Lecture Notes in Electrical Engineering 783

Amit Kumar
Sabrina Senatore
Vinit Kumar Gunjan *Editors*

ICDSMLA 2020

Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications

Springer

Lecture Notes in Electrical Engineering

Volume 783

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

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

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

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

China

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

India, Japan, Rest of Asia

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

Southeast Asia, Australia, New Zealand

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

USA, Canada:

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

All other Countries:

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

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

More information about this series at https://link.springer.com/bookseries/7818

Amit Kumar · Sabrina Senatore ·
Vinit Kumar Gunjan
Editors

ICDSMLA 2020

Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications

Springer

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

Vinit Kumar Gunjan
Department of Computer Science
and Engineering
CMR Institute of Technology
Hyderabad, Telangana, India

Sabrina Senatore
Department of Computer Engineering,
Electrical Engineering and Applied
Mathematics
University of Salerno
Fisciano, Salerno, Italy

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-16-3689-9 ISBN 978-981-16-3690-5 (eBook)
https://doi.org/10.1007/978-981-16-3690-5

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

Preface

The world is filled with data, lots and lots of data. It does not look like it is going to slow down anytime soon. Machine learning brings the promise of deriving meaning from all of that data. Data Science brings together computational and statistical skills and machine learning for data-driven problem solving. This rapidly expanding area includes deep learning, large-scale data analysis and has applications in e-commerce, search/information retrieval, natural language modelling, finance, bioinformatics and related areas in artificial intelligence.

Data Science is the area of study which involves extracting insights from vast amounts of data by the use of various scientific methods, algorithms and processes. It helps you to discover hidden patterns from the raw data.

It is an interdisciplinary field that allows you to extract knowledge from structured or unstructured data. This technology enables you to translate a business problem into a research project and then translate it back into a practical solution. The term Data Science has emerged because of the evolution of mathematical statistics, data analysis and big data.

This book is comprised of selected and presented papers of the International Conference on Data Science, Machine Learning and Applications (ICDSMLA 2020) held at Pune, Maharashtra, India. It consists of selected manuscripts, arranged on the basis of their approaches and contributions to the scope of the conference. The chapters of this book present key algorithms and theories that form the core of the technologies and applications concerned, consisting mainly of genetic algorithms, evolutionary computation, soft computing, machine learning, neural networks, pattern recognition, intelligent control, optimization, nonlinear system and control, system theory and control theory, video and image processing, Internet security, big data, cellular networks. This book also covers recent advances in medical diagnostic systems, computational intelligence and advanced computing.

Hyderabad, India — Amit Kumar
Fisciano, Italy — Sabrina Senatore
Hyderabad, India — Vinit Kumar Gunjan

Contents

Automatic Notes Generation from Lecture Videos

D. R. Pratheeksha, R. P. Shreya Reddy, and R. Jayashree

Abstract Videos are an unavoidable part in academia. In day-to-day lives of either students or professors, we notice that we tend to refer to many online resources, majority of which are videos, as part of our learning process. This paper proposes a method to generate notes and summarized PowerPoint Presentations for non-blackboard style (i.e. videos without usage of images, diagrams or graphical media) lecture videos. The different phases involved are audio extraction from the uploaded video, speech recognition, punctuating the transcript, applying Natural Language Processing (NLP) techniques on the generated transcripts, generating Portable Document Format (PDF) from the text and creating PowerPoint Presentation by extracting important points from the text using K-means clustering and LexRank extractive text summarization.

Keywords Speech recognition · K-means clustering · LexRank · Natural language processing · Extractive text summarization

1 Introduction

Videos have elements of both voice and vision making them very effective for knowledge sharing and efficient learning. A beginner who is learning to code or a highly experienced professor who wants to research a topic, would definitely refer the various videos available online to acquire knowledge. Even though videos are such a great medium for learning, users need to spend more time if they choose to watch videos as compared to referring to text documents or similar online resources. Hence there is a need for a tool which helps the users to obtain the summary of a lecture. Therefore we developed a model which allows the users to upload a lecture video and obtain the summarized content of the video in the form of PDF and PPT.

D. R. Pratheeksha (✉) · R. P. Shreya Reddy · R. Jayashree
Department of Computer Science and Engineering, PES University, Bengaluru, Karnataka, India
e-mail: jayashree@pes.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_1

2 Related Work

There are not many effective models proposed for preparing notes for educational purpose especially for speakers having an Indian English accent. Miller [1] proposes a RESTful service based on Python which performs text embedding using the BERT model and utilizes K-means clustering to select sentences that are similar to the centroid for extractive text summarization. Although BERT performs better than other NLP algorithms for sentence embedding, it requires fine tuning the model for a particular type of lecture to fetch better results. This solution also does not effectively fill in the gaps for missing context in the summary. Sariki et al. [2] proposed a model which selects important words, sentences, phrases and paragraphs from documents which efficiently summarizes the document. The proposed approach is a combination of the statistical method and the semantic information present in the document. It consists of the following modules: Keyword/Keyphrase Extraction and Scoring, NER (Named Entity Recognition), Cue Phrase and Filtering Sentences. Aswin et al. [3] propose a technique for generating subtitles automatically and semantic video summarization which generates the subtitles using speech recognition. NLP text summarization algorithms are applied on the generated subtitles. The performance of the model is boosted through Ensemble method using Weight based learning method and Intersection method. The output is a summarized video having the relevant subtitles and video frames. Sun [4] proposes the Interaction Canonical Correlation Network (ICCN) which learns the learns correlations between audio, video and text using "deep canonical correlation analysis (DCCA)". Saleem et al. [5] describe extractive text summarization by extracting key text segments i.e. sentences or paragraphs from the text, based on parameters like word frequency or cue words to locate the sentences to be extracted. Cue Phrases are used to identify important sentences. This paper considers following attributes for extractive text summarization: Sentence Position Attribute, Sentence Distance, End to End Attribute, Proper Noun Attribute, Upper-case Word Attribute, Keyword Attribute, Typeset Based Attribute, Pronouns and Influenced Word Attribute. The models proposed to implement are as follows: Neural network, feature Fusion, Cluster based method, Fuzzy logic, ranking approach, LDA based approach, Query based extractive text summarization and Graph based approach. Guenebaut [6] proposes a way to generate subtitles following standards by using speech recognition. Three steps are defined: audio extraction, speech recognition and subtitle generation using Sphinx-4. We notice that Sphinx-4 doesn't recognise voices with Indian English accent very effectively. Chen et al. [7] propose different approaches to automatically extract key terms from lectures including audio signals, ASR transcriptions and slides. Key terms are segregated as: key phrases and keywords and many methods are utilised to extract them in an order. Key phrases are extracted using right/left branching entropy and keywords are extracted by learning from the following feature sets: prosodic features, lexical features and semantic features from Probabilistic Latent Semantic Analysis (PLSA). Model learns using K-means (unsupervised method), AdaBoost and neural network (supervised methods).

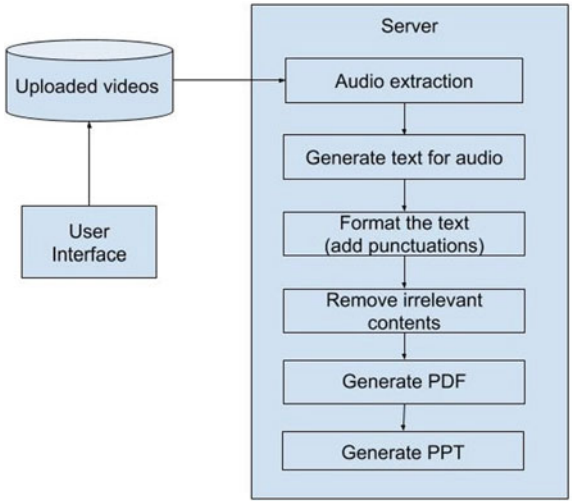
3 Proposed Method

The proposed model has the following main phases:

1. Upload mp4 video via User Interface.
2. Extraction of audio from the uploaded lecture video.
3. Speech recognition applied on the audio to generate text.
4. Punctuate the generated text and apply NLP methods to clean the data.
5. Implement extractive text summarization algorithm to extract the most relevant content and display the summarised content which can be edited on the UI.
6. Download the final PDF and PPT.

The data set used in the development of this model has been collected from the Impartus team at PES University. These are classroom lectures which have been recorded. We have used videos with non-blackboard style teaching i.e. videos for theoretical subjects, the content of which has been delivered without the usage of images, illustrations or any graphical medium. All the videos are in MP4 format. The videos have a primary speaker delivering content in Indian English accent. It also includes the voice of various students asking doubts and few segments without audio being recorded. NPTEL videos have been used for testing the accuracy of our model.

Fig. 1 System architecture



4 Procedure

4.1 Extraction of Audio and Speech Recognition

The video uploaded by the user must be in MP4 format. "ffmpeg" python framework is used to extract the audio from the video in .wav format. If video file contains data which can't be decoded, further processing will be stopped and raises an exception stating that audio file cannot be generated. The python speech recognition library (which in turn uses Google Web Speech API) is used to convert the audio into text. We segment the entire audio file into 3 min chunks. The library tries to recognise each chunk and writes recognised text to a consolidated file which is in .txt format. If an audio segment cannot be recognised because of noise or some other reason, it skips and tries to recognise the next 3 min audio segment. The recognized speech will be unformatted without any punctuations.

4.2 Punctuating the Text and Preprocessing

Punctuator, "a bidirectional recurrent neural network pre-trained model with attention mechanism for restoring missing inter-word punctuation in unsegmented text" is used to punctuate the text generated after speech recognition [8]. It adds punctuations such as period (.), comma (,), question mark (?), etc. Natural Language ToolKit (NLTK) is used to perform preprocessing of the punctuated text to remove stop words, stemming, tokenization, lemmatization, etc.

4.3 Dividing into Subtopics

Vector representation (word embeddings) is found for all sentences. Cosine similarities between sentence vectors are stored as a matrix. The similarity matrix is converted into a graph whose vertices are sentences and edges are similarity scores, for sentence rank calculation.

K-means clustering method is implemented to divide the entire text into clusters (i.e. subtopics) where each cluster has sentences having the highest cosine similarity with each other. The K-value i.e. number of clusters needs to be predefined. We have assumed K-value as 10 for quite a large text document.

4.4 Generating PDF and PPT

The text which has been divided into subtopics in the above step is displayed on the UI to allow the user to make changes if required. The summarised content with user's changes can be downloaded in Portable Document Format (PDF) using PyPDF2 python library. LexRank extractive text summarization algorithm extracts the most relevant sentences from the text which is later written to a PowerPoint Presentation (PPT) in the form of bullet points using python-pptx library. LexRank is an unsupervised graph based approach for automatic text summarization. The scoring of sentences is done by using the graph method. LexRank computes sentence relevance using the eigenvector centrality from a graph representation of sentences. An adjacency matrix for the graph representation of sentences composed of cosine similarity between the sentences. This method selects a centroid sentence which acts as the mean for other sentences in the document. The other sentences are ranked based on their cosine similarity score which helps to group sentences with same intend into a cluster.

5 Results

We observe that the performance of the model depends on the accuracy at which speech recognition works. Python's speech recognition library tends to give higher accuracy compared to other speech recognition frameworks .The recognized speech will be unformatted without any punctuation. Our model takes quite some time in recognising the speech and in punctuating the unstructured text generated from the speech recognition module. After summarized content is generated, user can edit it on the user interface and proceed to download the notes in the form of PDF and PPT. After running the test script on NPTEL videos, comparing our model generated text and the available transcripts, on an average the similarity is around 0.9.

The PPT will have the most important sentences from summarized content in the form of bullet points. The PPT can be customized to add different logo, number of slides, number of bullet points required per slide, etc. The generated PDF file will have the summarized content which has been divided into subtopics. Testing the accuracy of the generated notes for the Impartus videos has been done manually as there are no transcripts available for them. Testing has been done using NPTEL videos, which also include video wise transcripts available in PDF format. Cosine similarity has been used to find the similarity between notes generated using the proposed model and the downloaded transcripts.

6 Conclusion and Future Work

In this paper we have proposed a model for users to obtain the summarized content of a lecture video in the form of PDF or PPT without going through the hassles of watching long-duration videos. This model is mainly intended for non-blackboard style teaching rendered in English and gives effective results for the same. The techniques used are: audio extraction, speech recognition, punctuating generated text, preprocessing, K-means clustering to divide into subtopics and LexRank extractive text summarization method to extract the most relevant information from the text. The model can be further extended to work for black-board style teaching, different languages, ability to include diagrams/graphs in the summary, support different video formats, etc.

References

1. Miller D (2019) Leveraging BERT for extractive text summarization on lectures. arXiv:1906.04165
2. Sariki TP et al (2019) An adroit approach for extractive text summarization. Int J Eng Adv Technol (IJEAT) 8(5) ISSN: 2249-8958, June 2019
3. Aswin VB, Javed M (2019) NLP driven ensemble based automatic subtitle generation and semantic video summarization technique. In: Advances in Intelligent Systems and Computing, vol 1133. Springer, Singapore. https://doi.org/10.1007/978-981-15-3514-7_1
4. Sun Z, Sarma PK (2019) Learning relationships between text,audio,and video via deep canonical correlation for multimodal language analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol 34. pp 8992–8999. https://doi.org/10.1609/aaai.v34i05.6431
5. Saleem SM et al (2015) Study on text summarization using extractive methods. Int J Sci Eng Technol Res (IJSETR) 4
6. Guenebaut B (2009) Automatic subtitle generation for sound in videos. University WestIn: http://urn.kb.se/resolve?urn=urn:nbn:se:hv:diva-1784
7. Chen Y-N et al (2010) Automatic key term extraction from spoken course lectures using branching entropy and prosodic/semantic features. In: Spoken Language Technology Workshop (SLT), 2010 IEEE. IEEE, pp 265–270
8. Tilk O, Alumae T(2016) Bidirectional recurrent neural network with attention mechanism for punctuation restoration. In: Interspeech 2016, pp 3047–3051

Further Reading

9. Sheikh R et al, An approach towards generating subtitles automatically from videos by extracting audio
10. Mathur A, Saxena T, Generating subtitles automatically using audio extraction and speech recognition
11. Hearst MA, Multi-paragraph segmentation of expository text
12. Bahar P et al, A comparative study on end-to-end speech to text translation
13. Anjali et al, Subtitle generation: merging audio extraction, speech recognition
14. Choi Y et al, Pansori: ASR corpus generation from open online video contents

Correlation Between Code Smells for Open Source Java Projects

Inderpreet Kaur and Arvinder Kaur

Abstract The increasing demands and excessive use of computers and applications worldwide makes it challenging for the application and software developers to deliver the product on time and without faults. Looking at applications built for bugs, and thus minimizing the faults becomes tedious. Different researchers have extensively studied the maintenance of software and prediction of vulnerabilities. Code smells, and bug prediction is one way to find and locate the bugs through smell. These smells are beneficial in bug prediction if the smells are known beforehand. This paper studies the correlation for six code smells compiled for three open source java projects. The study results show that three smells are strongly correlated with code smells, namely God class, Brain method, and Dispersed Coupling.

Keywords Code smells · Bug prediction · Correlation · Machine learning · Open-source projects

1 Introduction

Today, every industry such as hospitality, banking, trading, education, travelling, e-commerce, healthcare, Government and defense etc. is dependent on software. The software's demanding functionalities and structure make it a challenge for developers to finish the project in time. While most of the time in software development is given to software coding, very less time is left for software testing phase.

Huge size and complex designs of the software makes them prone to software errors. It is a challenge for developers to minimize the software bugs and hence the failures so that the cost of software developed is low. Software Bug Prediction is the area where previous studies and histories helps the testers trace the bug in the current software [1–6].

I. Kaur (✉) · A. Kaur
USICT, GGSIPU, Delhi, India
e-mail: inderpreet.usict.phd@ipu.ac.in

A. Kaur
e-mail: arvinder@ipu.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_2

Developers have started using Code smell for Bug Prediction [7–10]. Code smell is the sign of complication that might exist in the system and can help in tracing the problem. The word code smell was invented by Kent Beck when Martin Fowler [11] first phrased it. Code smells are places where code maintenance can be problematic and thus requires proper consideration.

