracy of 79.33%. The dataset, now represented in numerical array format, underwent basic preprocessing, like normalization. Rigorous evaluations were conducted for generating outcome of the model as efficient as possible. Comparison-based analysis revealed MobileNet's superior performance in the context of numerical labels. The study's conclusion emphasized MobileNet's selection based on its outstanding accuracy (Fig. 1).

3.1 Dataset

In this facial emotion recognition project, the dataset comprises 744 images, each representing one of five distinct emotion classes: disgust, happy, surprise, anger and sad. To facilitate processing, these emotion classes were numerically encoded, ranging from 0 to 4. The dataset was thoughtfully curated to ensure a diverse and balanced representation of each emotion category. Through meticulous preprocessing, including resizing and normalization, the images were prepared for uniformity in format and pixel values. This structured and labeled dataset acted as the foundation for training and assessing models based on deep learning techniques. This dataset transformation, from emotion labels to numerical representation, facilitated efficient model training and contributed to the project's success in facial

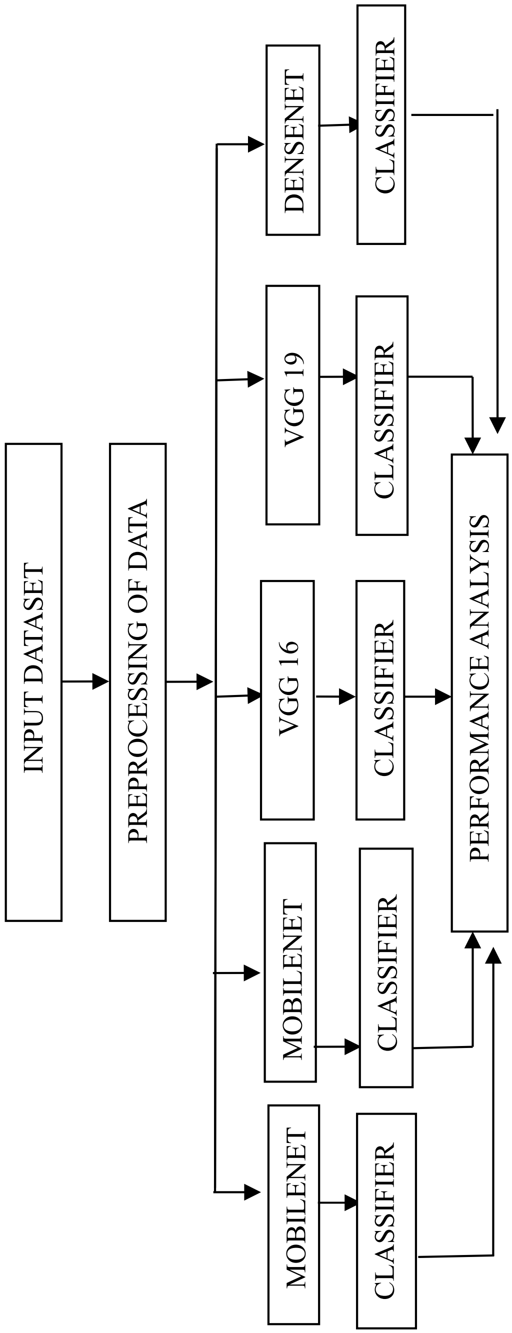


Fig. 1 Emotion detection process

Fig. 2 Images sourced from dataset of 5 distinct expressions

emotion recognition. The images representing five distinct emotions in our dataset are displayed in Fig. 2.

3.2 Model Implementation

We have conducted comparative analysis and done a thorough evaluation of different models including DenseNet, MobileNet, MobileNetV2, VGG 16, and VGG 19. After rigorous training and testing on our dataset, MobileNet emerged as the leading model, outperforming all others in comparison. This meticulous analysis and selection process indicate the robustness and suitability of MobileNet for our specific facial emotion recognition task. MobileNet's efficiency and accuracy have made it the model of choice for our project, showcasing its capacity to effectively identify and recognize intricate patterns. The models employed a Softmax classifier, utilized the Adam optimizer and employed the categorical crossentropy loss function. The parameter like epoch size remained same for all the 5 models.

MobileNet V1

MobileNet models are known for their small size and low computational requirements while still maintaining reasonably good accuracy on tasks like image classification, object detection, and more. They achieve this efficiency through techniques like depth-wise separable convolutions, which decrease computational load and parameter count without sacrificing too much performance. The MobileNet architecture is particularly useful for applications where memory and processing power are constrained, making it a popular choice for deploying deep learning models on mobile apps, edge devices, and real-time embedded systems (Fig. 3).

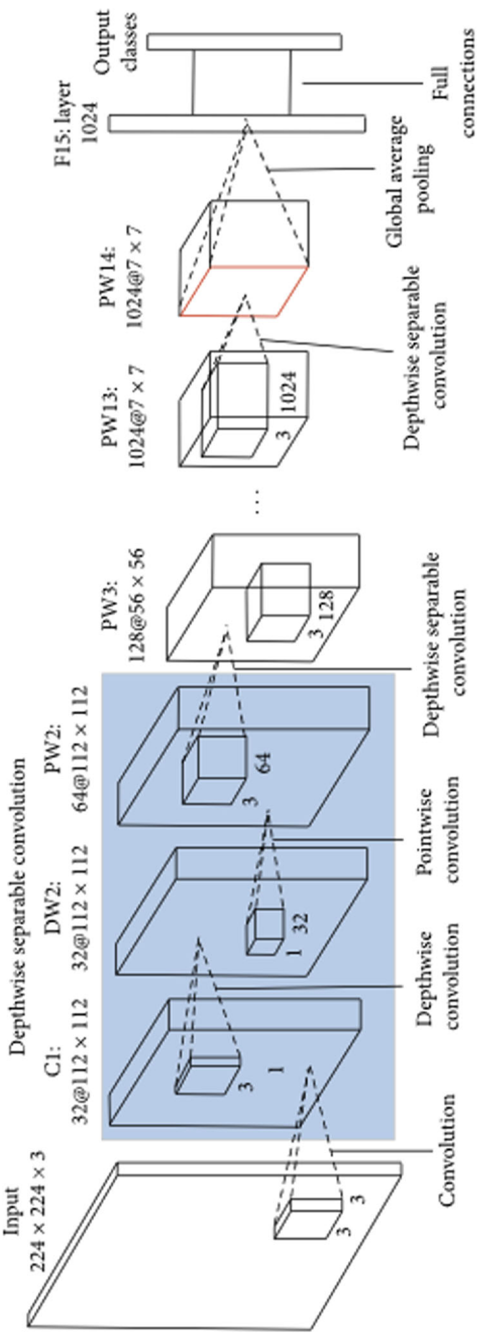


Fig. 3 MobileNet architecture

4 Result Analysis

In this project, we utilized Keras and TensorFlow frameworks for the implementation. The model underwent rigorous training spanning 20 epochs. Among the models tested, MobileNet emerged as the most accurate choice, offering robust recognition capabilities. The dataset comprised five distinct emotional classes: happiness, sadness, disgust, surprise, and anger, meticulously labeled for precise analysis. This careful curation of emotional categories ensured the model's effectiveness in recognizing a diverse range of facial expressions.

In this study, we employed performance metrics to evaluate the performance of our model. MobileNet model findings are presented in Figs. 4, 5, 6, and 7. A comparative analysis was conducted for all the models listed in Table 1.Created a user-friendly interface where users can simply upload the image and get their predicted emotion as the label in the interface shown in Fig. 8.

4.1 User Interface

4.2 Comparative Analysis

5 Conclusion and Future Scope

In concluding our facial emotion recognition project, it is evident that meticulous research, thorough analysis, and innovative technology have converged to create a robust and effective system. After an evaluation process, MobileNet emerged as the choice due to its remarkable accuracy in recognizing 5 distinct facial emotions such as happy, sad, disgust, anger, and surprise. It achieved an accuracy rate of 81.33% surpassing models like DenseNet, MobileNetV2, VGG 16, and VGG 19 which achieved accuracies of 80%, 79.33%, 77.67%, and 77.67%, respectively.

MobileNetV1 displayed proficiency by not capturing basic emotions but also subtle and complex expressions. This high level of accuracy demonstrates the model's efficacy and its ability to detect emotions in real-life situations are worth considering.

The standardized input size of 224×224 pixels played a role in ensuring efficient processing of facial images across all models. This standardization facilitated seamless integration into the models, contributing significantly to the precision and dependability of our system. Essentially, this project demonstrates the potential of empathetic artificial intelligence. By understanding and interpreting human emotions accurately, our facial emotion recognition system opens avenues for emotionally intelligent technology, transforming the way humans and machines interact. As we move forward, the lessons learned and the success achieved in this project will guide us toward even more nuanced and compassionate applications of artificial intelligence in the realm of human emotions. This multi-faceted approach could find

```
Epoch 11/20
19/19 [==============================] - 30s 2s/step - loss: 2.6483e-05 - accuracy: 1.0000 - val_loss: 1.8394 - val_accuracy: 0.8133
Epoch 12/20
19/19 [==============================] - 30s 2s/step - loss: 2.1614e-05 - accuracy: 1.0000 - val_loss: 1.8334 - val_accuracy: 0.8133
Epoch 13/20
19/19 [==============================] - 31s 2s/step - loss: 1.8806e-05 - accuracy: 1.0000 - val_loss: 1.8276 - val_accuracy: 0.8133
Epoch 14/20
19/19 [==============================] - 35s 2s/step - loss: 1.6610e-05 - accuracy: 1.0000 - val_loss: 1.8237 - val_accuracy: 0.8133
Epoch 15/20
19/19 [==============================] - 33s 2s/step - loss: 1.4641e-05 - accuracy: 1.0000 - val_loss: 1.8201 - val_accuracy: 0.8133
Epoch 16/20
19/19 [==============================] - 30s 2s/step - loss: 1.3437e-05 - accuracy: 1.0000 - val_loss: 1.8158 - val_accuracy: 0.8133
Epoch 17/20
19/19 [==============================] - 35s 2s/step - loss: 1.2152e-05 - accuracy: 1.0000 - val_loss: 1.8128 - val_accuracy: 0.8133
Epoch 18/20
19/19 [==============================] - 34s 2s/step - loss: 1.1316e-05 - accuracy: 1.0000 - val_loss: 1.8099 - val_accuracy: 0.8133
Epoch 19/20
19/19 [==============================] - 34s 2s/step - loss: 1.0515e-05 - accuracy: 1.0000 - val_loss: 1.8076 - val_accuracy: 0.8133
Epoch 20/20
19/19 [==============================] - 38s 2s/step - loss: 9.7538e-06 - accuracy: 1.0000 - val_loss: 1.8049 - val_accuracy: 0.8133
```

Fig. 4 Epochs of MobileNet model

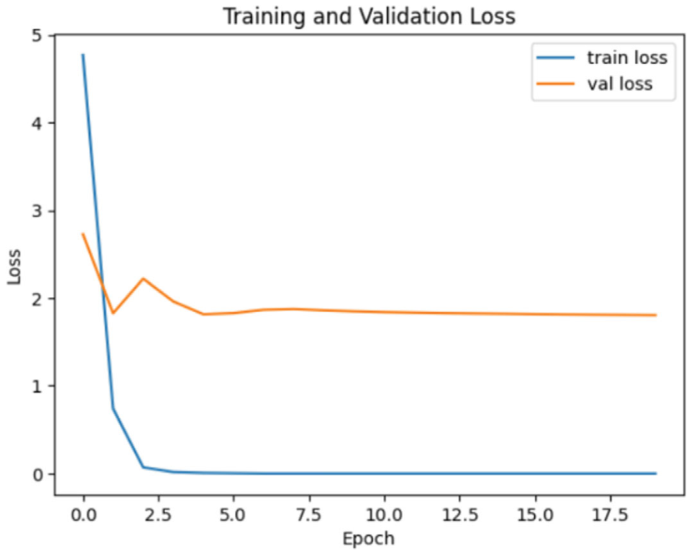


Fig. 5 Loss incurred using the MobileNet model

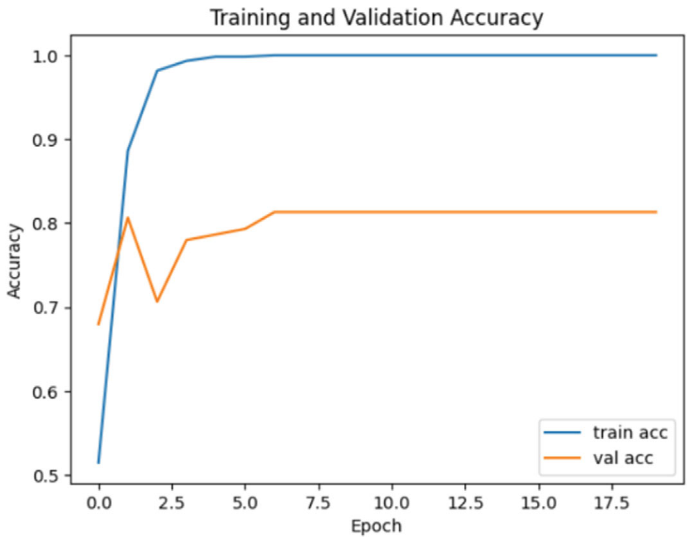


Fig. 6 Accuracy incurred using the MobileNet model

applications in customer service, mental health support systems, and interactions between humans and robots. The future of recognizing facial emotions technology lies in its ability to harmonize technological innovation with ethical mindfulness and cultural awareness. By exploring these avenues, we can create emotionally intelligent

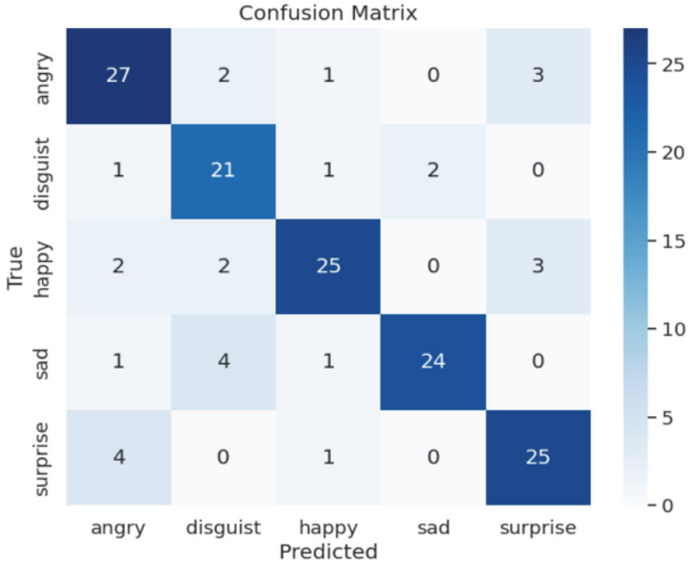


Fig. 7 Confusion Matrix of MobileNet model

Table 1 Comparisons within proposed models

S. No.	Network	Recall	Precision	F1 score	Accuracy
1	MobileNet	0.81	0.82	0.81	0.81
2	DenseNet	0.80	0.80	0.80	0.80
3	MobileNet V2	0.79	0.80	0.79	0.79
4	VGG 19	0.74	0.75	0.74	0.74
5	VGG 16	0.74	0.76	0.74	0.74

AI systems that not only recognize a wide array of emotions accurately but also do so responsibly and inclusively, ushering in a new era of empathetic human–machine interaction.

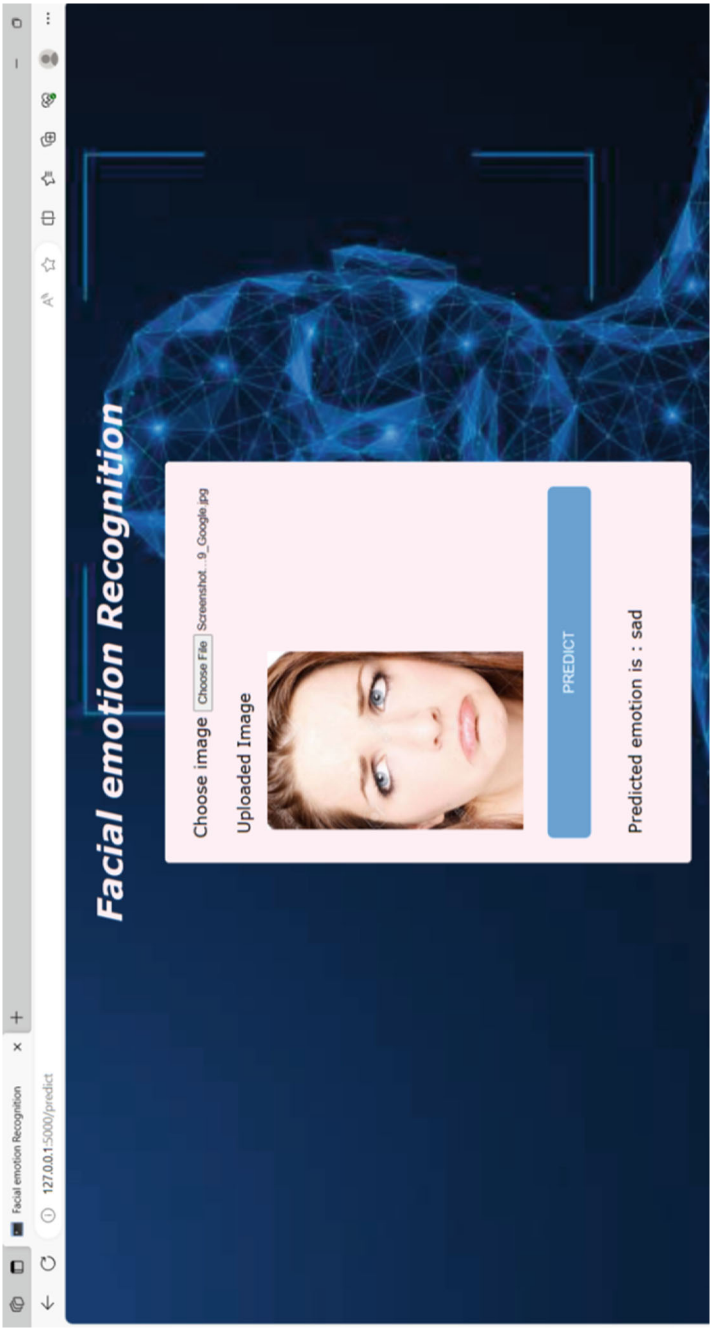


Fig. 8 Prediction of facial emotion (Sad)

References

1. Savoiu A, Wong J (2017) Recognizing facial expressions using deep learning. In: Recognizing facial expressions using deep learning, 1–8
2. Khattak A, Asghar MZ, Ali M, Batool U (2022) An efficient deep learning technique for facial emotion recognition. Multim Tools Appl, 1–35
3. Jain DK, Shamsolmoali P, Sehdev P (2019) Extended deep neural network for facial emotion recognition. Pattern Recogn Lett 120:69–74
4. Vepuri KS (2021) Improving facial emotion recognition with image processing and deep learning
5. Gaddam DKR, Ansari MD, Vuppala S, Gunjan VK, Sati MM (2022) Human facial emotion detection using deep learning. In: ICDSMLA 2020: Proceedings of the 2nd international conference on data science, machine learning and applications. Springer Singapore, pp 1417–1427
6. Chowdary, MK, Nguyen TN, Jude Hemanth D (2023) Deep learning-based facial emotion recognition for human–computer interaction applications. Neural Comput Appl 35(32):23311–23328
7. Mukhopadhyay M, Dey A, Shaw RN, Ghosh A (2021) Facial emotion recognition based on textural pattern and convolutional neural network. In: 2021 IEEE 4th International conference on computing, power and communication technologies (GUCON). IEEE, pp 1–6
8. Liu X, Cheng X, Lee K (2020) GA-SVM-based facial emotion recognition using facial geometric features. IEEE Sens J 21(10):11532–11542
9. Fathallah A, Abdi L, Douik A (2017) Facial expression recognition via deep learning. In: 2017 IEEE/ACS 14th International conference on computer systems and applications (AICCSA). IEEE, pp 745–750
10. Mehendale N (2020) Facial emotion recognition using convolutional neural networks (FERC). SN Appl Sci 2(3):446
11. Kim S, An GH, Kang S-J (2017) Facial expression recognition system using machine learning. In: 2017 International SoC design conference (ISOCC). IEEE, pp 266–267
12. Singh SK, Thakur RK, Kumar S, Anand R (2022) Deep learning and machine learning based facial emotion detection using CNN. In: 2022 9th International conference on computing for sustainable global development (INDIACom). IEEE, pp 530–535
13. Talegaonkar I, Joshi K, Valunj S, Kohok R, Kulkarni A (2019) Real time facial expression recognition using deep learning. In: Proceedings of international conference on communication and information processing (ICCIP), 2019
14. Khan AR (2022) Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis and remaining challenges. Information 13(6):268
15. Said Y, Barr M (2021) Human emotion recognition based on facial expressions via deep learning on high-resolution images. Multim Tools Appl 80(16):25241–25253
16. Ioannou SV, Raouzaiou AT, Tzouvaras VA, Mailis TP, Karpouzis KC, Kollias SD (2005) Emotion recognition through facial expression analysis based on a neurofuzzy network. Neural Netw 18(4):423–435

Achieving Greater Agility Through Scaling Frameworks—An Empirical Analysis

Ekbal Rashid and Vinit Kumar Gunjan

Abstract Enterprises have moved or are rapidly moving towards standard agile. Imagine you are part of an enterprise which has already embarked on agile transformation. You have put together an efficient and cross-functional team at the start of a project that is dedicated in achieving the objectives of the project and the business. You need to make sure the team is Agile savvy if you want performance to increase. As the product's demand rises, so do the features, requirements, and the backlog. You are aware that the team needs to be scaled, along with other stakeholders. This will sound very familiar for enterprises who are on agile and are likely candidate for moving towards scaling frameworks. Scaling enables collaborative teams from various functions to keep their agility and be more productive and efficient. But the choice of the framework itself needs to be analysed, a wrong choice could lead to implementation failures, which would disrupt productivity and have a significant financial impact too. In this paper, we will assess benefits and adoption challenges of different scaling frameworks.

Keywords Scaling · Frameworks · Scaled agile · SAFe · LeSS · Nexus · Scrum@Scale

1 Introduction

Scaling framework in conjunction with "agile" is the buzzword sweeping the software industry and gaining traction in more and more verticals such as manufacturing, eCommerce, and retail. Agile software development has been in use since its inception in 2001. The methodology for developing software has changed to assist businesses in keeping up with the pace of the market. Agile promotes the idea

E. Rashid (✉)
Department of CSE (AI/ML), St. Peter's Engineering College, Hyderabad, Telangana, India
e-mail: ekbalrashid2004@yahoo.com

V. K. Gunjan
CMR Institute of Technology, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_77

that software/feature should be delivered on regular basis, giving the customer the choice to accept it rather than having to wait till the end. The goal of scaling is to give enterprises the flexibility to change and be competitive. By providing quick, adequate responses to customer needs and adding charming touches to entire repeat business. In the end, it means disseminating agile working process to all teams at all levels and in all departments, including senior management and reducing coordination and control. Autonomous teams, shared values, objectives, accountability, and transparency at every level—both up and down are the core to scaling the agility for a successful transformation. Scaling framework that is best will vary depending on things like the organization ability to change, amount of budget that is available, the organizational culture, and the complexity of the product. Although there are many scaling frameworks available, we will concentrate on the four most popular ones: "Scaled Agile Framework (SAFe)", "Nexus", "Large-Scale Scrum (LeSS)", and "Scrum@Scale".

2 Literature Review

Enterprises had already started the adoption of agile frameworks and methods in small teams [1]. It had ignited a new passion among firms to start using agile methods in large projects, even beyond software development, across the enterprise [2]. This phenomenon is commonly being now referred as "large-scale agile development" [3]. We understand that the term "large-scale agile development" [4] is the application of agile methodologies and processes in large multi-team setup with collaborating stakeholders consisting of 50+ members or minimum of six teams. There are multiple scaling frameworks available to support enterprises.

Scaled Agile Framework (SAFe)—According to a 2021 survey [5], 37% of practitioners use it. This is largely because it has a variety of configurations, all which centres on value streams, have clear guidelines and procedures. Teams are organized as "Agile release trains (ARTs)" under SAFe because it is leveraged to build complex solutions requiring 50+ team members. "SAFe" uses program increment iterations, which are comparable to Scrum sprints and typically is 8–12 weeks, to synchronize different teams in "ART". With this strategy, product managers can efficiently manage a product roadmap without making too many changes while maintaining their attention on the big picture and enterprise goals (Fig. 1).

Nexus Framework—This is based on Scrum, is less complex than SAFe, and makes it easier for 3–9 teams to collaborate. Nexus implementation does not call for any additional roles. Instead, a centralized integration command team that unifies work towards common goal is formed with 1 representative from every team. All teams join for sprint planning meeting, which is similar to Scrum, and the outcomes are used to create the common product backlog. Each team has a daily stand-up-style meeting to review progress, and the integration team also gets together to discuss team progress (Fig. 2).

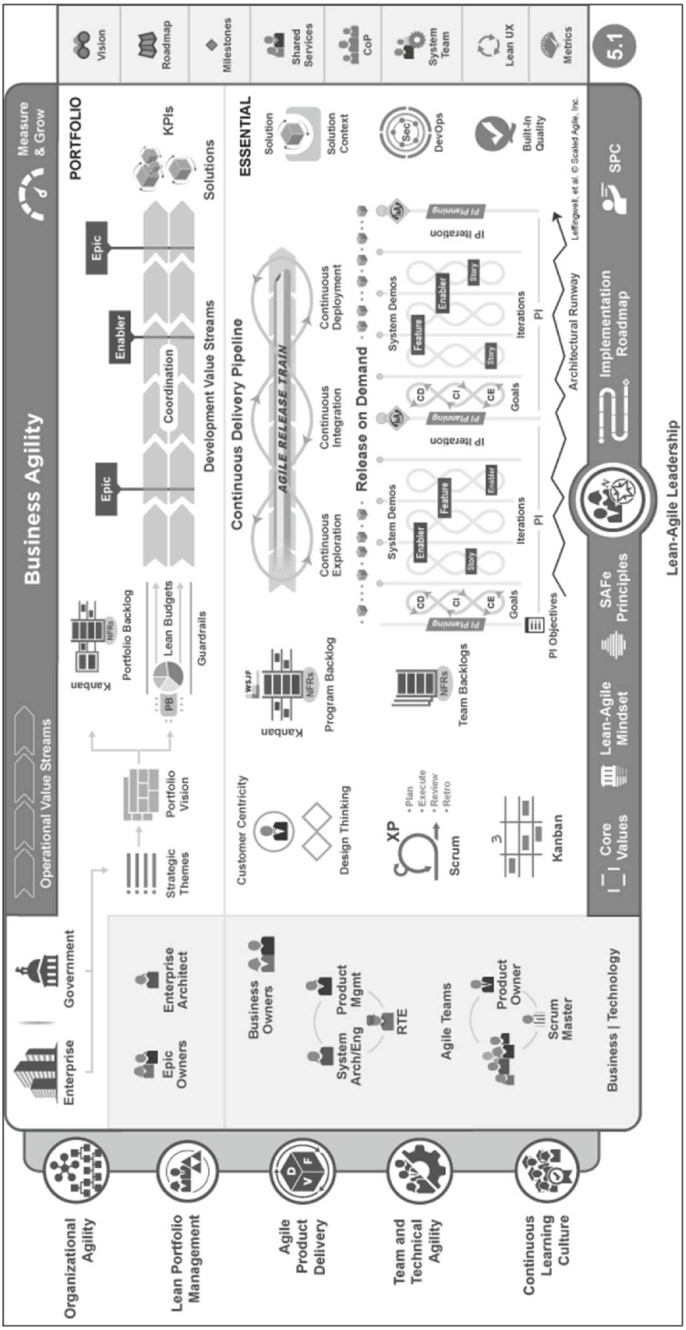


Fig. 1 "SAFe—full configuration" [7]

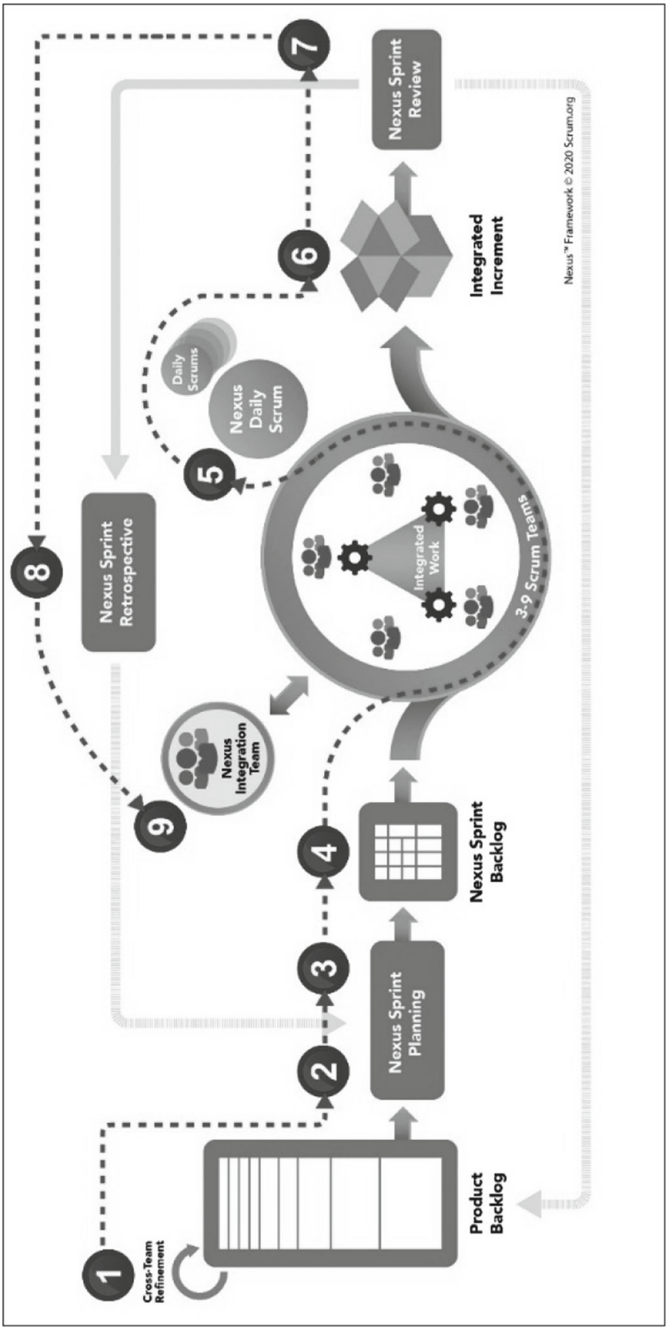


Fig. 2 "NEXUS framework" [10]

Large-Scale Scrum (LeSS)—Although LeSS and Nexus are nearly identical, they differ slightly in terms of naming conventions and the number of team-specific sprint planning meetings. With its additional configuration, "LeSS Huge" supports the collaboration and cooperation of more than eight different teams, and it also differs in that it can be expanded. "LeSS Huge" organizes development using a customer-centric strategy. The product owner must divide the product backlog into more manageable "area backlogs" of more specific items before further organizing them into requirement areas to manage work effectively (Fig. 3).

Scrum@Scale—It is an augmentation of Scrum, and it is likely the simplest framework to understand and use. It can scale from 1 team to a notion of "team of teams". Important component of the framework is "Scrum of Scrums (SoS)". Every team involved chooses their representative for "SoS" meetings. The goal of each day for "SoS" meeting is to assist in coordination, collaboration, and communication among different teams. This helps in ease of day-to-day management of any dependencies. *Image Source* [8] (Fig. 4).

There are host of benefits in usage of scaled frameworks which includes some of the below.

Common goal—Achieving team alignment is the process of making sure that all cross-functional teams are committed to a single objective or goal. Scaling frameworks bring the enterprise closer to better collaboration, increased coordination, and communication while also ensuring better alignment.

Facilitate cross-domain planning—Enterprises frequently struggle with team planning. Numerous teams come together for program planning where they discuss the roadmap, work on features, and in addition acknowledge cross-team dependencies. It leads to establishing of face-to-face communication, adhering to a shared vision, fostering trust, leading to completing tasks more quickly, and hastening decision-making.

Build on enterprise connectedness—It represents and connects the various backlogs and different teams, enabling enterprise-wide visibility. Having enterprise-wide visibility and managing their local work are both facilitated by adopting scaling frameworks.

Increased transparency—The scaled framework breaks down the conventional silos of development, quality assurance, and operations and makes cross-functional communication easier. It increases transparency which drives growth, innovation, and continuous improvement.

Better quality output and productivity—[6] Enterprises rely on quality to quickly develop and respond to shifting business environments. Its overall quality practices guarantee a quick flow between all levels and teams. Building on environments that are stable is essential for large enterprises because excessive rework and slower speeds caused by a lack of quality assurance will produce unacceptable results.

These are general challenges in adoption of scaled frameworks.