These code smells are strong and compelling instructions for the developers to trace and find the bad design areas which may otherwise lead to software failures [11]. Fowler in his studies found out that these smells have dependencies between them. While some dependencies show that even single conversion can remove the signs of these smells, the others might have shared common error on the origin. The way these smells are dependent on each other can help us find how they can be used.

The code smells are correlated, i.e. they have some dependencies amongst them. To statistically check the relation of these code smells, the correlation function is used. There are several correlation techniques, such as Pearson's correlation, Spearman correlation, and Kendall correlation [12]. This work uses Pearson's Correlation to find the dependencies between the smells. The primary outcome of correlation is a correlation coefficient which should range between -1 to $+1$. The higher the magnitude, the higher is the correlation. But if the extent has a negative sign, it means that they are inversely related.

Our research focuses on discovering the connection between six code smells, God Class, Data Class, Brain Process, Shotgun Surgery, Distributed Coupling, and Message Chains. Three open-source java projects have been taken from the JcodeOdor tool [13], Dr. Java, EMMA, and FindBugs [14] and their code smells have been collected.

This work is split into subsequent parts. The summary of relevant studies carried out in this context is given in Sect. 2. The research methodology used for the work performed, including the collected data sets and their code smell collection using the JCodeOdor tool, is given in Sect. 3. Section 4 reflects the outcomes of the study, while Sect. 5 gives the work's findings and future scope.

2 Related Study

In a study by Shantnawi and Li [15] found that code smells are highly related to fault proneness. Their work considered an open-source project Eclipse 2.1 with 6225 classes and had six seamless, namely God method, Refused Bequest, God class, Data class, Shotgun Surgery, and Feature Envy. Their findings indicate that shotgun surgery, God method and God class are positively associated with fault proneness.

Taba et al. [16] in their study on 21 projects namely nine versions of ArgoUML ($0.12-0.26.2(9)$) and 12 versions of Eclipse ($2.0-3.3.1(12)$) found out that files with a large number of bugs are associated with antipatterns. Their findings were based on the Wilcoxon rank-sum test to find the statistical significance and indicate that out of 21 projects, 17 projects had high correlation. Their analysis showed that antipattern-related metrics based on ANA, ACM, and ARl provide further insight into bug proneness over the code churn, PRE, and LOC of existing product and process metrics. ARL, which can help develop prediction models both inside and in cross-projects, was the best of the proposed metrics.

Hall et al. [17] researched the correlation between faults and five code smells, i.e., Data Clumps, Chains of Texts, Speculative Generality, Declarations of Switches and Middle Man. Three Eclipse, ArgoUML, and Apache Commons open-source projects were taken up, and a tool was developed to detect smells in them. The results of their analysis show that Data Clumps reduced the Eclipse and Apache faults, but increased the ArgoUML faults. In two programs, message chains raised flaws. Only in Eclipse could Speculative Generality minimize faults.

Jaafar et al. [18] researched the mutations and associativity of antipatterns between the evolution of these antipatterns and the propensity of faults. Their research was carried out on 27 launches of Mylyn, Rhino, and ArgoUML, three open-source projects. The study has twelve antipatterns under study, and the findings of the study indicate that antipatterns mutate due to structural changes from one type to another, and some of these mutations are more unsafe in terms of fault propensity.

A study developed by Zhang et al. [19] indicates that three kinds of smells have an immense effect on the susceptibility of faults. The analysis was carried out on 18 versions of the Apache Commons series. The research also shows that there was a very poor association between the three scents taken.

Results for the Ma et al. [20] shows that the efficiency of fault prediction is enhanced by refracting code smells. AntiSingleton, Blob, ClassdataShouldBePrivate, ComplexClass, LazyClass, LongParameterList, refusedParentBequest SwissArmyKnife based the analysis on eight code smells. Blob, RefusedParentBequest, Long Parameter list have major apportion, Of 8 Code smells three smells, there are significant opportunities for Blob, RefusedParentBequest, Long Parameter list to find and correct bugs.

The bug prediction results show that Random Forest was the best algorithm with the highest accuracy among Naïve Bayes and Logistic Regression in a study by Ubayawardana and Karunaratna [21] performed on 13 open-source java projects. Their findings also show that source code metrics alone are not sufficient for bug prediction, and if code smell metrics are taken into account, model training can achieve higher accuracy. With code smell metrics, even the cross-project bug prediction could be performed effectively.

Pecorelli et al. [22] conducted a study by ranking the code smells based on machine learning based on developers assigned critical values. The study was performed on nine open-source projects of Apache and Eclipse. The analysis was conducted on nine Apache and Eclipse open-source projects. God Class, Complex Class, Spaghetti Code, and Shotgun surgery were regarded as code smells in the four-class lever code smells. The results of the results show that with F-Measure, Random Forest was the best modelling technique in the range of 72–85%.

Our research focuses on the associativeness of code smells in three Java open-source projects, i.e., EMMA, Dr. Java, and FindBugs. The research is primarily carried out on the correlation of six code smells, including God Class, Data Class, Brain Method, Dispersed Coupling, Shotgun Surgery, and Message Chains. For the code smells, the coefficient of person correlation was evaluated.

3 Research Methodology

3.1 Datasets

This work has used three open source java projects Dr. Java, EMMA, and FindBugs, taken from SourceForge [14]. The code smell metrics have been gathered using the JcodeOdor [13]. JCodeOdor is a code smell detection tool and has precomputed metric values for 74 Systems. Six code smells God class, Data class, Brain Method, Shotgun Surgery, Dispersed Coupling and Message chains are gathered for the three datasets.

3.2 Code Smells

Code smells are the hint or sign that there might be a problem somewhere in the code. In his book [11], Fowler has defined different types of smells that could be there in a code. Developers can use these smells to trace the actual problem location, thus reducing the testing time. Six code smells in this study and is as under:

- **God Class**: A class with a large number of functions that increase code lines and thus create the class of God by creating tight coupling and thus creating a threat to code maintenance [11].
- **Data Class**: A class with techniques for either accessing or setting them (getters and setters). These are classes of containers and have no additional features and cannot be operated on their own [11].
- **Brain Method**: It refers to any bulky and complex technique that controls class intelligence [11].
- **Shotgun Surgery**: It is referred to as shotgun surgery [11] if a single change is made to multiple classes simultaneously.
- **Dispersed coupling**: A technique suffers from dispersed coupling if many techniques are called that are dispersed among many classes [11].
- **Message Chains**: This results in message chains if a request is made for an object that is dependent on another object and that object is further dependent on another object, etc. Changing these dependencies would require an amendment of the application [11].

Table 1 Strength of association for r

	God class	Data class	Brain method	Shotgun surgery	Dispersed coupling	Message chains
God class	1.00					
Data class	−0.01	1.00				
Brain method	0.08	−0.01	1.00			
Shotgun surgery	0.16	0.00	0.10	1.00		
Dispersed coupling	0.05	−0.01	0.34	0.12	1.00	
Message chains	0	0	0	0	0	1.00
Code smells	**0.59**	0.28	**0.62**	0.22	**0.53**	0

Source: Bold values represent the higher correlation

3.3 Correlation

Correlation is a statistical measure to analyze the relationship of various variables or the knowledge of the association of the variables in the data [12]. Correlation is put into terms through a correlation coefficient which ranges from − 1 to 1. If the variables are not related to each other, the value of the correlation coefficient is 0, if they are directly correlated, the value is positive whereas the coefficient is negative if they are indirectly correlated. In other words, if the value of one variable increases, the value of other variable increases, the correlation is positive. If the value of one variable increases, the value of the further variable decreases, the correlation is negative. There is no correlation if there is no dependence between the two variables that represent it. There are different types of correlation, depending on the type of data. For ordinal data, Spearman's (rho) correlation is evaluated while for interval or ratio level data, Pearson's (r) is used. Kendall correlation can be used for non-parametric tests [12]. This work uses Pearson's Correlation.

Table 1 shows the values of Pearson's correlation coefficient concerning the strength of the association.

4 Results

The three open-source projects, Dr. Java, EMMA and FindBugs, were considered for this work. The smells of the code were collected from JcodeOdor. For the six code smells, namely God Class, Data Class, Brain Method, Shotgun Surgery, Dispersed coupling, and Message Chains, Pearson's correlation was evaluated.

Table 2 Correlation coefficient for Dr. Java

	Coefficient, *r*	
Strength of association	Positive	Negative
Small	0.1–0.3	– 0.1 to – 0.3
Medium	0.3–0.5	– 0.3 to – 0.5
Large	0.5–1.0	– 0.5 to – 1.0

Table 3 Correlation coefficient for Emma

	God class	Data class	Brain method	Shotgun surgery	Dispersed coupling	Message chains
God class	1.00					
Data class	#DIV/0!	1.00				
Brain method	0.36	0	1.00			
Shotgun surgery	0.21	0	0.17	1.00		
Dispersed coupling	0.00	0	0.00	0.25	1.00	
Message chains	0	0	0	0	0	1.00
Codesmells	**0.59**	0	**0.62**	0.22	**0.53**	0

Source: Bold values represent the higher correlation

Tables 2 and 3 show the results obtained for correlation amongst code smells for Dr. Java and Emma, respectively. It can be seen from results that code smells are highly correlated with God class, Brain Method and Dispersed coupling i.e., if any of these three smells are present, there is a high possibility of code smells. While the correlation coefficient is 0.59 for God Class, it is 0.62 for Brain Method and is 0.53 for Dispersed coupling.

For Dr. Java and Emma, there is no correlation between Message chains and code smell. Data class smells, and code smells correlation coefficient for Dr. Java is 0.28 while there was no Data class smells in Emma.

Shotgun surgery smells are correlated with code smell with the coefficient of 0.22 in both the datasets.

Table 4 shows the correlation matrix for FindBugs dataset. Code smells for FindBugs dataset are positively correlated with God Class, Brain Method and Dispersed Coupling while Data class, Shotgun surgery, and Message chains are less correlated with code smell. Only for FindBugs dataset, Message Chains have a small correlation with code smell with coefficients 0.15.

Figure 1 shows the correlation of FindBugs dataset for various smells.

Table 4 Correlation coefficient for FindBugs

	God class	Data class	Brain method	Shotgun surgery	Dispersed coupling	Message chains
God class	1.00					
Data class	−0.01	1.00				
Brain method	0.03	−0.01	1.00			
Shotgun surgery	0.08	−0.01	0.02	1.00		
Dispersed coupling	0.17	−0.02	0.07	0.05	1.00	
Message chains	0.00	0.00	0.05	0.00	0.05	1
Codesmells	**0.59**	0.28	**0.62**	0.22	**0.53**	0.15

Source: Bold values represent the higher correlation

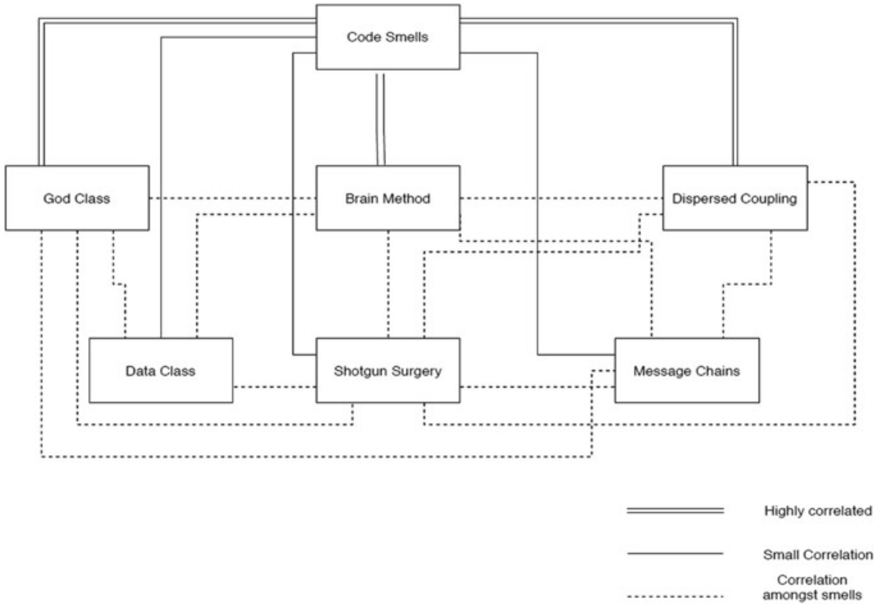


Fig. 1 Correlation representation for FindBugs

5 Conclusions and Future Scope

This research was carried out on three Java open-source projects collected from SourceForge, namely Dr. Java, EMMA, and FindBugs. Six Smells of Code, i.e. For the corresponding dataset, God class, Data Class, Brain Method, Shotgun Surgery, Dispersed Coupling, and Message Chains were compiled from JcodeOdor.

The Pearson correlation coefficient was calculated for the smells to determine the associativity between the different code smells. The following conclusions can be taken from the findings.

- For six code smells taken under study, it was found that God class, Brain Method and Dispersed coupling were positively associated with the code smells for all three datasets. Their increase in value positively affects the presence of code smells.
- Shotgun Surgery has a mild correlation with code smell for all three datasets.
- There were no code smells and data class smell correlation for Emma dataset.
- Message chain has no associativity with code smell for two datasets Dr. Java and EMMA. Whereas for FindBugs dataset, there is a mild correlation between code smells and Message chains.

References

1. Briand LC, Ẅust J, Daly JW, Victor Porter D (2000) Exploring the relationships between design measures and software quality in object-oriented systems. J Syst Softw 51(3):245–273
2. Kanmani S, Rymend UV (2004) Object-oriented software quality prediction using general regression neural networks. SIGSOFT Softw Eng Notes 29(5):1–6
3. Nagappan N, Laurie W (2005) Early estimation of software quality using in-process testing metrics: a controlled case study. In: Proceedings of the 3rd workshop on software quality. St. Louis, Mo, USA, pp 1–7
4. Aggarwal KK, Singh Y, Kaur A, Malhotra R (2006) Empirical study of object-oriented metrics. J Object Technol 5(8):149–173
5. Olague HM, Etzkorn LH, Gholston S, Quattlebaum S (2007) Empirical validation of three software metrics suites to predict fault-proneness of object-oriented classes developed using highly Iterative or agile software development processes. IEEE Trans Softw Eng 33(6):402–419
6. Tran HM, Le ST, Van Nguyen S, Ho PT (2020) An analysis of software bug reports using machine learning techniques. SN Computer Sci 1(1):4
7. Palomba F, Zanoni M, Fontana FA, De Lucia A, Oliveto R (2019) Toward a smell-aware bug prediction model. IEEE Trans Softw Eng 45(2):194–218, 1 Feb 2019. https://doi.org/10.1109/TSE.2017.2770122.
8. Palomba F, Zanoni M, Fontana FA, De Lucia A, Oliveto R (2016) Smells like teen spirit: improving bug prediction performance using the intensity of code smells. In: 2016 IEEE international conference on software maintenance and evolution (ICSME), Raleigh, NC, pp 244–255. https://doi.org/10.1109/ICSME.2016.27
9. Pritam N, Khari M, Kumar R, Jha S, Priyadarshini I, Abdel-Basset M, Long HV (2019) Assessment of code smell for predicting class change proneness using machine learning. IEEE Access 7:37414–37425
10. Catolino G, Palomba F, Fontana FA, De Lucia A, Zaidman A, Ferrucci F (2020) Improving change prediction models with code smell-related information. Empir Softw Eng 25(1):49–95
11. Refactoring: improving the design of existing code (1999). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA
12. https://www.statisticssolutions.com/correlation-pearson-kendall-spearman/
13. Fontana FA, Ferme V, Zanoni M, Roveda R (2015) Towards a prioritization of code debt: a code smell intensity index. In: Proceedings of the seventh international workshop on managing technical debt (Mtd 2015), 16–24. IEEE, Bremen, Germany. https://doi.org/10.1109/MTD.2015.7332620

14. Sourceforge: https://www.sourceforgenet/
15. R. Shatnawi, Li W (2006) An investigation of bad smells in object-oriented design. In: Third international conference on information technology: new generations (ITNG'06), Las Vegas, NV, pp 161–165. https://doi.org/10.1109/ITNG.2006.31
16. Taba SES, Khomh F, Zou Y, Hassan AE, Nagappan M (2013) Predicting bugs using antipatterns. In: 2013 IEEE international conference on software maintenance, Eindhoven, pp 270–279. https://doi.org/10.1109/ICSM.2013.38
17. Hall T, Zhang M, Bowes D, Sun Y (2014) Some code smells have a significant but small effect on faults. ACM Trans Softw Eng Methodol (TOSEM) 23(4):1–39
18. Jaafar F, Khomh F, Guéhéneuc Y, Zulkernine M (2014) Anti-pattern mutations and fault-proneness. In: 2014 14th international conference on quality software, Dallas, TX, pp 246–255. https://doi.org/10.1109/QSIC.2014.45
19. Zhang X, Zhou Y, Zhu C (2017) An empirical study of the impact of bad designs on defect proneness. In: 2017 international conference on software analysis, testing and evolution (SATE), Harbin, pp 1–9. https://doi.org/10.1109/SATE.2017.9
20. Ma W, Chen L, Zhou Y, Xu B (2016) Do we have a chance to fix bugs when refactoring code smells? In: 2016 international conference on software analysis, testing and evolution (SATE), Kunming, pp 24–29. https://doi.org/10.1109/SATE.2016.11
21. Ubayawardana GM, Karunaratna DD (2018) Bug prediction model using code smells. In: 2018 18th international conference on advances in ICT for emerging regions (ICTer), Colombo, Sri Lanka, pp 70–77. https://doi.org/10.1109/ICTER.2018.8615550
22. Pecorelli F, Palomba F, Khomh F, De Lucia A (2020, May) Developer-driven code smell prioritization. Int Conf Min Softw Repositories

Throughput Improvement in Energy Efficient Heterogeneous Wireless Sensor Network

M. Lakshmi and C. R. Prashanth

Abstract Designing an energy efficient scheme in Heterogeneous Wireless Sensor Network (HWSN) is a critical issue which degrades the network performance due to node failure. An Energy Efficient Clustering (EEC) algorithm is proposed by combining the rotation based clustering and energy saving scheme for avoiding the node failure and prolonging the network lifetime. In EEC, network is partitioned in to clusters and cluster head is selected on rotation based by considering the highest energy and residual energy of the sensor nodes. Other cluster members are accommodated in a cluster on the basis of Basic Cost Maximum flow (BCMF) so as to allow the cluster head for transmission. Carrier Sense Multiple Access (CSMA), a contention window based protocol is used at the MAC layer for detecting the collision and to provide channel access prioritization to HWSN of different traffic classes with reduction in End to End delay, energy consumption and improved throughput and Packet delivery ratio (PDR) and allowing the cluster head for transmission without depleting the energy. The simulation result presents that EEC is not only stable to broadcast the data to the destination in HWSN but also save energy of the nodes, thereby increasing the lifetime of the network. Simulation parameters of the proposed system such as Throughput, Energy, and Packet Delivery Ratio are obtained and compared with the existing system.

Keywords HWSN · Clustering · EEC · CSMA · Contention based protocol · PDR

1 Introduction

Wireless sensor network (WSN) [1] is the group of sensor nodes which are deployed randomly in a focused area over actively changing environments. These wireless sensor nodes are used for sensing, processing and forwarding the data to the intermediate nodes or neighboring nodes and also to the Base Station (BS) [2]. Since

M. Lakshmi (✉) · C. R. Prashanth
Department of Electronics and Telecommunication Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_3

these nodes are small devices and are limited with memory, low processing, low computation and small power unit that are battery powered.

The sensor nodes are scattered in a sensor field which are transceivers having the capability of collecting, routing the data to the sink using multihop infrastructure less architecture. The devices carryout computations using their processing capabilities to transmit the processed and required data.

1.1 Sensor Node

There are five components in a sensor node; a micro controller unit, a transceiver unit, a memory unit, a power unit and a sensor unit as shown in Fig. 1. The micro controller unit is responsible for handling different tasks such as data processing and controlling of other components in the sensor node.

The controller unit is having a small storage unit which is integrated in to the embedded board or on board memory. This unit also takes the risk of managing the procedures that make the sensor node for sensing, processing using algorithms and collaborating with the other nodes using wireless communication. Transceiver unit is the most power consuming unit where a sensor node communicates with other nodes and other parts of the WSN [3]. Random Access Memory, Read Only Memory and other memory types are used as temporary storage of sensed data. Power unit is a critical component for supplying node energy. Batteries are used for storing the power. Sensor unit is the main component in a wireless sensor node which provides information gathering capabilities from the physical phenomenon. This unit is responsible of gathering information from outside world like light, sound, temperature and others. There are two sub units i.e. a sensor and an analog to digital converter (ADC). The focus of these techniques is to achieve energy efficiency and maximization of network lifetime. Redundant messages are eliminated while forming the efficient clusters and selecting or reselecting the Cluster Head (CH). It is required to construct the topology for distributing the nodes uniformly in the clusters which makes the network efficient. Poor network performance is due to the excessive energy consumption while forming the clusters periodically and reselecting the CH [4–6].

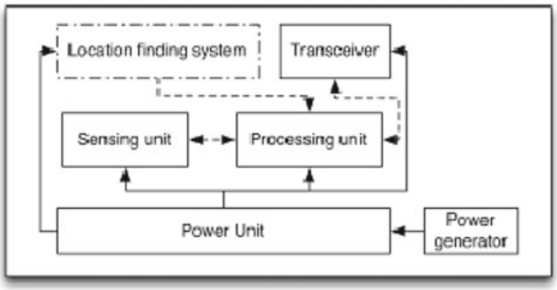


Fig. 1 Components of a sensor node

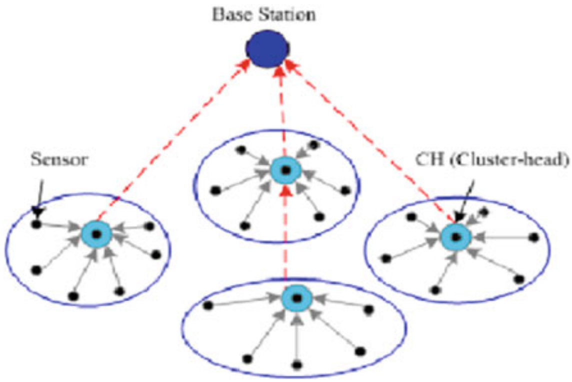


Fig. 2 Hierarchial network structure

1.2 Energy Saving Mechanisms in WSN

There are many energy saving mechanisms such as Energy efficiency through directional antennae, Duty cycling, Energy efficient scheduling, Scheduled rendezvous, Clustering, On-demand wakeup scheme and others. Due to high energy consumption, direct communication from sensor nodes to the BS or multichip communication towards BS is not practical which results in depletion or expiry of sensor nodes. Single tier or direct communication [7] is not appropriate for large scale network as it has many disadvantages like consumption of high energy, data duplication and quick expiry of farthest node. To overcome these disadvantages two tier communication is used by using hierarchical approach in which network is partitioned in to clusters. Cluster head is the leader node that is responsible for data aggregation and forwarding.

Two level hierarchy in the hierarchical network structure is used where cluster heads are at the upper level and members nodes at the lower level. CH receives the data periodically from lower level nodes. Energy consumption is high in CH nodes compared to member nodes as the data is being sent by these over long distances all the time [8]. After certain rounds, the CH may expire due to high energy consumption. So to balance the energy consumption, the role of CH is changed periodically in order to make sure load balancing. Figure 2 shows the communication in hierarchical network structure. Single hop communication takes place within a cluster where as multihop communication is between clusters which are named as intra cluster and inter cluster communication.

1.3 Routing Protocols

An algorithm for defining the exact route for a packet from source to the destination is nothing but the routing protocol. Due to resource constraints in WSN, routing is more challenging compared to mobile adhoc networks or vehicular adhoc networks [9]. So new routing mechanisms are very important in keeping in view the network architecture and application requirements. High energy consumption results due to negligence in route maintenance and frequent topological changes.

Nodes in the network are grouped in to clusters by using some algorithms and cluster head is elected on some criteria and held responsible for routing the packets. Two layer approaches is used in hierarchical routing and one layer is for sensing the physical phenomenon and other layer for routing. Higher energy nodes are used for collecting, aggregating and forwarding the data sensed by lower energy nodes [10]. Some of the clustering techniques are compared such as Low Energy Adaptive Clustering Hierarchy (LEACH), Hybrid Energy Efficient Distributed Protocol (HEED), Extended HEED, Distributed Weight-based Energy-efficient Hierarchical Clustering protocol (DWEHC).

1.4 Energy Efficiency in Mac Protocols in WSN

Significant amount of energy is saved in a multi hop network by the occurrence of routing using an efficient routing algorithm. Transmission power is dependent on the basis of distance and range, and energy is a major factor while route discovery. Sensor node has a component called Radio which is a major power consuming and it is controlled by Medium Access Control (MAC) Protocol. Network lifetime increases to a great extent by an efficient MAC Protocol. In addition, sharing of wireless medium is also controlled by MAC protocol [11]. Less power is consumed by a sensor node when it is in sleep mode.

One of the adaptive protocols called Sensor Protocols for Information via Negotiation (SPIN) [12] assumes all the nodes as potential base stations and disseminates all the information at each node to all other nodes. This protocol having variants are proposed to enhance the performance of the network. In Gradient Based Routing (GBR) [13], the difference between the height of a node and its neighbor node is calculated named as gradient value. Transmission of packet takes place on a link having largest gradient value. GBR working is on the basis of memorizing hops count while diffusing the entire network. Some authors have proposed a hierarchical routing algorithm which is energy aware routing for cluster based networks [14]. This method is based on the architecture of three tier. Here the clusters are formed by grouping the sensors and Time division multiple access (TDMA) MAC scheme is used to operate the network. Network flow approach [15] is introduced to overcome the problem of routing in sensor networks. The approach is better in maximizing the network lifetime by considering the link cost as a function remaining energy of the nodes and the transmission energy using the link.

1.5 Motivation

Some issues are found in the clustering protocols such as selection of faithful CH, Formation of energy efficient clusters, and management of network. Probabilistic approach is considered and linked to the criteria while selecting the CH. A node can be selected as CH by varying the selection of CH which in turn increases the internal overhead. Energy consumption in the network is increased when the node is selected as CH in the region and hence the performance of the network reduces. All the nodes connected in the network should have better internal parameter and connectivity. So the performance parameters are enhanced by selecting the better CH approach.

The work is contributed to overcome the issues related to consumption of energy. MAC layer has scheduling mechanism for designing a network that makes collision free and more energy conserving. Many MAC protocols have been designed to avoid collision on a medium and contention window based protocol CSMA plays an important role during data transmission which occurs while carrier sensing. In the network, clusters are formed and CHs of all the clusters are selected at a time on the basis of highest energy. Main idea is to select the CH for data transmission and other CH with less transmission and make them sleep. This methodology helps in achieving network lifetime by less consumption of node energy.

2 Literature Review

Nikolidakis et al. [16] proposed a system by designing a protocol which selects CH on the basis of routing energy cost by node with reference to the BS from the corresponding cluster. In the proposed system, Gaussian elimination method is used for selecting the low loss route which is most feasible to node for CH role. CH is elected for two levels. CH node that is near to the BS is selected for the first level and the one which is farther is selected as lower level CH. Lossy path is avoided by offering the suitable route discovery scheme, thereby enhancing the energy efficiency and prolonging the network lifetime. But this scheme has high possibility of abrupt stability breakage and hence reduction in lifetime. It has limitations on multihop routing. Routing table is very huge and demanded higher energy on the node respectively.

Rana et al. [17] proposed a system with an approach which divides the network in to multiple clusters that reduces communication overhead, and enhance energy efficiency. CH is selected by selecting less loss energy path with reference to nodes. Regular updating of energy information is required to select best suitable path but extra load is imposed in scalable network. Energy consumption increases if the CH is very far from BS. Wang et al. [18] proposed a system to overcome the demerits of centralized clustering approach used for designing and unsuitable for scalable network. Authors minimized the problem by the optimum number of clusters and utilizing the energy available. They designed a protocol by using the principle of

LEACH and considering the residual energy and also by calculating the death toll from BS. Before the completion of energy, CH is elected based on the available time period. In this system, stability time period and network lifetime is improved, thereby enhancing the energy utilization.

Li et al. [19] proposed a system by extending the Low Energy Adaptive Clustering Hierarchy (LEACH). They designed a system by using energy efficient clustering protocol for improving the remaining energy in the network thereby enhancing the network lifetime. The designed protocol considers the node residual energy and distance with respect to BS and CH is selected by considering the cluster members.Level-1 CH is a node which has higher distance with respect to BS and Level-2 CH has lower distance to BS. In level-1, data is collected from the cluster and transferred to other level-2 from CH for further communication to BS. Burden of transmission is minimized for CH in level-1 and energy is saved in the case of level-2. There is a systematic utilization of energy. Updating of parameters is needed at the cluster cycle each time. Threshold probability of CH selection differs from normal LEACH. This scheme has a disadvantage of less supportive to different capability nodes.

Liu et al. [20] proposed an algorithm called novel Quorum Time Slot Adaptive Condensing (QTSAC) based medium access control for minimizing the delay and energy efficiency of wireless sensor network. Duty cycle of the system is prolonged by condensing the Quorum time slot to the data transmission period on the basis of WSN characteristic and also this innovation reduced the network latency. Utilization of energy remaining in the area which is far from the sink increases the QTS for achieving the above mentioned parameters. Sert et al. [21] proposed a two tier distributed fuzzy logic based protocol (TTDFP) for the data aggregation efficiency improvement in multihop wireless sensor networks. Clustering is considered for efficient aggregation requirements like consumed energy. A network is divided in to clusters and data is transmitted from member nodes to the CHs, then packets are received at the base stations from CHs. This proposed TTDFP for extending the lifespan of multihop WSNs by considering the clustering efficiency and routing phases.

Alsaafin et al. [22] designed a system by Heterogeneous Aware Distributed Clustering (HADC) algorithm. They presented node heterogeneity in terms of energy for enhancing the network lifetime. CHs in the proposed system are selected based on the cost function by considering the residual energy and node load. By making the trade-off between the distance and node degree towards CHs, authors reduced the energy consumption and balancing the load in the network. Proposed system performance is compared with HEED and LEACH in terms of efficiency. Toor et al. [23] proposed a routing protocol named Energy Aware Cluster Based Multi-hop (EACBM) which reduces the energy consumption of sensor nodes by considering the concept of clustering and multihop communication. Additionally, sub clustering concept is considered in the proposed system for the sensor nodes which are not reachable and are not included in any of the cluster. This protocol is compared with the existing protocols LEACH, CEEC and SEP). Authors improved the network lifetime, stability and energy efficiency in heterogeneous WSNs.

3 Energy Consumption Model

Transmitter and Receiver are the main sections of communication. Energy is required for both the sections, a sensor node having transmitter section need energy to run some circuits such as radio electronics and amplifier. Whereas energy consumed by the receiver section in the packet over a distance at radio electronics. Calculation of energy consumed is based on the free space energy consumption model as the distance between the communication entities is less than the threshold distance. If the distance is more than the threshold distance, then multipath model is used. Data aggregation is a feature for the consumption of energy which is implemented to CH and the data collected by the cluster members is delivered to BS [24]. Energy consumed in the system is calculating by the parameters $L_{p,}$ a packet of length (number of bits) and distance d_s as shown in (1).

There are many significant parameters for processing the information. E_{el} is the consumed energy by transmitting or receiving electronics circuitry. Two energies abbreviated as E_T and E_R, are the transmitting and receiving energies consumed for processing a data packet L_P. A scheme of digital coding and modulation is used on which the parameters depend on. Other parameters considered in the system are E_{fsm} and E_{mpm}, energy needed for free space model and multipath model.