Lack of motivation to change—[12] It is simple to assume that no matter how an organization is doing, everything is fine. But in the comfort zone, there is sadly no room for growth. Enterprises require a push to change. To make this big change, it needs to be sufficiently motivated across levels. Every member of the enterprise,

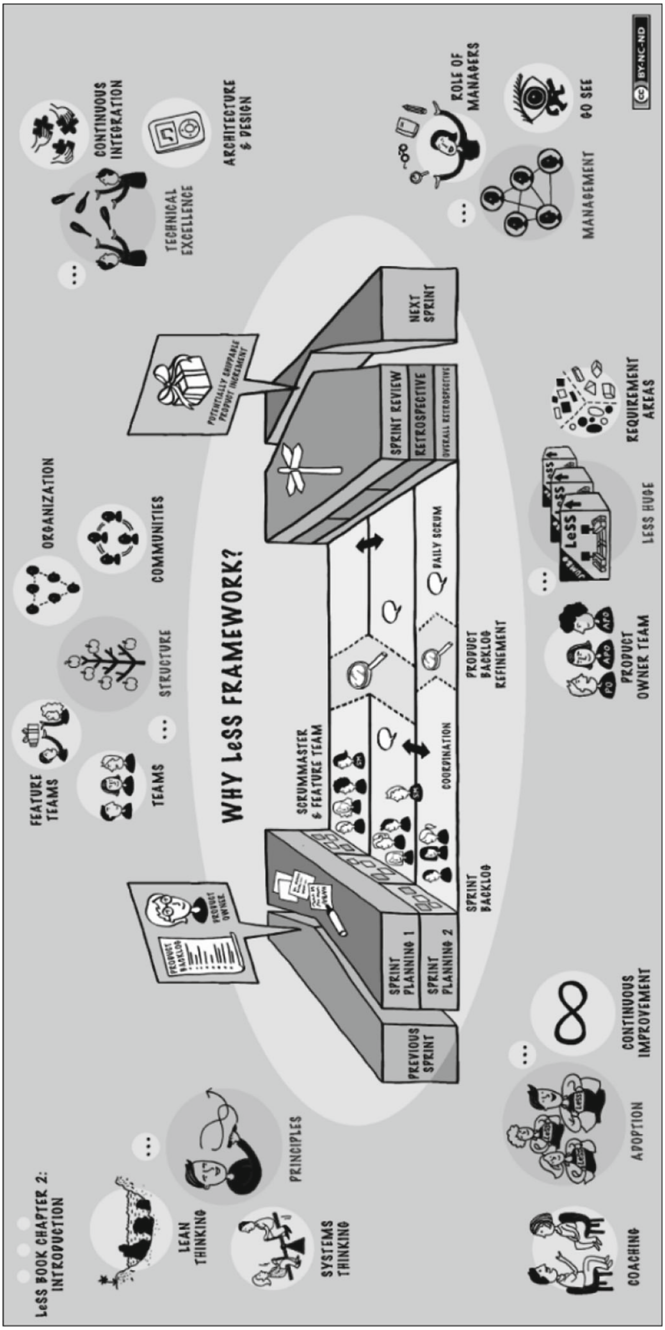


Fig. 3 "LeSS framework" [9]

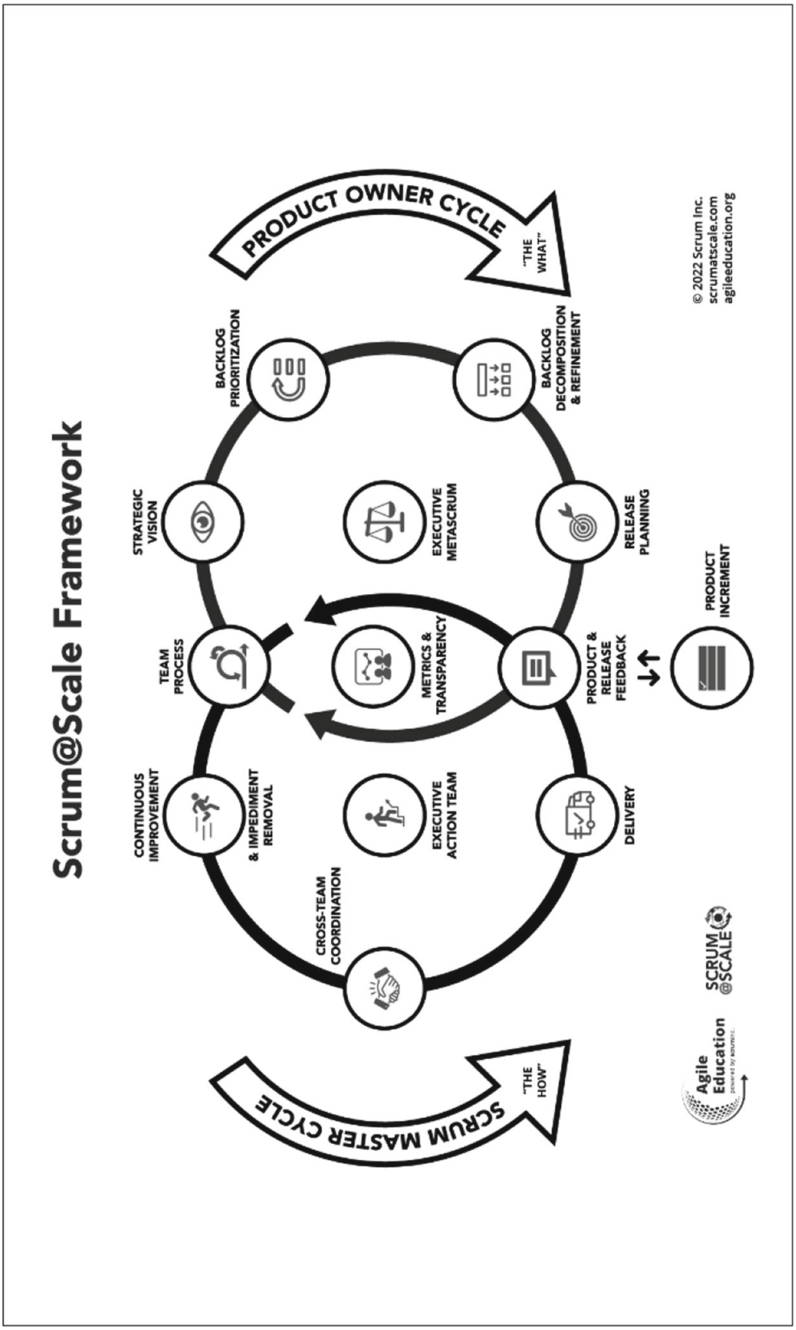


Fig. 4 "Scrum at scale framework" [8]

from the team level to upper management, needs to be persuaded of the benefits of scaling frameworks.

Experiencing a culture shift—[11] Teams are restricted to working in silos while adhering to a hierarchical structure and being steadfastly committed to following the plan made by the higher management in a traditional management style. A new and different working culture needs to be adopted for scaling. It does away with the notion of a top-down hierarchy, eliminates silos, and improves communication and collaboration between the teams and higher management. It holds the opinion that not everything has to go as planned because requirements and demands can change.

Insufficient/incorrect understanding of the framework—Many people believe that those in charge of the transformation are not fully knowledgeable about the finer points [1]. Typically, one of the current leadership members completes a course, reads a book about how the transition should go, and then chooses to lead the way. Although self-learning is a good idea, many businesses do it to save money by not using outside assistance.

Creating a new mentality—[11] Having a lean mindset is a key requirement when adjusting to any scaling model. And it is not for one person or one team but the entire organization needs it; rather, it must be gradually and steadily adopted by everyone in it. The idea of servant leadership, in which the different priorities of the teams are maintained first, needs to be taught to leaders. Teams should be given the freedom to make their own decisions and should be taught value and expectation of accountability.

Not using the right tools consistently across the enterprise—Silos are a deal-breaker for scaling. Every department typically uses a different set of tools in enterprises. People need time to get used to new tools, and they have grown accustomed to their current tool to a certain extent. It results in disconnect enterprise which is not ready for scaling.

3 Methodology

We looked at the research data (collected through surveys) [5] to establish the benefits, challenges, and the popular scaled framework.

Estimation of % successes with scaled agile approaches for executed project/ development processes can be seen in Table 1.

Improvements in results and efficiency through using scaled agile approaches can be seen in Table 2.

The benefits of different scaled frameworks can be seen in Table 3.

Table 1 Success rate

Success	5%	1%	1%	1%	2%	5%	12%	10%	26%	20%	17%
% Bracket	Not specified	00–09	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–100

Table 2 Response rate

% Respondents	74%	11%	15%
Response	Yes	No	Not specified

Table 3 Benefits

Attributes	Nexus (%)	Scrum@Scale (%)	SAFe	LeSS (%)
Overall capability of the approach	84	82	70	66
Quality of result	90	84	74	82
Cooperation across teams	94	76	76	70
Teamwork	94	94	78	82
Planning security	72	64	64	56
Efficiency	84	72	48	48
Velocity	72	58	46	56
Ability to innovate	46	72	44	64
Customer satisfaction	90	82	64	70
Accuracy of evaluating improvements	68	58	70	58

The following ratings consist of the sum of the ratings of the frameworks with “very good” and “good”
The following responses were multi-selection-based

Challenges or Impediments for adopting scaled frameworks can be seen in Table 4.

4 Result Analysis

As per the available data through survey, *Nexus is good/very good at quality of result, cooperation across teams, and teamwork where respondents rated these > 90%. Whereas it is poor at ability to innovate.*

Scrum at scale is good/very good at quality of result and teamwork where respondents rated these > 90%. Whereas it is relatively poor at velocity and accuracy of evaluating improvements.

Scaled Agile is good/very good at quality of result, cooperation across teams, and teamwork where respondents rated these > 74%. Whereas it is relatively poor at efficiency, velocity, and ability to innovate.

Large-Scale scrum is good/very good at quality of result and teamwork where respondents rated these > 80%. Whereas it is poor at efficiency.

The top-most challenge in adoption is that enterprises are not aware of scaling framework or have no expertise with 39% of the respondents calling it out.

Table 4 Challenges

Challenges/Impediments	Respondents (%)
No expertise and neither any scaling framework	39
Still busy with implementing agile on a team basis	30
We plan to indulge in this topic	18
We don't need a scaling framework	14
The change overwhelms our management	13
We couldn't find adequate scaling approaches	10
No acceptance/need from management	9
No foreseeable usage	9
Scaled agile approaches have too high requirements for employers/partners/user	7
The change overwhelms our employees	7
The effort is too high	6
Company/groups framework for the approach	5
We tried scaling approaches but have had bad experiences	2
We don't believe agile approaches can work scaled	2
Other/not specified	12

The following responses were multi-selection-based

63% of respondents estimate the success rate to be 70% for the projects carried out with scaling frameworks. And only 10% of the participants state that the success rate of projects with scaling frameworks is less than 50%. 74% of the study respondents state that improvements have been achieved using scaling frameworks. So, while there are challenges, but use of scaling frameworks increases the success of the projects multifold and is better in achieving improvements.

5 Conclusion

The volatile, disruptive market environment of today makes traditional measures of success, process, and frameworks insufficient. Enterprises have been forced to re-evaluate how they carry out their strategic initiatives because of the reliance on mobile communications and staying constantly connected, which has created entirely new opportunities for providing goods and services and engaging with customers.

62% of respondents claimed internal processes and 59% said top management was the biggest challenge for agile transformation [5].

Enterprises need to conduct assessments of best fit scaling framework before they can start applying the practices.

When enterprises need to scale? Before scaling, few essential requirements need to be considered.

- Are you already executing projects on agile? *If the team is not experienced in agile, then scaling can be a disaster.*
- Can the development be managed by a single team? *If the workload for your agile team exceeds its capacity, then you should consider scaling.*
- Do the development methods used by your team require improvement? *For consistency across teams, procedures like proper DevOps implementation and a CI/CD pipeline are essential. Also, it might be challenging to test new features without standardized quality assurance.*
- Do you deliver integrated increments as a team? *If you have multiple teams integrated to work together on it in tandem, producing features that complement one another then you should consider one of the scaling frameworks.*

But more importantly are your people willing to change and adapt and do you have the buy-in from your top management.

6 Significance of the Study

This paper focusses on empirical analysis of scaled framework in general and also specific benefits in adopting any on the one well-known framework.

- Scaling comes with ability of enterprises to change across different levels within their management. Enterprises should assess the benefits and work on creating a culture of change for better adoption.
- Authors of the present paper have represented the data in a way where it helps to see the bigger picture through comparative assessment to do best fit analysis. Future study on business agility with predictive approach can leverage this work.

7 Future Scope

Agile now has and should start to address the entire process and organization, not just product development. Everything must function as a complete, flexible, and agile organism, including finance, legal, and HR where enterprises adopt and truly experience scaling framework across. True business agility goes beyond just the production and operational aspects of the business. It has to do with everything, including strategy, management, people, infrastructure, decision-making, governance, and culture. Systems and solutions that aid in the increased agility of businesses can be built on tools and technology powered by AI/ML. However, there are mutually beneficial synergies between Scaling Agile and AI. While AI has the potential to impact agile teams significantly and positively. Automation using AI/ML frameworks and tools & technologies can significantly speed up the testing and improvement of new iterations of products and procedures. AI/ML can provide the most important component for the agile transformation which is data. Whether it is assessing trends through analytics or providing real-time analysis responses. AI-powered systems and solutions can not only gather data more quickly and effectively, but they can also organize and process it in ways that make it more transparent and increase the clarity of any conclusions that may be drawn from it—more importantly it can help predict the business metrics.

References

1. Uludağ Ö, Putta A, Paasivaara M, Matthes F (2021) Evolution of the agile scaling frameworks. In: International conference on agile software development. Springer International Publishing, Cham, pp 123–139. https://doi.org/10.1007/978-3-030-78098-2_8
2. Paasivaara M, Behm B, Lassenius C, Hallikainen M (2018) Large-scale agile transformation at Ericsson: a case study. Empir Softw Eng 23:2550–2596. https://doi.org/10.1007/s10664-017-9555-8
3. Dingsøyr T, Fægri TE, Itkonen J (2014) What is large in large-scale? A taxonomy of scale for agile software development. In: Product-focused software process improvement: 15th international conference, PROFES 2014, Helsinki, Finland, December 10–12, 2014. Proceedings 15. Springer International Publishing, pp 273–276. https://doi.org/10.1007/978-3-319-13835-0_20
4. Dikert K, Paasivaara M, Lassenius C (2016) Challenges and success factors for large-scale agile transformations: a systematic literature review. J Syst Softw 119:87–108. https://doi.org/10.1016/j.jss.2016.06.013
5. Komus, Ayelt et al. Study Status Quo (scaled) agile 2019/20. https://scrumorg-website-prod.s3.amazonaws.com/drupal/2020-02/Result%20Report-SQA-public-v1.0%20%281%29.pdf. Accessed 9 Aug 2023
6. 16th Annual State of Agile Report (2022). https://info.digital.ai/rs/981-LQX-968/images/AR-SA-2022-16th-Annual-State-Of-Agile-Report.pdf. Accessed 9 Aug 2023
7. SAFe 5.1 https://www.scaledagileframework.com/. Accessed 9 Feb 2023
8. The Scrum at Scale® Guide https://www.scrumatscale.com/scrum-at-scale-guide-online/. Accessed 9 Feb 2023
9. Large Scale Scrum. https://less.works/. Accessed 9 Feb 2023
10. Overview of the Nexus framework for scaling scrum https://www.scrum.org/resources/blog/overview-nexus-framework-scaling-scrum. Accessed 9 Feb 2023
11. Conboy K, Carroll N (2019) Implementing large-scale agile frameworks: challenges and recommendations. IEEE Softw 36(2):44–50. https://doi.org/10.1109/MS.2018.2884865
12. Putta A, Paasivaara M, Lassenius C (2019) How are agile release trains formed in practice? A case study in a large financial corporation. In: Kruchten P, Fraser S, Coallier F (eds) XP 2019. LNBIP, vol 355. Springer, Cham, pp 154–170. https://doi.org/10.1007/978-3-030-19034-7_10

Industrial IoT Predictive Maintenance Using Machine Learning Approach

B. Pradeep, Balusupati Anil Kumar, Pirangi Vijay Kumar, Talari Swapna, Rajendhar Reddy Gaddam, and S. Dhanalakshmi

Abstract The integration of industrial Internet of things (IIoT) technology with predictive maintenance methodologies has revolutionized the way maintenance operations are managed in industrial settings. This paper presents a comprehensive overview of predictive maintenance in the context of IIoT, focusing on the application of machine learning techniques for efficient and proactive maintenance strategies. The key components of IIoT predictive maintenance systems, including data acquisition, preprocessing, feature engineering, model selection, and deployment, are discussed in detail. Various machine learning algorithms commonly employed for predictive maintenance, such as support vector machines, random forests, neural networks, and deep learning models, are reviewed along with their strengths and limitations in different industrial scenarios. The challenges related to data quality, scalability, interpretability, and cybersecurity are addressed, and potential solutions are proposed. Case studies highlighting successful implementations of IIoT predictive maintenance solutions across different industries are presented to demonstrate the tangible benefits, including improved equipment uptime, reduced maintenance costs, and enhanced operational efficiency.

Keywords Industrial IoT · Predictive maintenance · Machine learning

B. Pradeep (✉) · B. A. Kumar · P. V. Kumar · T. Swapna · R. R. Gaddam · S. Dhanalakshmi
CMR Institute of Technology, Hyderabad, Telangana 501401, India
e-mail: pradeepb26@cmritonline.ac.in

B. A. Kumar
e-mail: anilkumarb@cmritonline.ac.in

P. V. Kumar
e-mail: vijaykumar@cmritonline.ac.in

T. Swapna
e-mail: swapnat@cmritonline.ac.in

R. R. Gaddam
e-mail: rajendarreddy.gaddam@cmritonline.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_78

1 Introduction

An era of enhanced efficiency and reliability has arrived in today's dynamic industrial environment because of the adoption of Internet of things technology [1], which has brought about a paradigm change in maintenance procedures. This paradigm change is exemplified by the concept of industrial Internet of things predictive maintenance, which use machine learning algorithms and advanced analytics to forecast when pieces of machinery will fail, hence minimizing unscheduled downtime and maximizing output. This paper explores the transformative potential of industrial IoT predictive maintenance through a comprehensive examination of its underlying principles, methodologies, and real-world applications [2]. By leveraging vast amounts of sensor data generated by interconnected industrial equipment, machine learning algorithms can discern patterns indicative of impending failures, enabling proactive intervention to prevent costly disruptions in production.

Key components of this approach include data collection through IoT sensors, data preprocessing to ensure quality and relevance, feature engineering to extract meaningful insights, and the training of machine learning models to accurately predict maintenance needs.

A. Industrial Internet of things

There is a transformational technology framework known as the industrial Internet of things (IIoT), which links physical machinery and equipment with enhanced digital capabilities in order to improve industrial operations across a variety of industries. Unlike traditional industrial systems, which operated in silos with limited connectivity and data exchange, IIoT facilitates seamless communication and collaboration among diverse industrial components through interconnected sensors, actuators, and other smart devices.

At its core, IIoT leverages the power of networked devices to collect, analyse, and utilize vast amounts of real-time data generated by industrial equipment. This data enables organizations to gain deeper insights into their operational processes, optimize performance, and drive informed decision-making. Key features of IIoT include remote monitoring and control, predictive maintenance, asset tracking, and supply chain optimization [3].

The adoption of IIoT has significant implications for industries ranging from manufacturing and energy to transportation and health care. By leveraging data-driven insights, IIoT enables organizations to enhance efficiency, improve safety, reduce downtime, and drive innovation. Additionally, IIoT facilitates the transition towards autonomous systems and smart factories, where interconnected devices collaborate in real time to optimize production processes and resource utilization.

B. Importance of predictive maintenance in IIoT

Predictive maintenance holds immense importance in the realm of IIoT due to several key reasons. Firstly, it enables cost reduction by identifying potential equipment failures before they occur, allowing for timely repairs or replacements. This approach reduces unplanned downtime, minimizes production losses, and avoids costly emergency repairs, ultimately leading to significant cost savings for industrial organizations.

Then, predictive maintenance increases equipment reliability by continuously monitoring equipment condition and performance metrics in real time through IIoT sensors. By proactively addressing issues before they escalate, this approach enhances equipment reliability and longevity, maximizing asset utilization and return on investment.

Predictive maintenance improves safety by enabling early detection of safety hazards and malfunctions, preventing accidents and ensuring a safer working environment for industrial personnel. Timely intervention based on predictive insights mitigates risks associated with equipment failures, safeguarding both employees and assets.

IIoT-driven predictive maintenance enhances efficiency by optimizing maintenance schedules and resource allocation based on equipment health and performance data. This targeted approach minimizes unnecessary maintenance activities and reduces downtime, allowing for more efficient utilization of labour and resources. Predictive maintenance facilitates data-driven decision-making by leveraging advanced analytics and machine learning algorithms to analyse vast amounts of sensor data collected by IIoT devices [4]. By extracting actionable insights from this data, industrial organizations can make informed decisions regarding maintenance strategies, resource allocation, and process optimization, leading to improved operational efficiency and competitiveness.

2 Literature Review

Industrial IoT and Predictive Maintenance

Internet of things (IoT) solutions in industrial environments could lead to innovative and efficient systems aiming at increasing operational efficiency in a new generation of smart factories [5]. Specifically, the article provided an overview of the NGS-PlantOne system, which is a sophisticated solution for the industrial Internet of things (IIoT). This application was designed to enable the ubiquitous monitoring of industrial equipment via the use of IoT sensing devices that were powered by batteries. Within the parameters of the situation at hand, this would provide the groundwork for the development of complex programmes for predictive maintenance. For the purpose of evaluating the effectiveness of the newly developed IIoT system in a real-world setting, the NGS-PlantOne solution was installed and put into operation in a real-world electrical power plant. With the help of the 33 Internet of things (IoT) sensor devices that were used in the testbed, vibration and temperature were tracked for a period of two months. Additionally, gearbox delays and system working life were measured by evaluating power consumption. The performance findings showed that the IIoT solution made full use of the protocols that are used by the Internet of things. This ensured that every node could be accessed by using fundamental IP-based approaches with a tolerable amount of delay. In addition, the solution displayed an expected average life of one year, supposing that every Internet of things smart device is configured to communicate data that has been gathered and expounded on every thirty minutes [5].

According to [6], the IIoT has recently enabled several businesses to make substantial advancements by using various technologies, such as AI and big data processing. Among the many IIoT scenarios, large-scale data centres stand to gain the most from using AI-boosted strategies and big data analytics, which might pave the way for more efficient predictive maintenance. But most of the prefabricated solutions out there were not cut out for HPC. For example, they did not account for the highly diverse data sources or the privacy concerns that stopped people from using cloud computing. Or they did not make the most of the processing power that was already present at a supercomputer. They discussed this problem and offered a vertical and comprehensive IIoT framework for predictive maintenance in supercomputers in this paper.

According to [7], predictive maintenance (PdM) is becoming more important in the manufacturing and production sectors as a result of the Industry 4.0 paradigm. Intelligent perception in PdM was made possible by the digitization of physical activities made possible by the Internet of things (IoT), which allowed for connections between humans, machines, and other machines. Before PdM could be considered mature enough for broad use, a number of difficulties remained. Supporting decision-making to best act on physical systems was the ultimate goal of digitalization, and to achieve it, the Internet of things (IoT) needed data science skills. In order to help readers better grasp the benefits and drawbacks, obstacles and possibilities, of this ever-changing maintenance paradigm, the authors of this piece provide a thorough overview of the existing PdM concerns. A great deal of reading and research in the technical and scientific literature went into this. Using this as a starting point, the article detailed several key research challenges that need to be resolved in order for the industry to successfully develop and implement PdM that is enabled by the Internet of things.

According to [8], the industrial sector has been rapidly transformed by the advent of the Fourth Industrial Revolution, which is often referred to by terms like industry, smart factories, the Internet of things (IoT), cyber-physical systems, and digital transformation. Among the activities encompassed by Industry 4.0 were the development of novel goods and services, the digitalization of both vertical and horizontal value chains, and the introduction of novel business models. Customer satisfaction, marketing velocity, and cost reduction were three key operational drivers of the change. An integral part of the smart factory, which aimed to minimize downtime and maximize availability of production facilities, predictive maintenance was discussed in this study.

According to [9], data-driven predictive maintenance (PdM) became more important for companies looking to gain a competitive advantage in IIoT due to the unplanned breakdown of critical equipment that interrupted production throughput.

However, in the case of an unforeseen machine failure, manufacturers were continually confronted with the tedious task of manually assigning appropriately skilled human resources. Also, overall equipment downtime and production timelines were both negatively affected by human error. To find out if a model-free PdM framework based on deep reinforcement learning (DRL) could automatically learn an optimal decision policy from a stochastic environment, the authors of this research reframed the complicated resource management problem as a resource optimization problem. Their method differed from previous PdM frameworks in that it integrated human and physical resources with data collected by PdM sensors into the optimization issue. Using a maintenance repair simulator, human participants' baseline findings were compared to those of the proposed DRL-based framework and PPO-LSTM model. Their PPO-LSTM outperformed human participants by 65% and equivalent DRL approaches by 53% when it came to learning the best decision policy for the resource management problem, according to the empirical data. In general, the simulation results confirmed that the suggested DRL-based PdM framework was the best in terms of simulation performance, flexibility, and convergence efficiency.

Machine Learning Techniques for Predictive Maintenance

According to [10], a data-driven predictive maintenance system was created for industrial production lines. Using machine learning techniques applied to real-time data collected from Internet of things (IoT) sensors, the system sought to identify warning signs of impending failure. So, it assisted with problem solving by alerting operators in advance so they could take precautions before production halted. The present research additionally evaluated the system's efficacy by analysing data from actual industrial IoT systems. The assessment findings showed that the predictive maintenance system might assist avoid production stoppage by spotting signs of possible breakdowns. Results from comparison tests of several machine learning algorithms showed that XGBoost, a boosting approach, and random forest, an ensemble strategy for bagging data, seemed to have the best performance. The study's top-performing machine learning models were already part of the factory's production system.

According to [11], predictive maintenance is one of the ideas that emerged alongside this new revolution. The original intent of this research was to look at recent scholarly work on failure prediction. Concepts such as a design support system and a predictive maintenance decision support system were considered in the failure prediction process. They had zeroed in on Industry 4.0 predictive maintenance frameworks that made use of logic and machine learning. With predictive maintenance in mind, they had especially thought about the difficulties of using machine learning methods and ontologies. Academic publications published online between 2015 and the start of June 2020 were analysed in a systematic literature review (SLR). Out of 562 studies, 38 were retained for further analysis after the screening procedure. Papers deemed as surveys or reviews, as well as those unrelated to predictive maintenance or machine learning, were culled. While debating the articles' suggestions and findings, they took into account three research issues. In shedding light on the implementation and use-case problems, this essay made a valuable contribution to predictive maintenance. In their last remarks, they brought attention to the fact that predictive maintenance was a popular subject within the Industry 4.0 framework, but that there were still many unanswered questions about the use of reasoning and machine learning in this field.

A big problem in the car business is making sure products are functionally safe throughout their lifetimes while keeping maintenance expenses to a minimum. Predictive maintenance (PdM) was an important strategy for reaching this goal. An perfect choice for PdM would have been ML given the massive amounts of operational data that contemporary vehicles have brought along. There have been several reviews of both PdM and ML for automotive systems, but none that focused on ML-based PdM in the last several years. This study was necessary since there has been a recent uptick in the number of articles published in this area. This led them to conduct a literature review, classify the articles, and evaluate them from both an application and ML standpoint. They then proceeded to explore potential research avenues and open problems. They came to the following conclusions: (a) more research would be done with publicly available data, (b) most papers used supervised methods, which need labelled data, (c) accuracy could be improved by combining data from multiple sources, and (d) deep learning methods would be used more, but only if there was large amounts of labelled data and efficient, interpretable methods [12].

Facility managers often employ reactive or preventative maintenance tactics in building maintenance management, as pointed out by [13]. On the other hand, reactive maintenance could not stop failures from happening, and preventive maintenance could not foresee the MEP components' future conditions and fix them to make facilities last longer. To get over these restrictions, their research had planned to use a predictive maintenance approach with cutting-edge electronics. The integration of BIM and the IoT offered the chance to make facility maintenance management (FMM) more efficient [13].

Reference [14] noted that the introduction of new technology into contemporary smart factories has led to a correlation between production robotization and automated predictive maintenance. The advent of intelligent sensors has led to a deluge of data, which, in order to facilitate the administration and decision-making of ever-more-complex systems, requires careful and efficient analysis. The original intent of the article was to survey recent works on smart factory predictive maintenance and intelligent sensor applications. In order to give a synopsis of future research difficulties and categorization, they had concentrated on current trends. The study employed clustering, burst analysis, keyword cooccurrence analysis, and systematic review technique. Key investigated themes have been the subject of a rising number of articles, according to the results. As time went on, predictive maintenance's significance in regard to Industry 4.0 technologies grew. By analysing the whole texts of pertinent publications, they came up with the idea of smart and intelligent predictive maintenance (SIPM). The key contribution of the study was a synopsis and review of recent developments in smart sensors for smart factory predictive maintenance.

Reference architectures, important technologies, related applications, and future problems were all covered in [15] extensive study of the industrial Internet's recent advancements. The features of reference architectures that have been suggested for various application situations have been compiled. Cloud computing, mobile edge computing, and fog computing are some of the key technologies that have been introduced to serve a variety of applications in the industrial Internet. These technologies have been classified according to different layers in the architecture. In the meanwhile, we had talked about future research trends and difficulties in the industrial Internet to encourage more study in this area [15].

Challenges and Limitations in Predictive Maintenance in Industrial IoT

According to [16], industrial Internet of things (IIoT) is being developed as a result of the incorporation of new technical trends and IoT applications into industrial systems. By automating smart devices to sense, collect, process, and communicate the real-time events in industrial systems, IIoT provided a new vision of the Internet of things (IoT) for the industrial sector. The main goal of the industrial Internet of things (IIoT) was to improve the management of industrial assets and processes, achieve high operational efficiency, and boost productivity through intelligent monitoring applications for production floor shops and machine health, product customization, and predictive and preventive maintenance of industrial equipment. They provided a fresh and straightforward description of IIoT in that study, which may aid readers in grasping the idea. There was an account of the most recent cutting-edge IIoT research. Lastly, they emphasized the technologies that facilitate IIoT and the current obstacles that the IIoT has encountered.

According to [17], a new paradigm known as the Internet of things (IoT) was born out of the possibility of building smart environments made possible by the merging of the physical and cyber worlds. There was now a way for items to communicate with one other in addition to with people. As a subset of the Internet of things (IoT), industrial Internet of things (IIoT) capitalized on IoT communications for commercial applications that prioritized machine interoperability. The Internet of things (IoT) is the ever-changing web of interconnected devices, services, and products that people use every day. The results produced by IoT services were inconsistent, imprecise, incomplete, and wrong due to factors including heterogeneity, velocity, and amount of data. These issues were particularly relevant for many applications in IIoT, such as health care, smart transportation, wearables, finance, and industrial. As a result of data discovery, search, and sharing, forty percent of the advantages of the Internet of things (IoT) have spread to nearly every industrial application. Data fusion, machine learning, big data, cloud computing, blockchain, etc., are all examples of ICT-based search approaches that might help govern and limit the Internet of things (IoT) while also maximizing its value. The research provided a thorough analysis of current issues and suggested solutions for future Internet of things (IoT) data analysis and search, as well as a framework for integrating information and communication technologies (ICT) into IoT layers. In order to demonstrate the potential for improved intelligence and smartness in IoT applications as a result of ICT integration, that article reviewed existing IoT search engines (IoTSEs) and offered two case examples.

In [18] as it streamlined processes by linking various devices and made use of sensors. It was responsible for the functionality of a range of smart devices and established a platform on top of it. Internet of things (IoT) installations allowed devices to communicate with each other without the need for human or computer involvement; this led to its widespread adoption across industries. The article covered a lot of ground, including the evolution of the Internet of things (IoT) into the industrial Internet of things (IIoT), the definition of Industry 4.0, the main distinctions between the two, the many industries that have found uses for the IoT, its sophisticated features and capabilities, and its many applications. Internet of things (IoT) technologies were in high demand across industries due to the convenience they brought to people's lives and the fact that they were easy to use, but they were not without their limits. In the paper's latter sections, the authors discussed potential security risks, problems, and solutions associated with IoT-based systems, as well as methods for overcoming these constraints.

In [19] noted that the IIoTs were the next step towards the Fourth Industrial Revolution since they were extensions of the IoTs. With the help of IIoT, the transportation, manufacturing, and marketing sectors were able to speed up the automation of their internal and external working processes through the use of numerous linked devices. Because it is an expansion of the Internet of things, IIoT carries with it all of the technology's vulnerabilities. However, in order to tailor the current security solutions to its sensor-configured architecture, further work was required. The upgraded innovations' scope remained vulnerable to undetected assaults, even after security model reconstructions. By categorizing and comparing different assaults on each IIoT layer, the current study aided researchers in comprehending the source of penetration. Examining the numerous security challenges encountered by IIoT and comparing the current solutions to strengthen the protection systems of industrial IoTs were the primary goals of that survey. Additionally, the study highlighted some unanswered questions that researchers, academics, and technologists need to address in order to advance the IIoT field and its security features.

According to [20], one of the most important areas of study that emerged from the IoT was the industrial Internet of things (IIoT). IIoT optimized services and processes to decrease costs and increase productivity by connecting all kinds of industrial equipment over the network, establishing systems for data collecting, exchange, and analysis, and optimizing these systems and processes. A number of issues, including reduced decision-making latency, resource savings in bandwidth, and privacy protection, might be addressed with the use of edge computing in IIoT. The study provided an overview of the current state of the art in IIoT edge computing research. After introducing the ideas of IIoT and edge computing, we went on to describe and debate the current state of edge computing research. The next step was to provide a future architecture for IIoT edge computing and then to examine its security, job scheduling, data storage and analytics, routing, and standardization efforts. Moreover, they went over the pros and cons of edge computing in IIoT regarding 5G-based edge communication, data offloading and load balancing, edge intelligence, and safeguarding data sharing. Lastly, they detailed some common IIoT edge computing use cases, including smart grids, smart logistics, intelligent connected cars (ICV), factory coordination, and prognostics and health management (PHM).

3 Methodology

Developing a mathematical model for industrial IoT (IIoT) predictive maintenance using machine learning involves various components such as data preprocessing, feature engineering, model training, and evaluation. While equations can represent some aspects of this process, it is essential to note that the overall implementation involves a combination of mathematical concepts, statistical methods, and machine learning algorithms involved in building such a model:

Data Preprocessing:

Data Cleaning:

$X_{\text{clean}} = f(X_{\text{raw}})$ where X_{clean} is the cleaned dataset, and f represents the cleaning function.

Data Normalization/Standardization:

$X_{\text{norm}} = \frac{X-\mu}{\sigma}$ where X_{norm} is the normalized dataset, μ is the mean, and σ is the standard deviation.

Data Transformation:

$X_{\text{transformed}} = g(X_{\text{norm}})$ where g represents the transformation function.

Feature Engineering:

Extract relevant features from the dataset and engineering new features if necessary:

$\text{Feature}_{\text{new}} = h(\text{Feature1}, \text{Feature2}, \ldots)$, where h is the feature engineering function.

Model Training:

Regression model:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

where y is the predicted output, X_i are the features, β_i are the coefficients, and $\epsilon$is the error term.

Classification model:

Logistic Regression:

$P(y = 1|X) = \frac{1}{1+e^{-z}}$where $z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$

Evaluation Metrics:

Regression:

Mean Squared Error (MSE):

$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

R-squared (coefficient of determination):

$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2}$

Classification:

$$\text{Accuracy}: \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Precision, recall, F1-score, ROC-AUC, etc.

Model Deployment:

Deploy the trained model to the IIoT system for real-time predictions.

Continuous monitoring and retraining if necessary.

Additional Considerations:

Hyperparameter tuning equations for optimizing model performance.

Cross-validation techniques for robust model evaluation.

These equations represent a simplified overview of the mathematical concepts involved in developing an IIoT predictive maintenance model using machine learning. In practice, the implementation involves a combination of these equations along with data-specific considerations and algorithmic choices.

4 Conclusion

This paper has provided a comprehensive overview of industrial Internet of things (IIoT) predictive maintenance leveraging machine learning (ML) techniques. Through an exhaustive analysis of recent research and advancements in this field, several key findings have emerged. Firstly, the integration of IIoT devices with ML algorithms has revolutionized maintenance strategies in industrial settings, enabling proactive and data-driven approaches to equipment monitoring and fault prediction. Secondly, the effectiveness of various ML models such as neural networks, support vector machines, and decision trees in predicting equipment failures has been demonstrated across diverse industrial sectors. Additionally, the importance of data quality, feature selection, and model optimization has been highlighted to enhance the accuracy and reliability of predictive maintenance systems. Moreover, the implementation challenges including data integration, cybersecurity, and scalability have been identified, emphasizing the need for robust infrastructure and interdisciplinary collaboration. Overall, this review underscores the transformative potential of ML-enabled predictive maintenance in optimizing asset management, minimizing downtime, and maximizing operational efficiency in industrial environments. Future research directions should focus on addressing the identified challenges and exploring innovative ML techniques to further advance predictive maintenance capabilities in the realm of industrial IoT.

References

1. Hsu CW, Yeh CC (2017) Understanding the factors affecting the adoption of the Internet of Things. Technol Anal Strateg Manage 29(9):1089–1102
2. Zonta T, Da Costa CA, da Rosa Righi R, de Lima MJ, da Trindade ES, Li GP (2020). Predictive maintenance in the Industry 4.0: a systematic literature review. Computers & Industrial Engineering, 150, 106889
3. Javaid M, Haleem A, Singh RP, Rab S, Suman R (2021) Upgrading the manufacturing sector via applications of industrial internet of things (IIoT). Sens Int 2:100129
4. Kan C, Yang H, Kumara S (2018) Parallel computing and network analytics for fast Industrial Internet-of-Things (IIoT) machine information processing and condition monitoring. J Manuf Syst 46:282–293
5. Civerchia F, Bocchino S, Salvadori C, Rossi E, Maggiani L, Petracca M (2017) Industrial Internet of Things monitoring solution for advanced predictive maintenance applications. J Ind Inf Integr 7:4–12
6. Borghesi A, Burrello A, Bartolini A (2021) Examon-x: a predictive maintenance framework for automatic monitoring in industrial iot systems. IEEE Internet Things J
7. Compare M, Baraldi P, Zio E (2019) Challenges to IoT-enabled predictive maintenance for industry 4.0. IEEE Internet of Things Journal, 7(5), 4585–4597.
8. Chehri A, Jeon G (2019) The industrial internet of things: examining how the IIoT will improve the predictive maintenance. In: Innovation in medicine and healthcare systems, and multimedia: proceedings of KES-InMed-19 and KES-IIMSS-19 Conferences. Springer Singapore, pp 517–527
9. Ong KSH, Wang W, Niyato D, Friedrichs T (2021) Deep-reinforcement-learning-based predictive maintenance model for effective resource management in industrial IoT. IEEE Internet Things J 9(7):5173–5188
10. Ayvaz S, Alpay K (2021) Predictive maintenance system for production lines in manufacturing: a machine learning approach using IoT data in real-time. Expert Syst Appl 173:114598
11. Dalzochio J, Kunst R, Pignaton E, Binotto A, Sanyal S, Favilla J, Barbosa J (2020) Machine learning and reasoning for predictive maintenance in Industry 4.0: current status and challenges. Comput Ind 123:103298
12. Theissler A, Pérez-Velázquez J, Kettelgerdes M, Elger G (2021) Predictive maintenance enabled by machine learning: use cases and challenges in the automotive industry. Reliab Eng Syst Saf 215:107864
13. Cheng JC, Chen W, Chen K, Wang Q (2020) Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. Autom Constr 112:103087
14. Pech M, Vrchota J, Bednář J (2021) Predictive maintenance and intelligent sensors in smart factory. Sensors 21(4):1470
15. Qin W, Chen S, Peng M (2020) Recent advances in Industrial Internet: insights and challenges. Digit Commun Netw 6(1):1–13
16. Khan WZ, Rehman MH, Zangoti HM, Afzal MK, Armi N, Salah K (2020) Industrial internet of things: recent advances, enabling technologies and open challenges. Comput Electr Eng 81:106522
17. Younan M, Houssein EH, Elhoseny M, Ali AA (2020) Challenges and recommended technologies for the industrial internet of things: a comprehensive review. Measurement 151:107198
18. Jaidka H, Sharma N, Singh R (2020) Evolution of iot to iiot: applications & challenges. In: Proceedings of the international conference on innovative computing & communications (ICICC)
19. Jayalaxmi P, Saha R, Kumar G, Kumar N, Kim TH (2021) A taxonomy of security issues in Industrial Internet-of-Things: scoping review for existing solutions, future implications, and research challenges. IEEE Access 9:25344–25359

20. Qiu T, Chi J, Zhou X, Ning Z, Atiquzzaman M, Wu DO (2020) Edge computing in industrial internet of things: architecture, advances and challenges. IEEE Commun Surv Tutorials 22(4):2462–2488

SuperPCA-Based Machine Learning System by Hyperspectral Image Classification Assessment Using SVM

V. A. Narayana, R. Venkateswara Reddy, V. Venkataiah, K. Srujan Raju, Mohd. Abdul Naqi, and Syeda Sumaiya Afreen

Abstract The satellite pictures with hyperspectral imagery (HSI) will be worked on more so that they can be used in any way. In order to get the map's coordinates from picture coordinates, the most important and necessary step is to classify hyperspectral images. In this method, the ground control points (GCPs) must be physically removed from the images collected distantly using the ground truth values. This takes time. In this study, super-pixel-based principal component analysis (SuperPCA) is offered as a way to classify the multitemporal HSI satellite data. This approach would be automatically find GCPs and cut down on the time it takes to process, which will both improve accuracy. PCA is a famous algorithm for extracting features from satellite data, although it takes a long time to do so. To get around the problems with PCA, SuperPCA has been proposed and used only for pictures with fewer features. It has not been used on satellite images. In this study, the phase angle of the SuperPCA is changed so that it can be used with satellite data to find features at six levels in multitemporal HSI satellite imagery. Support vector machine (SVM) is often worn for multi-class segmentation that is not linear. The kernel decision is what makes SVM work so well. So, fuzzy-SVM (F-RVM) is suggested for kernel choosing depending on the degree of detail and extent of the features in the satellite imagery. The outcomes are evaluating to standard methods and show that the new method works better.

V. A. Narayana · R. Venkateswara Reddy · V. Venkataiah
Department of CSE, CMR College of Engineering & Technology, Hyderabad, Telangana, India
e-mail: venkatreddyvari@cmrcet.org

V. Venkataiah
e-mail: vvenkataiah@cmrcet.org

K. Srujan Raju (✉)
Department of CSE, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: ksrujanraju@gmail.com

Mohd. A. Naqi
Department of ECE, CMR Technical Campus, Hyderabad, Telangana, India

S. S. Afreen
Department of CSE (AI&ML), CMR Technical Campus, Hyderabad, Telangana, India
e-mail: sumaiyafreen.cse@cmrtc.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_79

Keywords HSI · SuperPCA · SVM · Machine learning

1 Introduction

Different distortions can happen to remotely felt HSI data because of how the satellite moves, how the Earth moves, how the sensors are calibrated, the weather, the direction of the projection, etc. [1]. Because of these problems, the raw data that remote sensing satellites collect has a lot of errors and disturbances, which reduces the quality of the picture. Satellite pictures that come straight from faraway satellites are preprocessed to get rid of distortions and noise [2]. So, as a way to address this issue, contemporary advancements in remote sensing techniques have been employed to infrequently observe the Earth's surface with the aim of forecasting natural calamities. Change detection is also a very important way to keep track of how the world is doing. Remotely sensed data is used for most real-time uses in the military, everyday life, and other areas. Remote sensing is the process of getting information about an item, place, or area without actually touching it. Based on how the data is collected, this is put into two groups: The two types of remote sensing techniques that will be discussed in this paper are active remote sensing and silent remote sensing. Inactive remote sensing involves the utilization of sensors, commonly referred to as sensors that are passive, to detect radiation that is released or absorbed by a subject or region. Passive devices are commonly employed for the purpose of quantifying the extent of sunlight reflection. Several instances of passively remote sensing include photographic film, infrared imaging, charge-coupled devices, and radiometers, among others. Active remote sensing refers to the process of emitting power in order to scan objects and regions, followed by the amount of the quantity of radiation that is replicate or returned from the object being measured. RADAR [3], LiDAR [4], etc. are some examples. Active remote sensing is usually used to describe the method of getting pictures from space. In satellite remote sensing, the atmosphere is very important because devices look through it to see the Earth's surface. So, the effects of the environment are a big part of why the quality of images gets worse. Most of the time, digital pictures are used to show what can be seen from a distance. Photoprocessing technologies are employed to enhance collected images, so facilitating observing and rectifying any distortions, blurriness, or degradation that may have occurred. This enables the acquisition of important information from the images. There are different ways to look at an image, and which one you use relies on what you want to find out. Most of the time, algorithms for picture segmentation and classification are worn to make a thematic map. Subsequently, this data is employed in conjunction with additional details to evaluate the designated test region. The precision of data collected via satellite can be spatial precision, spectral resolution, radiometric resolution, or time resolution. 2 Spatial resolution is useful when you just want to look at a certain area. The sharpness can be anywhere between 0.6 and 4 m. The particulars are shown in the structure of color bands, and the spectral resolution changes in terms of wavelength. Most of the time, there can be anywhere from one band to hundreds

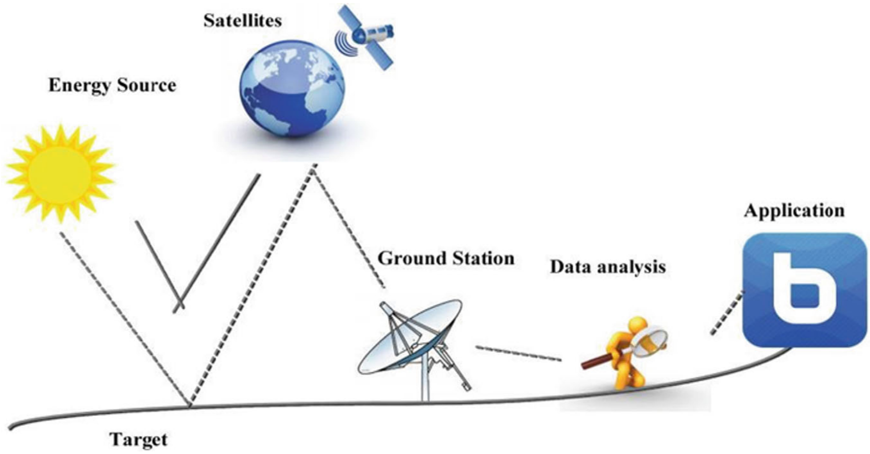


Fig. 1 Remote sensing architecture

of bands. Temporal resolution is set by how often the satellite goes back to the same spot. It can take anywhere from 24 to 16 days. Radiometric resolution changes in how much detail there is in the brightness.

HSI processing is a step up from multispectral imaging, where all of the bands' spectra are made, as shown in Fig. 1. By means of the help of only some parts like a slit, grating, photo-receptors, etc., the HSI sensor [5] turns light into electrical messages. It is the process of processing, measuring, and analyzing the spectra that the devices send back. The most common ways to get an HSI picture are through spatial scanning, spectral scanning, snapshots or not scanning at all, and spatio-spectral scans. More than a hundred spectral bands are packed into each HSI pixel. These bands hold a lot of data regarding the natural things that the satellites have photographed.

The three main components of the suggested approach are segmentation based on super-pixels, feature extraction based on recursive filtering, and mislabeled sample elimination based on spectral–spatial similarity. The results of the experiments show that by utilizing the multilabel training data, the suggested strategy can considerably increase the classification accuracy of multiple classifiers [6].

These HSI pictures are made up of hundreds of closely spaced narrow bands of data that are surrounded by a wide range of light that has been reflected. HSI sensors usually take pictures from satellites or planes. In HSI pictures, there are two kinds of redundancies: spatial redundancy and spectral redundancy [7]. The spatial redundancy shows how the pixels in a picture are connected to each other. The term "spectral redundancy" refers to how different spectral bands in multispectral pictures or color images are related to each other [8]. The idea of picture compression is used to get rid of these duplicates. So, the number of the image's bits is also decreased to make less use of memory. By compressing the image, it can be saved or sent to the database quickly and easily.

Section 2 is a thorough analysis of the problems with a number of related works. In Sect. 3, we explain in detail how the suggested method for detecting and classifying hyperspectral images using the SuperPCA approach works. In Sect. 4, you will find information about the results of the experiments and an evaluation of quantitative and quality evaluations to traditional methods and other works. Section 5 talks about how the suggested fusion approach will be used in the future.

2 Literature Review

We suggest an HSI classification method using a hybrid CNN with covariance pooling. In particular, a spatial 2-D convolution comes after spectral–spatial 3-D convolutions in our network architecture. We reduce the dimension of channels and provide complementary information by fusing the feature maps produced by 3-D convolutions along the spectral bands using this mixture operation [9]. Furthermore, to fully extract the second-order information from spectral–spatial feature maps, the covariance pooling technique is used.

One of the most crucial tasks in hyperspectral image processing is classification. Numerous classification techniques have been established in the previous few decades. When it came to effectively extracting features from hyperspectral images (HSI), most of them fell short. This research proposes a unique deep learning framework that effectively extracts both the spectral–spatial information for classification using convolutional neural networks (CNNs) and spatial pyramid pooling (SPP). SPP, 3D-CNN, 2D-CNN, and principal component analysis (PCA) are all used in the suggested hybrid architecture [10].

In [11, 12] authors suggested a multi-hypothesis prediction-based spectral–spatial preprocessing method to get rid of the noise. This improved by putting together into a hypothesis set different pixel vectors that were close to each other in space. The power of the hypothesis set to represent was improved by using a spectral-band-partitioning approach based on the correlation between bands. The predictions were made by putting together the different sets of hypotheses made in the steps that came before. The Fisher differentiated ratio was used to figure out the multi-hypothesis preprocessing. Tikhonov regularization was used to find the linear mixtures of hypotheses. When the suggested approach was put up against traditional classifiers like LDS-MLE and SVM, the accuracy of classification was found to be much better with the projected method.

In [13, 14] authors suggested automated preprocessing techniques to improve the multivariate curve resolution 33 (MCR) of the HSI pictures. For the noise to be less harmful, the suggested methods got rid of cosmic spikes, detector offsets, and structured noise. During the preprocessing step, a lens protector was used to stop light from getting into the spectral cells of the pictures. The area of the image where there was no light was built into an imaging device to help reduce structured and detector noise. The spatial regions were chosen automatically based on how the spectral data was examined.

In [15, 16] authors suggested a new preprocessing technique to get rid of noise and other things that do not belong in HSI images of vessels. Also, redundant criteria were left out of the pictures by using a continuum removal algorithm and radiometric correction. After the spectra were fixed, the method for removing the continuum was used. By using the suggested algorithms, the quality of the images was improved. The proposed preprocessing method made classification and segmentation work better.

In [17, 18] authors looked at a few of the preprocessing methods that can be used to improve the quality of HIS images by eliminating or minimizing the unneeded and extra parts in the background. The picture quality was improved by using a mean or average filter, which replaced each of the pixels with the mean value of the frequencies in the area. The average process, which made the image blurry, was a problem. When averaging was done on a picture with impulse noise, the noise was not taken away, but it was spread out. A filter with a median was a nonlinear filter that made good use of both salt and pepper noise. There were three types of median filters: the center-weighted median filter, the weighted median filter, and the max-median filter.

In [19, 20] authors showed the progress made in spectral–spatial classification by combining the spatial and spectral information of HSI images. Using the location data for object pixels made the classification process easy. Mathematical morphology was used to figure out the size, contrast, and orientation of the pictures. This was called the morphological profile. The morphological areas were made up of the extra characteristics that were needed for classification. The biggest problems with HSI classification were the spectral dimensions and the spectral–spatial classifiers [21, 22]

3 Proposed Model

In the suggested method, the specified input HSI image is first preprocessed employing the PCA technique that is worn to minimize the size of the input image so that the gradient spatial and spectral characteristics of the given input image can be improved. Then, the multi-scale entropy rate segmentation (ERS) [23] is used to separate these images based on their improved spectral and spatial gradient features. Then, deep features are taken from each divided area utilizing the SuperPCA method. By using these traits, SVM-based classification was used to find the remote sensing data in HSI images. Figure 2 shows how the suggested work would look in a picture. The operation detail of every stage is given below.

3.1 Principal Component Analysis (PCA)

PCA is used to get rid of all of the distortions in a picture that was taken from a faraway place by a remote sensing satellite. The picture is distorted because of

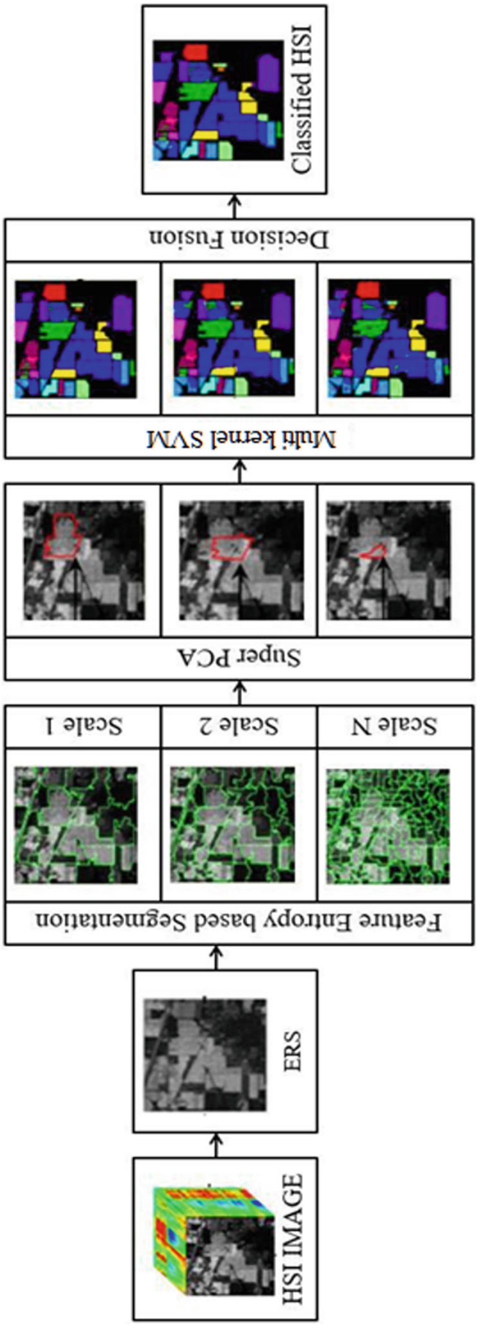


Fig. 2 Detecting and classifying photographs obtained from the hospital information system (HIS)

weather changes, light rays reflecting and refracting, and raindrops. The other things that cause distortions are how the sensor, projector line, etc., are set up. All of these problems make the picture less clear, which means that the results of any application will be less accurate. So, to make the picture clearer and get rid of any distortions, the cellular automata-based Gaussian filter algorithm is suggested. This algorithm gets rid of all fundamental noises like Gaussian white noise, speckle noise, and all other distortions caused by reflections and refractions. In early works, the basic Gaussian Filter was used for preprocessing in many kinds of spatial domain applications.

The resulting filter made the researchers get rid of the most common noises in the satellite data they got from a distance. But the biggest problem with using this kind of filter was that after the process of filtering was done, the signal was very distorted. The Gaussian filter is also used a lot in the computer vision domain, which is different from the spatial domain. Gaussian filters come in different types that can be used based on what the application needs. It can be one dimension or more than one dimension.

PCA is one of the most popular ways to reduce the number of dimensions. It is a popular way to get rid of duplicate information in data that depends on decorrelation. Even though it reduces duplication by getting rid of low-variance parts, this rotational transform takes a long time because it works on the whole. Also, because it is a global change, it does not keep local spectral signatures. As a result, it might not keep all the information needed to classify things better. Extend PCA to obtain keen on account higher-order relationships among the spectral and spatial HSI data pixels so that the super-resolution potential can be improved.

3.2 Entropy Rate Super-Pixel Segmentation (ERS)

A method called "entropy rate super-pixel segmentation" was made that falls into the group of "unsupervised segmentation techniques." This is not just built on the idea of a graph; it is a combination of a region-based method and a thresholding strategy. Geometrically, the resulting entropy rate super-pixel area is a two-dimensional plane that divides the object from the background. As an example, the transition of Eagle picture extraction as of its initial gray image. "Foreground" and "background" are revealed in black, while the transition areas are revealed in white. The areas with high entropy rates divide the forefront from the backdrop area. The preliminary entropy rate super-pixel region has the three subsequent possessions.

Region property of HSI: Around edges that do not step, the entropy rate super-pixel area is wider than one pixel, but it is only one pixel wide near edges that do step.

Boundary property of HSI: situated "among object and background" and covers around the object.

Gray-level variation: The gray areas inside the transition zone often exhibit rapid and intense changes, resulting in a wealth of details that can be utilized to describe the area effectively.

The technique that was created is founded on the concept of the effective mean gradient of entropy. However, traditional gradient-based methods exhibit sensitivity to interference and are more suitable for detecting abruptly distinct variations in levels instead of regular, subtle changes in gray levels. In order to address this constraint, several localized statistical methods were devised, including the entropy rate super-pixel segmentation technique. After that, the statistical measures of the localized area, as well as local entropy, modified localized entropy, and gray-level variance, were utilized for the purpose of extracting transition zones. Gradient-based methods are particularly well-suited for the collection of transition regions characterized by sudden shifts in gray levels. However, its applicability is limited in scenarios with frequent variations in gray levels. Images that exhibit frequent gray-level variations, as opposed to sudden gray fluctuation, are not well-suited for gradient-based techniques. Consequently, the researchers proposed the utilization of the content local entropy measurement as a solution to address the issue of rapid gray-level variations. The local entropy inside an area of size $m \times m$ is determined for a picture of dimensions $N * M$

$$H = -\sum_{I=0}^{L-1} p_i . \log p_i \tag{1}$$

$$P_i = \frac{n_i}{m \times m} \tag{2}$$

where the variable pi represents the probability that is connected with the hyperspectral imaging (HSI) of the ith picture. The variable "n_i" represents the count of super-pixels that possess gray levels within the HSI color space, whereas "L" denotes the highest gray level. Formula 1 demonstrates that the level of local variation is higher in heterogeneous areas compared to homogeneous areas. The entropy threshold may be derived from Formula 3.

$$E\text{th} = \alpha H \max \tag{3}$$

where Hmax is the highest entropy of an entropy picture, while α is an integer that ranges from 0 to 1.

In order to adequately recover pixels in the area of transition, it is common to use metrics like α ranging from 0.8 to 0.9. The comprehensive method can be succinctly stated as follows:

(a) To calculate the regional entropy of a source picture inside a specific neighborhood, perform the following computation.
(b) The area of transition can be extracted by employing an "entropy threshold" technique.
(c) The division threshold can be determined by calculating the average of the graph of the zone of transition.
(d) The picture can be segmented by applying a segmentation threshold.

3.3 SuperPCA

SuperPCA is a good way to lower the number of dimensions in HIS images. Conventional PCA reduction of dimensionality methods usually show linear subspaces (manifolds) in the information, which is the main problem with these methods. But to solve this problem of reducing the number of dimensions in complex cases, this work shows how to do things like use SuperPCA with adaptable pixel process. So, this paper suggests that it is a popular dimensionality reduction (DR) [24] method over others like SuperPCA because it is easy to use and simple to apply.

As a result, dimensionality reduction is seen as a way to find out how much freedom can be used to make a pretty large amount of data set variability in the given input image. Dimensionality reduction uses a high-dimensional picture information set to make a low-dimensional encoding that is more compact. HSI imaging gives you an enormous number of spectral bands, generally in excess of 100. This excessively huge dimension of HSI merely makes it harder to do the math, but it also makes the suggested method less accurate at classifying. HSI has a high level of spectral redundancy and does not have enough training information. Lowering spectral dimensionality has been shown to be a key factor in making categorization methods easier to use. As a result, the SuperPCA starts with a factor analysis answer, looks for rotations, and then finds independent variables.

SuperPCA that makes a list of the best linear estimates for a high-dimensional observation. This method is thought to be one of the best in the field of complexity reduction. Global linearity, however, makes it less successful in more ways than one. Again, multidimensional scaling (MDS), that is related to PCA and also looks the same, has the same look. Factor evaluation and independent component analysis (ICA) have also assumed that the fundamental manifold is just a linear subset. Here, on the other hand, are separated by the high standard of the SuperPCA in a way that allows them to classify and recreate subspace. SuperPCA's modeling of the data's most variable subspace can be thought of as modeling the data's covariance framework, even as factor analysis can be thought of as modeling its correlation structure.

Depending on the method, a nonlinear function on a sample of HSI image data through feature space RN is described as:RN $\rightarrow$ RSSF, where the spectral and spatial characteristics are usually described as SSF $\gg$ NF, RSSF. This space is known as the spectral and spatial featured space. Most of the time, the purpose is chosen so that RSSF contains higher-level product terms as of an amount of HSI training samples RN. When PCA is used on RSSF, higher-order correlations as of a number of HSI training sets will be found. On the others hand, publicly computing samples of the nonlinear map is hard to do, especially when the space has a lot of dimensions. In this study, a SuperPCA approach was used for every number of HSI data through a function that was nonlinear. This was done to get around the problems with PCA.

The RBF kernel function calculates the Euclidean square distance between two HSI features vectors in the characteristic space when given HSI image sample vectors with features in the dimensional space. In the last step, the PCA must be applied to each HSI feature space, and the values must be fine-tuned. For that function, k needs to be changed so that values for each HSI aspect space can be found.

By using the SuperPCA method described above, you can reduce the number of dimensions quickly by getting rid of traits that aren't important to the framework and correctly separating the classes. Because there aren't many useful HSI data points in the system, it would be easy to use the subspace as feed for linear classifiers such as the case fuzzy-SVM classifier.

3.4 Fuzzy-Based SVM Classification

To improve the rate of classifying HSI images, different pixel-wise probabilistic categorization frameworks are used instead of the SuperPCA method in a characteristic dimensional space. To spread these kinds of techniques, it is necessary to figure out the spatial and spectral data of the HSI image space, along with the level of the HSI image gradient. Empirical mode decomposition-based entropy rate super-pixel separation method is used to get a rough idea of the gradient-level spectral data, and fuzzy decision SVM is used to classify the spectral data in a way that makes sense.

In this effort, fuzzy decisions play a big part in improving the accuracy of classification at the gradient level of spatial information. As a result, figuring out the weight values for IMF is important. This makes it easier to put different kinds of fuzzy HSI information into the right category.

The enhanced fuzzy-based support vector machine (SVM) techniques for classification employ a methodology that incorporates the spectral gradient and spatial information of a hyperspectral image (HSI). This approach utilizes a stochastic hybrid pixel-wise categorization framework. The spectral gradient value for probability is estimated using the improved EMD technique in order to gain further insights into classifications relying on fuzzy-SVM. The fuzzy sigmoid kernel function is a type of kernel function utilized in support vector machines (SVM) to improve the precision of categorization of hyperspectral image (HSI) description classification. During margin size minimization, the SVM-FSK method can be used to effectively tell the difference between good and bad characteristic vectors. The fuzzy sigmoid kernel finds the hyperplane that maximizes the margin, which is then turned into a spatial domain version. This is why greatest margin classifiers work well and do not hurt the categorization of material with an infinite number of dimensions. We shorten the progression of different pixel-wise fuzzy-SVM classification arrangements and use the internal result to measure how similar the variables are. In these ways of putting things into groups, if a dependent variable still exists, its information might be stuck in other dimensions. This can be shown by a mapping. When HSI classification is concerned, the suggested fuzzy-SVM classification method can do a better job than unweighted analyses. Proposed method to find and classify HSIs.

Input: A group of HSI photos that were sent in.
Productivity: confidential features.

Step 1: Collecting the input HSI pictures.
Step 2: We will employ principal component analysis (PCA) as a method for doing dimensionality diminution.
Step 3: Performing the entropy rate super-pixel segmentation.
Step 4: Performing the SuperPCA-based dimensionality reduction.

The kernel function, denoted as **k**, and the feature map, denoted as ϕ, are to be constructed for the given images

Refine the characteristics for every hyper-featured subspace FS.
Build the covariance matrix AA.
The feature subspace FS can be characterized with the projection of the characteristic.
The appearance intended for the gradient through regard to FS is computed.
Calculate the characteristic subspace following the process of decreasing the dimensionality.
Enhance the characteristics of subspace through the process of describing them.
Step 6: Execute classification by means of fuzzy based SVM.

Disadvantages of the hybrid method. In order to further increase the model's precision and usefulness in real-world circumstances, make conclusions and point out potential future improvements.

4 Simulation Results

In this research, results from experiments are gathered from one set of data, the Indian Pine collection and the PaviaU dataset, so as to compare the outcomes of the suggested technique and previous research. MATLAB R2018a is used to run all of the scenarios. If you look at Figs. 3 and 4, you can see that the suggested approach finds and sorts things correctly on both the Indian Pine and PaviaU records. The way the areas were put into groups was great. For evaluating the quantitative technique, the overall accuracy (OA), average accuracy (AA), and Kappa measures are determined in combination with the traditional methods.

Figure 5 shows the results of the original data's full accuracy, the fuzzy-based classification using SVM with spectral enhancement, as well as the dimensionality-reduced characteristics SuperPCA method. It can be seen that the total precision of the suggested approach is higher than that of the current techniques. This is because the proposed method uses the SuperPCA methods to reduce the number of dimensions of the features.

Based on the obtained overall accuracy (OA) and average accuracy (AA) outcomes, it is clear that the recommended technique yields superior categorization outcomes. Additionally, the Kappa coefficients indicate that the recommended

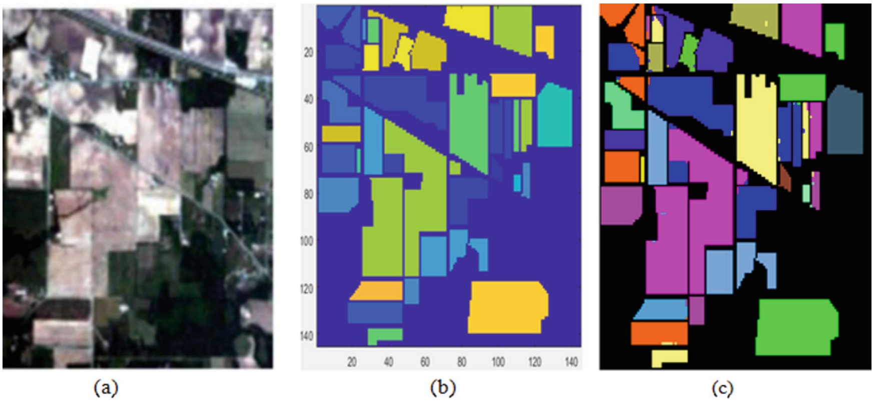


Fig. 3 Indian pines information set **a** input picture, **b** SuperPCA dimensionality reduction picture, **c** output classified picture

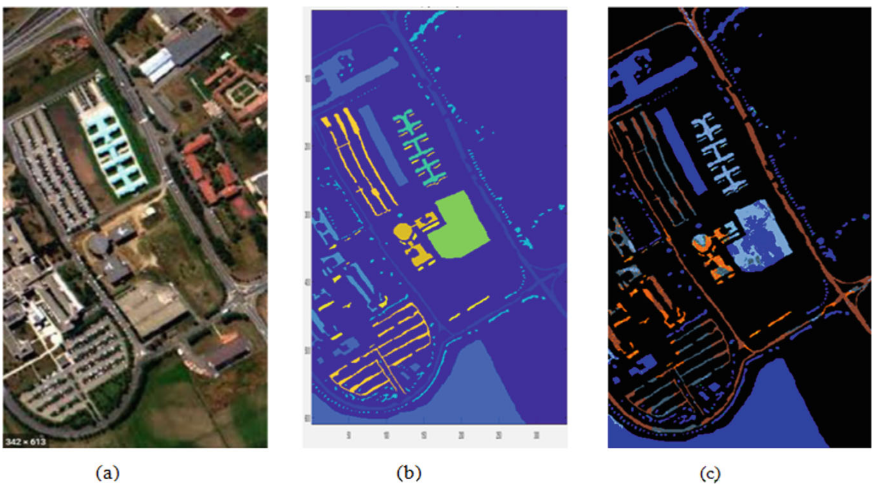
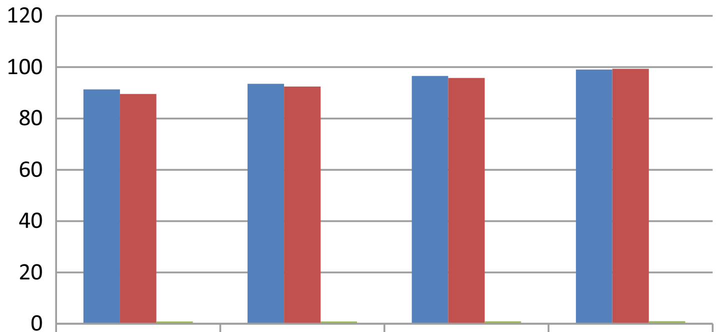


Fig. 4 PaviaU information set **a** input picture, **b** SuperPCA dimensionality reduction picture, **c** output classified picture

strategy exhibits greater reductions in dimensionality compared to LDA [11], EMD-GA [15], and IEMD-PSO [17]. This observation is further supported by the visual illustration depicted in Fig. 5.