For calculating energy, the free space power loss model is considered if the distance d between the source and destination is less than d_0 controlled by power amplifier with the information. Transmitter with free space model need energy which is given as

$$E_T(L_P d_s) = L_P E_{el} + L_p E_{fsm} d_s^2 \quad d_s \ll d_0 \tag{1}$$

Energy calculation using multipath model is given below in which distance d_s is greater than the threshold distance.

$$E_T(L_p d_s) = L_p E_{el} + L_p E_{mpm} d_s^4 \quad d_s \gg d_0 \tag{2}$$

Consumption of energy by the Receiver is,

$$E_R(L_p d_s) = L_p E_{el} \tag{3}$$

d_0 can be calculated as,

$$d_0 = \sqrt{E_{\mathrm{fsm}}} / \sqrt{E_{\mathrm{mpm}}} \tag{4}$$

4 EEC Network Model

EEC is an energy efficient clustering in the proposed network model which has sensor nodes deployed randomly on N×N sensing layout and uses heterogeneity with three-level node energy. After deployment, all the nodes and BS becomes static which is predefined. Links between each node for communication is considered to be static. CH is responsible in the sensing network for forwarding the data collected to BS directly. Data messages that are supervised or unsupervised transacted through wireless links. The assumptions made in the network includes, the similar capabilities possessed by the nodes have unique id and different energy levels. Network is divided in to clusters based on the node population and has BS is stationary and located at the middle of the network field. BS has stable power supply without any memory restraints, energy or computation.

4.1 Cluster Formation and Cluster Head Selection

The process of clustering and information transmission consists of three stages such as Cluster arrangement and cluster head selection, accumulating information, collecting and transmitting information. On the basis of residual energy of all the nodes, cluster head is selected and node density as per the protocol.

Highest residual energy node is chosen as cluster head, which transmits the data across the cluster head. All the sensor nodes receive broadcast message by their respective cluster heads. Sensor nodes select the cluster head based on the received signal energy and based on the maximum received energy the decision is sent to CH. Acknowledgement is sent to all sensor nodes regarding the CHs decision. Generation of TDMA signal and broadcasting is scheduled once the clusters are determined. Finally, the energy levels of all the nodes are checked for proceeding to the next round and the clusters having zero energy are removed from the network. After forming the clusters, the cluster members in each cluster gather information and send to their respective CH. Once the data is collected by the CH, the information transmission stage occurs by sending the aggregated data to the base station. The individuals in the group spend energies for gathering, and sending the information to their CHs.

Energy required for completing one round by CH is estimated as,

$$E(CH_{load}, R\_1) = CH_{(EI,RI,k)} + CH_{(E1,A1,NCM)} + CH_{(E1,T1,d1BS,k)} \tag{5}$$

In the above equation, E_1, R_1, T_1, A_1 and N_{CM} abbreviate energy, receive aggregation and cluster member nodes in a respective cluster in the first round of transmission. E (CHload,R1) is the energy required by CH for completing one round of transmission. CH $_{(E1,R1,k)}$ is the energy consumed by CH for receiving k-bits message. CH $_{(E1,A1,NCM)}$ stands for energy needed to aggregate the data from Cluster members

N_{CM}. CH $_{(E1,T1,d1BS,k)}$ indicates the consumption of energy to transmit the aggregated data to the base station.

$$CH_{(E1,R1,NCM)} = N_{CM}X(kXE_{el}) \tag{6}$$

$$CH_{(E1,A1,NCM+1)} = (N_{CM} + 1)X(kXE_A) \tag{7}$$

$$CH_{(E1,T1,d1BS,k)} = KXE_{el} + \begin{Bmatrix} kX \in_{fsm\ X} \ d_1^2 BS & if\, d_{1BS} < d \\ kX \in_{mpm\ X} \ d_1^4 BS & otherwise \end{Bmatrix} \tag{8}$$

d has the threshold value.

The above Eqs. (6), (7) and (8) calculate the energy required for receiving the data from the members in the cluster, aggregating the data with its own data, and sending the aggregated data to the BS. CH does estimation of number of communication rounds before it dies.

The estimation of number of communication round is done by CH successfully before it dies and given as,

$$CH_{E\_R} = floor\{(E_{iE\_CH})/E(CH_{load}, R_{\_1})\} \tag{9}$$

In the above Eq. (9), CH_{E_R} is the number of rounds that CH completes before it dies, and E_{iE_CH} is the initial energy when the node becomes CH respectively. Death of the CH is calculated as $= E_{Ti}/E_{Ri}$, where E_{Ti} is the total energy and E_{Ri} is the energy required to complete one round by CH.

Dynamic threshold for CH-Rot is calculated by using Eq. (9) by estimating CH_{E_R}. If $CH_{E\_R} > 1$, then $CH_{E\_R}-1$ is completed successfully by CH and again CH-Rot is triggered for selecting CH. The CH is selected among cluster members on the basis of residual energy. If CHE_R < 1, then re-clustering is done. Cluster head rotation is given below in Eq. (10),

$$CH - Rot = \begin{Bmatrix} CH_{ER} - 1 & if\, CH_{ER} > 1 \\ RE - clustering & Otherwise \end{Bmatrix} \tag{10}$$

5 Proposed Protocol

Proposed system is built using energy efficient clustering protocol. A main objective of the protocol is to enhance the energy efficiency in HWSN and improving the remaining energy, hence reducing the energy dissipation which leads to increased lifetime and throughput. In HWSN, the node is selected as CH so that Cluster members CM communicate with BS in an efficient way without giving extra load to other nodes in the cluster. Based on some parameters, the CH is selected and all the nodes update their parameter information to BS at the end of cluster cycle. At the initial stage, all the nodes are randomly deployed over the network; network is divided in to many sections based on the density in the form of clusters.

In the proposed work, the node is selected for the role of CH based on the parameters available with the sensor node. Parameters such as initial energy, residual energy and hop count with respect to the base station is considered. For selecting CH, initial energy must be high enough for cluster management activity for advertising, formation of frames, and transaction with respect to the base station. After cluster round, the remaining energy of a node must be high enough to handle the role of CH in the next upcoming rounds. Scalable network must be handled for multihop communication by CH after selection.

In this approach, energy efficient clustering algorithm is introduced which Consists of following stages: Arrangement of clusters, Decision of group head, Accumulation of information and collection, and transmission of information It is necessary for selecting CHs based on initial energy, residual energy and one hop count. All the nodes of different energy levels are deployed randomly over a network. Clusters are formed on the basis of single hop distance and neighbors are calculated using Euclidian distance formula. Energy levels of each node are calculated as explained in Energy consumption model. Once energy levels of all the nodes are calculated, first CH is found using high initial energy. In the next step, all other CHs are calculated based on the energy and also by giving a condition of not including the CH already found in other clusters. Totally 6 CHs are found in the network. During transmission, particular cluster is considered; CH is transmitting the data packets to the BS. Whereas other CHs remain static as transmission is not taking place within those clusters. Once the energy of CH that is transmitting the data gets depleted, CH rotation takes place.

6 Simulation Parameters

The parameters set for implementation are highlighted below and simulation work having performance validation are explained.

Throughput: Amount of throughput per cluster round is defined as the number of data sent towards base station from the sensor nodes over the cluster round. So it is the total number of data packets delivered at the destination node per unit time. It is measured in packets/second or bits/second.

Energy Efficiency: Energy consumed per unit of successful communication is referred to as Energy efficiency. It is defined as the ratio of throughput to the energy consumed.

Packet Delivery Ratio (PDR): PDR is the ratio of data packets delivered successfully to a destination to the total number of data packets sent by the sender. It can be calculated by considering the number of data packets sent and number of received

Table 1 Simulation parameters

S. No.	Parameter symbol	Description	Value
1	$N \times N$	Network area	1300 m * 1300 m
2	N	Number of nodes	100
3	*MAC*	MAC protocol	IEEE 802.11
4	E_0	Initial energy of nodes	0.5–1.5 J
5	L_p	Data packet size	4000 bits
6	E_{el}	Radio electronics energy	50 nJ/bit
7	E_{fsm}	Free space energy	10 pJ/bit/m^2
8	E_{mpm}	Amplification energy	0.0013 pJ/bit/m^4
9	d_0	Threshold distance	87–87.7 m
10	*Packet size*	Packet size	512

data packets. Set of parameters utilized in implementing the proposed work for the purpose of simulation is given in Table 1.

7 Simulation Results and Discussion

In this section, implementation and results of the simulated work and the validation with protocols like HEED and Rotating Energy Efficient Clustering for Heterogeneous Devices (REECHD) with reference to parameters are focused. In Heterogeneous Wireless Sensor Network, nodes are deployed randomly and the network is divided in to clusters based on the population of nodes. BS is placed at the centre of the network and the node 49 with position (145, 125). Number of nodes in the network is 100, and all the nodes are assigned with different energy levels.

7.1 Network Initialization and Neighbour Discovery

The proposed work uses EEC algorithm having two phases, clustering phase and operation phase. The network is densely populated with 100 nodes having different energy levels. As shown in Fig. 3 all the nodes are placed randomly in a network area. In a flat grid. The parameters such as antenna type, link layer, queue type, routing protocol and interface type are set up.

The neighbor nodes are found out by Euclidian distance by considering the position of nodes i.e. (x,y) co-ordinates. The transmission range is set to 200 m and the nodes fall in those range are considered. Each node should send Hello messages to find the position of nodes.

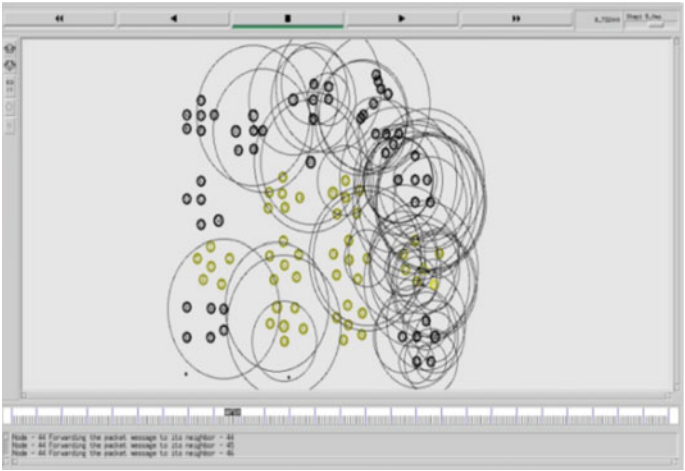


Fig. 3 Node deployment and neighbor discovery

7.2 Cluster Formation and CH Selection

The Clusters are formed on the basis of one hop distance as shown in Fig. 4 which is equally sized clusters. Totally 100 nodes are deployed and 16 clusters are formed. Based on high energy of nodes, CH is selected and other member nodes in the cluster have the feature of BCMF. Communication overhead of CH is reduced by these CMs as maximum flow of data packets takes place.

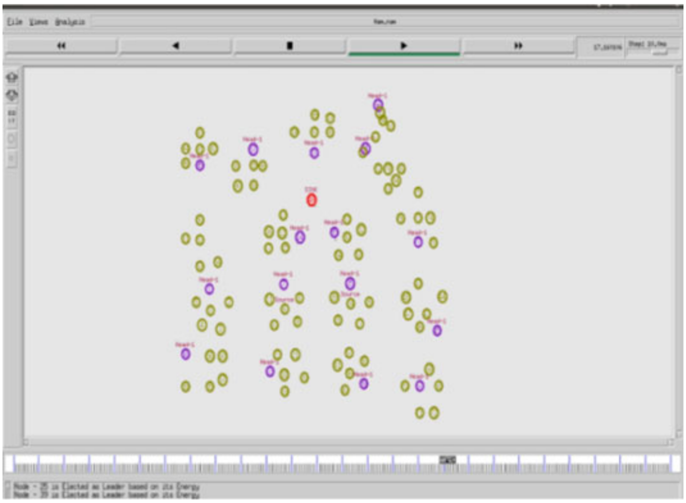


Fig. 4 Formation of cluster and CH selection

8 Performance Analysis

Proposed system is implemented using NS-2.35 and compared with HEED and REECHD in order to compare the network performance in terms of parameters such as energy efficiency, throughput, and PDR. HWSN is considered by deploying 100 nodes on a grid size considered. Following are the parameters considered in the proposed system with graph and values. Clustering is done for all the nodes and Energy calculation is done as explained in the above algorithm section. Initial energy is considered between 0.5 and1.5 J. Source and destination nodes are fixed by considering the Sink as destination node.

8.1 Energy Graph

As shown in below Fig. 5, proposed system uses a clustering algorithm which calculates the CH based on the high energy. CHs of all clusters are calculated subsequently by neglecting the cluster whose CH is already found. CHs of all clusters in the network are identified.

During transmission, particular cluster is considered; CH is transmitting the data packets to the BS. Whereas other CHs remain static as transmission is not taking place within those clusters. This saves energy of all other clusters as they are not part of transmission. CH rotation takes place when the energy of CH which is transmitting gets depleted. Consumption of energy reduces during CH selection and transmission.

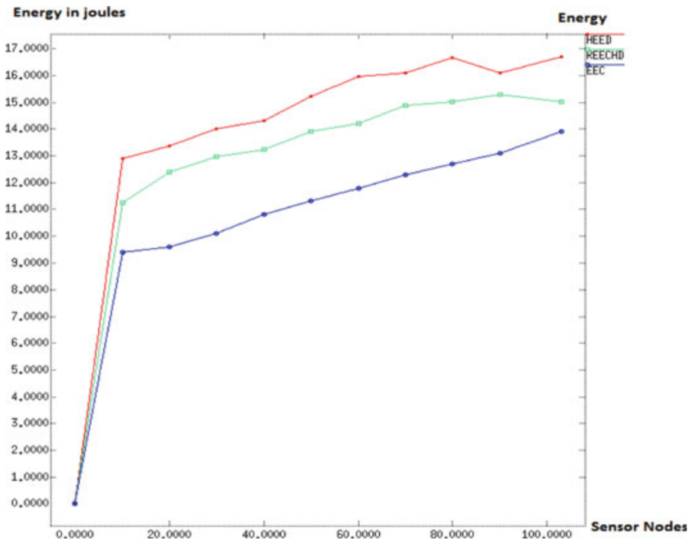


Fig. 5 Energy graph

Table 2 Energy values of proposed and existing systems

Sensor nodes	Energy consumption (in joule)		
	HEED	REECHD	EEC (Proposed)
0	0	0	0
20	13.38	12.38	9.6
40	14.32	13.25	10.8
60	15.98	14.20	11.8
80	16.67	15.01	12.7
100	16.72	15.02	13.9

Table 2 shows the values of the proposed system of energy consumption during transmission and compared with the existing system with HEED and REECHD. Energy consumption values are taken by considering number of sensor nodes on X-axis and Energy in joule along Y-axis. Consumption of energy in the EEC system is 13.9 J where as energy consumed is more in the existing systems.

8.2 *Packet Delivery Ratio*

Figure 6 shows PDR of both existing and proposed system. It is the ratio of no. of packets obtained at the receiver to the total no. of packets including the packet drop. In the proposed system, the PDR is achieved with better value as the MAC

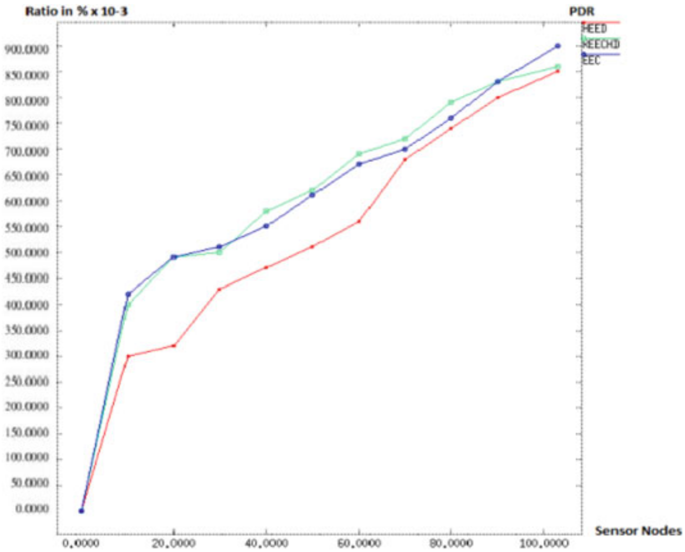


Fig. 6 Packet delivery ratio

Table 3 PDR values of proposed and existing systems

Sensor nodes	PDR (Ratio in % $\times 10^{-3}$)		
	HEED	REECHD	EEC (Proposed)
0	0	0	0
20	0.32	0.49	0.49
40	0.47	0.58	0.55
60	0.56	0.69	0.67
80	0.74	0.79	0.76
100	0.85	0.86	0.90

layer is tuned up with parameters such as CSMA for detecting the collision and the contention window is maximized for faster transmission. Physical transmission range is also extended for making the transmission faster. The packet drop is minimized compared to the existing one.

Below Table 3 shows the values of the proposed and existing systems of PDR. The proposed system achieves better packet delivery ratio of about 90%, where as existing system has 85 and 86%. The graph is plotted taking sensor nodes and PDR on X-axis and Y-axis.

8.3 Throughput

The no. of data packets delivered from source to destination in a specific time period is the Throughput. It is affected by many factors such as routers and cables, congestion, and packet loss. In the proposed system, the number of packets sent is 590. The MAC layer is enhanced with contention window based algorithm which tunes some of the parameters like contention window size. Figure 7 shows the throughput graph of both proposed and existing systems by plotting sensor nodes and throughput in kb/s.

Table 4 have throughput values with comparison with HEED and REECHD. In the EEC system, decongestion method is incorporated in MAC layer. Better throughput is achieved in EEC as the system is initiated with data rate 11 mbps and frequency 2.4 GHz. EEC system achieves by sending 370 data packets to the sink node.

9 Conclusions

An important research area is saving energy in a communication media. Many existing systems having energy efficient mechanisms have been surveyed. It is very difficult to employ a particular type of mechanisms to save energy due to wide range of network applications. MAC uses many approaches to improve the performance by saving the energy. In the proposed system, Energy Efficient Clustering (EEC)

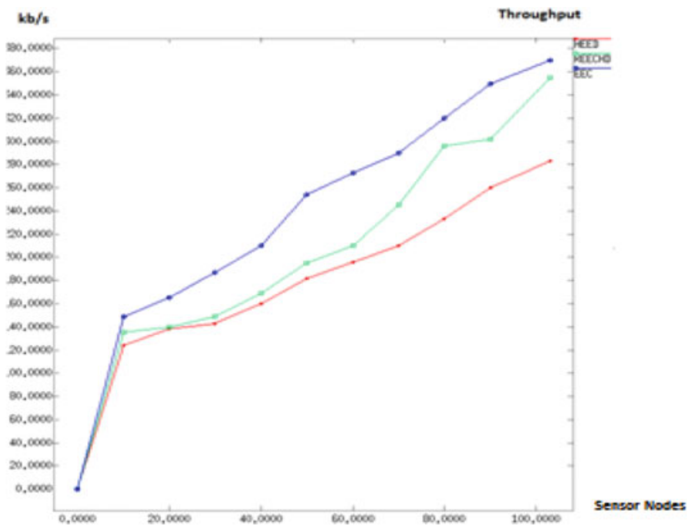


Fig. 7 Throughput

Table 4 Throughput values of proposed and existing systems

Sensor nodes	Throughput (kbps)		
	HEED	REECHD	EEC (Proposed)
0	0	0	0
20	138	140	165
40	160	169	210
60	196	210	273
80	233	296	320
100	283	355	370

is presented for HWSNs which is clustering protocol that uses CH rotation. EEC operates in electing the leader and rotation by combining the node higher energy and node induced work. Intra traffic communication is reduced by this CH selection strategy thus extending the network lifetime. Contention window based algorithm is enhanced in MAC layer for decongestion thereby increasing the throughput and reducing the delay. EEC is more energy efficient compared to HEED and REECHD. In future, many strategies to experiment various member selections for the formation of cluster by incorporating unequal sized clustering can be done.

References

1. Anna Forster (2016) Introduction to wireless sensor networks. Wiley: Technology and Engineering. pp 1–192. https://doi.org/10.1007/978-1-4939-2468-4_1
2. Jan B, Farman B, Javed H, Montrucchio B, Khan M, Ali S (2017) Energy efficient hierarchical clustering approaches in wireless sensor networks: a survey. Wirel Commun Mob Comput 17(2):1–14. https://doi.org/10.1155/2017/6457942
3. Kaur K, Singh M, Singh G (2017) A survey on energy efficient hierarchical clustering algorithm for wireless sensor networks. Int J Eng Comput Sci 6(6):21849–21853. https://doi.org/10.5120/ijca2016907896
4. Jia D, Zhu H, Zou S, Hu P (2016) Dynamic cluster head selection method for wireless sensor network. IEEE Sens J 16(8):2746–2754. https://doi.org/10.1109/JSEN.2015.2512322
5. Farman H, Javed H, Ahmad J, Jan B, Zeeshan M (2016) Grid-based hybrid network deployment approach for energy efficient wireless sensor networks. J Sens 16(2):1–14.https://doi.org/10.1155/2016/2326917
6. Meng X, Shi X, Wang Z, Wu S, Li C (2016) A grid-based reliable routing protocol for wireless sensor networks with randomly distributed clusters. Ad Hoc Netw 51(3):47–61. https://doi.org/10.1155/2016/2326917
7. Sudarshan TV, Manjesh BN (2015) A survey on heterogeneous wireless sensor networks. Int J Eng Res Technol 4(4):1303–1306. https://doi.org/10.17577/IJERTV4IS041398
8. Nayak P, Devulapalli A (2016) A fuzzy logic-based clustering algorithm for WSN to extend the network lifetime. IEEE Sens J 16(1):137–144. https://doi.org/10.1109/JSEN.2015.2472970
9. Zin SM, Anuar NB, Kiah MLM, Pathan ASK (2014) Routing protocol design for secure WSN: review and open research issues. J Netw Comput Appl 41(1):517–530. https://doi.org/10.1016/j.jnca.2014.02.008
10. Mann PS, Singh S (2017) Energy-efficient hierarchical routing for wireless sensor networks: a swarm intelligence approach. Wireless Pers Commun 92(2):785–805. https://doi.org/10.1007/s11277-016-3577-1
11. Banđur D, Jakšić B, Banđur M, Jović S (2018) An analysis of energy efficiency in wireless sensor networks (WSNs) applied in smart agriculture. Comput Electron Agric 156(1):500–507. https://doi.org/10.1016/j.compag.2018.12.016
12. Heinzelman WR, Kulik J, Balakrishnan H (1999) Adaptive protocols for information dissemination in wireless sensor networks. In: Proceedings of the 5th ACM/IEEE mobicom conference, pp 174–185. https://doi.org/10.1016/j.compag.2018.12.016
13. Schurgers C, Srivastava MB (2001) Energy efficient routing in wireless sensor networks. MILCOM Proceedings communications for network-centric operations: creating the information force, vol 1 pp 357–361. https://doi.org/10.1109/MILCOM.2001.985819
14. Younis M, Youssef M, Arisha K (2002) Energy-aware routing in cluster-based sensor networks. In: Proceedings 10th IEEE international symposium on modeling, analysis and simulation of computer and telecommunications systems, pp 129–136.https://doi.org/10.1109/MASCOT.2002.1167069
15. Chang J-H, Tassiulas L (2004) Maximum lifetime routing in wireless sensor networks. IEEE/ACM Trans Networking 12(4):609–619. https://doi.org/10.1109/TNET.2004.833122
16. Nikolidakis SA, Kandris D, Vergados DD, Douligeris C (2013) Energy efficient routing in wireless sensor networks through balanced clustering. J Algorithms 6(1):29–42. https://doi.org/10.3390/a6010029
17. Rana K, Zaveri M (2013) Synthesized cluster head selection and routing for two tier wireless sensor network. J Comput Network Commun 13(3):1–11. https://doi.org/10.3390/a6010029
18. Wang MY, Ding J, Chen W, Guan W (2015) Search: a stochastic election approach for heterogeneous wireless sensor networks. IEEE Commun lett 19(3):443–446.https://doi.org/10.1109/LCOMM.2015.2391100
19. Li H, Liu J (2016) Double cluster based energy efficient routing protocol for wireless sensor network. Int J Wireless Inf Netw 23(1):40–48 https://doi.org/10.1007/s10776-016-0300-9

20. Liu Y, Ota K, Zhang K, Ming M, Naixue X, Anfeng L, Long J (2018) QTSAC: an energy-efficient MAC protocol for delay minimization in wireless sensor networks. IEEE Access pp 8273–8291. https://doi.org/10.1109/ACCESS.2018.2809501
21. Sert SA, Alchihabi A, Yazici A (2018) A two-tier distributed fuzzy logic based protocol for efficient data aggregation in multihop wireless sensor networks. IEEE Trans Fuzzy Syst 26(6):3615–3629. https://doi.org/10.1109/TFUZZ.2018.2841369
22. Alsaafin A, Aghbari ZA, Khedr AM (2018) Heterogeneous-aware distributed clustering for wireless sensor networks. In: International conference on electro/information technology, pp 0012–0017. https://doi.org/10.1109/EIT.2018.8500313
23. Toor AS, Jain AK (2018) A new energy aware cluster based multi-hop energy efficient routing protocol for wireless sensor networks. In: International conference on smart energy grid engineering, pp 133–137. https://doi.org/10.1109/SEGE.2018.8499464
24. Micheletti M, Mostarda M, Piermarteri A (2018) Rotating energy efficient clustering for heterogeneous devices (REECHD). In: International conference on advanced information networking and applications, pp 213–220. https://doi.org/10.1109/AINA.2018.00042

Deep Learning Models for Rubik's Cube with Entropy Modelling

B. V. Amrutha and Ramamoorthy Srinath

Abstract Rubik's Cube is a classic and standard puzzle which has a very large state space of 4.3×10^{19} different states in $3 \times 3 \times 3$ cube. However, there is only one terminal state. Recently, there are few contemporary solutions to solve Rubik cube which exploit Machine Learning Techniques. Our goal is to explore and generate Reinforcement Learning, CNN and LSTM techniques for sequence learning of Rubik cube solution with entropy modelling. The entropy is maximum when the cube is in unsolved state and minimum when it's in solved state. We perform an Entropy traversal modelling for Rubik cube solution.

Keywords Deep reinforcement learning/CNN and LSTM techniques · Sequence learning · Entropy modelling

1 Introduction

Rubik's cube is a successive move puzzle invented by Erno Rubik in 1974 [1]. The cube has six faces and has nine squares on each of these faces (Front, Back, Right, Left, Up, Down) the faces have six fixed and different colours (Red, white, orange, blue, yellow and green). The standard $3 \times 3 \times 3$ cube contains 6 center pieces, 12 edges and 8 corners. The goal of solving the puzzle is to move all the edges and corners into correct positions. The six faces can be rotated in 90 degrees and 180 degrees in any direction to reach the goal state which aligns all 9 squares of same colour onto same faces. The moves that are performed on Rubik cube form a Mathematical structure which is called Group. Rubik cube has a very huge state space around 4.3×101^9 configurations. However, there is only one final goal state.

Rubik cube has 8 corners and 12 edges. These 8 corners can be arranged in 8! ways. Each can be arranged in 3 orientations hence we can have 3^8 possibilities for each permutation of the corner pieces. But only 1/3rd of the permutations has correct corner piece rotations in the Rubik cube. The edges can be arranged in 12! ways. Each edge can be arranged in 2 orientations; hence we can have 2^{12} possibilities

B. V. Amrutha (✉) · R. Srinath
Department of Computer Science, PES University, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_4

of edge permutation. But only 1/2 of the edge permutations have the correct edge-flipping as the original cube orientation. And only 1/2 of these have correct cube re-arrangement parity. Hence, we arrive at the huge distinct initial configuration (4.3×10^{19}) of unsolved state in the Rubik cube [2]. The input to the Rubik cube's solving algorithm can be any one of these configurations. The resulting solution contains same colored squares on all the six sides of the cube. This is universally accepted solution for Rubik cube.

There are few contemporary solutions to solve the cube which exploits machine learning Techniques (MLT) which include Neural Nets, Deep-Heuristic learning and Learned Guidance Function. We have used efficient Deep Learning models like Reinforcement Learning, Convolutional Neural Networks and LSTM for Rubik cube solution. We have studied the suitability of these models in solving Rubik's cube. However, there are few contemporary solutions available to solve the cube using Machine Learning Techniques in which the entropy modelling traversal is not considered. Where entropy is maximum for all unsolved states (scrambled cube) and entropy is zero for solved state (unscrambled cube).

The rest of the paper is organized as follows: Sect. 2 presents related works, Sect. 3 is about proposed methodology, Sect. 4 is experimental results and discussions. The last Sect. 5 presents conclusion and future work.

2 Related Works

This study incorporates investigation on different algorithms used to solve Rubik Cube. We have also studied the mathematical structure of Rubik cube. The Rubik cube operates as a group [3]. The group operation includes a concatenation of sequences of cube moves or rotations on a face of the cube. The popular Sing master notation is representation of cube group [3]. The Rubik cube solution is also studied as temporal sequence [4] learning with multiple actions. The main challenge of this problem is that the domain includes a very large state space. The Application Mechanism [5] focuses to give future scope for building solutions to the combinatorial optimization problems such as Rubik cube Fridrich Method algorithm, layer first law algorithm and Korf algorithm are used to solve Rubik cube.

Various Machine Learning Techniques have been exploited to solve Rubik cube. A deep neural network to solve the RL problem [6] and include a training model to predict the value function for a (state, action) pair. The RL model [6, 7] describes the development of agents that can teach themselves to solve the problem in huge complex domains. The approaches of neural networks [8] in deriving solution for combinatorial problems and few games like Go and Poker. The trained Deep Neural Network (DNN) [9] Regression model which is used as a heuristic for $A*$ search algorithm was only limited to simple heuristic functions. The Learned Guidance Functions (LGFs) [10] is used to solve Rubik cube as a sequence move problem. The different approaches like direct policy discovery (which is a form of reinforcement learning) and genetic programming (GP) [11] for the standard $3 \times 3 \times 3$ Rubik

Cube. The drawback is that this work is limited to only 5 twists in order to facilitate comparison. The scope needs to be increased for a greater number of twists.

The popular algorithm "Approximate Policy Iteration" (API) [12] runs repetitively between two policies named tree search slow policy, and neural network fast policy. The Rubik cube is also studied as an Evolutionary Computations (EC) problem [13, 14] because of its huge search space. The set of goals to find the methods were studied for generating optimal solutions [15] to the Rubik's Cube in an efficiently with respect to both space and time. We explored how neural nets can give good actions in the Rubik Cube setting when provides with very less domain knowledge [10].

3 Proposed Methodology

This section explains the architectural structure of the study. And the explanation of modules in detail. It gives an insight to the design discussions in each phase of the study. Figure 1 shows the Deep Learning Architecture. The proposed method includes the following phases:

3.1 *Data Structure Generation*

The data structure for the cube includes 8 corner piece orientation, 12 edge orientation and 6 center piece orientation. Also, we have defined the rotation of cube by representing in x, y and z axis. The 6 faces of Rubik cube, Front (F), Right (R), Up (U), Down (D), Left (L) and Back (B) have colors like white, orange, blue, yellow, red and green. The 18 moves are considered to solve the cube are. They are the 90°, 180° and 270° rotations on each face of the cube.

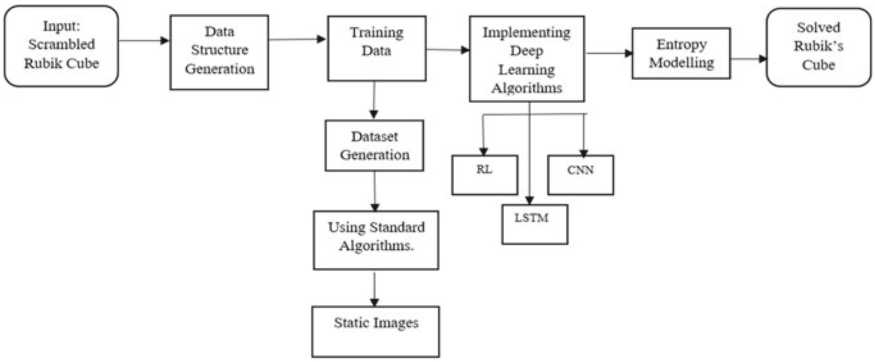


Fig. 1 Deep learning architecture

3.2 Data Set Generation Using Standard Algorithms

The training set is critical to construct an ML model. The training quality depends on the input data set we provide. We have to generated the data set using some of the above-mentioned standard algorithms. Thistlethwaithe and Kociemba algorithms [1] were suitable for our approach. These algorithms take a 54 (squares on Cube i.e. 9 squares on each of the 6 faces) length valid string from 4×10^{19} state space which is generated by considering the data structure i.e. the edge and cube square configurations, as input. The algorithms give the solution moves to reach the final end configuration. By considering every move, the next cube configuration is obtained. Hence, we get a sequence of cube strings required from any initial configuration to the final state. The encoding of the 6 cube faces and 18 moves are done suitably. At the end of the 54-length cube configuration one of the 18 moves is also saved. The dataset we have considered includes 100,000 valid initial configurations of the Rubik cube. And we have generated the sequence of valid moves and its cube configurations as a dataset using standard algorithms. This unique encoded dataset is been used for training deep neural networks.