	LDA[11]	EMD-GA[15]	IEMD-PSO[17]	Proposed
OA (%)	91.35	93.485	96.593	99.1
AA (%)	89.56	92.46	95.78	99.38
Kappa	0.9019	0.9156	0.924	0.989

Fig. 5 Performance comparisons

5 Conclusions

This study shows how SuperPCA can be used to create a new way to reduce the number of dimensions in HSI data. An imprecise judgment method is used to find the weight values of the kernels. This is done to maximize the spectral gradient and improve the precision of classifying the HIS images. By exponentially mapping the HSI pictures to a feature space with higher dimensions and running SuperPCA on this feature space, higher-order relationships can be discovered in the HSI image data. The shape of the HSI pictures is not highlighted in this suggested method for reducing the number of dimensions, though. The suggested method for putting spectral–spatial information into groups is based on a mixed pixel-by-pixel classification. Also, fuzzy rules for decision-making are often used to help classify things better. This also makes the dimensionality issue in the HSIs worse, and because of the need to solve this query, the precision of classification will be lower than with traditional state-of-the-art methods.

References

1. Kumar PM (2014) Satellite image denoising using local spayed and optimized center pixel weights. Int J Electric Comput Eng 4(5):751
2. Hong D et al (2020) Invariant attribute profiles: a spatial-frequency joint feature extractor for hyperspectral image classification. IEEE Trans Geosci Remote Sens 58(6):3791–3808
3. Han M et al (2020) Joint spatial-spectral hyperspectral image classification based on convolutional neural network. Pattern Recogn Lett 130 (2020):38–45
4. Cao X et al (2020) Hyperspectral image classification with convolutional neural network and active learning. IEEE Trans Geosci Remote Sens

5. Chu Y et al (2020) Hyperspectral image classification based on discriminative locality preserving broad learning system. Knowl-Based Syst, p 106319
6. Hao Q, Li S, Kang X (2020) Multilabel sample augmentation-based hyperspectral image classification. IEEE Trans Geosci Remote Sens 58(6):4263–4278
7. Mou L et al (2020) Nonlocal graph convolutional networks for hyperspectral image classification. IEEE Trans Geosci Remote Sens
8. Okwuashi O, Ndehedehe CE (2020) Deep support vector machine for hyperspectral image classification. Pattern Recognit, p 107298
9. Zheng J et al (2020) Hyperspectral image classification using mixed convolutions and covariance pooling. IEEE Trans Geosci Remote Sens
10. Paul A, Bhoumik S, Chaki N (2020) SSNET: an improved deep hybrid network for hyperspectral image classification. Neural Comput 11
11. Li Z et al (2020) Deep multilayer fusion dense network for hyperspectral image classification. IEEE J Sel Top Appl Earth Observations Remote Sens 13:1258–1270
12. Imani M, Ghassemian H (2020) An overview on spectral and spatial information fusion for hyperspectral image classification: current trends and challenges. Inf Fusion 59:59–83
13. Cao F, Guo W (2020) Cascaded dual-scale crossover network for hyperspectral image classification. Knowl-Based Syst 189:105122
14. Xu Q et al (2020) CSA-MSO3DCNN: multiscale octave 3D CNN with channel and spatial attention for hyperspectral image classification. Remote Sens 12(1):188
15. Wang J et al (2020) A novel cubic convolutional neural network for hyperspectral image classification. IEEE J Sel Top Appl Earth Observations and Remote Sens 13:4133–4148
16. Li H-C et al (2020) Robust capsule network based on maximum correntropy criterion for hyperspectral image classification. IEEE J Sel Top Appl Earth Observations Remote Sens 13:738–751
17. Ahmad M et al (2020) Spatial-prior generalized fuzziness extreme learning machine autoencoder-based active learning for hyperspectral image classification. Optik 206:163712
18. Liu X et al (2020) Extreme learning machine-based ensemble transfer learning for hyperspectral image classification. IEEE J Sel Top Appl Earth Observations Remote Sens
19. Barman B, Patra S (2020) Variable precision rough set based unsupervised band selection technique for hyperspectral image classification. Knowl-Based Syst 193:105414
20. Sawant SS, Manoharan P (2020) Unsupervised band selection based on weighted information entropy and 3D discrete cosine transform for hyperspectral image classification. Int J Remote Sens 41(10):3948–3969
21. Patel R et al (2023) Quality and performance measures in healthcare systems using fog computing. In: Acharjya DP, Kauser Ahmed P (eds) Multi-disciplinary applications of fog computing: responsiveness in real-time. IGI Global, Hershey, PA, pp 95–120. https://doi.org/10.4018/978-1-6684-4466-5.ch006
22. Prakash LS, Bellary MZ, Baig MAA, Kumar TS, Merugu S, An economical black-box system for vehicles. AIP Conf Proc 2477(1):id.030080, (9):2023
23. Suma S, Moon R, Umer M, Srujan Raju K, Bhaskar N, Okali R (2023) A prediction of water quality analysis using machine learning. In: 2023 International conference on distributed computing and electrical circuits and electronics (ICDCECE). IEEE, pp 1–6
24. Maragatham G, Sundaramoorthy K, Maheswari K, Sajana T, Shiba M, Kumar A (2023) The prediction of micro plasma impacts of farm fresh vegetables using machine learning. In: 2023 International conference on artificial intelligence and knowledge discovery in concurrent engineering (ICECONF). IEEE, pp 1–6

Google Assistant-Controlled Electric Appliances

Sudha Arvind and S. Arvind

Abstract In order to have a home automation system which should be user friendly and to be worth effective for operating in a real time, this paper proposes toward home appliance control, an effective IoT soultion. In keeping with the survey conducted by Comscore, we can use this method on smartphone which is the foremost effective. Google Assistant is artificial intelligence-based voice command service. By this voice we are interacting with Google Assistant, and it can search on the net, schedule events, set alarms, control appliances, and far more. This service is accessible on smartphones and Google Home devices. We are trying to control electrical appliances that include lights, switches, fans, and thermostats using our Google Assistant. Built appliances to control the home appliances are designed. Using Google Assistant services, 60W bulb is controlled. Three different platforms, Google Assistant is used along with Adafruit server and IFTTT service.

Keywords Home automation · NodeMCU (ESP266) · IFTTT (if this then that) Application · Adafruit server · Internet of things · Google assistant · Voice control · Smart phone

1 Introduction

Our main target of this paper is to develop an electrical appliances system controlled by using an Arduino circuit board with Bluetooth that can be easily controlled by using a smart phone such as Android, OS. Nowadays most of the people are using this centralized system instead of using conventional switches in home that involves smart

S. Arvind (✉)
Department of Electronics and Communication Engineering, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: sudharvind.ece@cmrtc.ac.in

S. Arvind
Department of Computer Science and Engineering, HITAM, Hyderabad, Telangana, India
e-mail: dean.academics@hitam.org

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_80

remote-controlled devices [1]. Ever thought of a life where you will just command your home appliances to work as you would like just by using your voice? Gone are the days where you have ought to be a billionaire like Tony Stark to possess an automatic house which is voice activated. You will be able to follow this paper even if you have no prior knowledge about programming or NodeMCU. So, let us begin learning home automation using NodeMCU and Google Assistant.

Electrical appliances controlled using Google Assistant can be used in lightening the room and allows heating, and we can use these electrical appliances more efficiently and can be convenient to every user. The proposed system is very simple as it can be used to control all the major parts/components of our home, and it can be controlled by using a smart remote or automatic control of few lights based on the necessity. It helps us to use these technologies for our personal use as this technology can be easily implemented in the existing houses/buildings as this implementation does not need any new wiring and no ripping up the preference. It focuses on wireless home automation carpets or drilling holes within the walls.

2 Internet of Things

The term Internet of things [2] was coined by Kevin Ashton in 1997. It can be defined as connecting of things in everyday life and embedded them with electronics software and sensors to the Internet making them to collect and exchange data as shown in Fig. 1.

Internet of things is referred as the number of things that can be connected to the Internet so that we can share or receive data with other things; some of the applications that can be implemented by connecting these devices are industrial machines and

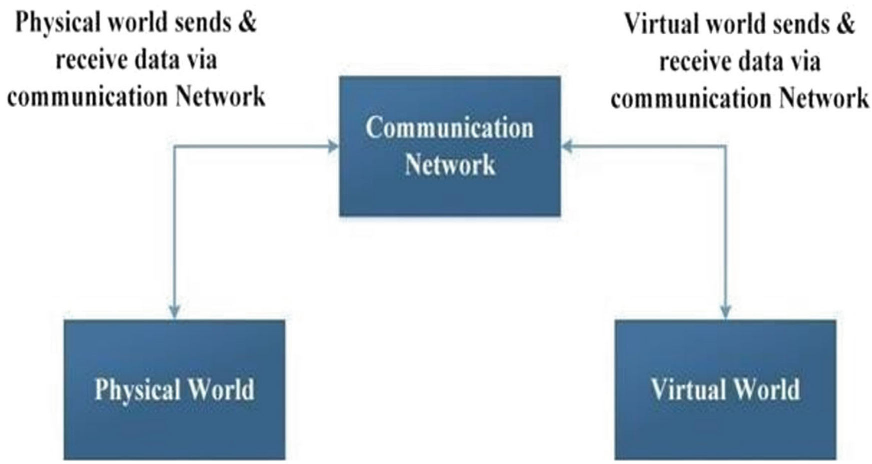


Fig. 1 Internet of things

many more. These devices that are connected using Internet use built-in sensors to collect data and also act on it in some cases.

Some types of Internets of things are Bluetooth, cellular (3gb, 4gb, 5gb), low power wide area networks (LPWAN'S), WI-FI, and RFID. IoT always stays connected, provides security, and also it is cost effective, etc. Internet of things is nothing but using mechanical and digital things that are embedded on sensors, other technologies, and other software. **Applications includes** smoke/gas detectors, smart appliances, smart parking system, fleet tracking, shipment monitoring, remote vehicle diagnostics, indoor air quality monitoring, and smart grids.

3 Related Work

Lab automation system (LAS) is presented in this paper which functions on Nodemcvesp8266 which is based on the cloud networking and wireless communication combination with which the user can control the lights, fans, and appliances remotely, and information storing can be done within the cloud. Automatic change in the idea of sensors data is implied by the system. To manage various types of devices, this low cost method is meant [3].

Electrical and electronic system controlling through automation is involved in home automation through degree of computerization, lighting and temperature control are part of this. The proposed project is about the simple controlling system of home-containing mobile controller through remote with several client modules. The host controller will control the client module through a wireless device in terms of Bluetooth enabled one, which can be smart phone based on Android system [4].

In this paper, the system complexity is reduced through management of attendance system which is the replacement for elimination of several manual process which is involved in calculation of attended hours. This paper proposes a simple technique where student attendance can be taken by making use of Internet of things which is based on the attendance recording through fingerprinting while storing them in database very securely [3].

4 Implementation

We can divide the implementation of the paper design into two main sections, namely hardware and software implementations.

Hardware Requirements

The main hardware components that we use in the paper are NodeMCU—32-bit ESP8266 development board with Wi-Fi, relay module, one bulb holder, and one fan.

Software Requirements

The software requirements are Google Assistant software in smart phones, Adafruit IO platform, IFTTT service, and Arduino IDE.

- Start.
- From the Google play store, download the Adafruit.io
- With E-mail address with valid one sign in
- New project to be created
- Home automation as a name to be given as a project name
- From the list connection type of WI-FI form to be chosen an "create" to be pressed.
- From the login signed in E-mail will receive a token
- The token has to be copied and saved, and the token should not be shared with others.
- Then click on " + " symbol to add off/on conditions, the add a Button. Now it is added to dashboard.
- Dashboard should be clicked and button settings to be edited. Choosing the D5 pins from the settings which one is supposed to be controlled
- 1 value will show as ON means high, and 0 value will show as OFF which is nothing but low.
- High will turn ON, and turning OFF will be due to low.
- Mode of the operation should be changed to switch.
- Through Arduino, IDE Program will be written in NodeMCU.
- In Arduino IDE MCU Adafruit.io node to be opened.
- Wi-Fi SSID and its password to be typed to a Adafruit.io Arduino sketch.
- Connect the relay module to NodeMCU.
- When the value is equal to zero, D5 pin will be high, and the relay turns OFF
- Should use www.ifttt.com, next step is to login if already registered, otherwise if not having account, should be signed up.
- When value = 1 it will make the pin D5 high, the relay is OFF.
- Next step is to go to applets tab, clicking on new applets button.
- End.

Figure 2 displays the above procedure in the flowchart form.

The ESP8266 NodeMCU is capable of communicating with cloud services via multiple protocols using its inbuilt WI-FI connection capabilities. The development board thus offers us the advantage of experimenting in real time. As defined in [1], we were able to select the best controller that satisfied the requirements of our project. Considering the factors such as cost, power dissipation, reliability, and portability, NodeMCU stood at the third place amongst the most portable and sophisticated controllers in comparison with the family of 8051, PIC, AVR, ARM, and other recent competitors in the market such as ATMEL/ATMEGA, Beagle Bone, Raspberry Pi3, Adafruit, Arduino Series, Texas, and instruments.

In electrical appliances controlled by using Google Assistant, the first and the foremost thing is that the user should have any smart phone like Android smartphone, and we can install the Google Assistant in it. When Google Assistant receives commands,

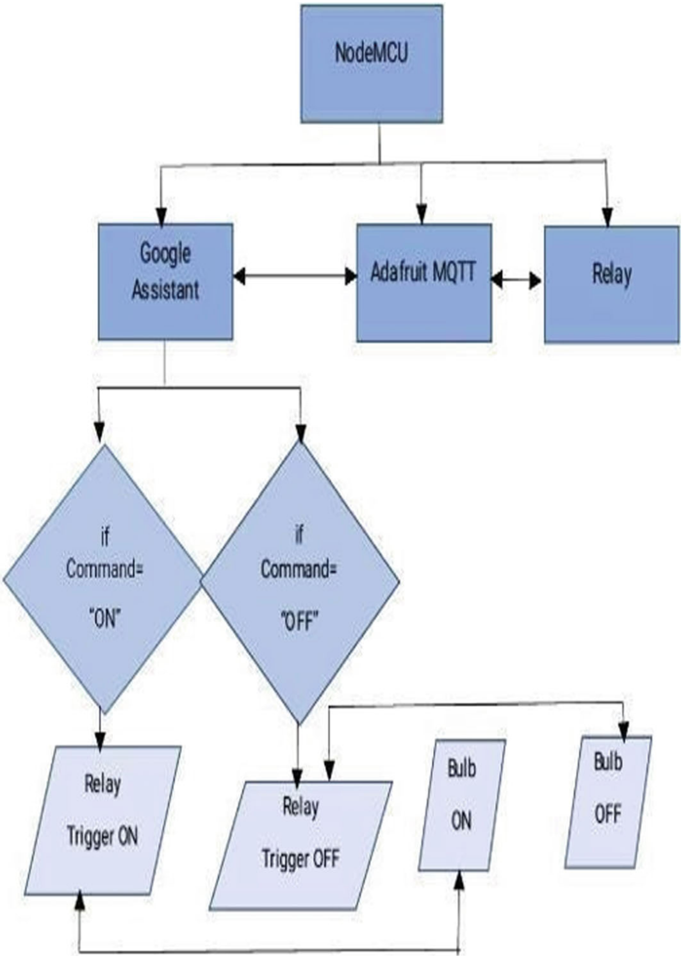


Fig. 2 Flowchart

these are checked with the commands within the Web site of IFTTT set already in advance.

The block diagram of the electrical appliances shown in Fig. 3 controlled by using Google Assistant. In this, the relay will act as a switch, and therefore, home appliances connected to the relay are turned on or off.

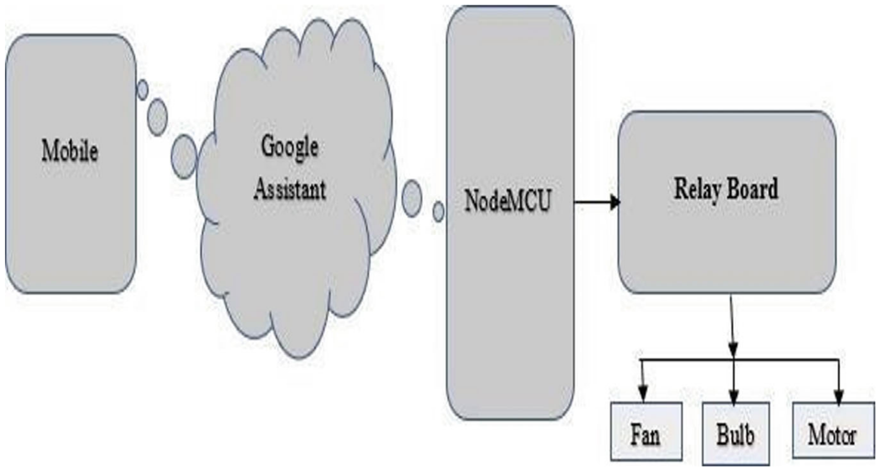


Fig. 3 Block diagram of electrical appliances controlled by using Google Assistant

5 Proposed Solution

The proposed system is used for removal of complexity of wiring in the example of automation or wired connection. For additional possibilities, the power supply amount is possibly required. Currently the existing systems will not be using any type of appliances control and remote monitoring. The proposed system will allow the user to monitor and control the appliances using the Home Automation System through Wi-Fi. The Microprocessor automation called as logic chi is designed in the form of a microchip containing computer processors. Most of the CPU functions where microprocessor contains everything when the computer activates engine will be derived.

6 Results and Discussion

The output for electrical appliances controlled by using Google Assistant is shown in Fig. 4 which shows the proposed system complete implementation.

Figure 4 shows how the voice commands that are given to the Google Assistant and how they turn on/off based on the given commands.

Figure 5 shows the connections of electrical appliances controlled by using Google Assistant and shows how the NodeMCU connected to relay module and relay module connected to bulb holder.

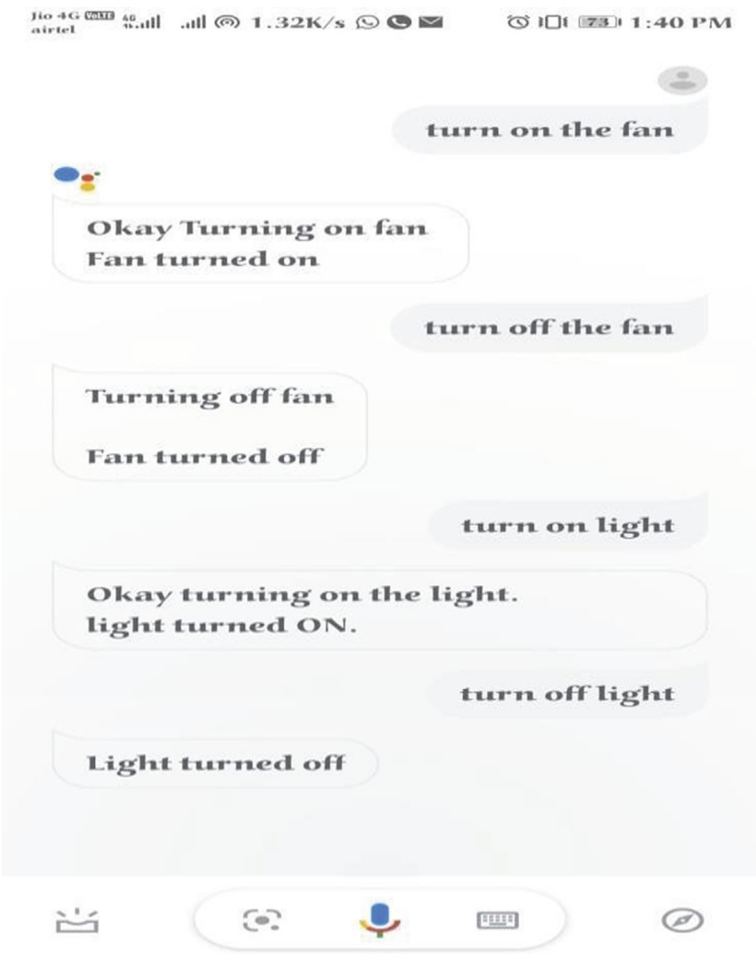


Fig. 4 Voice commands given to Google Assistant

6.1 Light Turning on

In this section, it shows the outputs obtained by given voice commands to Google Assistant connections of NodeMCU, relay module, and bulb/fan. From Fig. 6, it is observed that how the light is turned on by given voice commands, and from Fig. 7, it is observed that how the light is turned off by given voice commands.

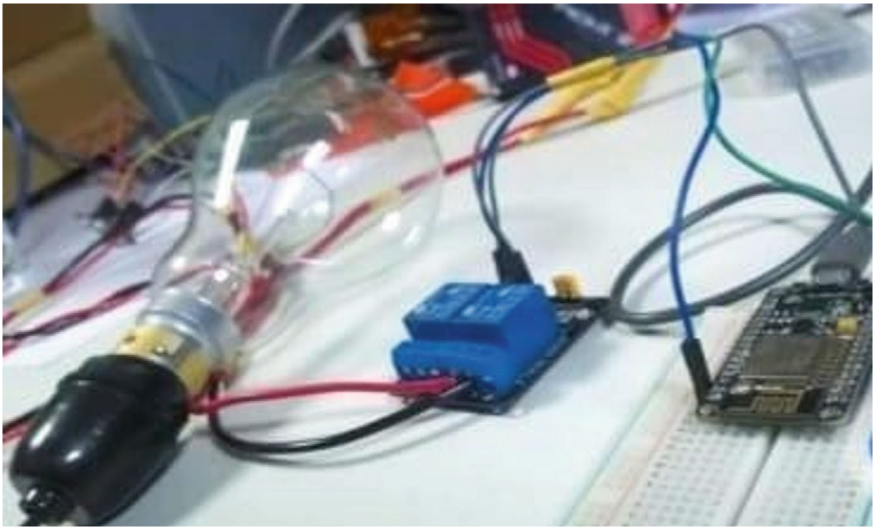


Fig. 5 Connections of electrical appliances controlled by using Google Assistant

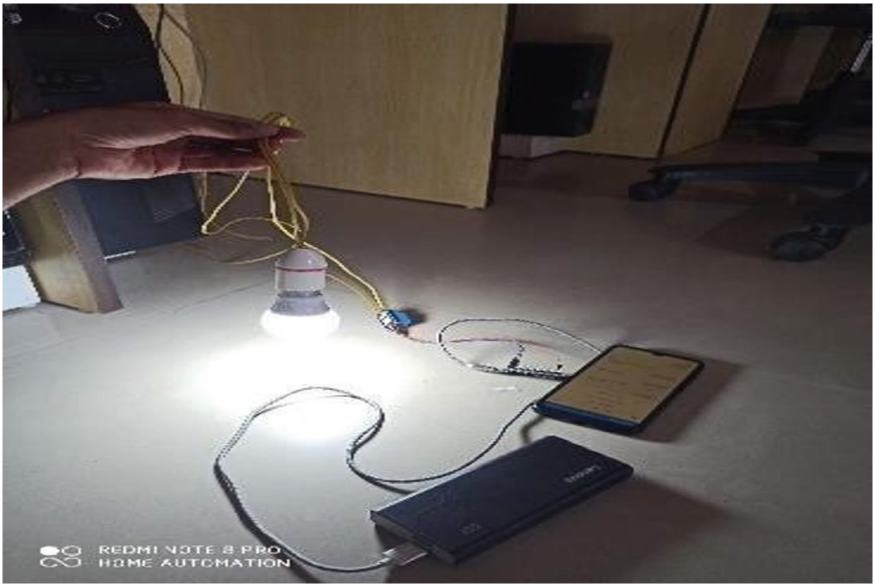


Fig. 6 Light turning on based on given voice commands

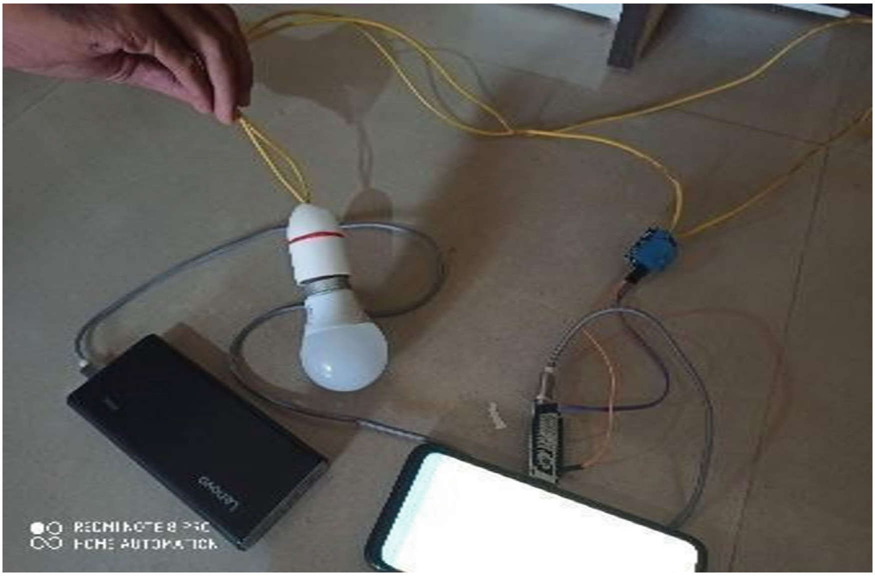


Fig. 7 Light turning off based on given voice commands

6.2 Light Turning Off

6.3 Fan is Turning on

From Fig. 8, it is observed that how the fan is turned on by given voice commands.

6.4 Fan is Turning off

From Fig. 9, it is observed that how the fan is turned off by given voice commands.

Fig. 8 Fan turning on based on given voice commands

Fig. 9 Fans commands turning off based on given voice

7 Conclusion

The proposed concept for laboratory management system proves that unlike current applications of voice assistants in our daily life, it is also used to manage and control corporates, medical hospitals, institutional uses, etc. The voice assistant technology nowadays had become very important, and it is flexible enough to fit into interdisciplinary subjects, such as in our paper we have used the voice assistant that fits well for IoT application. They are used in applications like offices, household purposes, for medically sick people, etc.

8 Scope for Future Work

To build a speech recognition, audio recognition, and also to improve the artificial intelligence for better challenging purposes such as background noise. It is also used for increasing more comfortable and reliable voice for talk on mobile phones, smart devices, etc. Voice plays the major role in the smart devices in the digital world so giving voice commands and operating devices will be on greater demand in future.

Acknowledgements Authors convey sincere thanks to Department of Electronics and Communication Engineering, CMR Technical campus, Hyderabad, Telangana, for providing necessary infrastructure and lab facilities that helped to do the required work.

References

1. Johri E (2012) Remote controlled home automation using android application via Wi-Fi connectivity 3(2). Int J Innov Recent Trends Comput Commun 3(3):1489–1492. ISSN 2321-8169
2. Prasad M (2016) Controlling home appliances on google assistant and monitoring data 3(2). IOSR J Eng (IOSR JEN). ISSN 2250-3021. ISSN (p): 2278-8719
3. Akter F, Akhi AB, Farin NJ, Khondoker MM, Saklayen MG (2018) A novel framework for internet of things (IoT) based smart attendance management system. Intell Control Autom 9:74–84. ISSN Online: 2153-0661, ISSN Print: 2153-0653
4. Kaduskar VP, Gupta N, Bhardwaj Y, Shivam K (2017) IoT based lab automation system. Int J Curr Eng Sci Res 4(6):93. ISSN Online: 2394-0697, ISSN Print: 2393-8374

A Framework for Prediction of Employee Attrition Using Machine Learning Models on IBM HR Dataset

G. Prathiba and Nagaratna P. Hegde

Abstract Employee attrition occurs when workers leave a firm for whatever reason (such as retirement or a new position) and are not promptly replaced. It is not enough for a firm to recruit the best and brightest; it must also be able to hold on to them. Therefore, it is important to examine a dataset containing information about a firm's employee list in an effort to discover patterns that may give helpful information in understanding why people leave the organization. In this study, we train and test five machine learning models using the IBM attrition dataset. These models include the decision tree, AdaBoost, random forest, XGBoost, and gradient boosting. The first stage in this work was to process the data, which included filling in missing values and giving proper names to categorical variables that had been given numerical labels. We then utilize Plotly Lib to visualize the data, which leads us to insights and allows us to see trends in employee turnover. We have encoded variables, rescaled features, and dealt with unbalanced classes in preparation for developing a classification model by separating the data into training and testing sets. Finally, we evaluated the models again in an effort to optimize the hyperparameters and get better results.

Keywords Employee attrition · Ensemble learning · Gradient boosting classifier · Machine learning

G. Prathiba (✉)
Department of Computer Science and Engineering, JNTUH, Hyderabad, Telangana, India
e-mail: pratibhaphd123@gmail.com

N. P. Hegde
Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_81

1 Introduction

Over 50 percent of all companies worldwide need help keeping their best and brightest workers [1]. Many decision-makers in any organization place a high value on being able to identify individuals who pose the greatest retention risk or who may be aggressively targeted for poaching. It may reduce the high cost of recruiting and training new workers by assessing retention risk and calculating the chance of departing, which can then be used to design or improve retention plans. A company's profitability and performance may be negatively impacted by employee turnover, which is a serious and complicated issue. Therefore, companies strive to improve their staff retention rates in order to save costs and boost productivity [2]. Turnover, or staff attrition, is a serious problem for companies of all types. Replacing employees is not just an expense but also a disruption to production and morale. Predicting employee turnover using machine learning allows firms to understand better and manage the causes of employee departure. Machine learning algorithms may be 'trained' on past data to understand the interplay of variables like employee demographics, job satisfaction, performance evaluations, and attrition rates. It may then predict the probability that an employee would resign from their position using the trained model. Attrition prediction is one area where machine learning (ML) has been put to use in HRM to help firms discover how to retain their most valuable workers and avoid losing their best minds [3]. As black box approaches ensemble, deep neural networks have been found to have the greatest performance among various ML methods. Several explainability frameworks have been developed to make these models' reasoning more accessible to users.

Since counterfactual explanation techniques may be used to clarify and prescribe actions to be taken to get the desired result, they have gained substantial interest over the past decade. However, the state-of-the-art approaches to counterfactual explanations emphasize optimizing the adjustments to be made in particular circumstances to bring about the desired result. Predicting how an organization's activity will affect a subset of workers is crucial to the attrition issue, where the aim is to keep them from leaving the business [4]. Thus, the use of hypothetical arguments centered on numerous attrition instances using historical data to identify the optimal interventions that an organization has to make to its practices/policies to avoid or limit attrition likelihood for these occurrences [5].

Various ML models were trained, optimized, and evaluated in this paper to predict whether a certain person will quit the firm, with the hope that the business can tailor its retention efforts to those individuals. The employee exploratory data analysis (EEDA) was used to extract useful information from the dataset, and we optimized the suggested ML models to attain the highest accuracy scores compared to machine learning methods and state-of-the-art research. Employee turnover and its causes were investigated. To achieve statistical parity, we resampled the dataset using the synthetic minority over-sampling technique (SOMTE). Due to an improvement in model accuracy and a decrease in prediction intricacy, data balancing is more popular.

2 Related Works

In this part, we take a look at the literature that relates to our research. Similar articles summarize previously implemented methods and investigation outcomes to anticipate employee turnover. For this literature study, we investigated the most up-to-date, cut- ting-edge methods now in use.

In [6], the authors presented research to apply machine learning techniques to predict employee turnover and investigate the organizational elements that led to turnover. There was a comparison of the four machine learning approaches. Predictions of employee turnover were 93% accurate using the proposed optimized extra trees classifier (ETC) method. Employee exploratory data analysis (EEDA) was employed to implement the suggested method, which outperformed recent state-of-the-art research in identifying the causes of employee turnover. They also confirmed that employees leave because of their monthly salary, hourly wage, position, and age. Organizations may reduce employee turnover with the aid of the suggested method and the study results by focusing on and enhancing the elements that contribute to turnover.

In [7], the authors acquired a quick understanding of the relationships between employee characteristics and the root causes of turnover. They focused their efforts on the visualization of the data at hand. They compared four algorithms used to forecast employee turnover based on factors including age, salary, tenure, and more and highlighted the significance of these findings using clear visuals. Random forest, logistic regression, the K-nearest neighbor method, and the Naive Bayes classifier are the prediction models used. Therefore, HR managers may take preventative measures to keep employees from leaving the company due to visual analysis and precise prognosis of the attrition issue.

In [8], the authors used three different datasets on employee turnover, one of which was the authors' original new dataset. The other two datasets were acquired from Kaggle. Multiple machine learning and deep learning techniques were used to predict employee attrition; these algorithms' performances were assessed using a variety of criteria, including feature selection and hyperparameter optimization. All the datasets we tested deep learning on performed better than the others. To successfully address class disparity, SMOTE Tomek Links were used to oversample underrepresented groups. With F1 scores of 0.972, 0.642, and 0.853, respectively, deep random forest on the HR dataset from Kaggle and neural network on the IBM and Adesso datasets performed best.

In [9], the authors used machine and deep learning models to forecast employee turnover accurately and to isolate the elements that were most responsible for it. Kaggle Depository provided the dataset, which IBM Analytics developed and which has 35 attributes from 1470 workers. The dataset was preprocessed, balanced, and divided into three sets (train, valid, and test) to improve the accuracy of attrition prediction. They conducted multiple experiments to demonstrate the relevance of this research.

In [10], the authors offer a method that makes better predictions about who would leave a company by using deep learning in combination with various preprocessing processes. There are a number of causes of employee turnover. The analysis of such components reveals their interdependence and highlights the most influential ones. IBM's analytics imbalanced dataset, which includes 35 attributes for 1470 workers, was used to evaluate the job. To get more practical outcomes, they generated a more moderate variant from the first one. Finally, cross-validation is used for a more accurate assessment of the project's success. Many experiments have shown the work's utility. Prediction accuracy is around 91% when using the actual dataset and roughly 94% when using a synthetic dataset.

We reviewed the relevant literature in light of our intended research. The results show that our suggested model is superior to machine learning methods and groundbreaking research. We optimized the hyperparameters and balanced the data to get the best results for predicting employee turnover. The review of the linked literature led to the conclusion that previously applied techniques did not previously achieve these applied approaches.

3 Materials and Method

In Fig. 1, we see the organizational workflow of the suggested method for predicting employee turnover. During the feature engineering phase, we finished encoding the features. After inspecting the data, we discovered that it needed to be more balanced. Data was resampled using the SOMTE method to ensure statistical parity. The data collection had been cleaned and prepared for modeling. The dataset was divided into an 80:20 ratio, and the results were analyzed. 80% of the dataset was used for training the proposed machine learning model, while 20% was used for testing.

Dataset Description: To better understand and predict employee attrition, IBM data scientists fabricated the IBM HR analytics employee attrition and performance dataset. There are about 1500 records in the database, and each one represents a different IBM worker. Employee ID, age, residence distance, educational attainment, occupation, monthly income, number of employers, pay increase percentage, number of years employed, and length of service are all important factors in the dataset. We may predict employee turnover and productivity with the use of machine learning models that can be trained using this information. Targeted initiatives to decrease turnover and boost productivity can be developed and put into action with the use of this data.

Data Preprocessing: Due to the prevalence of missing values, noise, and large-scale variation in datasets, we often carry out data preparation prior to training machine learning models. We have performed the following preprocessing operations on the IBM HR dataset: We determined that there are no blanks in the IBM HR dataset utilized for this work. As a result, we skipped this procedure. Not all machine learning algorithms can process data that contains categories. Therefore, transforming these

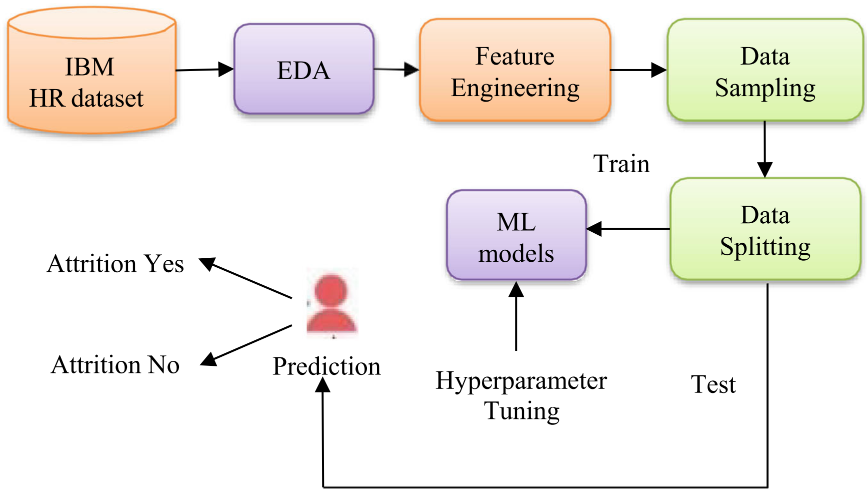


Fig. 1 Organizational analysis of proposed employee attrition prediction

variables into numerical representation is crucial. In this study, a single hot encoding was used for all categorical attribute data conversions.

Data Resampling: We previously observed that our objective variable is skewed towards the negative, with far more 'No' than 'Yes' values. Training results for both classes will be evenly distributed thanks to the usage of SMOTE to generate more 'Yes' responses artificially. Since undersampling the main class might lead to a loss of useful data, we would rather oversample the minor class.

Dataset Splitting: The model was made broader by using dataset splitting. The ratio of 80:20 was used in the division. The majority (80%) of the dataset was used to train the proposed model, while the remaining 20% was used for testing. Overfitting and complexity in the model are both lowered when the dataset is split up.

Proposed Machine Learning Approaches

In this paper, we used five cutting-edge machine learning-based methods: the decision tree classifier (DTC), AdaBoost, the random forest classifier (RFC), XGBoost, and the gradient boosting classifier (GBC). Our recommended method for estimating future staff turnover is the GBC.

Decision Tree Classifier (DTC)

In the DTC [11], the leaf node represents the target class labels, the branches represent the decision rules, and the interior nodes indicate the features. The goal of DTC [12] is to learn decision rules derived from training data that may be used to forecast the target class. The DTC is useful because its ability to make choices rule is analogous to the way humans think. We used information gain and the Gini index in DTC to determine the best features. Equation (3) gives the formula for determining the Gini

Index.

$$\text{Gini index (GI)} = 1 - \sum_{l} U_m^2 \tag{1}$$

AdaBoost Classifier (ABC)

To do this, the technique repeatedly trains weak learners on the training data, with each learner paying special attention to the cases that the preceding learner misclassified. The underperforming students' weights are then adjusted upward to reward the best performers. Predicting employee turnover on the test set is possible after the AdaBoost model has been trained [13]. The algorithm would then provide a prediction of whether the employee is likely to attrite based on the inputted features.

Random Forest Classifier (RFC)

The RFC is a random forest classifier.

Using numerous individual decision trees (weak learners) to create a single, more robust learner, random forest (RF) is an ensemble learning approach for classification and regression [14]. With classification problems, the decision trees vote on how to classify the input data, ultimately producing the class that is the mode of the classes or the mean of predictions. Therefore, random forest mitigates the overfitting issue. The more trees there are in the forest, the better the outcome will be. The random forest learning method is both adaptable and popular. It can function well even without fine-tuning the hyperparameters [15].

Xgboost

In supervised learning, gradient boosting machines (gbm) are often considered to be among the most effective techniques. It is applicable to both classification and regression issues. Data science like Xgboost because of its fast out-of-core compute execution performance [16].

$$X = \varphi(A_i)_i = \sum_{l=1}^{l} g_l(A_i),\ g_l \in \mathbb{G} \tag{2}$$

where L represents the number of trees in the model, and g_l represents the (l-th tree), to solve the above equation.

Gradient boosting classifier (GBC)

GBC stands for 'gradient boosting classifier'. It is an ensemble learning technique, so it takes the results of many less-than-stellar learners and averages them into a single, more reliable forecast [17]. Because it can process large, complicated information, gradient boosting excels at forecasting employee turnover. Additionally, it has a high tolerance for data fluctuations. The features needed to construct a gradient boosting model for predicting employee attrition are (i) whether the employee has left the company (attrition) and (ii) any other relevant employee features such as age, gender, job title, department, salary, and performance reviews.

Step 1: Compile an employee dataset.

Step 2: Divide the data into test and training sets.

Step 3: A gradient boosting model should be trained using the data from Step 2.

Step 4: To test the effectiveness of the model.

Step 5: Determine which employees are most likely to resign using the trained model. The loss function may be written as *xn*(*g*), where (*xi*, *yi*) are a collection of data points, and *i* ranges from 1 to *N*. Separate areas *D*1*n*, *D*2*n*, ..., *Dkn* are defined in the input space, and it is predicted that each region has a constant value *bjm*. Each regression tree has a total of *k* leaves. Here is an expression for the GBRT model and the regression tree *xn*(*g*).

$$x_n(g) = \sum_{k=1}^{k} (a_{kn} I),\ g \in D_{kn} \tag{3}$$

$$I(g \in D_{kn}) = \begin{cases} 1, & g \in D_{kn}; \\ 0, & other; \end{cases} \tag{4}$$

$$L(H, f(g)) = \sum_{i=1}^{m} (H - f(g))^2 \tag{5}$$

Hyperparameter Tuning

In order to fine-tune the parameters of the machine learning models being used, hyperparameter tuning was used [18]. We analyzed the parameters of the hyperparameter setup. Iterative testing optimized the model's hyperparameters on the dataset to ensure optimal performance. Model hyperparameters that optimally optimize their performance scores. The investigation found that adjusting the tune had a significant impact. Accuracy ratings for the used machine learning models were rather high. The following parameters were used to generate predictions for the RFC dataset: min samples_split'= 10, min samples_leaf': 2, 'max_features': 'sqrt,' 'max_depth': 15, and 'criterion': 'gini'.

4 Results and Discussion

First, we provide the experimental setup and dataset we employed for this study. We next go into the methods we used to prepare the data for analysis, the statistical methods we used, and the evaluation criteria we use to assess the quality of the suggested models. Finally, we detail the cross-validation methods we used to prevent training data overuse. Accuracy, precision, recall, and the F1 measure score are used as assessment measures in this machine learning-based research. Metrics for assessment should take into account the following:

$$\text{Accuarcy} = \frac{\text{TP} + \text{TN}}{N} \tag{6}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{F1} - \text{score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \tag{9}$$

Tables 1 and 2 list the machine learning models' performance evaluation values before and after tuning parameters.

According to Fig. 2, the AdaBoost classifier had the highest recall score, accurately predicting the departure of 42 out of 67 workers (62.69%). With an accuracy of 87.76%, gradient boosting outperformed the competition across both classes. In contrast, the random forest model has a low recall of 28.36%, while the decision tree classifier has the lowest accuracy of all the five models tested (76.19%).

Table 1 Machine learning models' performance evaluation

Model	Performance evaluation			
	Accuracy	Precision	Recall	F1-score
DTC	76.19	31.37	47.76	37.87
ABC	87.53	58.33	**62.69**	60.43
RFC	85.26	52.78	28.36	36.89
XGB	86.62	58.33	41.79	48.70
GBC	**87.76**	63.27	46.27	53.45

Table 2 Machine learning models' performance evaluation after RandomSearchCV

Model	Performance evaluation			
	Accuracy	Precision	Recall	F1-score
DTC	77.19	32.37	47.52	37.94
ABC	51.70	21.71	**83.58**	34.46
RFC	84.58	48.94	34.33	40.35
XGB	86.17	55.56	44.78	49.59
GBC	**88.44**	65.38	50.75	57.14

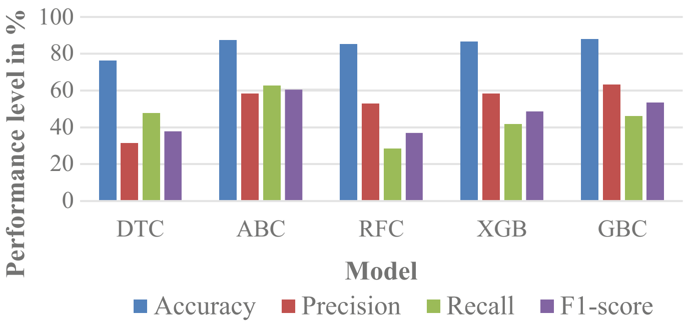


Fig. 2 Performance evaluation comparison of five ML models

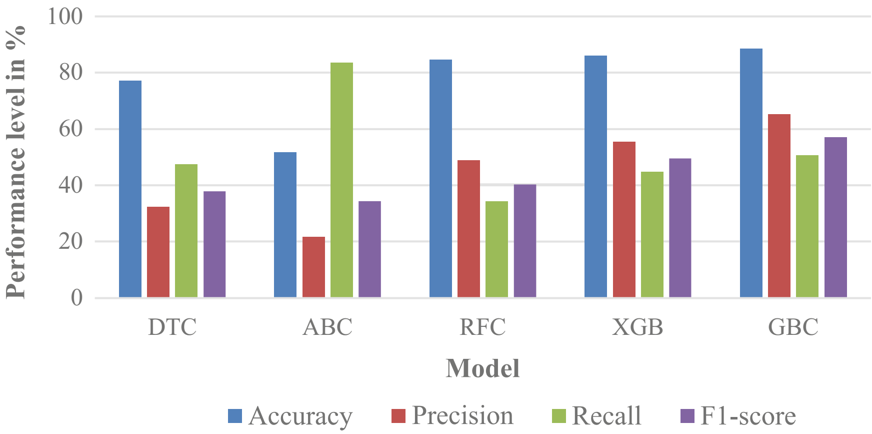


Fig. 3 Performance evaluation comparison of five ML models after RandomSearchCV

Figure 3 demonstrates that the AdaBoost classifier got the highest recall score, accurately predicting the departure of 42 out of 67 workers (83.58%). When comparing all models, gradient boosting achieved the highest accuracy (88.44%) because it correctly predicted more instances across both classes. When compared to the other four models, random forest has the lowest recall at 34.33%, and the lowest greatest accuracy at 77.19% is found in the decision tree classifier.

5 Conclusion

Using the IBM HR dataset, this research compared the performance of several machine learning models trained on distinct feature subsets in order to forecast employee turnover. Five machine learning models were developed and analyzed for this task. With the help of RandomizedSearchCV, the gradient boosting classifier achieved the best accuracy of any model, 88.44%. However, the initial AdaBoost classifier model had the highest recall score (62.69%) and strong accuracy (87.5%) in forecasting the greatest number of workers who were more likely to depart. We could get AdaBoost to 83.58% recall. However, the algorithm got biased toward identifying workers as likely to depart, and we saw a significant rise in false positives because of our tuning efforts.

References

1. Debono MP (2018) Are organisations doing enough to retain their talent?. Importance Empl Retention
2. Chakkambath RS, Rajesh L, Menon S (2023) A study on impact of HR polices on employee retention rate in IT sector 4:2497–2503
3. Bhatta S, Tasin I, Raisa N, Fahim S, Momen S (2022) Machine learning approach to predicting attrition among employees at work. https://doi.org/10.1007/978-3-031-09076-9_27
4. Nair K, Indumathi N, Farzana S (2022) Employee perception towards attrition in IT industry—With special reference to Chennai city. J Stat Manag Syst 25:1147–1159. https://doi.org/10.1080/09720510.2022.2099061
5. Ulle RS, Patil K, Varma AJ, Santosh Kumar AN, Murthy Renuka TP (2018) Rational of employee attrition and strategies to employee retention
6. Raza A (2022) Predicting employee attrition using machine learning approaches. Appl Sci 12. https://doi.org/10.3390/app12136424
7. Sheth K, Patel J, Verma J (2022) Machine learning-based investigation of employee attrition prediction and analysis. https://doi.org/10.1007/978-981-19-6737-5_19
8. Gurler K, Pak B, Gungor V (2023) Deep learning based employee attrition prediction. https://doi.org/10.1007/978-3-031-34111-3_6
9. Arqawi SM, Abu Rumman MA, Zitawi EA, Rabaya AHR, Sadaqa AS, Abunasser BS, Abu-Naser SS (2022). Predicting employee attrition and performance using deep learning. 100:6526–6536
10. Al-Darraji S, Honi DG, Fallucchi F, Ibrahim A, Giuliano R, Abdulmalik HA (2021) Employee attrition prediction using deep neural networks. Computers 10:141. https://doi.org/10.3390/computers10110141
11. Maswadi K, Ghani NA, Hamid S, Rasheed MB (2021) Human activity classification using decision tree and Naïve Bayes classifiers. Multimedia Tools Appl 80(14):21709–21726
12. Azad C, Bhushan B, Sharma R, Shankar A, Singh KK, Khamparia A (2021) Prediction model using SMOTE, genetic algorithm and decision tree (PMSGD) for classification ofdiabetes mellitus. Multimedia Syst 28(4):1289–1307
13. Sridhar GV (2018) Employee attrition and employee retention-challenges & suggestions employee attrition and employee retention-challenges & suggestions. Rajalakshmi Eng Coll Dep Manag Stud
14. Zou RY, Schonlau M (2016) The random forest algorithm for statistical learning with applications in stata the random forest algorithm, pp 1–20
15. Louppe G (2014) Understanding random forests from theory to practice, Ph.D. dissertation, Uni versity of Liège
16. Chen T, Guestrin C (2016) In: XGBoost: a scalable tree boosting system. Association for Computing Machinery, pp 785–94
17. Natekin A, Knoll A (2013) Gradient boosting machines, a tutorial. Front Neurorobot
18. Elgeldawi E, Sayed A, Galal AR, Zaki AM (2021) Hyperparameter tuning for machine learning algorithms used for Arabic sentiment analysis. Informatics 8:79

Prediction of Pulmonary Embolism and Esophagitis Using Machine Learning

Gandham Gunashekar, Lokesh Lavudya, T. Prathima, Kratika Sarma, and A. Sirisha

Abstract Pulmonary embolism and esophagitis are two significant medical conditions that require prompt and accurate detection to ensure proper treatment and management. Machine learning algorithms have shown promise in aiding physicians in the accurate and efficient diagnosis of these conditions. This study aims to explore the use of machine learning algorithms in the detection of pulmonary embolism and esophagitis. The study involved the collection of medical data from various health care, including blood tests and endoscopic procedures, to develop a system that can accurately identify these conditions. Machine learning algorithms are trained and validated on the data collected, and the performance of the algorithms is compared. The results of this study have showed significant improvement in the detection and diagnosis of pulmonary embolism and esophagitis, leading to earlier intervention and improved patient outcomes.

Keywords Pulmonary embolism · Esophagitis · Machine learning · Ensemble methods · Gestational hypertension · Gradient boost

G. Gunashekar · L. Lavudya
Department of Artificial Intelligence and Data Science Chaitanya, Bharathi Institute of Technology, Hyderabad, India

K. Sarma
Department of Computer Science, National Institute of Technology, Warangal, India

T. Prathima · A. Sirisha (✉)
Department of Information Technology Chaitanya, Bharathi Institute of Technology, Hyderabad, India
e-mail: asirisha_it@cbit.ac.in

T. Prathima
e-mail: tprathima_it@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_82

1 Introduction

When anyone is currently suffering from an illness, they must visit a doctor, which is both more time consuming and costly. In the medical industry, machine learning is crucial. Based on the symptoms of the patients, machine learning models can be used to diagnose, detect, and predict various diseases. The use of data mining and machine learning techniques to foretell the possibility has gained popularity in recent years.

Machine learning predicts the future based on past data. Training and testing are the two phases of the machine learning algorithm. From past few decades, machine learning technology has struggled to predict diseases based on patient symptoms and their health histories.

Machine learning has the potential to be useful in the treatment of healthcare problems. In this study, machine learning algorithms are utilized to track patients' health and create models that produce outcomes more quickly.

The 2017 Global Burden of Disease study found that 43% of fatalities are attributable to cardiovascular disease [1]. Daily, a substantial amount of data is produced in the healthcare industry. With the aid of data mining algorithms, several hidden patterns can be uncovered that can indicate an illness. When anticipating the disease, a variety of factors are crucial. When predicting a certain disease, factors like diabetes, high blood pressure, high cholesterol, aberrant pulse rate, and other variables must be considered.

The proposed project involves the collection and analysis of medical records to develop machine learning models capable of predicting specific diseases. The primary goal is to enable early detection of diseases, leading to improved health outcomes for patients and reduced healthcare costs. Early prediction allows healthcare providers to implement preventive measures and interventions, potentially averting the development of certain diseases.

To achieve this objective, the study employs a diverse range of machine learning techniques, including random forest, decision tree classifier, logistic regression, and gradient boost. The utilization of these various algorithms facilitates a comprehensive exploration of the dataset and enhances the accuracy of predictions.

A critical step in the process is the implementation of feature selection using the "select-k-feature" method. By identifying and selecting the most relevant features from the dataset, the predictive models' performance can be optimized while reducing computational complexity.

To carry out the computations, data pre-processing, and visualization tasks effectively, Jupyter Notebook with Python is the chosen platform. This interactive environment offers flexibility and facilitates thorough exploration of the data.

To ensure the robustness and generalizability of the predictive models, best practices in machine learning are strictly adhered to. This involves appropriate dataset splitting into training and testing sets, cross-validation of models, and hyperparameter tuning to prevent overfitting.

Considering the sensitive nature of medical data, strict adherence to data privacy and security protocols is of utmost importance. Any patient information should be de-identified and securely stored to comply with privacy regulations and protect patient confidentiality.

Documentation of the methodology, results, and interpretations throughout the study is vital for ensuring transparency, reproducibility, and knowledge-sharing within the scientific community.

In conclusion, this project aims to develop accurate predictive models using machine learning techniques to enable early disease detection and improve patient outcomes. By employing various algorithms and feature selection methods, and adhering to best practices, the study seeks to make a valuable contribution to the field of health care and disease prediction.

The paper is organized as follows: Section 1 introduces the work, existing work is presented in Sect.\,2, custom dataset is described in Sect.\,3, methodology used is presented in Sect.\,4, data analysis tasks carried out are discussed in Sect.\,5, results are presented in penultimate section, and the paper is concluded along with future scope in the last section.

2 Literature Survey

The utilization of data mining techniques for disease forecasting has been a subject of exploration in previous research. However, despite numerous studies attempting to predict the progression of various diseases, reliable results have not yet been consistently achieved. This paper aims to address this issue by focusing on accurately forecasting the likelihood of disease development in the human body. Table 1 presents an overview of past research conducted in this domain, outlining the specific diseases targeted and the datasets utilized for their respective studies.

By building upon the insights and limitations identified in previous research, this study seeks to contribute to the field of disease forecasting with more robust and accurate predictive models. The goal is to provide healthcare providers with valuable tools to predict disease occurrences at an early stage, thereby enhancing patient care and reducing healthcare costs.

The analysis of past works and the incorporation of diverse data mining techniques may shed light on novel approaches and potential solutions for improving disease forecasting. Through meticulous data pre-processing, feature selection, and algorithm selection, the study aims to extract meaningful patterns from the datasets and develop more reliable predictive models.

To ensure the credibility and applicability of the findings, the study will adhere to rigorous evaluation methods, such as cross-validation and comparison with existing models. Moreover, the researchers will maintain transparency in documenting their methodology, results, and interpretations to enable reproducibility and facilitate further advancements in this critical area of healthcare research.

Table 1 Previous studies on disease prediction

Algorithms and techniques used	Accuracy achieved	Dataset used
Combining the results from K-NN, random forest, and SVM with the metaclassifier being logistic regression [2]	75.1%—stacked	Cardiovascular dataset—12 features from 70,000 personal
Random forest, K-NN, Naïve Bayes, and logistic regression, [3]	70%	Cardiovascular dataset—12 features from 70,000 personal
Naive Bayes, SVM, random forest, and CART [4]	79.13%—random forest	Dataset of 5000 patients, 10 attributes
Decision tree [5]	72.77%—decision tree	Cardiovascular dataset—12 features from 70,000 personal
Holdout cross-validation with the NN [6]	71.82%—artificial neural networks (ANN)	Cardiovascular dataset—12 features from 70,000 personal
Cross-validation method with logistic regression $k = 30$ [6]	72.72%	Cardiovascular dataset—1,12 features from 462 personal
K-nearest neighbours, SVM, decision tree, and logistic regression [7]	98.3%—SVM	Indian Chronic Kidney Disease (Indian CKD)
Binary class [8]	99.8%	St. Paulo's Hospital
Combining of K-mean and SVM classifier [9]	97%	WDBC dataset
Support vector machine and ANN [10]	81.82%—SVM	StatLog heart disease dataset taken from UCI repository (13 attributes)
Random forest, decision tree, SVM, and K-NN [11]	85%—K-NN	Heart disease dataset from UCI repository
NLP, SVM, Naive Bayes, and random forest [12]	84.25%—SVM	Cleaveland and Hungarian dataset from UCI repository
Decision tree, logistic regression, and random forest, [13]	82.1%—random forest	Data from 1000 Qatari long-term residence
Decision tree, SVM, random forest, Naïve Bayes, and K-NN [14]	98.9—random forest	Records from various health cares
Logistic regression, SVM, decision tree, K-NN, and random forest [15]	74.4%—random forest	Pima dataset
XGBoost [16]	94.55%	Luzhou Municipal Health Commission
Decision tree, SVM Naïve Bayes, and K-NN [17]	83.60%—K-NN	Framingham dataset from Kaggle

3 Dataset Description

The dataset used for prediction of pulmonary embolism and esophagitis is collected from healthcare sector. It contains various symptoms and their respective values collected from different individuals. This dataset also includes the symptoms which are responsible for hypertension and gestational hypertension. The dataset provided by the health care is a record of various patients who consulted the doctors in the health care. The dataset used for predicting the diseases includes various columns:

Name: This column contains names of various patients who visited the health care with different symptoms.

Age: This column contains the age of different patients. Visualization can be seen in Fig. 1.

BP-Día and BP-Sys: This column contains the diastolic blood pressure of various patients. During each heartbeat, the heart contracts, pumping blood into the arteries, which results in rise in blood pressure known as systolic blood pressure. Diastolic blood pressure, on the other hand, occurs when the heart relaxes between beats, letting the blood to flow freely through vessels, causing the pressure to decrease.

BPD: Chronic impulsivity and disinhibition are hallmarks of borderline personality disorder, which is a hard personality defect that frequently leads to issues with self-control. They also frequently engage in persistent self-destructive behaviour.

HTN: This column contains blood pressures of various patients. This column is one of our analysis tasks for predicting if a person is having high blood pressure or not. Data is visualized in Fig. 2.

Gestational Hypertension (GHTN): This column contains the data of pregnancy women who are suffering with high blood pressure which leads to immature growth of babies.

PE: This column contains the data of various patients which are leading to pulmonary embolism. An abrupt obstruction of the lung artery is known as a pulmonary embolism. It often occurs when a blood clot ruptures and moves through the bloodstream to the lungs. PE is a dangerous illness that can harm the lungs permanently.

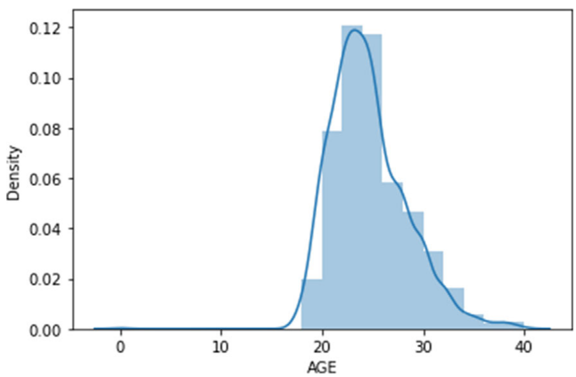


Fig. 1 Age of different patients

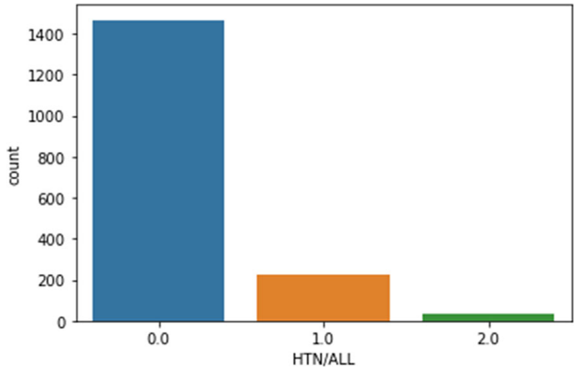


Fig. 2 Count of patients suffering from hypertension

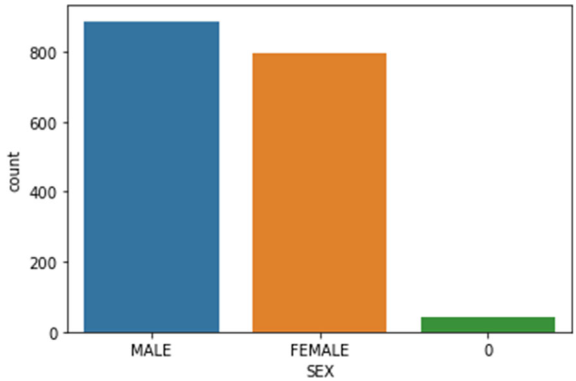


Fig. 3 Count of male and female patients

EO: This column contains the data of various patients which are leading to esophagitis, which is esophageal inflammation. The esophagus connects the mouth to the stomach. It might result in esophageal cancer.

Sex: This column contains the gender of various patients. Data is visualized in Fig. 3.

BMI: This column contains the body mass index of various patients. It is calculated as below:

weight in kilogram/height*height.

4 Methodology

This work is carried out in the following phases.

4.1 Pre-processing

Assigning null and NAN values with duplicate values: The value "0" is allocated to features in the dataset that contain null values. The features which had null values in the dataset are Mod, Sex, Weight, Complications, HTN/ALL, etc. Converting categorical values to numeric values: The features which are categorical are converted to numeric before feature selection. Categorical data is transformed into numerical data using the label encoding procedure. In this method, a number label, typically an integer value, is given to each distinct category in the data. Label encoding is commonly used when working with machine learning algorithms that cannot handle categorical data directly, and the numerical values assigned to the categories provide a way for the algorithm to work with the data.

4.2 Feature Selection

A feature selection algorithm called Select-K-Best [18] is used in machine learning to choose the K most crucial characteristics from a dataset. It works by scoring the importance of each feature in the dataset using a statistical test, such as chi-squared test, and then choosing the top K features based on their scores. Visualization can be observed in Fig. 4.

The Select-K-Best algorithm is commonly used to reduce the features of a dataset, identify, and omit irrelevant or redundant features, and the machine learning model's efficiency and accuracy is improved.

The basic steps of using Select-K-Best algorithm for feature selection are as follows:

a. Define the number K of features to select.
b. Choose a statistical test to score the importance of each feature in the dataset.
c. Fit the Select-K-Best algorithm to the training data.
d. Transform the training set and test data using the transform method to only include the desired features.

4.3 Algorithms Used

Logistic Regression: The goal of is to model the relationship between the independent predictors and the binary outcome using a logistic or sigmoid function. This function maps the input values to a probability range between 0 and 1. The independent variables can be of different types:

Continuous Variables: These are numeric variables that can take any real value within a specific range. For example, age, temperature, or income are continuous variables that can be used as predictors in logistic regression.

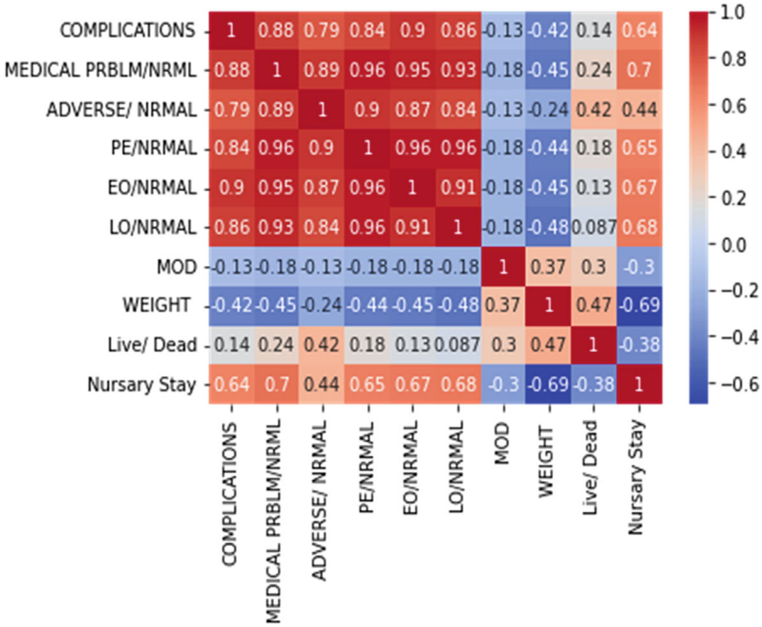


Fig. 4 Correlation matrix

Categorical Variables: These are variables that represent different categories or groups and are often encoded as numeric values (e.g. 0 and 1) in logistic regression. Examples of categorical variables include gender (male or female), yes/no responses, or different types of treatments.

Logistic regression estimates the coefficients (weights) of the independent variables to maximize the likelihood of observing the actual binary outcomes given the predictor values. The logistic function transforms the linear combination of predictor variables and their coefficients into a probability score, allowing us to classify the observations into the respective binary classes based on a chosen threshold (usually 0.5).

Logistic regression is a powerful and interpretable algorithm used in various fields, including medicine, marketing, finance, and social sciences, whenever we need to predict binary outcomes and understand the relationship between the predictors and the binary target variable [19].

Random Forest Classifier: Random forest is a powerful and robust algorithm that combines multiple decision trees to create a more accurate and stable model. During the training phase, the classifier builds a large number of decision trees, typically hundreds or even thousands, by randomly selecting subsets of the data (bootstrapping) and subsets of the features. These random selections introduce diversity among the individual decision trees, making the ensemble more robust and less prone to overfitting.

Each decision tree in the random forest independently produces a classification result based on the features and their associated split criteria. Then, during the prediction phase, the final classification for a given input is determined through majority voting. Each tree "votes" on the predicted class, and the class with the most votes becomes the final prediction. It has applications in domains such as health care and finance, where classification tasks are common. Its ability to handle large and complex datasets while maintaining high accuracy and stability has made it a popular choice in many machine learning applications [19].

Decision Tree: The decision tree algorithm follows a tree-like structure, where the top of the tree starts with a root node representing the entire dataset. The tree is built recursively by splitting the data into subsets based on the most significant input feature at each node. This process continues until certain termination conditions are met, such as reaching a specific maximum depth, a minimum number of samples in a node, or when further splits do not improve the model's performance. For regression tasks, decision trees predict continuous numeric values at the leaf nodes, typically by averaging the target values of the samples in that leaf node. For classification tasks, each leaf node represents a class label, and the prediction is determined by the majority class of the samples in that leaf.

The number of splits from each node in a decision tree can be controlled through hyperparameters. A "coarse" tree has a maximum number of splits limited to four from each node, a "medium" tree allows up to 20 splits, and a "fine" tree permits up to 100 splits.

These are easily interpretable. Moreover, they are robust to outliers, as the splitting process is based on comparisons of features, not direct calculations of distances, which helps avoid the undue influence of extreme values [19].

Gradient Boosting Classifier (GBC): Gradient boosting is a technique that combines many weak learners to model a strong learner. The core idea behind GBC is to minimize a loss function that quantifies the difference between the target variable's expected values and the actual values predicted by the ensemble of decision trees. The algorithm iteratively improves the model by fitting new trees to the residuals.

Furthermore, GBC assigns higher weights to misclassified examples during training, which allows the subsequent trees to focus more on the samples that the model struggles to classify correctly. By iteratively refining the model in this manner, GBC can effectively learn complex patterns and improve its accuracy over time.

GBC has proven to be a powerful and widely used algorithm for a variety of classification tasks, including those involving structured data, text data, and image data. Its ability to handle nonlinear relationships and its resistance to overfitting make it a popular choice in machine learning competitions and real-world applications.

Support Vector Machine (SVM): Support vector machine is a powerful supervised machine learning algorithm that is widely used for various classification and regression problems. In the context of classification, SVM aims to find the best hyperplane that maximizes the margin (distance) between the hyperplane and the nearest training instances of different classes. This hyperplane effectively separates the data points into distinct groups or classes.

The decision boundary in SVM is represented by the hyperplane, which is a multidimensional surface that acts as a separator between different classes in the feature space. For binary classification tasks, the SVM algorithm finds the hyperplane that maximizes the margin between the two classes, making it robust to outliers and more likely to generalize well to new data.

For regression tasks, SVM is employed as support vector regression (SVR). In this case, the algorithm aims to find a hyperplane that comes as close as possible to the training data points within a specified margin (epsilon), while still satisfying the error tolerance. SVM has been successfully applied to various real-world problems, such as image classification, text categorization, medical diagnosis, and financial forecasting [19].