3.3 Deep Learning Models (RL/CNN/LSTM) Implementation

Deep RL, CNN and LSTM algorithms have been used to build the model and obtain Rubik cube solution.

The CNN and LSTM models which we have implemented are for multi-classification problem. The different CNN training models are considered for training purpose. First model considered is 1-D CNN model which has an input vector of 54×1. Here 54 represents the number of squares in the cube which is given in 1-D format. Second model considered is 2-D CNN model which has an input vector of 9×6. Here 9×6 matrix represents the 9 squares on the 6 faces of the cube. Third model considered is 2-D CNN model which has an input vector of 9×9. When the cube is opened, we can observe that it can be inserted in 9×9 matrix with altering one of the faces to any corners of the 9×9 matrix. This model was considered to understand the importance of spatial adjacency in CNN models. Lastly, the N—Pipeline 2-D CNN model which includes an input array of 9×9. And this model trains 'N' CNN models in a pipeline considering the randomly selected rows in data following the Gaussian distribution. The output of CNN algorithms is the probability of one of the 18 moves required to give the instantaneous move of a cube configuration.

There are two different LSTM models are considered for training which are similar to CNN. First one is LSTM model having an input vector of 54×1. The next model is a LSTM model includes an input vector of 9×6. The input is a sequential input for LSTM model. It saves the sequences at the time of training. The output of LSTM model is the likelihood of one of the 18 moves required to give the instantaneous move of a cube configuration.

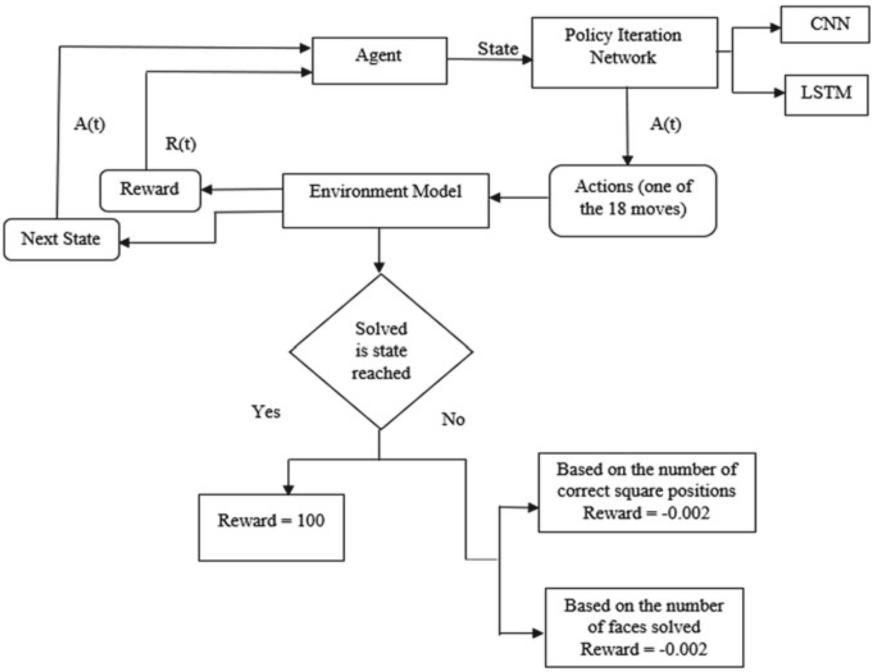


Fig. 2 Flow diagram for RL

The RL model is designed to reach and save the maximum cumulative reward. We have used the Q-learning Table approach. Figure 2 shows the RL flow diagram. The state in the model is different cube configurations. And the actions are a list of 18 moves. The state, action pair is saved based on the maximum reward calculated. To find the action i.e. the move for the cube configuration we are using the above CNN and LSTM model. The environment decides the reward based on the present state and next state of cube configurations. The whole model is trained for a number of episodes. At the start of each episode a random state is considered and fed into a Neural Net obtain an action. Then calculate the reward and continue the episode until the solved state is reached. Calculate the cumulative reward and update the details in Q-table. We find the accuracy based on the maximum rewards over number of episodes.

3.4 Entropy Modelling Traversal

We can define Entropy as "Expected number of bits needed to encode a class". The Shannon's formula to calculate the Entropy is as follows:

$$H(X) = -\sum_{i=0}^{N-1} P_i \log_2 P_i$$

The Entropy of an unsolved Rubik cube is maximum, which decreases to 0 when the Rubik cube is solved. We have observed the trend in entropy as it solves the cube in our algorithm. We have designed an algorithm which calculates the Entropy based on the squares in correct position and squares which are in incorrect position of the Rubik Cube.

4 Results and Discussion

The training accuracy is calculated based on the categorical accuracy values, since it's a multi-classification problem. This accuracy evaluates the index of the maximal true value and it is equal to the index of the maximal predicted value. To test the efficiency, we have automated the test cases by using the next state data structure and choose a random move out of the 18 moves. These test cases are scrambled for a certain number of times from the goal state. Based on the number of test cases solved from different models we can infer the comparison between models.

Tables 1 and 2 shows the number of solved cubes out of 100 with different CNN and LSTM models. These 100 test cases were chosen randomly and stored in Json file and fed to each of the models to observe the comparison. We can infer the better performance in Pipelined 2D models compare to other CNN models from the results shown in the table. And LSTM performed better than CNN model.

The RL model is trained by obtaining the next instantaneous state from CNN and LSTM models. Also, we have considered different number of scrambles while de-

Table 1 Comparison of results between CNN models

No. of scrambles sequence	CNN (54 × 1)	CNN (9 × 6)	CNN (9 × 9)	CNN pipelined-10 models in	CNN pipelined-18 models in
1	100	100	100	100	100
2	92	100	100	100	100
3	71	89	91	94	94
4	38	51	74	77	78
5	31	43	46	51	51
6	20	31	41	45	46
7	6	12	20	24	25
8	7	12	17	19	19
9	5	6	10	11	12
10	4	7	13	13	13

Table 2 Comparison of results between LSTM models

No. of scrambles	LSTM (54 × 1)	LSTM (9 × 6)
1	100	100
2	100	100
3	98	100
4	92	95
5	68	77
6	58	65
7	40	49
8	29	32
9	16	24
10	15	21

Table 3 Training accuracy of RL model

No. of scrambles	CNN	LSTM (9 × 6)	LSTM (9 × 9)
1	1.0	1.0	1.0
2	0.995	1.0	1.0
3	0.90	0.9819	0.9953
4	0.6565	0.8535	0.8875
5	0.4381	0.61	0.7019
6	0.2056	0.3542	0.4575

fining the random state at every episode ranging from 1 to 6. The Table 3 shows the accuracy calculated with Q-learning approach with different models and scrambles.

We have considered the entropy traversal graph by plotting the number of steps involved in following the sequence to solve the cube as x-axis and the entropy values at each step as y-axis. Figure 3 shows a graph with a sample cube configuration and a graph which is plotted by considering the average of entropy values for different solved configurations in the dataset. These graphs show the trend where the entropy is reducing in the initial stages of solving the cube. Later it increases for maximum values and then reduce to zero (solved cube). The variation is observed in the graph plotted against a dataset because all the cube configurations don't have exactly same number of steps to solve the cube it might vary; we can observe how entropy increases at the final steps of solving cube before entropy drops to zero.

5 Conclusion and Future Work

We have implemented the data structure for Rubik cube. We have generated training data sets including sequence of valid moves to solve the Rubik cube with 100,000 initial cube configurations. We have implemented Deep Learning Algorithms like

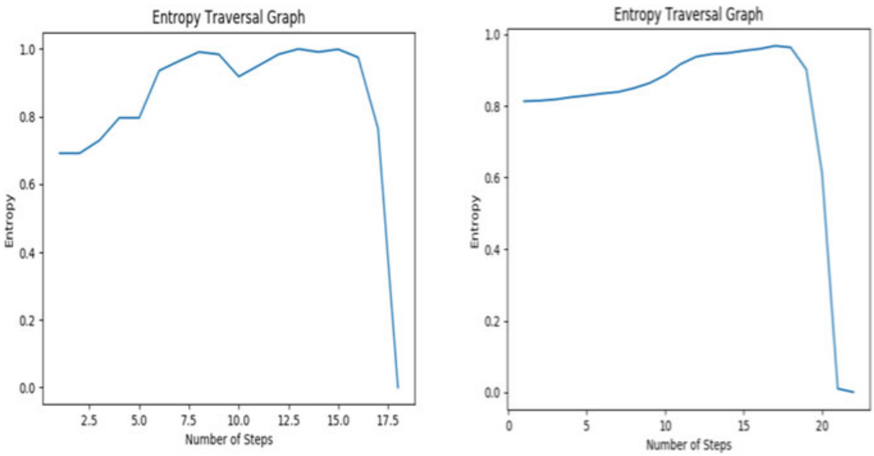


Fig. 3 Entropy graph for single cube configuration and for a datset respectively

Reinforcement Learning, Convolutional Neural Network and LSTM models to obtain Rubik cube solution and showcased the comparison between these models and the efficiency in each of them. We performed the study on suitability of Deep learning models to solve the Rubik cube. However, CNN and LSTM were successful in giving instantaneous moves of the Rubik cube with 100% accuracy. The RL model performed better when LSTM model was chosen to select the actions.

We also performed Entropy modelling traversal for solving Rubik cube. The observation was made such that the trend was always reducing the entropy in the initial stages of solving the cube, later it was forcefully increased to maximum values and then drops to zero. For future work we can concentrate on improving the proficiency of different models considered and also implement Transfer learning prototype to solve real world problems using Rubik cube solution derived.

References

1. Kaur HH (2015) Algorithms for solving the Rubik's cube. CSC, KTH
2. The mathematics of the Rubik's Cube, Introduction to group theory and permutation puzzles. MIT (2009)
3. Daniels L (2014) Group theory and the Rubik's Cube. Lakehead Universityc
4. Lichodzijewski P, Heywood M (2010) The Rubik cube and GP temporal Sequence learning: an initial study*. Genet Program Theor Pract
5. Zeng D-X, Li M, Wang J-J (2018) Overview of Rubik's cube and reflections on its application in mechanism. Chin J Mech Eng
6. Bowman NW, Guo JL, Jones RMJ (2018) BetaCube: a deep reinforcement learning approach to solving 2 × 2 × 2 Rubik's cubes without human knowledge. Stanford University
7. McAleer S, Agostinelli F, Shmako A, Baldi P (2018) Solving the Rubik's cube without human knowledge. University of California

8. Brunetto R, Trunda O (2017) Deep Heuristic-learning in the Rubik's cube domain: an experimental evaluation. Charles University in Prague
9. Tawfik A, Shaham R, Reznik L, Bielous G (2018) Introduction to artificial intelligence 2018 final project—Learning heuristics for the Rubik's Cube
10. Johnson CG (2018) Solving the Rubik's cube with learned guidance functions. University of Kent
11. Smith R (2016) Evolving policies to solve the Rubik's Cube: experiments with ideal and approximate performance functions. Dalhousie University
12. McAleer S, Agostinelli F, Shmako A, Baldi P (2019) Solving the Rubik's Cube with approximate policy iteration. Int Conf Learn Representations (ICLR)
13. Abo El-Maaty AM, Fayek MB (2017) Observations on exploration and exploitation effects on solving Rubik's Cube using evolutionary strategies. Int Computer Eng Conf (ICECO)
14. El-Sourani N, Hauke S, Borschbach M (2010) An evolutionary approach for solving the Rubik's Cube incorporating exact methods. In: EvoApplications 2010: applications of evolutionary computation
15. Cheeseman A, Teutenberg J (2001) Towards optimal solutions for the Rubik's cube problem. COMPSCI 765 FC Adv Artif Intell
16. Irpan A (2016) Exploring boosted neural nets for Rubik's Cube solving. University of California

Detecting Diabetic Retinopathy Using Deep Learning Technique with Resnet-50

Viraj Jiwane, Anubhav DattaGupta, Arunkumar Chauhan, and Vidya Patil

Abstract Diabetes mellitus, also known as diabetes, is a metabolic disease which may cause high blood sugar, blurred vision and several other complications. Patients already suffering from diabetes have a greater probability to develop Diabetic Retinopathy. Most prominently in the blood vessels of the tissue residing in the rear portion of the eye gets damaged causing vision problems. This disease is especially prevalent in India. Test for Diabetic Retinopathy involves a dilated eye exam, which is time consuming and requires specialized doctors and distinct equipment which may not be easily accessible to the people especially in rural areas. Predicting the possibility of Diabetic Retinopathy utilizing fundus image is considered an effectual analysis subject by many. In this paper a method is explored to diagnose Diabetic Retinopathy using fundus images and deep learning to accurately tell the patient whether he has a Diabetic Retinopathy. Features used from the eye fundus images are soft exudate and hard exudate along with optic disc. A website is also made to deploy the model and promote mass diagnosis at a faster pace.

Keywords Fundus images · Deep learning · Image processing · ResNet

1 Introduction

Diabetes often leads to Diabetic Retinopathy (DR) which causes aberrations in the retina, with severe cases resulting in blindness. Most notable symptoms of this abnormality don't arise till the disease reaches a severe state. With small changes to the retinal capillaries [3]. Eventually in due time both hard exudates & retinal edema are followed by rise in permeability of the capillary walls. Enfeebled blood vessels spill lipid formations which form hard exudates this causes clogs with the blood vessels leading to micro infarcts in the retina called exudates. When a noteworthy amount

V. Jiwane (✉) · A. DattaGupta · A. Chauhan
Department of Computer Engineering, Maharashtra Institute of Technology, Pune, India

V. Patil
School of Computer Engineering and Technology, MIT World Peace University, Pune, India
e-mail: vidya.patil@mitwpu.edu.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_5

of hemorrhages, exudates or microvascular malformations are discovered, Diabetic Retinopathy is presumed to be in nonproliferative state [4]. Year 2000 saw India's population figure affected by diabetes mellitus (DM) rise to 31.7 million as stated by WHO. If the current trend is followed then by 2030 affected population is set to further rise and become 79.4 million which is by far greatest in the world [5]. While non increasing Diabetic Retinopathy in itself cannot directly affect vision it triggers macular oedema or macular ischaemia which result in rapid vision loss at any phase [6]. Early detection encourages right treatment being received which can prevent sight-threatening conditions [7]. The objective of this paper is to explore a new method of detecting Diabetic Retinopathy using deep learning specifically Resnet50 [8] making use of the features such as soft exudate, hard exudate along with optic disc.

2 Literature Survey

There has been extensive work done in the field of Computer aided diagnosis of fundus images. Many classical and adaptive methods have been applied in the recent past to detect DR. Some of the recent works with classical methods have shown an accuracy as high as 90% [24]. Deep learning methods have shown very promising results. Chakrabarty and Chatterjee presented a model using CNN and Relu that gave 100% accuracy [9] Shah et al. showed that Artificial Intelligence (AI) would not require help of an eye doctor for detecting presence of Diabetic Retinopathy. This would promote Remote screening and increase the amount of patients that can be scanned in a particular time period revolutionizing telescreening, especially where people have poor access to specialized health care [10]. Dharwadkar et al. in a study to identify the flowers consisting of high semantic features found that Deep Learning based neural networks gives a better accuracy in comparison with a Simple neural network [11]. Survey of Identification of Diabetic Retinopathy Detection within Fundus Images is as shown in Table 1.

Alongside detecting hard exudates, it is also essential to detect soft exudates by employing some texture analysis methods. Doing so will aid in increasing the reliability of the solution. The simple neural network is ineffective. A neural network trained over a small dataset cannot give acceptable results. The minimal presence of knowledge prohibits the model to generalize and learn hidden features and patterns in the dataset and so this leads to higher losses and lower probabilities of getting accurate answers. Deep neural networks face challenges when the classes in the dataset are imbalanced. The gap between the population of a certain class from other classes makes the model biased towards the said class. This challenge is faced by every application of DNN in every domain. Few papers showed high accuracy using Deep Learning Techniques using relatively small datasets.