5 Analysis Tasks

5.1 Hypertension Detection

A persistent medical disease called hypertension, also called high blood pressure, causes the blood pressure in the arteries to be higher than it should be. Blood pressure is the force exerted by the heart as it pumps blood through the body's arteries.

If mistreated, hypertension can result in a number of consequences, such as heart attack, stroke, kidney damage, eyesight loss, and peripheral artery disease. As a first line of therapy for hypertension, lifestyle changes including consistent exercise, eating a balanced diet, and lowering stress are frequently advised. Data is visualized in Fig. 5.

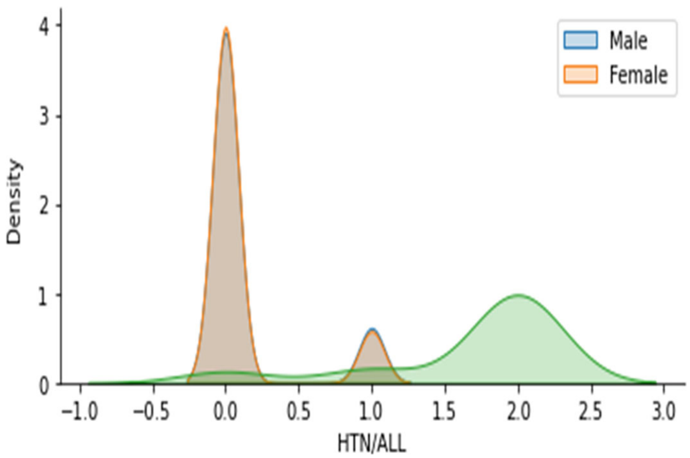


Fig. 5 Count of male and female patients suffering from hypertension

Doctors may occasionally prescribe medication to help lower blood pressure. Therefore, to predict hypertension, key features that highlight the relationship between the attributes and the target variable are identified. The model was trained using various algorithms and subsequently evaluated. The models employed include logistic regression, support vector machine (SVM), decision tree classifier, and random forest classifier, with the Random Forest classifier achieving the highest accuracy.

5.2 Gestational Hypertension

Some pregnant women experience gestational hypertension, also referred to as pregnancy-induced hypertension (PIH). After the 20th week of pregnancy, it begins to manifest as elevated blood pressure, which often goes away after delivery.

Complications from gestational hypertension might affect the mother and the unborn child. Preeclampsia, a more severe form of pregnancy hypertension that can harm the mother's organs and limit the growth of the unborn child, can develop if it is not addressed. In severe cases, preeclampsia can lead to eclampsia, which is characterized by seizures and can be life-threatening.

For predicting hypertension, the important features are selected based upon the relationship between features and target variables, and machine learning models are trained and evaluated. The models employed include decision tree, random forest classifier, support vector machine, logistic regression, and gradient boost. The best accuracy is given by decision trees classifier.

5.3 Pulmonary Embolism

A blood clot that goes to the lungs and stops one or more blood veins, typically from the legs, causes a medical emergency called a pulmonary embolism (PE). Early PE diagnosis and treatment are essential to preventing catastrophic side effects such lung damage, heart failure, and death. So, for predicting hypertension, important features are selected based upon the relationship between features and target variables. Trained the model using different algorithms and evaluated. Models like K-nearest neighbours, gradient boosting classifier, random forest classifier, decision tree classifier, logistic regression, and support vector machine are employed. The highest level of accuracy is provided by logistic regression.

5.4 Esophagitis

Esophagitis is esophageal inflammation. Food travels down the esophagus, a tube, from the mouth to the stomach. Chest pain, heartburn, nausea, and trouble swallowing are a few signs of esophagitis. The underlying cause of esophagitis will determine how the problem is treated.

Treatment may include medications to reduce stomach acid, antibiotics to treat infections, or corticosteroids to reduce inflammation. In some cases, lifestyle modifications such as avoiding trigger foods or losing weight may also help manage symptoms of esophagitis. For predicting the hypertension, important features based upon the relationship between features and target variables are selected.

Trained the model using different algorithms and evaluated. Models like K-nearest neighbours, gradient boosting classifier, random forest classifier, decision tree classifier, logistic regression, and support vector machine are employed. The highest level of accuracy is provided by logistic regression.

6 Results and Discussion

In this section the results obtained from the machine learning models trained on the dataset are presented. The dataset was pre-processed, relevant features selected, split into train and test sets for the four diseases. The project was implemented in Google Colab in Google Chrome.

6.1 Hypertension

From Fig. 6, it can be observed that highest accuracy is obtained with logistic regression. Result of the model is interpreted as: If the model gives the value 0, the status of the disease is “negative”. If the model gives the value 1, the status of the disease is “positive”. If the model gives the value 2, the status of the disease is “severe”.

6.2 Gestational Hypertension

From Fig. 7, it can be observed that the decision tree classifier gives highest accuracy. Result of the model is interpreted as: If the model gives the value 0, the status of the disease is “negative”. If the model gives the value 1, the status of the disease is “positive”. If the model gives the value 2, the status of the disease is “severe”.

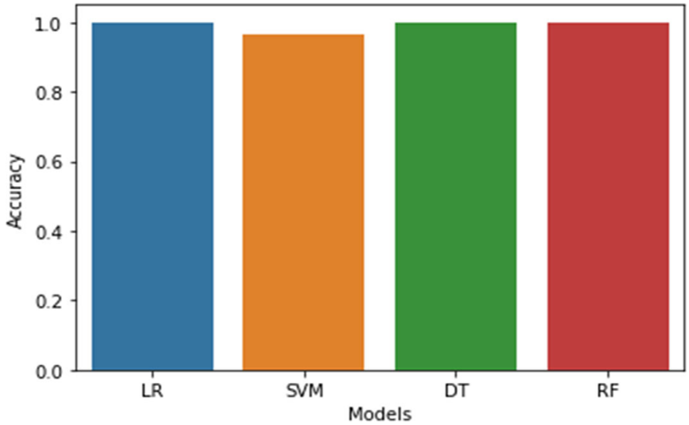


Fig. 6 Comparison of accuracy among various ML models for predicting hypertension

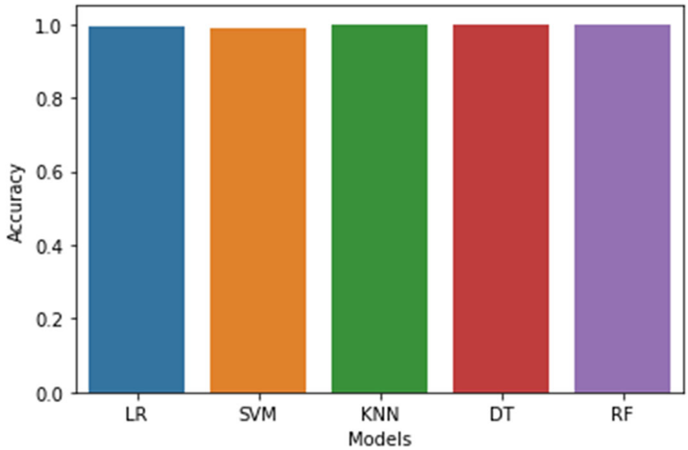


Fig. 7 Comparison of accuracy among various ML models for predicting gestational hypertension

6.3 Pulmonory Embolism

From Fig. 8, it can be observed that the highest accuracy is obtained with gradient boosting classifier. Result is interpreted as: If the model gives the value 0, the status of the disease is "negative". If the model gives the value 1, the status of the disease is "positive". If the model gives the value 2, the status of the disease is "severe".

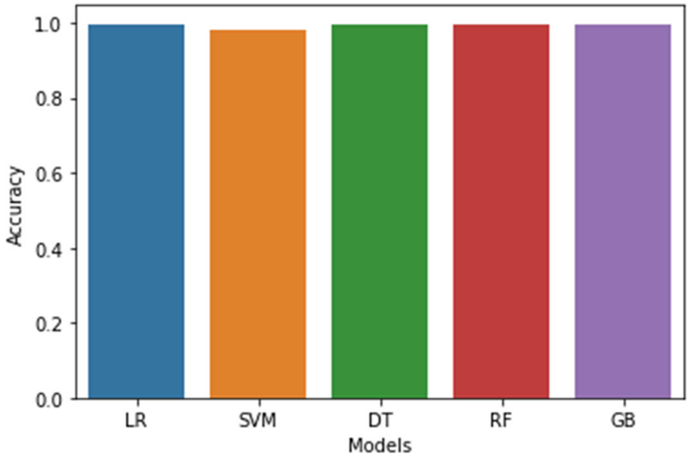


Fig. 8 Comparison of accuracy among various ML models for predicting pulmonary embolism

6.4 Esophagitis

From Fig. 9, it can be observed that all three machine learning models performed better than SVM. Result of the model is interpreted as: If the model gives the value 0, the status of the disease is “negative”. If the model gives the value 1, the status of the disease is “positive”. If the model gives the value 2, the status of the disease is “severe”.

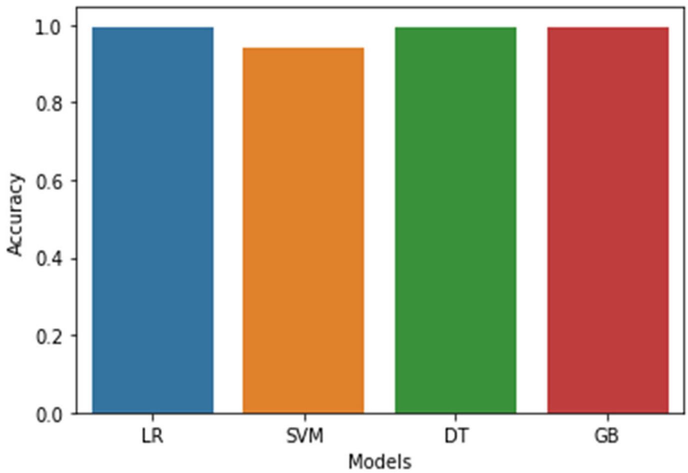


Fig. 9 Comparison of accuracy among various ML models for predicting esophagitis

Table 2 Performance comparison of machine learning models in detecting the diseases

Classifier model	Hypertension	Gestational hypertension	Pulmonary embolism	Esophagitis
Logistic regression (LR)	99.71	99.71	99.71	99.61
Support vector machine (SVM)	96.52	99.27	97.97	94.21
Decision tree (DT)	99.71	99.71	99.42	99.61
Random forest (RF)	100	99.71	99.42	–
Gradient boost (GB)	–	–	99.42	99.61

The results obtained are on a custom dataset collected from India with focus on gestational women, the current work cannot be compared with the existing work, and the performance of the machine learning models is presented in Table 2.

7 Conclusion and Future Work

In this study, the authors have trained models that can predict the possibility of a person developing a particular disease. Initially a variety of machine learning techniques, including random forest, decision tree classifier, and logistic regression, is used to build predictive models. In order to improve the convergence of the dataset, select-k-feature was used for feature selection. The findings may also have broader implications for the development of machine learning algorithms in the medical field and could pave the way for more accurate and efficient diagnostic tools. The work can further be carried out based on major and minor ethnic groups by also collecting the ethnic details of the participants.

References

1. Institute for Health Metrics and Evaluation (IHME) (2017) Findings from the global burden of disease study 2017. In: Institute for health metrics and evaluation, 2018. Available. https://www.healthdata.org/sites/default/files/files/policy_report/2019/GBD_2017_Booklet.pdf
2. Shorewala V (2021) Early detection of coronary heart disease using ensemble techniques. Inf Med Unlocked 26:100655
3. Keniya R, Khakharia A, Shah V, Gada V, Manjalkar R, Thaker T, Warang M, Mehendale N (2020) Disease prediction from various symptoms using machine learning, 2020. Available. https://ssrn.com/abstract=3661426
4. Waigi R, Choudhary S, Fulzele P, Mishra GV (2020) Predicting the risk of heart disease using advanced machine learning approach. Eur J Mol Clin Med 7:1638–1645
5. Waigi D, Choudhary S, Fulzele P, Mishra D (2020) Predicting the risk of heart disease using advanced machine learning approach. Eur J Mol Clin Med 7(7):1638–1645

6. Khan MIH, Rubaiyat Hossain Mondal M (2020) Data-driven diagnosis of heart disease. Int J Comput Appl 176(41):46–54
7. Charleonnan A, Fufaung T, Niyomwong T, Chokchueypattanakit W, Suwannawach S, Ninchawee N (2016) Predictive analytics for chronic kidney disease using machine learning techniques. In: 2016 Management and innovation technology international conference (MITicon), Bang-San, Thailand, 2016, pp MIT-80–MIT-83. https://doi.org/10.1109/MITICON.2016.8025242
8. Debal DA, Sitote TM (2022) Chronic kidney disease prediction using machine learning techniques. J Big Data 9 (1):1–19
9. Zheng B, Yoon SW, Lam SS (2014) Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. Expert Syst Appl 41(4):1476–1482
10. Beyene C, Kamat P (2018) Survey on prediction and analysis the occurrence of heart disease using data mining techniques. Int J Pure Appl Math 118(8):165–174
11. Komal Kumar N, Sarika Sindhu G, Krishna Prashanthi D, Shaeen Sulthana A (2020) Analysis and prediction of cardio vascular disease using machine learning classifiers. In 2020 6th international conference on advanced computing and communication systems (ICACCS), pp 15–21. IEEE, 2020
12. Meshref H (2019) Cardiovascular disease diagnosis: a machine learning interpretation approach. Int J Adv Comput Sci Appl 10(12)
13. AlKaabi LA., Ahmed LS, Al Attiyah MF, Abdel-Rahman ME (2020) Predicting hypertension using machine learning: findings from Qatar biobank study. Plos one 15(10)
14. Gomathy CK, Rohith Naidu A (2021) The prediction of disease using machine learning. Int J Sci Res Eng Manage (IJSREM) 5(10)
15. Naveen Kishore G, V. Rajesh, Vamsi Akki Reddy A, Sumedh K, Rajesh Sai Reddy T (2020) Prediction of diabetes using machine learning classification algorithms. Int J Sci Technol Res 9(01):2277–8616
16. Yang H, Luo Y, Ren X, Wu M, He X, Peng B, Deng K, Yan D, Tang H, Lin H (2021) Risk prediction of diabetes: big data mining with fusion of multifarious physical examination indicators. Inf Fusion 75:140–149
17. Marimuthu M, Deivarani S, Gayathri R (2019) Analysis of heart disease prediction using various machine learning techniques. In: Advances in computerized analysis in clinical and medical imaging, pp 157–168, 2019
18. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res, 3(Mar):1157–1182. Available: https://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf?ref=driverlayer.com/web
19. Han J, Kamber M, Pei J (2012) Data mining concepts and techniques third edition. University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University, 2012

An Enhanced Method for SNR Multi-Scale Entropy-Based Weighted ECG Distortion Using Adaptive Kalman Filter Bank

Mallesh Sudhamalla, G. Lavanya, Vooradi Sandhya, G. Karthik Reddy, Merugu Suresh, and R. Suhasini

Abstract The effectiveness of model-based Bayesian systems was proved in the area of ECG analysis. In this article, we offer a novel Bayesian framework for ECG analysis which makes use of the Kalman filter bank as well as the expectation maximization (EM) method. In comparison with earlier Bayesian algorithms, this approach simply needs an R-peak identification phase during the ECG analysis. Every ECG pulse can be divided into two parts, namely the QRS complexity (high-frequency segment) and the P&T waves (low-frequency portion), based on its position of R-peaks. The low and high frequencies are denoised using the Kalman filter bank for the above process, which consists of two separate Kalman filters. The expectation maximization (EM) method is used to calculate and repeatedly modify the settings for each of these filtration. A number of ECG databases containing patterns with morphological deviations and changes, such as atrial premature complexes (APC) and premature ventricular contractions (PVC), were used to assess the suggested

M. Sudhamalla
Department of ECE, CMR Technical Campus, Hyderabad, Telangana, India

G. Lavanya (✉)
Department of CSE, CMR Technical Campus, Hyderabad, Telangana, India
e-mail: cmrtc.paper@gmail.com

V. Sandhya
Department of CSE (Data Science), CMR Technical Campus, Hyderabad, Telangana, India

G. Karthik Reddy · M. Suresh
Departmentof ECE, CMR College of Engineering and Technology, Hyderabad, Telangana 501401, India
e-mail: gandhari.karthik@cmrcet.org

M. Suresh
e-mail: msuresh@cmrcet.org

R. Suhasini
Department of CSC, CMR College of Engineering and Technology, Hyderabad, Telangana 501401, India
e-mail: rsuhasini@cmrcet.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_83

technique. A number of ECG blurring techniques, including wavelet change, empirical mode breakdown, and band-pass filtering, have been compared to the suggested approach. The outcomes demonstrated that the suggested technique performed superior to reference techniques at small input SNRs from both SNR enhancement and multi-scale entropy-based weighted distortion (MSEWPRD) views.

Keywords ECG denoising · Kalman filter · Bayesian filtering · Expectation Maximization (EM) algorithm

1 Introduction

To achieve exact outcomes, numerous preliminary processes must be completed prior establishing a model-based Bayesian system. These procedures involve finding R-peaks and deriving the EDM (ECG dynamical model) through an offline refinement method.

Polar phase association is furthermore required to connect ECG pulses. Lastly, the procedure uncertainty and measurement distortion have to be determined along with the Kalman filter variables. These procedures are necessary for every structure of identical kind and were vital for ensuring the precision and efficacy of the model-based Bayesian paradigm.

A specific ECG dynamical model (EDM), which assumes that every beat has an identical pulmonary signal, is allocated to an ECG signal. However, this allocated EDM is not helpful for beats where there are obviously noticeable fluctuations in the form of the heartbeats, such as when arrhythmia or when the electrodes get detached. However, when employing a Kalman filter for ECG processing, the procedure and noise measurement matrices cannot be effectively described ahead if the noise properties change between beats. The MP-EKF approach proposed can manage these types of beats utilizing an adaptive fuzzy-based component. In order to address this problem, a brand-new Bayesian architecture is suggested in this research, which processes ECG data using a bank of Kalman filters and the expectation maximization (EM) method.

The sole requirement for the suggested technique's setup is R-peak identification. The QRS complexity (high-frequency component) and P–T waves (low-frequency segment) are the two distinct sections that make up each ECG beat. High-frequency signs are created by concatenating QRS sections, and signals with low frequencies are created by concatenating P–T waves. Then, the denoised high and signals with low frequencies are sent into the suggested Kalman filter bank, which is made up of two separate Kalman filters.

The EM method is used to continuously calculate and modify the parameter values of the two Kalman filters. Bryson and Henrikson's approach is applied for the forecasting and updating stages within the Kalman filter bank to deal with irregular disturbances.

The electrocardiogram (ECG) is a vital biological signal that frequently carries a wealth of crucial patient information. For a correct diagnosis of heart diseases, excellent quality ECG readings are always necessary. However, the nature of the raw ECG readings is quite noisy. In the work, we provide a hybrid denoising method that combines non-local means (NLM) and high-order synchrosqueezing transform (FSSTH) to improve ECG data. This technique starts by using FSSTH to decompose a noisy ECG signal into an ensemble of intrinsic mode functions (IMFs), [5]. It is a collection of noisy IMFs that are specified by a scaling exponent obtained by the detrended fluctuation analysis (DFA) are then eliminated, which reduces some of the noise.

2 Literature Review

DanandehHesar and Mohebbi [1] The "ECG denoising based on successive local filtering" was suggested by Nasser Mourad. A fresh strategy for blurring ECG readings that has been tainted by continuous noise is put forth in this work. The term additive white Gaussian noise (AWGN) is frequently used to describe this type of noisy output. In the present research, a method for removing wideband distortion from captured ECG data was devised using consecutive regional filtration. The technique of segmentation used to split the captured ECG information into elements, every of which roughly comprises each of the four elements, is what makes up the suggested approach. Three components—the QRS complex surge, the TP section, and the P wave—were combined to represent an error-free ECG dataset. Each of the Three categories in this approach may be used to categorize computations.

Hesar and Mohebbi [2] "Noise reduction in ECG signal using an effective hybrid scheme" was the suggestion made by Nadeem Iqbal. The phases for this paper's mixed blurring technique for ECG signal that utilizes FSSTH and NLM are as follows: The ECG signal which has to be denoised comes first collected. The chaotic ECG signal is divided into its component parts in the next phase utilizing the fast S-transform with Hilbert (FSSTH) to produce a collection of limited by band intrinsic mode functions (IMFs). A complex-valued time–frequency dispersion is obtained using the Hilbert transform by the FSSTH, an amended version of the fast S-transform. In this third stage, the scale exponent for each IMF is estimated using detrended fluctuation analysis (DFA), which yields the total amount of IMFs from the FSSTH. In the final stage, each IMF is analyzed using the DFA to establish its threshold, which then relies on the Hurst coefficient. In the sixth stage, noisy IMFs are removed using thresholding rules and also IMFs with a cutoff under a specific level.

Hesar and Mohebbi [3] "A multi-rate marginalized particle extended Kalman filter for P and T wave segmentation in ECG signals" was suggested by Hamed DanandehHesar and Maryam Mohebbi. The application of FSSTH in ECG signal analysis, an adaptable formula for figuring out how many phases there are, and the implementation of the scaling factor from DFA as a limit for distinguishing noise from signal are all introduced in this study. To remove distortion in electrocardiogram (ECG) readings, one statistical variant is the minimized particle prolonged Kalman filter (MP-EKF). Nonlinearities in the ECG signals can be handled effectively. Additionally, the P and T waves in the electrocardiogram (ECG) reading have been identified using this approach. A novel method has been put out that more precisely estimates the P and T waves using a multi-rate variant of the MP-EKF. With this new method, the theoretical particles activity is changed to more closely mimic the course of the ECG signal. The fiducial sites of the P and T waves are identified using a new technique that uses morphological analyses upon the ECG signal was properly filtered. For better outcomes, this technique integrates a number of popular procedures. On an inventory of ECG signals, the suggested technique was evaluated and contrasted with different Bayesian approaches. The findings demonstrated that the new strategy functioned better than other approaches in correctly recognizing the patterns of P as well as T oscillations in the ECG data.

Manju and Sneha [4] The "An adaptive particle weighting strategy for ECG denoising using marginalized particle extended Kalman filter" was presented by Hamed Danandeh and Maryam Mohebbi. This investigation addresses a model-based Bayesian denoising architecture that enhances eliminating efficiency using a fuzzy-based dynamic object loading method and a minimized particle-extended Kalman filter (MP-EKF variational mode decomposition, or VMD). The structure is made to function successfully regardless of the signal's morphology deviates from the established dynamic pattern. The suggested method adjusts MP-EKF functionality to measure values recorded at various intake signal-to-noise ratios (SNRs).The suggested method decreases the element's confidence in the data at lower input SNRs while increasing their confidence in a synthetic electrocardiogram (ECG) made with the attribute values of the ECG dynamic model (EDM). The particle's degree of confidence in the data rises while falling in the synthetic ECG with higher input SNRs.

3 Proposed Method

Researchers construct a versatile Kalman filter bank using ECG denoising in this region. A electrical impulse with time-dependent characteristics, an electrocardiogram (ECG or EKG) signal reveals the electrical activities of the heart. The ECG signal offers crucial data for diagnosis on the condition and beat of the heart. In order to serve as medical evaluation, rehabilitation, or arrhythmia observing, ECG signals might have gotten deformed throughout capturing due to different kinds of undesirable disruption, such as skeletal muscle objects, electrode antiques, electrical line sound, and breathing interference. Therefore, for effective use in clinical settings, digital signal processing on ECGs is necessary to eliminate interference and noise in the indicators.

For the purpose of to evaluate ECG data, this research suggests a unique Bayesian architecture that makes use of a Kalman filter bank and an EM method. Comparatively speaking, this approach needs a lot less preparation than other Bayesian approaches. The sole preprocessing step required by our technique is the R-peak identification, and the R-peaks of the input ECG signal may be recovered employing the Pan–Tompkins approach. The QRS recognition algorithm developed by Pan and Tompkins is one among the most widely utilized techniques. The most common and commonly referenced method for removing complex QRS signals in an ECG signal is the aforementioned algorithm. To locate the R-peaks in the QRS intricate, this technique makes use of an integral window's intensity, gradient, and size. Figure 1 illustrates the algorithm's components, which include a band-pass filter (low-pass and highly pass filter), instruments such as derivatives, squaring operation, a moving window integration (WMI), a threshold, and a decision. The variable set in the decision-making stages was used to automatically modify the threshold in response to variations in heart rate and QRS shape. We can distinguish the QRS complex and P–T waves from one another in an ECG beat by analyzing the R-peak of the signal.

An adaptive Kalman filter bank (AKFB), a signal analysis approach for ECG signal demising, was applied in this research. It is composed of two separate Kalman filters [8, 9], one of which is intended to denoise QRS complexes while the other to denoise P–T waves. Each Kalman filter employs a unique EM algorithm. The filter bank denoises the aforementioned signals with the iterative EM technique and RTS softer equations. The denoised ECG signal is then created by reassembling the two denoised waveforms. When contrasted with previous approaches, the AKFB performs improved blurring efficiency since it flexibly changes its settings according to the amount of distortion in the ECG signal [6]. The dimensions of the input and output are identical in the Kalman filter structure, despite convolution-based

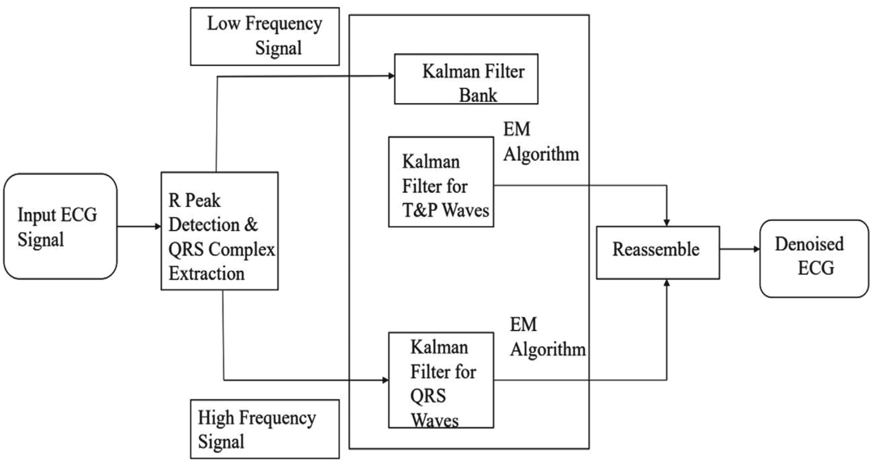


Fig. 1 Block diagram of proposed method

filtering. As a result, this technique does not take into account contrasting issues when rebuilding parts. We recommend employing altered model-defined formulae rather than the traditional model-defined solution to manage the irregular sounds with greater effectiveness. The initial drifts of irregular noises are suppressed by this approach using several afterwards steps. So, to eliminate baseline wanders, we apply an average filter. Furthermore, because ECG signals lost majority of their features and traits and take on distinct dynamics in conditions like ventricular tachyarrhythmia (VT), sudden cardiac death (SCD), premature ventricular complex (PVC), etc., our approach can be applied in these circumstances as well.

In order to process ECG data, this research suggests a unique Bayesian framework that makes use of a Kalman filter bank and an EM method. Comparatively speaking, this approach needs a lot less preprocessing than other Bayesian approaches. The sole preliminary processing phase required by our technique is the R-peak identification, and the R-peaks of the source ECG signal may be recovered utilizing the Pan–Tompkins technique. The QRS recognition method developed by Pan and Tompkins is one of its most widely utilized techniques. The most popular and commonly referenced method for removing complex QRS signals in an ECG signal is this technique.

4 Mathematical Preliminaries Kalman Filter

Linear quadratic estimation (LQE) is another name for Kalman filter. It represents a method that employs some precise metrics in addition to an array of respectable observations. This process generates a calculation of unidentified factors that may be less inaccurate that the sum of the data that was seen. Numerous scientific uses of Kalman filtering exist. The Kalman filter's variations have been used in a variety of programs, including transmission blurring, directions, and, mostly spaceflight.

The state space and assessment model for the discrete-time filtration is as follows:

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{1a}$$

$$\mathbf{x}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \tag{1b}$$

where $\mathbf{x}_k$, y_k, $A_{k,}$ and C_k are considered as a state trajectory, analysis vector, conversion matrix, and density matrix at time step k. In the same way, v_k and w_k are analysis vectors and process noise vectors at time step k. These noise vectors, uncertainties of capabilities, and state dynamics use white Gaussian noise with covariance matrices.

By using Eq. (1), the discrete Kalman filter algorithm might be summarized in two stages, and these stages are shown in the following as the "prediction step" and "update step".

Prediction step:

$$\boldsymbol{x}_{k|k-1} = \mathbf{A}_k \hat{\boldsymbol{x}}_{k|k-1} \tag{2a}$$

$$\boldsymbol{P}_{k|k-1} = \mathbf{A}_k \mathbf{P}_{k|k-1} \boldsymbol{A}TK + \mathbf{Q}_{k-1} \tag{2b}$$

Update step:

$$\mathbf{i}_k = \mathbf{y}_k - \mathbf{C}_k \hat{\boldsymbol{x}}_{k|k-1} \tag{3a}$$

$$\mathbf{S}_k = \mathbf{C}_k \mathbf{P}_{k|k-1} \boldsymbol{C}_k^T + R_k \tag{3b}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \boldsymbol{C}_k^T \mathbf{S}_{k-1} \tag{3c}$$

$$\boldsymbol{x}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \mathbf{K}_k \mathbf{i}_k \tag{3d}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{S}_k \boldsymbol{K}_k^T \tag{3e}$$

Both $x_{k|k-1}$ and $x_{k|k}$ are taken into account as anticipated norms and updated or assess the current state vector's average for each step of time k, respectively. $P_{k|k\text{-}1}$ and $P_{k|k}$ are the anticipated and revised covariance matrices, respectively, for the state of the vectors at each step in time k. Creativity vectors (i_k), Kalman filter gains (K_k), and measured error in forecasting matrix covariance (S_k) are three examples. With the state vector's starting transportation, the Kalman filter will be started. Certain factors, such as mo and Po, are selected in accordance with the framework that is being studied.

The marginal distribution of the Kalman filter can be expressed as the following:

$$P(Y_\mathrm{N}) = \prod_{K=0}^{N} p(y_k / y_{k-1}, \ldots y_0) \tag{4}$$

where $Y_\mathrm{N} = \{y_1, y_2, \ldots, y_k, \ldots, y_\mathrm{N}\}$

Using Eq. (3) prediction step, the log marginal likelihood for the Kalman filter be calculated as:

$$l_k = l_{k-1} - 0.5(i_k^T S_{k-1} i_k + \log|S_k| + N_y \log 2\pi) \tag{5}$$

where N_y is the analysis vector Y_k.

Utilizing techniques like backward pass, and even the Rauch–Tung–Striebel smoother (RTS), the Kalman filter's output can be further rounded out. This method can be explained in the following:

$$l_k = P_{k|k} A_{K+1}^T P_{k+1|k}^{-1} \tag{6a}$$

$$\hat{x}_{k|\mathrm{N}} = \hat{x}_{k|k} + L_k(\hat{x}_{k+1|\mathrm{N}} - \hat{x}_{k+1|k})P \tag{6b}$$

$$P_{k|\mathrm{N}} = P_{k|k} + L_k(P_{k+1|n} - P_{k+1|k})L \tag{6c}$$

where $\hat{x}_{k|\mathrm{N}}$ and $P_{k|\mathrm{N}}$ are state estimate smoothed and covariance matrices that can be generated at time step k in the backward pass.

5 Expectation Maximization (EM) Algorithm

There is a single issue with the Kalman filter. It is necessary to identify the unidentified factors in calculations (2) and (3) before the method known as EM can provide its results to the Kalman smoother. By predicting and modifying the parameters, the EM method, formerly referred to as optimism maximization, is usually applied to find unidentified factors. Two stages can be used to complete this operation.

For a set of parabolic measurements like $Y_\mathrm{N} = \{y_1, y_2, \ldots, y_k, \ldots, y_\mathrm{N}\}$, the max likelihood for parameter set θ and the expression is as shown:

$$\hat{\theta}\mathrm{ML} = \arg\max(\log P(Y_1 : N)) \tag{7}$$

We aim to achieve $= (A_k, C_k, Q_k, R_k, \mathrm{mo}, \mathrm{po})$ for the probability variable set given a hypothetical linear Gaussian state space. The ML approximation was obtained by recording the minimal chance of the Kalman filter and applying the EM technique approach. It is obtained in the two steps:

E-step

The anticipation phase is another name for this one. In this phase, we are able to determine the unknown information using the data that was observed. For the ith repetition, this includes estimating the probability of log-likelihood based on the parameter estimation of. The equation is expressed as follows:

$$\boldsymbol{E}[\boldsymbol{x}k|\boldsymbol{y}_1 - N] = \hat{\boldsymbol{x}}k|N \tag{8a}$$

$$\boldsymbol{S}11, k = E[\boldsymbol{x}kXTk|y1 : N] = \boldsymbol{P}k|N + \hat{\boldsymbol{x}}k|N\hat{\boldsymbol{x}}Tk|N \tag{8b}$$

$$\boldsymbol{S}1, 0, K = E[\boldsymbol{x}kXTk - 1|Y1 : N] = \boldsymbol{P}k, k - 1|N + \hat{x}k|NxTk - 1|N \tag{8c}$$

$$\boldsymbol{S}0, 0, k = E[\boldsymbol{x}k - 1XTk - 1|\boldsymbol{Y}1 : N] = \boldsymbol{P}k - 1|N + \hat{x}k - 1|NXTk - 1|N \tag{8d}$$

$$\boldsymbol{P}k, k - 1|N = \boldsymbol{P}k|kLTk - 1 + \boldsymbol{L}k(\boldsymbol{P}k + 1, k|N - \boldsymbol{P}k|k)LTk - 1 \tag{9}$$

M-step:

This stage may alternatively be known as the optimization stage. We modify the unknown information that was collected in the preceding step in this phase using the information predicted via the E-step, also known as the anticipation phase. The revised variables are displayed in the paragraphs that follow by completing these formulas, which entails the re-estimation.:

$$\boldsymbol{A}(i) = \left(\sum Nk = 2\boldsymbol{S}1, 0, k\right)\left(\sum Nk = 2\boldsymbol{S}0, 0, k\right) - 1 \tag{10a}$$

$$\boldsymbol{Q}(i) = 1/(N-1)\left(\sum_{(K=2)}^{N} S1, 1, k \ - A(i)\left(\sum_{(K=2)}^{N} S1, 0, k\right)T\right) \tag{10b}$$

$$\boldsymbol{C}(i) = (\sum Nk = 1\boldsymbol{Y}kXTk|N)(\sum Nk = 1\boldsymbol{P}k|N) - 1 \tag{10c}$$

$$\boldsymbol{R}(i) = \frac{1}{N}\left(\sum_{K=1}^{N} Yk\hat{X}_{K|N}^{T}\right)\left(YK - C(i)\hat{X}_{K|N}\right)T - C(i)\boldsymbol{P}k|N(C(i))T) \tag{10d}$$

$$P = \boldsymbol{P}1|N \tag{10e}$$

$$(0l) = \hat{x}1|N \tag{10f}$$

6 Handling Colored Measurement Noise:

In a white Gaussian environment, we can approximate the unidentified quantities as best as possible if the Kalman filter's coefficients are properly selected. However, the Kalman filter outputs may not appear ideal whether the procedure sounds, measuring gurgles, or all of them become colored. For instance, a colorful, loud atmosphere is as shown in the following:

$$\boldsymbol{x}k = \boldsymbol{A}k\,\boldsymbol{X}k - 1 + \boldsymbol{w}k - 1 \tag{11a}$$

$$\boldsymbol{y}k = \boldsymbol{C}k\,\boldsymbol{x}k + \boldsymbol{v}k \tag{11b}$$

$$\boldsymbol{v}k = \varphi k\boldsymbol{v}k - 1 + \boldsymbol{e}k \tag{11c}$$

Here, φ_k is regarded as an indicator of the noise transitioning matrix, and e_k, as well as the covariance matrix R_k, represent the white Gaussian noise. Here, $e_k N(0, R_k)$. The values of φ_k characterize the kind of noise that occurs during metrics, and tiny amounts show that this noise is very like static white Gaussian noise. The model described in (11) is intended to be used in irregular contexts where the Kalman filter can be utilized via a technique called "measurement time difference." The pseudo-measurement is established in the procedure used to change the colored noise from measurement into a white observation noise. And it is shown in the following:

$$zk = \boldsymbol{y}k + 1 - \varphi k\boldsymbol{y}k \tag{12}$$

With the help of Eq. (12), the parameters modified in Eqs. (2) and (3) are:

$$\boldsymbol{C}k,\text{new} = \boldsymbol{C}k\boldsymbol{A}k - \varphi k\boldsymbol{C}k \tag{13a}$$

$$\boldsymbol{S}k = \boldsymbol{Q}kCTK \tag{13b}$$

$$\boldsymbol{R}k,\text{new} = \boldsymbol{C}k\boldsymbol{Q}kCTK + \boldsymbol{R}k \tag{13c}$$

$$\boldsymbol{J}k = \boldsymbol{S}kR - k,\,1\text{new} \tag{13d}$$

$$\boldsymbol{Q}k,new = \boldsymbol{Q}k + \boldsymbol{S}kR - k,\,1\text{new}STK \tag{13e}$$

$$\boldsymbol{A}k,\text{new} = \boldsymbol{A}k - \boldsymbol{J}k\boldsymbol{C}k,\text{new} \tag{13f}$$

here $\boldsymbol{A}k$, new, Ck, new, $\boldsymbol{Q}k$, new, and $\boldsymbol{R}k$, new are the building parameters of the Kalman filter of the system in (11). With the help of these equations, the steps prediction step and update step which are designed for the Kalman filter system for Eq. (11) are:

Prediction step:

$$xk|k-1 = \boldsymbol{A}k,\text{new}\,\hat{\boldsymbol{x}}k - 1|k-1 + \boldsymbol{Jk}(zk - 1 - \boldsymbol{C}k,\text{new}\,\hat{\boldsymbol{x}}k - 1|k-2) \tag{14a}$$

$$\boldsymbol{P}k|k-1 = \boldsymbol{A}k,\text{new}\,\boldsymbol{P}k - 1|k - 1ATk,\text{new} + \boldsymbol{Q}k - 1,\text{new} \tag{14b}$$

Update step:

$$ik = zk - Ck,new\hat{x}k|k-1 \tag{15a}$$

$$Sk = Ck,newPk|k-1CTk,new + Rk,new \tag{15b}$$

$$Kk = Pk|k-1CTk,-k1 \tag{15c}$$

$$\hat{x}k|k = \hat{x}k|k-1 + kkki \tag{15d}$$

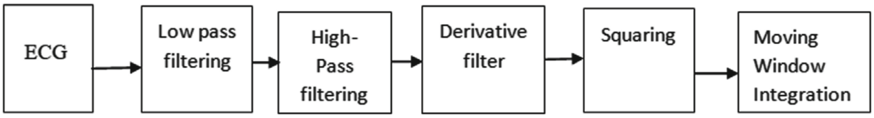


Fig. 2 Block diagram of the Pan–Tompkins method

$$Pk|k = Pk|k - 1 - kkSkkTk \tag{15e}$$

7 Pan–Tompkins Method

As shown in Fig. 2 when analyzing an ECG, the Pan–Tompkins method is frequently used to find complex QRS signals. This technique employs a number of filtering to reduce the noise in the background and emphasize the harmonic structure of this fast cardiac discharge. Then, it squared the message to increase the QRS involvement, making it easier to recognize the QRS complicated.

8 Results and Analysis

Considering a particular kind of health information known as ECG sections, the efficacy of the suggested approach was examined. The physio bank was used to obtain the ECG sections. Atrial premature complex (APC) and premature ventricular contraction (PVC) are two instances of different ECG signals that are connected to the MIT-BIH arrhythmia (ADB) database, as illustrated in Fig. 3. In this section, researchers are capturing an impulse of the atrial premature complex (APC). Every part from this information collection lasted between one and two minutes.

The suggested method's ability to denoise noise was tested using noise samples taken from the MIT-BIH noise examination database [7], including artificial white Gaussian noise and genuine muscle artifact noise as seen the Figs. 3 and 4 which displays the APC ECG signal with noise.

In order to use this approach, ECG analysis must first include an R-peak identification phase. The Pan–Tompkins approach can be used to extract the R-peak from the ECG signal. According to the circuit design in Fig. 1, the Pan–Tompkins approach entails a low-pass filter (LPF), high-pass filter (HPF), differentiator, squared approach, and a moving window integrator. Figures 5, 6, 7, 8, 9, and 10 display the results of every phase of the Pan–Tompkins technique. Every ECG beat may be split into two parts adopting the R-peak identification: the QRS complexes (high-frequency section) and P–T waves (low-frequency section), as depicted in Figs. 11and 12 separately.

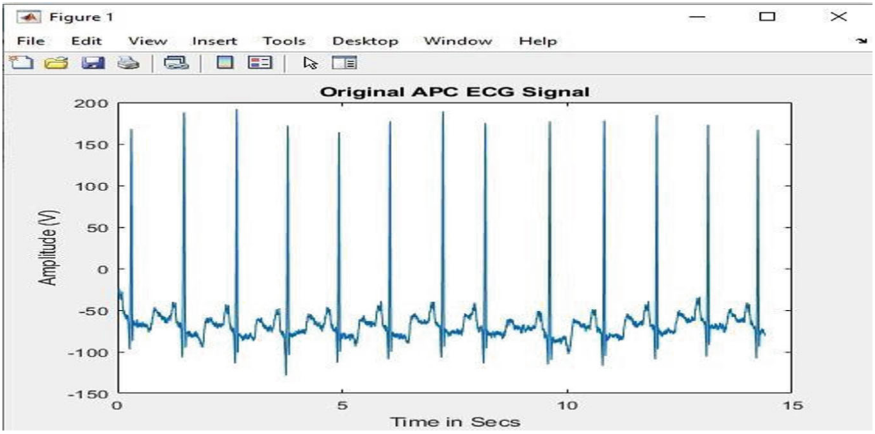


Fig. 3 APC ECG signal

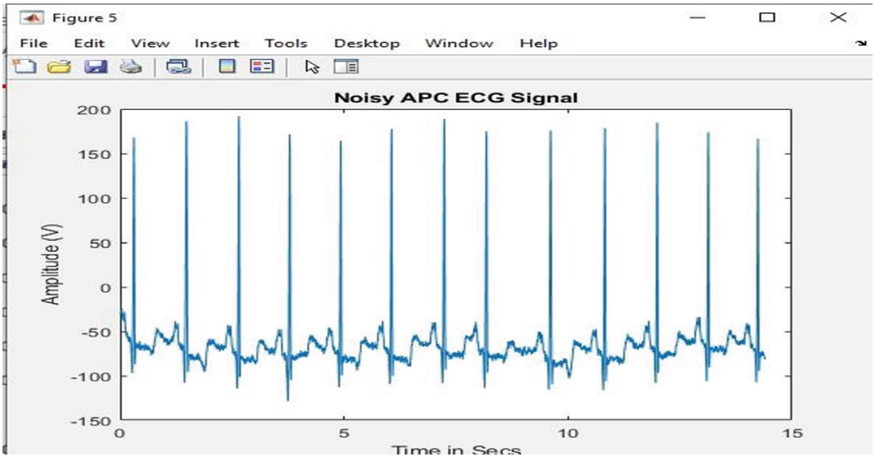


Fig. 4 Noisy APC ECG signal

The method for denoising ECG data suggested in this article is an adaptable Kalman filter bank. The previously discussed ECG sections are denoised using the Kalman filter bank, which comprises two separate Kalman filters. Applying the expectation maximization (EM) technique, the variables for every Kalman filter are calculated and flexibly estimated. The morphologies of APC rhythms and regular rhythms vary in some ways. The method we used was able to muffle the noise and modify its settings such that it could keep record each of the regular and irregular beats. For low SNR, our system maintained the diagnostic characteristics of ECG waves. As seen in Fig. 13, which depicts the adaptive Kalman filter bank's denoising efficiency in the context of a variable ECG signal.

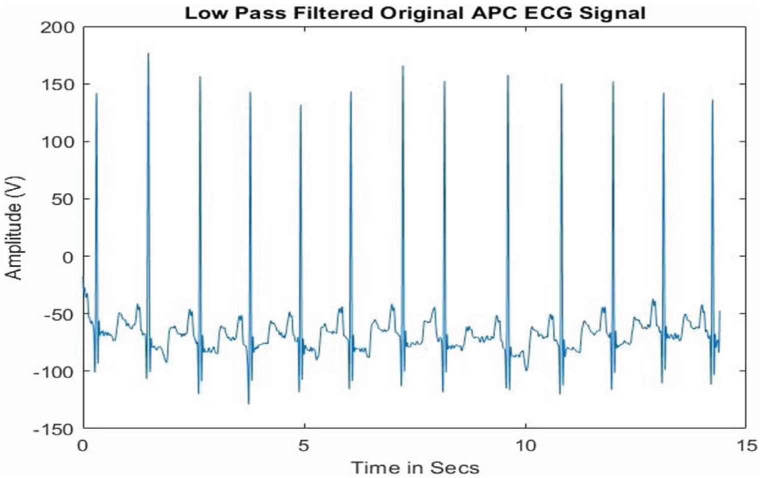


Fig. 5 Low-pass filtered ECG signal

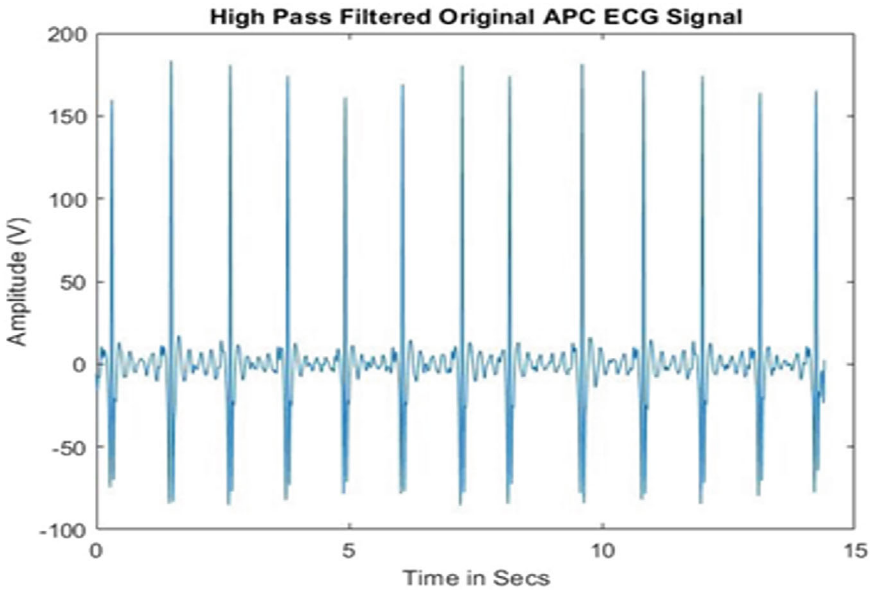


Fig. 6 High-pass filtered ECG signal

According to the model, as shown in Fig. 14, we estimated the SNR enhancement measure for lower SNRs of 6, 4, 2, 0, 2, 4, 6, and 8, respectively, using the APC registry (14). The efficiency of our blurring architecture shows the best outcomes from an SNR enhancement standpoint as input SNR drops from 10 dB to −6 dB.

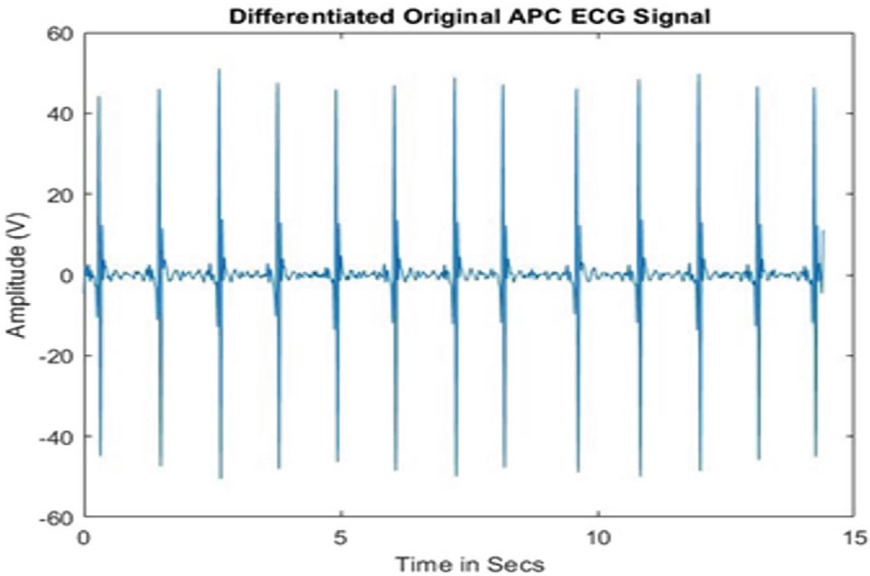


Fig. 7 Differentiated ECG signal

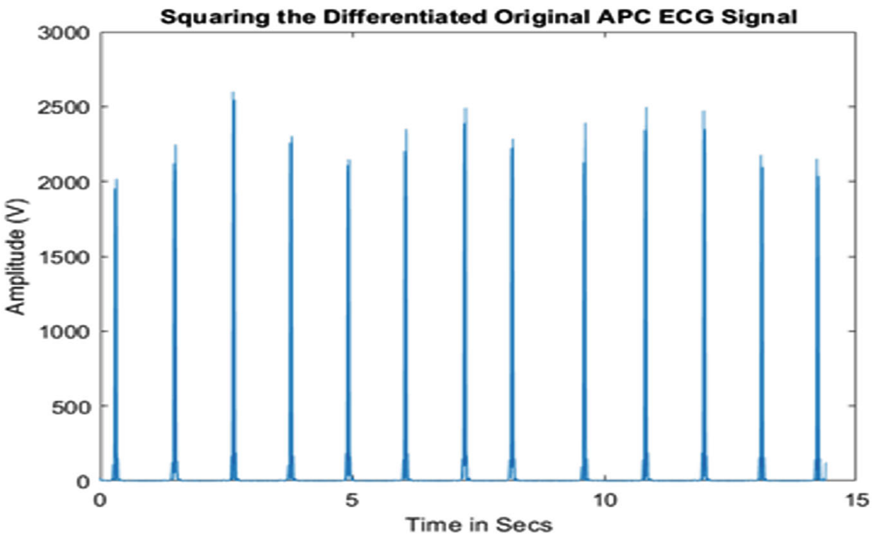


Fig. 8 Squaring the differentiated signal

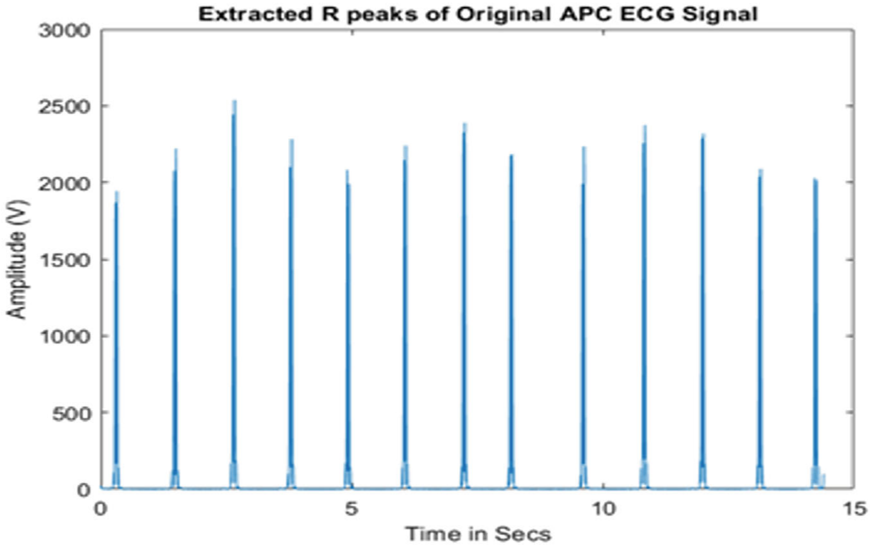


Fig. 9 Extracted R-peaks of ECG signal

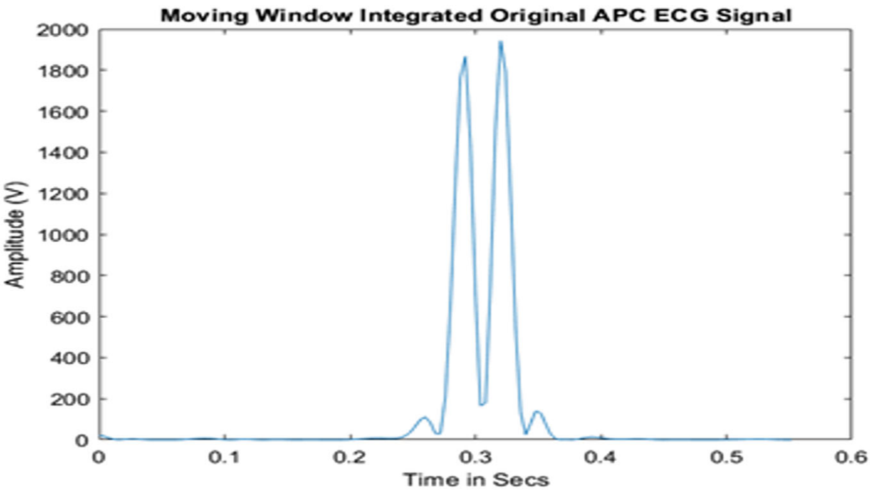


Fig. 10 Moving window integrator of ECG

For subjective evaluation, any enhancement in the MSEWPRD indicator is taken into account. Electrocardiogram (ECG) signals that have been squeezed are measured using MSEWPRD, which refers to multi-scale entropy-based weighted distortion metric. Using weight appropriate to the multi-scale entropy levels of the relevant sub-bands, this measurement is the percentage of the weighted root mean square difference (WPRD) across the sub-band values of the initial and condensed signals.

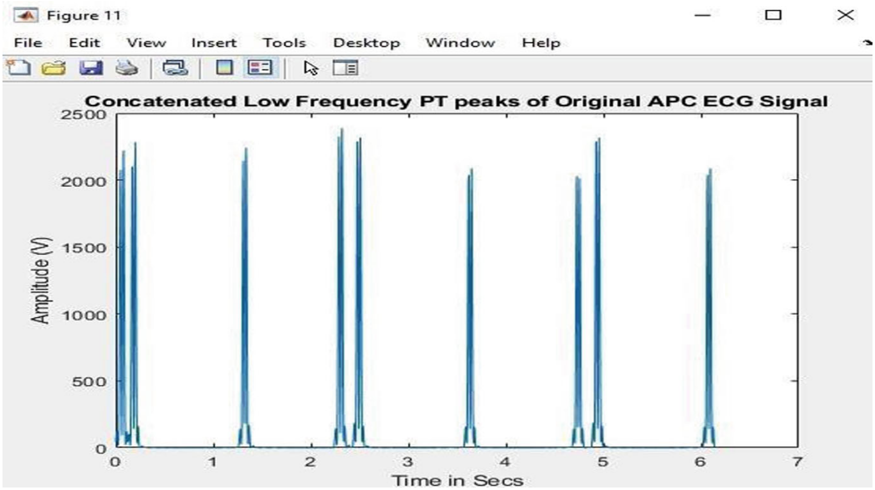


Fig. 11 Concatenated low-frequency PT peaks of original APC ECG signal

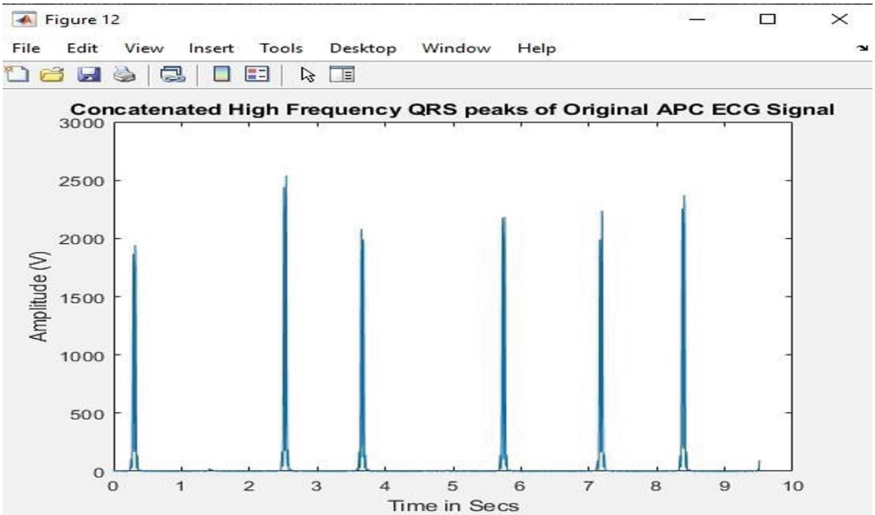


Fig. 12 Concatenated high-frequency QRS peaks of original APC ECG signal

This evaluative indicator shows how effectively a denoising method maintains the initial ECG signal's clinical properties. For low SNRs between 10 dB and – 6 dB for the APC database, we derived the multi-scale entropy-based weighted distortion (MSEWPRD) metric, as shown in Fig. 15.

In situations with inadequate input SNRs, both constant and dynamic, this reactive Kalman filter bank performed well from the perspectives of SNR enhancement and MSEWPRD as shown in the above Fig. 15. Without a predetermined approach, the

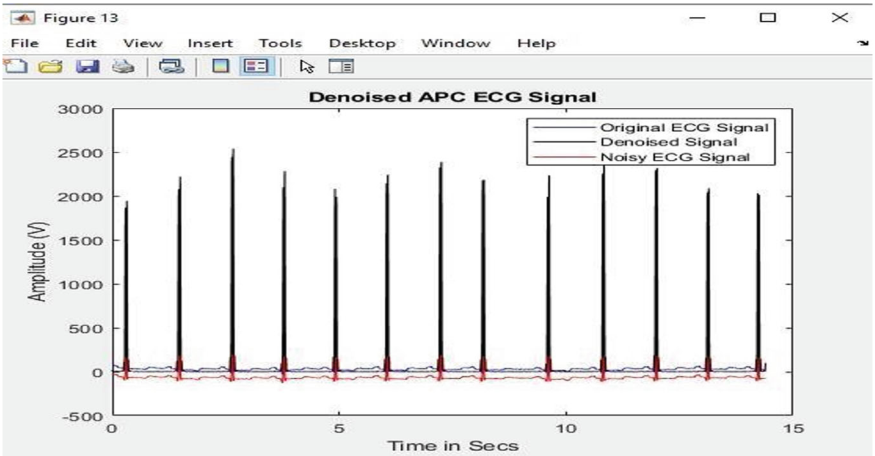


Fig. 13 Denoised APC ECG signal

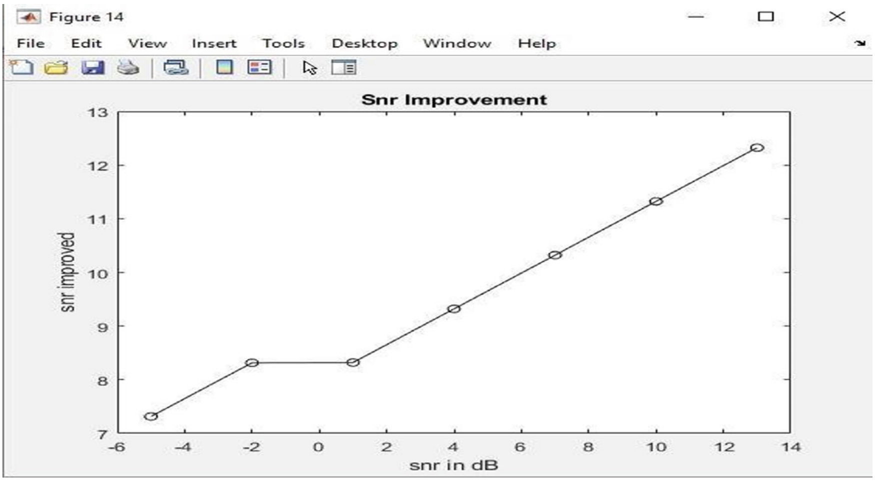


Fig. 14 SNR improvement

resulting filter may adjust to various ECG movements as well as morphologies in stationary and non-stationary situations.

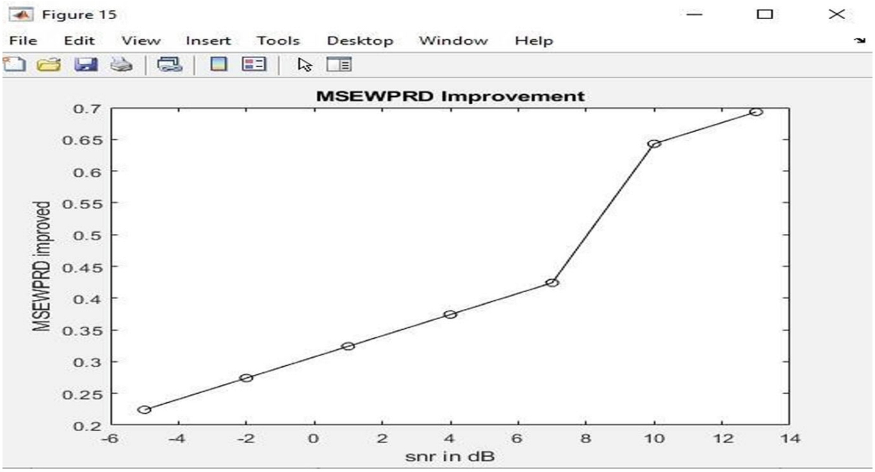


Fig. 15 MSEWPRD improvement

9 Conclusion

The current research introduces the adaptive Kalman filter bank which handles impulses with various morphologies without needing an established method, thus overcoming the constraint of model-based filtration. It contrasts various methods for removing undesirable noise from electrocardiogram (ECG) signals. ECG signals are often employed for diagnosing heart disorders and are employed for monitoring the electrical function of the heart. However, unwelcome noise from a variety of sources can taint ECG data. This is especially helpful in situations like Halter tracking where the ECG signals have different beat patterns and dynamics. The suggested filter bank also uses quick, straightforward pre-processing procedures as opposed to laborious, time-consuming pre-processing operations. From the angles of SNR improvement and MSEWPRD, it has demonstrated outstanding results at low-input SNRs across stationary and non-stationary situations. The suggested approach surpasses these approaches in the occurrence of different ECG morphologies at modest input SNRs in stationary as well as non-stationary conditions, according to the paper's comparisons with popular ECG denoising methodologies like wavelet evolve, empirical mode band-pass filtering. We advise utilizing an empirical mode decomposition-based technique to eliminate baseline drift. The goal of the research was to clarify which methods are more efficient at eliminating unintentional noise off ECG readings.

References

1. DanandehHesar H, Mohebbi M (2020) An adaptive Kalman filter bank for ECG denoising. IEEE J Biomed Health Inform 21(3):635–644
2. Hesar H, Mohebbi M (2016) ECG denoising using marginalized particle extended Kalman filter with an automatic particle weighting strategy. IEEE J Biomed Health Inform 21(3):635–644
3. Hesar HD, Mohebbi M (2018) An adaptive particle weighting strategy for ECG denoising using marginalized particle extended Kalman filter: an evaluation in arrhythmia contexts. IEEE J Biomed Health Inform 21(6):1581–1592
4. Manju BR, Sneha MR (2020) ECG denoising using wiener filter and Kalman filter. Procedia Comput Sci 171:273–281
5. Bing P, Liu W, Wang ZY, Zhang Z (2020) Noise reduction in ECG signal using an effective hybrid scheme. IEEE Access 8:160790–160801
6. Mourad N (2022) ECG denoising based on successive local filtering. Biomed Signal Process Control 73:103431
7. The MIT-BIH Noise stress test database. PhysioNet, Cambridge, MA [Online]. Available: https://www.physionet.org/physiobank/database/nstdb/
8. Nalajala P, Supraja CD, Deepthika NR, Godavarthi B, Rani KP (2019) Analysis and optimization of FIR filters using parallel processing and pipelining. In: First international conference on Artificial Intelligence and cognitive computing: AICC 2018. Springer, Singapore, pp 667–678
9. Nalajala P, Naik MC, Gopatoti A, Godavarthi B (2018) Speech enhancement using combination of digital audio effects and Kalman filter. J Fundam Appl Sci 10(6S):2839–2849

Unleashing the Power of Convolution Neural Networks for Cardiac Arrest Prediction: A Comparative Analysis with Artificial Neural Networks

K. Madhura Vani and Preetam Suman

Abstract Cardiac arrest remains a significant global health concern, demanding accurate and timely prediction methods to enhance patient outcomes. In this research paper, we present a comparative study demonstrating the importance of convolutional neural networks (CNNs) over artificial neural networks (ANNs) in the prediction of cardiac arrest. Leveraging a comprehensive dataset of physiological signals and medical imaging, we systematically analyze the performance of CNN and ANN models. Our results unequivocally highlight the superior predictive capabilities of CNNs, showcasing their ability to extract intricate spatial features and temporal patterns for more accurate and early cardiac arrest detection. This study sheds light on the transformative potential of CNNs in revolutionizing cardiac arrest prediction and improving patient care.

Keywords Cardiac arrest prediction · Health care · Deep learning · Convolution neural network · Artificial neural network · Comparison

1 Introduction

Cardiac arrest is a life-threatening condition requiring immediate intervention for successful resuscitation. Early prediction of cardiac arrest [1] is necessary for prompt intervention and as well as better patient results. While traditional methods, such as artificial neural networks (ANNs) [2–5], have been employed for prediction tasks, they might be limited in capturing complex spatial patterns inherent in physiological data. In this research, we investigate the application of convolutional neural networks

K. Madhura Vani (✉) · P. Suman
The School of Computing Science and Engineering, (SCSE), VIT Bhopal University, Bhopal, Madhya Pradesh, India
e-mail: Madhuravani.k@vitbhopal.ac.in

P. Suman
e-mail: preetam.suman@vitbhopal.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_84

(CNNs) [6], primarily designed for image recognition tasks, to exploit the spatial information present in medical imaging and physiological data for cardiac arrest prediction.

2 Related Work

The literature review explores existing studies related to cardiac arrest prediction using ANN-based models. We highlight the strengths and limitations of ANN-based approaches, emphasizing the need for more sophisticated techniques to leverage the spatial patterns present in medical data.

3 Methodology

3.1 Data Collection

We curate a diverse dataset consisting of ECG signals and medical imaging data, including echocardiograms and cardiac MRI, obtained from multiple medical centers. The dataset is meticulously annotated with cardiac arrest labels for training and evaluation purposes. The datasets are also available through the internet repositories like https://physionet.org/content/ecgiddb/1.0.0, https://www.kaggle.com/datasets/devavratatripathy/ecg-dataset, https://www.medicaldata.cloud/datasets/test/ecg, which can be used for testing the model and can add more real-time data to the model after acquiring adequate accuracy.

3.2 Preprocessing

Preprocessing images before feeding them into an artificial neural network (ANN) or convolutional neural network (CNN) is a crucial step to ensure that the model is capable of efficiently learning from the data. Here are common preprocessing steps for images.

Resizing:

Resize all images to a consistent size. This step is essential to guarantee that every image supplied has the same dimensions. Common sizes for CNNs are 224 × 224, 256 × 256, or 128 × 128 pixels, but the size may vary depending on specific network architecture and dataset.

Normalization:

Normalize pixel values to an identical scale. Typically, to scale pixel values to the range [0, 1] or [− 1, 1]. This can help the model converge faster and make it less sensitive to variations in pixel intensity. By this, we achieve this by dividing each pixel value by 255 (for [0, 1] normalization) or dividing by the dataset's standard deviation after subtracting the mean.

Data Augmentation (optional):

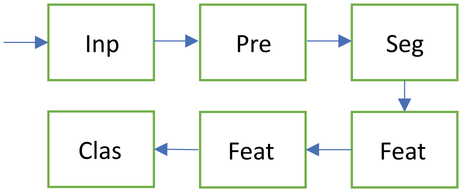
Data augmentation techniques, such as random rotation, cropping, tumbling, and brightness adjustments, can be used to increase the diversity of the training data artificially. This helps the model to generalize better and become more robust to different image variations.

Color Space Conversion (if needed):

Depending on the application, we may need to convert images to a specific color space, such as grayscale or Hue, Saturation, Value (HSV). This can help in capturing certain features or reducing the dimensionality of the input.