Table 1 Survey of identification of diabetic retinopathy detection within fundus images. Following are the common conclusions made from the survey of recent work as shown above

Reference number	Technique's used	Advantages of techniques	Accuracy	Future scope/gaps	Dataset
[12]	Resnet50, Inceptionv3, Xception, Dense121, Dense169	Proposed ensemble model trained on a large dataset has better performance in comparison to other modern methods. It is also able to detect all stages of diabetic retinopathy (DR)	80.08%	Failed to handle database bias. Specific models are to be trained for particular stages after which ensemble should be applied to the outcome to increase accuracy at early stages of DR	Kaggle and DiaretDB1
[13]	CNN, Relu, Max-pooling	Validation accuracy obtained is 100%	100%	Currently the model is applied to a dataset of 30 HRF images more images should be incorporated for variations	High-resoluti on fundus (HRF) image database
[14]	SVM	Suggested tree-based feature Selection gives a noteworthy improvement in Precision and Recall over other existing approaches using irrelevant features-building	88.3%	–	DIARET-DB

(continued)

Table 1 (continued)

Reference number	Technique's used	Advantages of techniques	Accuracy	Future scope/gaps	Dataset
[15]	Inception V3, Inception-Re snet-V2 and Resnet152	These models were amalgamated which obtained superior results	Inception V3 = 87.91%, Resnet 152 = 87.20%, Inception-Resnet-V2 = 86.18%, Integrated Model = 88.21%	In case a dataset is biased the trained model will have the same bias due to lack of weight assignment to optimal features	Beijing Tongren eye center patients independent database constructed
[9]	CNN	100% Accuracy validation and recall of 1.0 was obtained by feed-forward convolutional neural network—SVM hybrid model	100%	As the current model uses a small dataset. Working of this technique on a larger dataset can be optimized to achieve descent set of results	High resolution fundus (HRF) image database
[16]	Recursive region growing algorithm	Identification of hard exudates within fundus images with acceptable sensitivity and accuracy is achieved	Sensitivity = 0.87, F-Score = 0.78 Positive predictive value (PPV) = 0.76 for hard exudate lesion level detection	Performance can be refined by selection of proper features w.r.t exudates and machine learning based approaches should be used	DiaretDB1
[17]	CNN, SVM, ggNet, Alexnet, InceptionNet GoogleNet, DenseNet, and Resnet,	With the help of fine tuning CNN classification process computation time is reduced	Resnet + SVM = 9 5.83% InceptionV3 or VGGNet type-19 = 95.24%	A sizable amount of data and classes can be used to broaden the training process	Messidor database
[18]	AlexNet, VGG-16 and SqueezeNet,Customized5 layered CNN model	Superior results are achieved by customized 5 layered CNN models	AlexNet = 93.46%, VGG-16 = 91.82%, Customized CNN = 98.15%	–	The MESSIDOR

(continued)

Table 1 (continued)

Reference number	Technique's used	Advantages of techniques	Accuracy	Future scope/gaps	Dataset
[19]	CNN	Proposed work shows using deep neural network while detecting diabetic retinopathy is feasible	74%	Testing on diverse datasets necessary to boost the results	DIRETDB0,D IARETDB1

3 Proposed System

Symptoms (like hard exudate, optic disc etc.) need to be detected as well as classified into 5 ranges of severity from the fundus image. Two different modules are employed. First is a Feature Extraction module for detecting symptoms in the image and second one is an Fundus Images Classifying module to classify the image into various severity ranges. A deep neural network is used for this purpose as shown in Fig. 1.

3.1 Feature Extraction in Fundus Image

Classification: Resnet-50 is used as a base model.

Detection: For symptom detection Faster R-CNN architecture is used. Faster R-CNN is an object detection neural network architecture. This architecture is better in performance, speed and number of parameters from its predecessors because the architecture shares convolutional layers with RPN trying to learn convolutional features and thus learning the localization of the bounding boxes [20].

Algorithm Before processing starts, images are resized down keeping the aspect ratio approximately constant. Within object detection models before actual training of the model starts, Faster R-CNN takes one of the inputs as ground truth boxes. So we need to generate these boxes using a GUI image annotation tool called Labeling. This tool saves ground truths as xml files. Faster R-CNN has 3 conceptual parts.

1. A classifier: Resnet 50 is used as a classifier and weights are initialized with resnet-50 coco dataset. Coco is one of the largest datasets for object detection and segmentation trained over more than 3, 30,000 images.

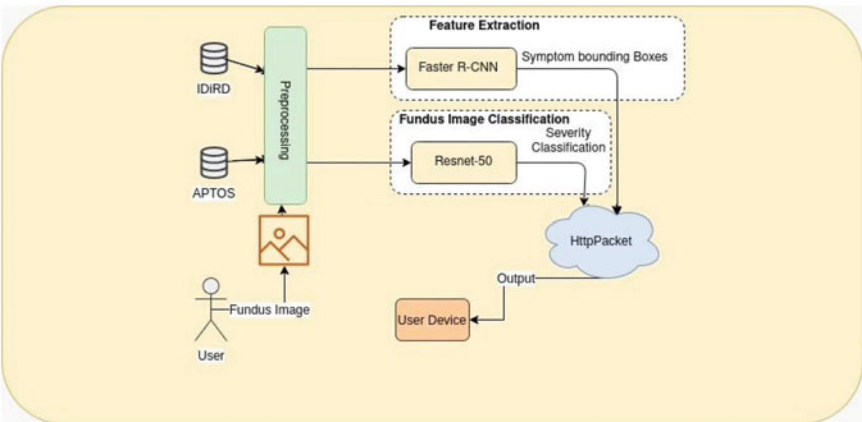


Fig. 1 High level system architecture of the proposed solution

2. A Regional Proposal Network: it is a shallow network of 4 convolutional layers responsible for generating various regions of interest. ROI is generated using a technique called anchor boxes. Where every anchor is a pixel in the image. Various ROI's are produced centered on these anchors. Three (3) different sizes and 4 different shapes are chosen for our model. Thus 12 (3×4) ROIs are generated for every anchor which leads to generation of a huge number of ROIs. Before feeding them to the next network most of the ROI's are removed. By manual observation of the dataset, it is found that no image in the dataset had symptoms in more than 50 locations. So we restricted the number of boxes going in the next layer to 100, which reduces computation time for the next layer significantly.
3. A Detection layer: It takes input from both feature maps generated by classifiers and ROIs generated by RPN and tries to classify images in each ROI. Resnet-50 is chosen as a classifier for this purpose. RPN and this layer are the only 2 layers that are trained. L2 Regularizer and softmax functions are used for score conversion. As the model gets trained we reduce the learning rate by a factor of 10 after every 1000 (approx) epochs. As an output of object detection we got bounding boxes (N $\times$ vector of size 4), labels (size $\times$ Nx1) and scores (N $\times$ 1). Bounding boxes contain diagonal corner coordinates of the rectangle localizing a symptom and the label has the prediction of which symptom was detected. And score has a probability value of the confidence with which the detection is correct.

3.2 Fundus Image Classification

The images are resized to 575×575 as opposed to conventional 224×224 because increasing the resolution a little improves the accuracy [21]. The images were given as an input into the Resnet-50 classifier, which is the most popular image classifier as it very effectively, tackles the problem of vanishing gradient by using identity connections [22]. Proposed model has a global average 2D pooling layer at end of resnet-50 network followed by a dense layer and a dropout layer of decay probability 0.5 and another dense layer and another drop out layer of decay 0.5 and a output layer of size 5 at the end [23]. All the layers are trainable. For training we used initial learning rate as 10–4 and then reduced learning rate by factor of 0.5 every time the model gets in a plateau for 3 epochs. Binary cross-entropy is used as a loss function and softmax as a final layer activation function. Validation loss is also monitored for early stopping of training of the model, if the model stays in the plateau for more than 5 epochs. On prediction we got a one hot encoded vector of size 5 which contains the prediction (classified range of severity). A web application is developed with a very intuitive UI for the proposed model. The user needs to upload a fundus image and the image will be sent to the server through http request. On the server side the image gets resized to 575×575. Both the Feature Extraction and Fundus Images Classifying models run simultaneously with the same input image. By multithreading the time taken to process has been reduced from 5 min to 113 s.

4 Results

This section shows the results obtained by implementing the system. Figure 2 shows a graph of training loss and accuracy. We observed higher accuracy (close to 0.9) and lower loss (below 0.4) but with very slight model over fitting.

4.1 Symptoms Identification

Different symptoms of Diabetic Retinopathy are identified like soft exudate, Hard Exudate, hemorrhage, Optic Disc, Microaneurysms etc. and bounding boxes are drawn around it with its confidence. The Result of symptom identification is as shown in Table 2.

Fig. 2 Training loss and accuracy

Table 2 Result of symptom detection

Category	Label	Precision	Recall
0	Optic dis	1.0	0.944444
1	Hard exudate	1.0	0.000781
2	Haemorrhage	1.0	0.000790
3	Soft exudate	0.8	0.036364
4	Microaneurysm	NaN	0.00000

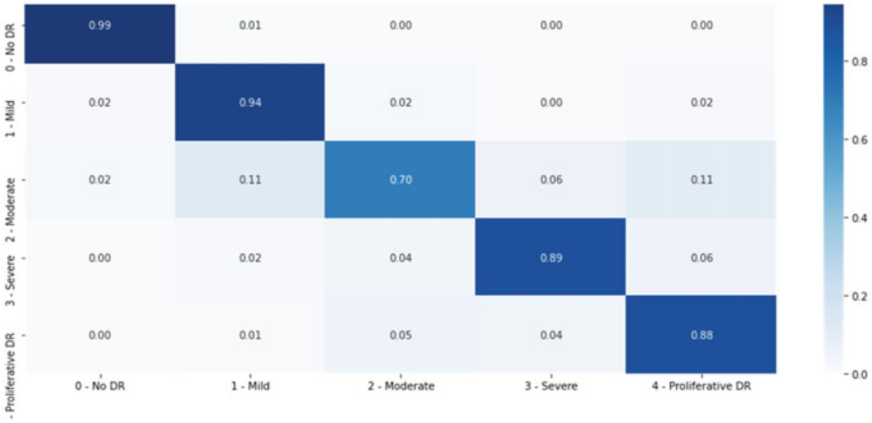


Fig. 3 Classification matrix

Table 3 Performance of matrices of proposed solution

	Precision	Recall	Jaccard score
Micro	0.9393773894046968	0.9393773894046968	0.8856848609680742
Macro	0.9047717580233817	0.8571614510440508	0.78531579319679
Weighted	0.9408622954749201	0.9393773894046968	0.89062825128829

Table 4 Balanced accuracy score of various models with different no of trainable layers

Model	First 2/3rd layers trainable	First 4/5th layers trainable	All layers trainable
Resnet50	0.498975332134017446	0.6476205959503115	0.8795748788451
Resnet152 V2	0.20281366111874588	0.2	0.3
VGG19	0.2	0.2	0.2

4.2 Disease Severity Classification

In Disease severity classification we identified the severity class to which fundus image belongs. Results of classification are as given in Fig. 3. It shows Accuracy of 0.94 over 3662 Images (Tables 3 and 4).

5 Conclusion and Future Scope

Paper shows implementation of a Diabetic retinopathy recognition system based on Faster R-CNN and Resnet50 for Object detection and Image classification respectively. About 94% accuracy is achieved for severity detection of ‘Diabetic retinopathy'. A very user friendly and intuitive GUI is also developed for the proposed system. The GUI can be used by a patient or a doctor. Proposed model shows that diagnosis of this very dangerous disease can be brought to every household. The system takes 113 s to generate the complete result. If this can be accelerated by either better architecture, techniques, models, dataset or training time, the system can be used for real time prediction by connecting it to a fundus camera. A few things which can be worked upon are the difficulties in detecting microaneurysms and hemorrhages. Future scope also includes reducing over fitting of the model by training over a larger dataset or an ensemble of different datasets. A variation in exposure, focal length, and resolution will ultimately lead to generalization of the features of the symptoms and the image as a whole.

References

1. Watson S (2018) Diabetes: symptoms, causes, treatment, prevention, and more. Healthline, 04 Oct 2018. https://www.healthline.com/health/diabetes (accessed 16 June 2020)
2. Diabetic retinopathy—Diagnosis and treatment—Mayo Clinic, May 2018. Accessed: 16 June 2020. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/diagnosis-treatment/drc-20371617
3. Nentwich MM (2015) Diabetic retinopathy—ocular complications of diabetes mellitus. World J Diabetes 6(3):489. https://doi.org/10.4239/wjd.v6.i3.489
4. Saha SK, Xiao D, Bhuiyan A, Wong TY, Kanagasingam Y (2019) Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: a review. Biomed Signal Process Control 47:288–302. https://doi.org/10.1016/j.bspc.2018.08.034
5. Gadkari SS, Maskati QB, Nayak BK (2016) Prevalence of diabetic retinopathy in India: the all India ophthalmological society diabetic retinopathy eye screening study 2014. Indian J Ophthalmol 64(1):38–44
6. Manjramkar M (2018) Survey of diabetic retinopathy screening methods. In: 2018 2nd international conference on trends in electronics and informatics (ICOEI). https://doi.org/10.1109/icoei.2018.8553843
7. Website. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy (accessed 16 June 2020)
8. Wu Y, Hu Z (2019) Recognition of diabetic retinopathy based on transfer learning. In: 2019 IEEE 4th international conference on cloud computing and big data analysis (ICCCBDA). https://doi.org/10.1109/icccbda.2019.8725801
9. Chakrabarty N, Chatterjee S (2019) An offbeat technique for diabetic retinopathy detection using computer vision. In: 2019 10th international conference on computing, communication and networking technologies (ICCCNT). https://doi.org/10.1109/icccnt45670.2019.8944633
10. Shah P, Mishra D, Shanmugam M, Doshi B, Jayaraj H, Ramanjulu R (2020) Validation of deep convolutional neural network-based algorithm for detection of diabetic retinopathy—artificial intelligence versus clinician for screening. Indian J Ophthalmol 68(2):398. https://doi.org/10.4103/ijo.ijo_966_19
11. Dharwadkar S, Bhat G, Subba Reddy NV, Aithal PK (2017) Floriculture classification using simple neural networks and deep learning. In: 2017 2nd IEEE international conference on recent trends in electronics, information and communication technology (RTEICT). https://doi.org/10.1109/rteict.2017.8256671
12. Qummar S et al (2019) A deep learning ensemble approach for diabetic retinopathy detection. IEEE Access 7:150530–150539. https://doi.org/10.1109/access.2019.2947484

13. Chakrabarty N (2018) A deep learning method for the detection of diabetic retinopathy. In: 2018 5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON). https://doi.org/10.1109/upcon.2018.8596839
14. Huda SMA, Asiful Huda SM, Ila IJ, Sarder S, Shamsujjoha M, Ali MNY (2019) An improved approach for detection of diabetic retinopathy using feature importance and machine learning algorithms. In: 2019 7th international conference on smart computing and communications (ICSCC). https://doi.org/10.1109/icscc.2019.8843676
15. Jiang H, Yang K, Gao M, Zhang D, Ma H, Qian W (2019) An interpretable ensemble deep learning model for diabetic retinopathy disease classification. In: 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC), vol 2019, pp 2045–2048
16. Palavalasa KK, Sambaturu B (2018) Automatic diabetic retinopathy detection using digital image processing. In: 2018 international conference on communication and signal processing (ICCSP). https://doi.org/10.1109/iccsp.2018.8524234
17. Qomariah DUN, Tjandrasa H, Fatichah C (2019) Classification of diabetic retinopathy and normal retinal images using CNN and SVM. In: 2019 12th international conference on information and communication technology and system (ICTS). https://doi.org/10.1109/icts.2019.8850940
18. ur-Rehman M, ur-Rehman M, Khan SH, Abbas Z, Danish Rizvi SM (2018) Classification of diabetic retinopathy images based on customised CNN architecture. In: 2019 Amity international conference on artificial intelligence (AICAI). https://doi.org/10.1109/aicai.2019.8701231
19. Arora M, Pandey M (2019) Deep neural network for diabetic retinopathy detection. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon). https://doi.org/10.1109/comitcon.2019.8862217.
20. Website. https://arxiv.org/pdf/1506.01497
21. Zheng L, Zhao Y, Wang S, Wang J, Tian Q (2016) Good practice in CNN feature transfer (01 Apr 2016)
22. Wang F et al (2017) Residual attention network for image classification. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). https://doi.org/10.1109/cvpr.2017.683
23. Bhalekar M, Sureka S, Joshi S, Bedekar M (2020). Generation of image captions using VGG and ResNet CNN models cascaded with RNN approach. https://doi.org/10.1007/978-981-15-1366-4_3
24. Amin J et al (2016) A review on recent developments for detection of diabetic retinopathy. Scientifica 2016:6838976. https://doi.org/10.1155/2016/6838976

Restoration of Rician Corrupted MR Data Using Improved Hybrid Model

Vedant Shukla, Prasad Khandekar, and Arti Khaparde

Abstract Restoration of MR Images using a hybrid model consisting of two stages is implemented. Application of Fourth Order PDE and the application of relaxed median filter are pre and post stages respectively. The results were calculated using standard parameters like PSNR, SSIM and RMSE. It was found that the hybrid mm gives acceptable results when compared with standard non-linear filters when applied to T1 weighted Thorax Images. For Brain MR Images, the hybrid model filter performs better for a finite range of input noise level ($\%N$). The algorithm was implemented on synthetic images of T1 weighted Thorax and Brain MRI using MATLAB 2015a.