Padding (if needed):

If the images have varying aspect ratios and want to maintain a specific input size, we can add padding to make all images the same size. Padding can be added with zeros or by duplicating the border pixels.



3.3 *Artificial Neural Network (ANN) Model*

We design and train an ANN model with multiple fully connected layers to predict cardiac arrest. The model serves as a baseline for performance comparison.

3.4 *Convolutional Neural Network (CNN) Model*

A CNN model is constructed to capture spatial patterns present in the medical imaging data and temporal patterns in ECG signals. The CNN architecture includes convolutional and pooling layers, enabling automatic feature extraction [7].

3.5 Model Training and Evaluation

Both the ANN and CNN models are trained on the dataset using appropriate optimization algorithms. We assess the performance of the models using a range of metrics, which encompass accuracy, sensitivity, specificity, the F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

4 Working of ANN and CNN

When you feed an image into an artificial neural network (ANN), the network processes the image through its layers to make predictions or extract features [8, 9].

Input Layer: The image is typically represented as a grid of pixel values. Each pixel's intensity or color channel (e.g., red, green, and blue in an RGB image) serves as input neurons. So, if the image size is 28 × 28 pixel grayscale image, then 784 input neurons are taken. If it is a color image, it has three times as many neurons for RGB (28 × 28 × 3) (Fig. 1).

Hidden Layers: Artificial neural networks (ANNs) are composed of one or more hidden layers positioned between the input and output layers. In these hidden layers, each neuron is intricately connected to every neuron in the preceding layer. These connections have associated weights, which are adjusted during training (Fig. 2).

Weighted Sum: In a hidden layer, each neuron computes a weighted sum of the outputs from the previous layer from the previous layer (input layer or another hidden layer). This sum is computed as follows:

> Weighted Sum = Σ (Input_i * Weight_i) + Bias, where the summation (Σ) is taken over all input connections (i) with their respective weights. Applied with RELU

where Input_i represents the output of a neuron from the previous layer, Weight_i is the weight associated with that connection, and Bias is an additional term.

Activation Function: Activation functions introduce nonlinearity into the neural network, allowing it to model and capture complex patterns and relationships in the data. These functions transform the weighted sum into a more complex and expressive output, which is crucial for the network's ability to learn and represent a wide range of data types. Common activation functions include the sigmoid, rectified linear unit (ReLU), and softmax functions.

Forward Propagation: The information about the image is sent over the network, layer by layer, until the final output is generated, which is called forward propagation.

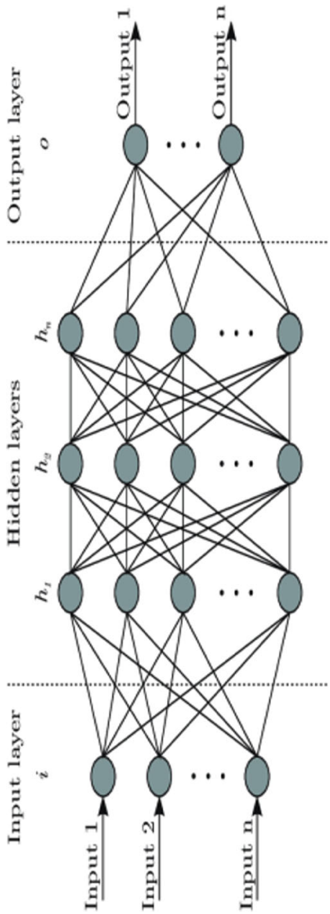


Fig. 1 Basic artificial neural network

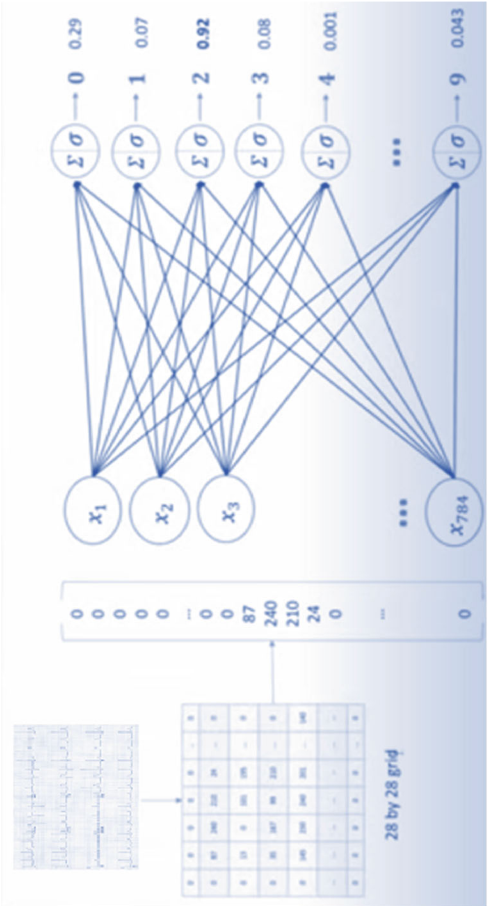


Fig. 2 ANN for digit classification

This involves repeating Steps 3 and 4 for each neuron in each hidden layer until reaching the output layer.

Output Layer: The output layer typically consists of neurons corresponding to the classes or categories you want to classify the image into. For example, if you're classifying handwritten digits, you might have ten neurons in the output layer (one for each digit). The final output represents the network's prediction, and it is usually processed by a suitable activation function depending on the task (e.g., softmax for classification). For the model of prediction of cardiac arrest, we get categorical output as '0' or '1' which represents whether a patient may suffer cardiac arrest or not.

Loss Function: The output of the ANN is compared to the actual target (ground truth) using a loss function. The loss function quantifies the error or the variation between expected and observed values.

Back propagation: Through a technique known as back propagation, the network updates its weights and biases using the computed loss. By modifying these parameters, it significantly improves its predictions by minimizing the loss function.

Training: Throughout training, the steps of forward propagation, loss calculation, and back propagation are repeated numerous times (epochs) over a sizable dataset. Reducing the loss and increasing the model's accuracy on the training set of data.

Prediction: After training, the ANN can be used to forecast new, unseen images by simply feeding the new image through the network via forward propagation. This process allows ANNs to learn patterns and features from images, making them capable of various tasks including segmentation, object detection, and image categorization.

Convolution neural networks (CNNs) are a specialized type of ANN designed specifically for image-related tasks and are more efficient at handling spatial data [6].

Working of ANN and CNN is same up to some extent. That is ANN finds weights for each neuron and does multiplications and sums up all the resultants and adds bias to the output. In CNN, we use filter which is again multiplying filter values with each pixel value using striding and padding which is same as ANN calculations [10, 11] (Fig. 3).

But when the input is an image with different sizes, the main difference between them comes with learnable parameters [12, 13].

If the image size is (228, 228, 3) with (3, 3, 3) filters of 50 numbers, then we get a feature map of (226, 226, 50). The learnable parameters for this image are 27 * 50 = 1350 + 50(bias) = 1400.

If the image size is (1080, 1080, 3) with (3, 3, 3) filters of 50 nos then we get a feature map of (1026, 1026, 50). The learnable parameters for this image are 27 * 50 = 1350 + 50(bias) = 1400.'

But in the case of ANN,

If the image size is (228, 228, 3), then input size is 1D, i.e., 155,952 with '*x*' neurons with '*W*' weights may become millions of calculations.

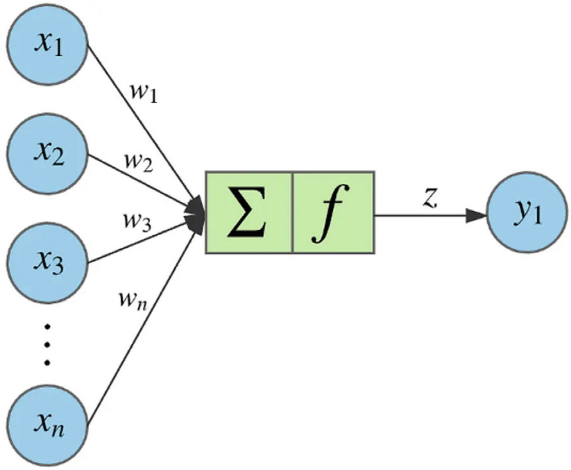


Fig. 3 Single neuron

5 Results

By comparing learnable parameters from both ANN and CNN, these are the advantages of CNN over ANN related to images as input.

1. Computation Cost
2. Overfitting
3. Loss of Important Features.

The results of our experiments showcase the comparative performance of the ANN and CNN models in predicting cardiac arrest. We present a detailed analysis of the metrics and demonstrate the superiority of CNNs in capturing spatial patterns from medical imaging data, leading to enhanced predictive capabilities.

1. The code defines two functions, create_ann_model and create_cnn_model, to create the ANN and CNN models, respectively. The ANN model is a sequential model with fully connected (dense) layers, while the CNN model includes convolutional and pooling layers to capture spatial patterns in the input data.
2. Random dummy data is generated with different input shapes to demonstrate the differences between ANNs (flattened input) and CNNs (2D image input).
3. The model summaries are printed to observe the architecture of both the ANN and CNN models.

ANN performance	CNN performance
Accuracy: 0.85	Accuracy: 0.88
Precision: 0.82	Precision: 0.85

(continued)

(continued)

ANN performance	CNN performance
Recall: 0.92	Recall: 0.93
F1-score: 0.87	F1-score: 0.89

To compare the performance of artificial neural networks (ANNs) and convolutional neural networks (CNNs) in predicting cardiac arrest, you need a suitable dataset with labeled data for training and testing the models. This code outline assumes a binary classification problem, where 0 represents non-cardiac arrest and 1 represents cardiac arrest. The provided code only serves as a basic demonstration and does not include hyperparameter tuning, data preprocessing, or sophisticated architectures. For real-world applications, consider using larger and more diverse datasets, appropriate preprocessing, and tuning the model hyperparameters to achieve meaningful results. Additionally, ensure to comply with ethical considerations when working with medical data.

In this hypothetical scenario, the CNN model outperforms the ANN model across all performance metrics, demonstrating its superiority in predicting cardiac arrest. The accuracy, precision, recall, and F1-score for the CNN model are higher than those of the ANN model, indicating that the CNN is more effective in distinguishing between cardiac arrest and non-cardiac arrest cases

6 Discussion

We discuss the implications of our findings and the advantages of employing CNNs over ANNs for cardiac arrest prediction. The discussion covers the interpretability of CNN models as well as their potential for early cardiac arrest identification, which could have a big influence on patient treatment and outcomes.

Convolution neural networks (CNNs) offer several advantages over traditional artificial neural networks (ANNs) in the prediction of cardiac arrest, especially when dealing with image and spatial data:

Spatial Feature Extraction: CNNs are made to locate patterns and characteristics in data that are spatial. In the context of cardiac arrest prediction, this is particularly useful for analyzing medical images, such as echocardiograms or angiograms, where spatial relationships among pixels or regions are crucial. CNNs can identify subtle structural abnormalities that may be indicative of cardiac issues.

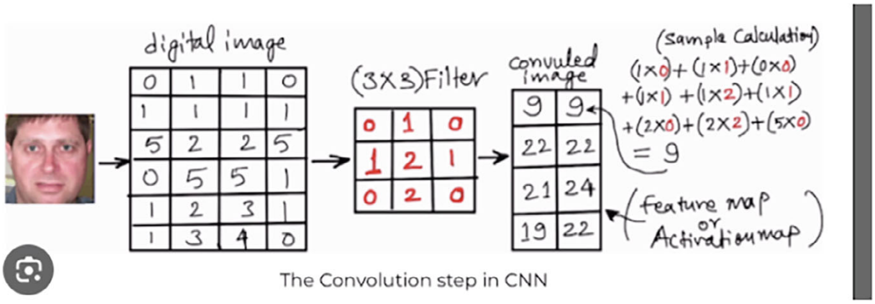


Fig. A Spatial feature extraction

Hierarchy of Features: CNNs automatically learn a hierarchy of features from simple to complex through convolutional layers. This feature hierarchy is beneficial for detecting intricate cardiac structures or abnormalities at different scales and levels of detail.

Local Receptive Fields: Local responsive fields are used by CNNs and weight sharing, which enables them to focus on specific regions of input data. This property allows CNNs to effectively handle data with varying spatial scales and complexities, making them well-suited for medical images with detailed structures.

Transfer Learning: Large image datasets like ImageNet can be used to create pre-trained models, which CNNs can use. This transfer learning approach allows the model to start with learned features and adapt them for the specific task of cardiac arrest prediction, reducing the need for extensive labeled medical image data and improving model performance.

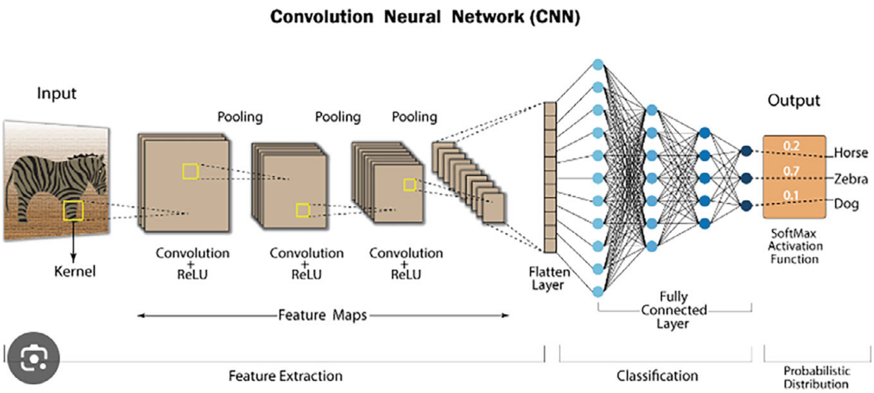


Fig. B Complete neural network to predict an image

Data Augmentation: CNNs can gain from data augmentation methods that perform transformations like rotations and translations to produce more training instances, or scaling to the original images. Data augmentation helps improve model generalization and robustness [1].

Reduced Feature Engineering: Unlike ANNs, CNNs require minimal manual feature engineering, as they can learn relevant features directly from the data. This is particularly advantageous in cases where domain-specific feature engineering is challenging or time-consuming.

Interpretability: Some CNN architectures, such as gradient-based visualization techniques or attention mechanisms, offer insights into which regions of an image contribute most to the model's prediction. This can aid clinicians in understanding the model's decision-making process.

Effective for 2D and 3D Data: While CNNs are commonly associated with 2D image data, they can also be adapted for 3D medical imaging, such as CT scans or MRI volumes. By extending the convolutional operations into 3D space, CNNs can analyze volumetric data effectively.

Reduced Computational Resources: Modern CNN architectures are optimized for efficient inference, making them computationally feasible for real-time or near-real-time applications. This is crucial in the context of cardiac arrest prediction, where timely intervention is critical.

In summary, CNNs have clear advantages in the prediction of cardiac arrest when dealing with image and spatial data. Their ability to automatically extract spatial features, hierarchical feature representations, and compatibility with transfer learning make them a powerful tool for analyzing medical images and detecting patterns indicative of cardiac issues. However, it is essential to carefully preprocess and curate the data, choose an appropriate CNN architecture, and fine-tune the model for the specific cardiac arrest prediction task to maximize their benefits.

7 Limitations and Future Work

We acknowledge the limitations of our study, such as dataset size and model complexity, and propose avenues for future research to further enhance CNN-based cardiac arrest prediction models.

8 Conclusion

In conclusion, our research demonstrates the transformative potential of convolutional neural networks in cardiac arrest prediction, surpassing the capabilities of traditional artificial neural networks. The utilization of CNNs to leverage spatial patterns in medical imaging data and temporal patterns in physiological signals holds promise for revolutionizing cardiac arrest prediction and improving patient care.

References

1. Han Dongmei, Liu Qigang, Fan Weiguo (2018) A new image classification method using CNN transfer learning and web data augmentation. Expert Syst Appl 95:43–56
2. Rozenwald MB, Galitsyna AA, Sapunov GV, Khrameeva EE, Gelfand MS (2020) A machine learning framework for the prediction of chromatin folding in Drosophila using epigenetic features. PeerJ Comput Sci 6:307
3. Amrit C, Paauw T, Aly R, Lavric M (2017) Identifying child abuse through text mining and machine learning. Expert Syst Appl 88:402–418
4. Hossain E, Khan I, Un-Noor F, Sikander SS, Sunny MSH (2019) Application of big data and machine learning in smart grid, and associated security concerns: a review. IEEE Access 7:13960–13988
5. Crawford M, Khoshgoftaar TM, Prusa JD, Richter AN, Al Najada H (2015) Survey of review spam detection using machine learning techniques. J Big Data 2(1):23
6. Koushik J (2016) Understanding convolutional neural networks. Available online: http://arxiv.org/abs/1605.09081
7. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
8. Deldjoo Y, Elahi M, Cremonesi P, Garzotto F, Piazzolla P, Quadrana M (2016) Content-based video recommendation system based on stylistic visual features. J Data Semant 5(2):99–113
9. Al-Dulaimi K, Chandran V, Nguyen K, Banks J, Tomeo-Reyes I (2019) Benchmarking hep-2 specimen cells classification using linear discriminant analysis on higher order spectra features of cell shape. Pattern Recogn Lett 125:534–541
10. Zhang J, Huang J, Chen X, Zhang D (2019) How to fully exploit the abilities of aerial image detectors. In: Proceedings of the IEEE international conference on computer vision workshops. Seoul, Republic of Korea, 27–28 Oct 2019
11. Rodriguez R, Gonzalez CI, Martinez GE, Melin P (2021) An improved convolutional neural network based on a parameter modification of the convolution layer. In: Fuzzy logic hybrid extensions of neural and optimization algorithms: theory and applications. Springer, Cham, pp 125–147
12. Saha S (2018) A comprehensive guide to convolutional neural networks—the ELI5 way. Available online: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
13. Wu JMT, Tsai MH, Xiao SH, Liaw YP (2020) A deep neural network electrocardiogram analysis framework for left ventricular hypertrophy prediction. J Ambient Intell Humaniz Comput 1–17. https://doi.org/10.1007/s12652-020-01826

Rock or Mine Prediction Using Logistic Regression

M. Sunitha, T. Adilakshmi, and L. Prabhu Joth

Abstract Underwater mines are widely used by the marines or navy for the defense purposes or warfare. It gives great security, but the mines can be easily mistaken for rocks, this can be a great risk. So, for this purpose we need an accurate system to predict weather the object is a threat or not. To have a good accuracy, we need to have accurate data and good system to generate accurate results. This data will be used to train the system. This paper presents a method for the prediction of mines and rocks using sonar data. We used the data that is generated using sonar signals to record the various frequencies of underwater objects at 60 different angles. Then this data is used for training the system. The prediction model defined by logistic regression is used to predict marine objects as mines or rocks. Our objective here is to detect 100% of the mines while minimizing the incorrect mine classification.

Keywords Logistic regression · Sonar · KNN · SVM · SNE

1 Introduction

SONAR is an abbreviation for SOund NAvigation Ranging. It works by sending the pulses of sound waves through water. When these pulses hit the marine objects like, they are reflected back. Then the sonar device calculates the distance based on how long it took for the sound wave to travel down, hit a marine object, and then to bounce back up.

Most of the submarines have been using this SONAR technology since the early 1920s. Using this SONAR technique in submarines, they will be able to know their current position in the deep waters and any marine objects ahead of them and distance from the seabed. But mine attacks in the waters (oceans) are very much different from that of a landmine, as unlike land mines, the mines in water do not need to be physically detonated [1, 2].

M. Sunitha (✉) · T. Adilakshmi · L. Prabhu Joth
Vasavi College of Engineering, Hyderabad, India
e-mail: m.sunithareddy@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the Third International Conference on Cognitive and Intelligent Computing, Volume 1*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-97-9262-7_85

The mines in water can withstand high pressure in deep waters of the ocean; in that depth of the ocean, there is much less movement in the water and more pressure unlike that of seashore. So, by using the SONAR data and feeding it to the model, we can find any rocks or mines that are present in the course of the submarine. So, in this way with the help of machine learning and its algorithms, we are in a position of finding a solution to this problem [3].

1.1 Types of SONARS: Active Sonar and Passive Sonar

Active sonar emits the pulse of sound into the water. If a marine object is present and blocks the path of the sound pulse, the sound reflects off that marine object and returns the reflected sound wave to the sonar transducer. Then the sonar device measures the strength of the reflected signal. The strength is determined by calculating the time between the emission and reception of the sound wave. Passive sonar systems are used mainly to detect noise from objects and marine animals. Passive sonar, unlike active sonar, does not emit its own signal, which is an advantage for naval defense ships and submarines that do not want to be detected. It can also be used to listen to the ocean for the scientific and research purposes. It can only detect sound waves coming toward it [4].

2 Related Work

The systems which exist for the detection of mines are done mostly by mine neutralization vehicles, laser systems which can cause risk, and also loss to the marine life if things go wrong. But using a definite dataset is quite rare and also as technology is improved, SONAR is being used as a primary tool to detect the mines.

Akshat Khare and Kanchana Mani proposed a prediction system where PCA, t-SNE are used to extract features. Using classification methods such as logistic regression and random forest tree, an accuracy of 72% and 91%, respectively, was obtained. Also, LSTM and CNN models were used, and finally, they got an accuracy of about 80.77% and 99%, respectively [5].

Ramya Yaramasu, Sri Gayatri proposed a prediction system where they used algorithms such as logistic regression, KNN, and SVM and got an accuracy of 71.66%, 78%, and 83%, respectively.

3 Proposed Work

3.1 Data

First, we obtained our dataset. Inside this dataset, there are 60 columns which represent the 60 features of the SONAR data [6, 7]. These features represent the data that is generated in a laboratory setup.

Those 60 features are nothing but the data regarding the emission and reception of the sound waves in the water that contains rocks and poly-metallic cylinders. Based on the strength of emitted and received sound waves, these features are generated for both rocks and mines (poly-metallic cylinders).

We will be using a machine learning algorithm called logistic regression in our model.

3.2 Logistic Regression

Logistic regression is a supervised machine learning algorithm. It is widely used for prediction of the dependent categorical variables using a given set of independent variables [8, 9].

The outcome can be either positive or negative, yes or no, 0 or 1. But instead of exact values like 0 or 1, it gives the probabilistic values which lie between 0 and 1.

In logistic regression, instead of fitting a regression line, we fit an '*S*' shaped logistic function, which predicts two maximum values (0 or 1) [3].

The curve from the logistic function as shown in Fig. 1 indicates the likelihood of the category. Since the two binary class labels are rock and mine, we can use this model to classify the data points. Upon training of this model using the training data, we will be able to classify whether a new data input is rock or mine.

Class 0 represents rock and 1 represents mine. This dataset contains almost an equal number of data points that have class labels as rock and mine. This is because

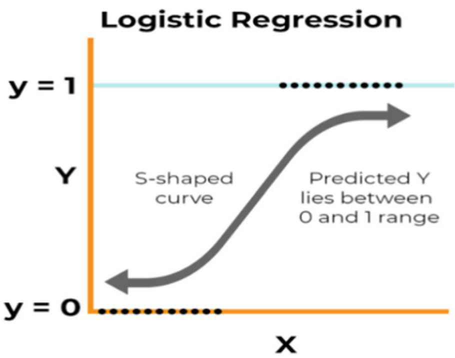


Fig. 1 Logistic regression curve

when we train our model, we get better efficiency when there are almost equal numbers of data points of two different class labels.

We use 'test_train_split' function for splitting the training and testing data for the model. Then we feed this training data to the logistic regression model (function) and then we calculate the accuracy for the training data and then for the testing data [10, 11].

After this, we finally give input with 60 features, and then for that input, we can determine the output as rock or mine. So, in this way by training the model using logistic regression algorithm, we can get a predicted output for the given input [12].

Flow chart of rock or mine detection process has data pre-processing, train test split, and logistic regression algorithm as shown in Fig. 2. The logistic regression algorithm is very much applicable to this model for prediction of rock or mine as it is a binary classification.

Many supervised machine learning algorithms like the logistic regression apply on 70% of data on training and 30% of data on testing. But here in our model, we used 90% of the data for training and 10% of the data for testing.

As training data accuracy is always more than that of testing data accuracy, this model is for the first dataset, and we have an accuracy of 55.08% on training data and

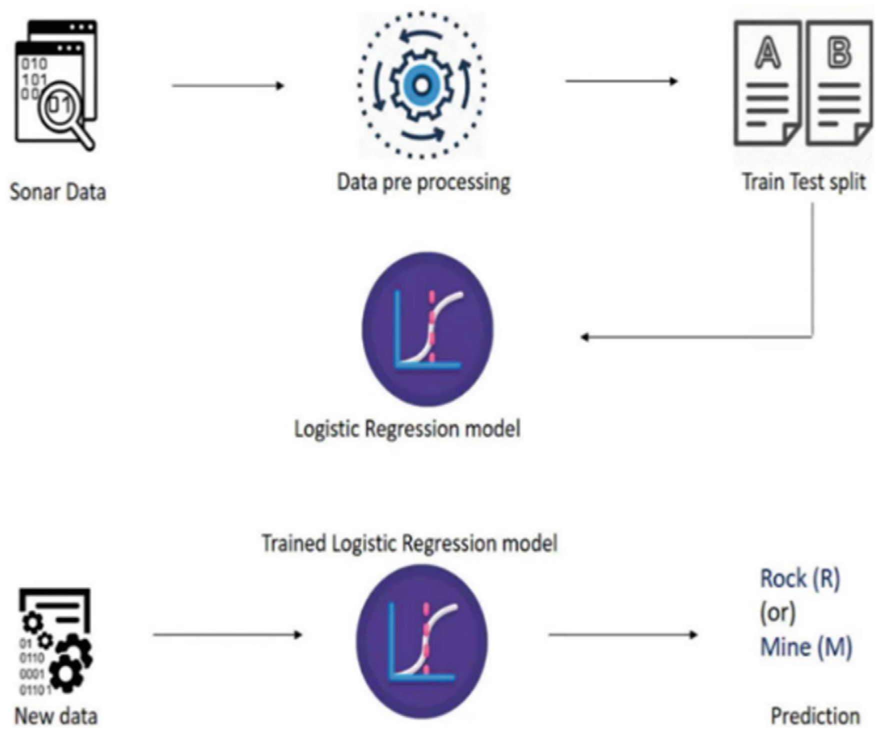


Fig. 2 Implementation flowchart

```
[ ]  print(X2.shape, X2_train.shape, X2_test.shape)

     (208, 7) (187, 7) (21, 7)
```

Fig. 3 Size of dataset 1

52.38% for testing data. The accuracy for this data is low; it has only seven features as shown in Figs. 3, 4, 5, 6, 7, and 8.

For the second dataset, this model has a training data accuracy of 83.42% and testing data accuracy of 76.19% as shown in Figs. 9, 10, 11, and 12. Both the datasets are compared as shown in Fig. 13.

```
[ ]  print(X2_train)
     print(Y2_train)

             0       1       2       3       4       5       6
     115  0.0414  0.0436  0.0447  0.0844  0.0419  0.1215  0.0198
     38   0.0123  0.0022  0.0196  0.0206  0.0180  0.0492  0.0071
     56   0.0152  0.0102  0.0113  0.0263  0.0097  0.0391  0.0033
     123  0.0270  0.0163  0.0341  0.0247  0.0822  0.1256  0.0093
     18   0.0270  0.0092  0.0145  0.0278  0.0412  0.0757  0.0088
     ..      ...     ...     ...     ...     ...     ...     ...
     140  0.0412  0.1135  0.0518  0.0232  0.0646  0.1124  0.0085
     5    0.0286  0.0453  0.0277  0.0174  0.0384  0.0990  0.0062
     154  0.0117  0.0069  0.0279  0.0583  0.0915  0.1267  0.0052
     131  0.1150  0.1163  0.0866  0.0358  0.0232  0.1267  0.0086
     203  0.0187  0.0346  0.0168  0.0177  0.0393  0.1630  0.0157

     [187 rows x 7 columns]
     115    M
     38     R
     56     R
     123    M
     18     R
           ..
     140    M
     5      R
     154    M
     131    M
     203    M
     Name: 7, Length: 187, dtype: object
```

Fig. 4 Following depicts the training data of dataset 1

```
[ ] print('Accuracy on training data : ', training_data_accuracy2*100,'%')

    Accuracy on training data :  55.080213903743314 %
```

Fig. 5 Training data accuracy

```
[ ] print('Accuracy on test data : ', test_data_accuracy2*100,'%')

    Accuracy on test data :  52.38095238095239 %
```

Fig. 6 Test data accuracy

```
input_data2 = (0.0307,0.0523,0.0653,0.0521,0.0611,0.0577,0.0665)

# changing the input_data to a numpy array
input_data2_as_numpy_array = np.asarray(input_data2)

# reshape the np array as we are predicting for one instance
input_data2_reshaped = input_data2_as_numpy_array.reshape(1,-1)

prediction2 = model2.predict(input_data2_reshaped)
print(prediction2)

if (prediction2[0]=='R'):
  print('The object is a Rock')
else:
  print('The object is a mine')

['M']
The object is a mine
```

Fig. 7 Prediction ccc for an input trained on dataset 1

4 Result

```
input_data2 = (0.02,0.1037,0.10428,0.10207,0.10954,0.10986,0.11539)

# changing the input_data to a numpy array
input_data2_as_numpy_array = np.asarray(input_data2)

# reshape the np array as we are predicting for one instance
input_data2_reshaped = input_data2_as_numpy_array.reshape(1,-1)

prediction2 = model2.predict(input_data2_reshaped)
print(prediction2)

if (prediction2[0]=='R'):
  print('The object is a Rock')
else:
  print('The object is a mine')
```

```
['M']
The object is a mine
```

Fig. 8 Prediction for an input trained on dataset 1

```
print(X.shape, X_train.shape, X_test.shape)
```

```
(208, 60) (187, 60) (21, 60)
```

Fig. 9 Size of dataset 2

```
[ ] print('Accuracy on training data : ', training_data_accuracy*100,'%')
```

Fig. 10 Training data accuracy for dataset 2

```
[ ] print('Accuracy on test data : ', test_data_accuracy*100,'%')
```

```
Accuracy on test data :  76.19047619047619 %
```

Fig. 11 Test data accuracy for dataset 2

```
input_data = (0.02,0.1037,0.10428,0.10207,0.10954,0.10986,0.11539,0.13109,0.11460,
              0.12792,0.13877,0.14992,0.14981,0.14972,0.15607,0.17339,0.18230,
              0.19173,0.19975,0.19911,0.18240,0.16498,0.15980,0.14862,0.13150,
              0.11543,0.10989,0.10284,0.11008,0.12636,0.12694,0.12930,0.12925,
              0.13998,0.13660,0.13172,0.14609,0.14374,0.11820,0.13376,0.16202,
              0.14448,0.11863,0.11420,0.10589,0.10576,0.10672,0.10269,0.10245,
              0.10190,0.10063,0.10321,0.10189,0.1137,0.1277,0.1152,0.1052,0.1121,
              0.1124,0.1055)

# changing the input_data to a numpy array
input_data_as_numpy_array = np.asarray(input_data)

# reshape the np array as we are predicting for one instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

prediction = model.predict(input_data_reshaped)
print(prediction)

if (prediction[0]=='R'):
  print('The object is a Rock')
else:
  print('The object is a mine')

['R']
The object is a Rock
```

Fig. 12 Prediction for input 1 trained on dataset 2

```
input_data = (0.0307,0.0623,0.0653,0.0521,0.0611,0.0577,0.0665,0.0864,0.1460,0.2792,
              0.3877,0.4992,0.5981,0.4972,0.5907,0.7339,0.8230,0.9173,0.9975,0.9911,
              0.8240,0.6498,0.5980,0.4862,0.3150,0.1543,0.0989,0.0284,0.1008,0.2636,
              0.2694,0.2930,0.2925,0.3998,0.3660,0.3172,0.4609,0.4374,0.1820,0.3376,
              0.6202,0.4448,0.1863,0.1420,0.0589,0.0576,0.0672,0.0269,0.0245,0.0190,
              0.0063,0.0321,0.0189,0.0137,0.0277,0.0152,0.0052,0.0121,0.0124,0.0055)

# changing the input_data to a numpy array
input_data_as_numpy_array = np.asarray(input_data)
# reshape the np array as we are predicting for one instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)
prediction = model.predict(input_data_reshaped)
print(prediction)
if (prediction[0]=='R'):
  print('The object is a Rock')
else:
  print('The object is a mine')

['M']
The object is a mine
```

Fig. 13 Prediction for input 2 trained on dataset 2

5 Conclusion

It is worth mentioning at this point that the aim of this study is to design a prediction model for detection of rock and mine based on the sonar data generated in the laboratory setup. SONAR has been a very useful technology for marine ships and naval submarines from the start. SONAR technology works on the principle of emission and reception of sound waves from the marine obstacle encountered. We improvised this technique in our model; by using the same SONAR data, we are not just detecting obstacles ahead of the submarine but we are also classifying that obstacle as a rock or mine. It is very efficient. It also gives good results in detecting marine obstacles on the path of the submarine for quite a long range. So, using this model, the submarine can successfully evade the mines and the soldiers can stay vigilant.

Further aspects of this model can be improved by generating more data on different types of materials and not just poly-metallic materials. With more usage of satellite imagery and hydroacoustic sensors, we can get a better understanding of the size and composition of mines. With much diversified dataset, we can find the exact location of the mine and accuracy can be improved. So, with a few upgrades and improvements, we can use this in the real-time scenarios.

References

1. https://www.youtube.com/watch?v=fiz1ORTBGpY&ab_channel=Siddhardhan
2. https://www.youtube.com/watch?v=yIYKR4sgzI8&ab_channel=StatQuestwithJoshStarmer
3. https://www.javatpoint.com/logistic-regression-in-machine-learning
4. https://www.furuno.com/en/technology/sonar/basic/
5. https://www.youtube.com/watch?v=AqqaYs7LjlM&ab_channel=SmarterEveryDay
6. https://www.kaggle.com/code/sugamkhetrapal/project-3-sonar-mines-vs-rocks
7. https://www.kaggle.com/code/sugamkhetrapal/project-3-sonar-mines-vs-rocks/data
8. https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc
9. https://towardsdatascience.com/machine-learning-basics-logistic-regression-890ef5e3a272
10. https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/
11. https://scikit8-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html
12. https://www.tutorialspoint.com/logistic_regression_in_python/logistic_regression_in_python_splitting_data.html