Keywords Fourth order PDE · Relaxed median filter · T1 weighted MR data

1 Introduction

MR Data is naturally corrupted by Rician noise. Restoration of such data becomes important. Presence of Rician noise affects the post-processing of molecular medical images.

Restoration of clinical molecular images such as CT Scan, Angiography, Ultrasound, MRI, MRA, etc. becomes important for analysis. Many filtering algorithms are developed in the last few decades for molecular images restoration. Although, it should be noted that staircase effect exists even in absence of noise [17]. A generic image denoising filter using a mask template (5×5) based on fuzzy logic to eliminate speckle noise [6]. Saladi et al. in 2016 proposed a Spatially Adaptive Non-Local Means Filter (SANLM) [11].

V. Shukla (✉) · P. Khandekar · A. Khaparde
Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India
e-mail: prasad.khandekar@mitwpu.edu.in

A. Khaparde
e-mail: arti.khaparde@mitwpu.edu.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_6

Fourth Order PDE is one method used to eliminate noise while preserve edges [16]. Liu et al. in [12] improved the speckle compensation algorithm which is an improvement over YK model. Fourth Order PDE is one method used to eliminate noise while preserve edges [12]. Ling et al. worked on hybrid model based on a linear combination of anisotropic diffusion and median filter [15]. Lai et al. proposed local and non-local Structured Oriented Total Variation for noise suppression [7]. In 2017, Khmag et al. worked on the estimation of Gaussian noise using local statistics. U like most methods discussed, a small patch of image is selected and the high frequency components are eliminated [9]. Sameera et al. reviewed all the existing methods of filtering medical images including MR images [3].

This research paper mainly focuses on the restoration of Rician distributed MR images using a hybrid model proposed by Rajan et al. [14], which was an improvement provided to the article by Ling and Bovik [15]. In this article, the anisotropic diffusion is applied, due to which staircase effect occurs. It is then filtered out using a median filter. Whereas, in [14], the anisotrpic filter is replaced by a fourth order PDE to avoid staircase effect and is followed by a relaxed median filter. The algorithm is implemented on T1w Brain and Thorax MR images.

This research article is divided into five sections. This section is the overall introduction, Sect. 2 overviews the background including noise distribution in MR images, the hybrid model with algorithm. Section 3 discusses the performance of hybrid model on T1w Thorax and Brain MR images using standard parameters, while Sect. 4 draws conclusion and summarization of the experimental results. Section "Acknowledgements" acknowledges the synthetic MR images used for carrying out this experimentation.

2 Background

2.1 Rician Distribution in Molecular Imaging

Molecular imaging is method used to non-invasively monitor the biochemical processes inside living organism in real time. Molecular Imaging is performed extensively in the field of radiology. MR Imaging has different modalities such as T1 weighted, T2 weighted, Proton density, Diffusion weighted etc. In this research paper, we focus on T1 weighted MR Imaging Modalities.

The Probability Distribution Function (PDF) of a Rician distributed image is expressed as:

$$p_M(M|A, \sigma_n) = \left[\frac{M}{\sigma_n^2}\right] \exp\left[-\frac{M^2 + A^2}{2\sigma_n^2}\right] I_o\left(\frac{AM}{\sigma_n^2}\right) u(M) \tag{1}$$

Here, $I_o(\cdot)$ is the 0 order Bessel Function of the first kind and $u(\cdot)$ is the Heaviside step function.

2.2 Application of Hybrid Filter in MR Imaging

For restoration of MR Images, Non-linear filtering is used. Non-Linear Fourth Order PDE is used for restoration by using a model of You and Kaveh as:

$$\frac{\delta u}{\delta t} = -\Delta^2[c(|\Delta^2 u|)\Delta^2 u] \tag{2}$$

$\Delta^2 u$ is the Laplacian of the image u. The Fourth Order PDE proves to be a good filter. The relaxed median filter is used to remove large spike noise. It is applied locally to a 9*9 window and the output can be mathematically shown as:

$$Y_{(ij)} = RM_{\alpha,\omega} W_{(ij)} \tag{3}$$

Here,

$Y_{(ij)} = X_{(ij)};\ if\, X_{(ij)} \in [[W_{ij}]_\alpha [W_{ij}]_\omega]$ and

$Y_{(ij)} = [W_{ij}]_m;\ otherwise.$

The relaxed median has two bounds, lower bound α and upper bound ω which a sublist in window $[W_{ij}]_{(\cdot)}$ which are need not be filtered. The hybrid model can therefore is given in [14] as:

$$u_{(ij)}^{(n+1)} = RM_{\alpha,\omega} u_{(ij)}^{n} - \Delta t \Delta^2 g_{(ij)}^{n} \tag{4}$$

3 Experimental Results

Experiments of restoration of Rician distributed T1 weighted synthetic Brain MR Images from [1] and Thorax images [2] is performed. The clean images are corrupted synthetically by Rician distribution at different noise levels ranging from 0.1 to 10. The hybrid model was applied and the results were compared with un-corrupted and median filtered images using parameters such as PSNR and SSIM.

3.1 Hybrid Model Applied to Thorax MR Images

Thorax is the part of human anatomy between the neck and the abdomen. Thorax imaging is usually done in case the subject suffers from frequent chest pains. The original, corrupted and filtered thorax images are as shown in Fig. 1.

Table 1 depicts the relation between %N and PSNR; where PSNR 1 is that of the Rician corrupted and original images, PSNR 2 of hybrid model filtered and original image and PSNR 3 of the median filtered and original image. Ideally PSNR should be ∞, i.e. Noise should be 0. PSNR 1 can be observed to be degrading from 86.52 to

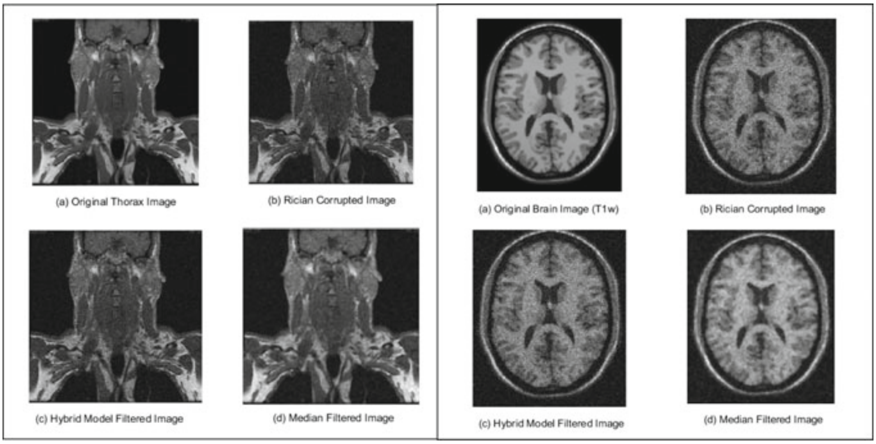


Fig. 1 **a** Original image [1]. **b** Synthetically Rician corrupted image. **c** Hybrid model filtered image. **d** Median filtered image

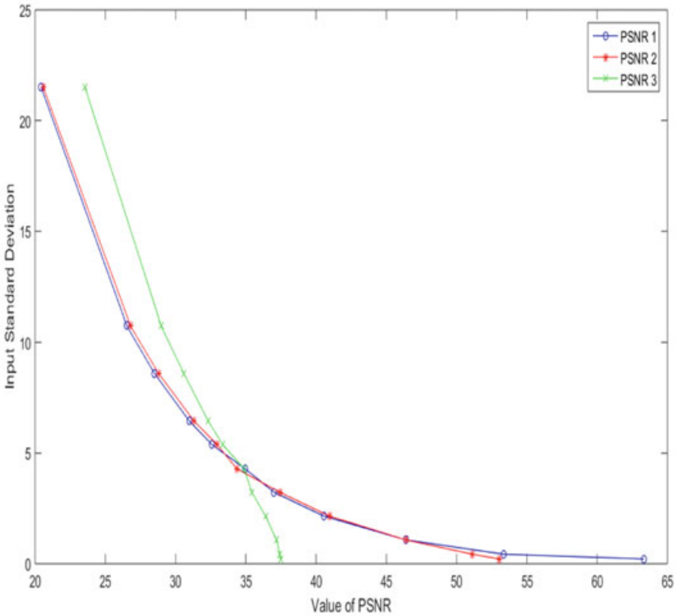


Fig. 2 Plot of noise level versus standard deviation (Table 1)

27.49 within the input noise levels of 0.1–10. The hybrid model gives better results than median filter for low noise levels. The same can be observed from the PSNR plot shown in Fig. 2.

Table 1 Input and calculated σ for T1 weighted image

%N	PSNR 1	PSNR 2	PSNR 3
0.1	86.52	51.25	37.05
0.2	66.76	51.20	37.01
0.5	52.88	50.09	36.98
1	47.91	47.18	36.88
1.5	44.40	44.57	36.78
2	41.92	42.33	36.52
2.5	39.97	40.47	36.28
3	38.40	38.92	35.98
4	35.79	36.30	35.23
5	35.77	36.28	35.20
10	27.49	27.79	30.17

In Table 1, In the noise level range of $0.1 < \%N < 1$, it can be observed that PSNR 2 shows better results than PSNR 3 but the unfiltered image has a higher PSNR, i.e. on application of filtering at low SNRs, the image is getting further degraded. Hence for such low noise levels, the hybrid and median filters should not be used. For noise levels $1 < \%N < 5$, the hybrid model starts showing better results than unfiltered or median filtered image but for $\%N \leq 10$, the median filter performs better than hybrid filter. Although, our concern is only the specific noise level images ($1 < \%N < 5$). A similar trend can be seen in the readings of SSIM and RMSE of the images as shown in Table 2.

One major image quality index is RMSE (Root Mean Square Error). RMSE is often used as a measure of difference of predicted and calculated value of a variable in a data set. In Table 2, RMSE follows the same results as obtained by PSNR in Table 1.

3.2 Hybrid Model Applied to Brain MR Images

In Fig. 1, the Brain MR synthetic image is shown in (a), ideally with 0 noise. This image is synthetically corrupted with Rician noise as shown in b. The hybrid model and median filter are applied on the corrupted image as shown in (c) and (d) respectively.

Table 3 depicts the relation between %N and PSNR; where PSNR 1 is that of the Rician corrupted and original images, PSNR 2 of hybrid model filtered and corrupted image and PSNR 3 of the median filtered and original image. Ideally PSNR should be ∞, i.e. Noise should be 0. PSNR 1 can be observed to be degrading from 63.35 to 20.44 within the input noise levels of 0.1–10. The hybrid model gives better results than median filter but for low noise levels ($N \leq 2$). Beyond this noise level, median

Table 2 Noise level versus RMSE of thorax and brain images

%N	SSIM 1	SSIM 2	SSIM 3
0.1	0.999	0.996026874	0.94416822
0.2	0.999865249	0.995894508	0.944150441
0.5	0.996577028	0.994290926	0.943564794
1	0.988076378	0.987399465	0.940679686
1.5	0.973385277	0.974989389	0.934875036
2	0.95290366	0.956925381	0.924012548
2.5	0.928565749	0.934823639	0.910838339
3	0.901612739	0.910033036	0.894814941
4	0.842080954	0.854229135	0.860977506
5	0.841978695	0.854302241	0.860943175
10	0.546536115	0.562242711	0.694472541
%N	RMSE 1	RMSE 2	RMSE 3
0.1	0	0.6973	3.5996
0.2	0.1164	0.7047	3.5998
0.5	0.5779	0.7928	3.6134
1	1.0506	1.1180	3.6420
1.5	1.5478	1.5228	3.7137
2	2.0401	1.9451	3.7891
2.5	2.5679	2.4264	3.9129
3	3.0755	2.8954	4.0804
4	4.1337	3.9005	4.4341
5	5.1980	4.9243	4.8864
10	10.8107	10.4422	8.0014

Table 3 Input and calculated σ for T1 weighted image

%N	PSNR 1	PSNR 2	PSNR 3
0.1	63.35	53.04	37.49
0.2	53.35	51.12	37.42
0.5	46.38	46.39	37.18
1	40.56	40.94	36.40
1.5	37.02	37.43	35.43
2	34.93	34.35	34.86
2.5	32.60	32.96	33.36
3	30.97	31.31	32.31
4	24.49	28.79	30.58
5	26.52	26.78	28.97
10	20.44	20.59	23.55

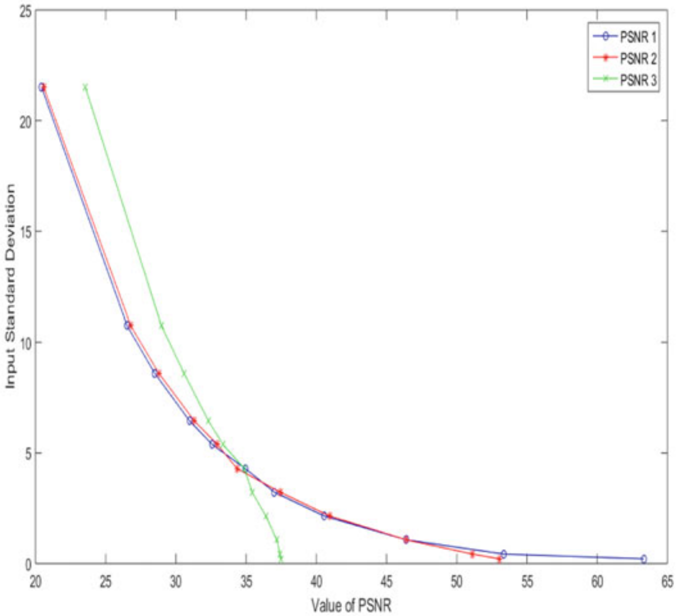


Fig. 3 Plot of noise level versus standard deviation (Table 3)

filter gives better results. The same can be observed from the PSNR plot shown in Fig. 3. The (SSIM and RMSE for the images is also computed as shown in Table 4. SSIM of the images show a similar trend as SSIM for ($N \leq 1.5$) shows negligible improvement.

4 Conclusion

For filtering of MR Images, the hybrid model and median filter are applied. The fourth order PDE and the relaxed Median filter are applied to T1 weighted Thorax and Brain MRI. In case of T1 weighted thorax MR image, it was observed that for $\%N > 1$, the hybrid model filter consistently gives better results than median filter. A similar trend was observed for T1 weighted Brain MR Images but for $\%N \leq 2$ the median filter gives better results than hybrid model filter. Also, it should be noted that for $\%N < 1$, filtering the images is not required as the application of filters further degrades the image.

Table 4 Noise level versus RMSE of thorax and brain images

%N	SSIM 1	SSIM 2	SSIM 3
0.1	0.999495601	0.999051667	0.985824173
0.2	0.9908204	0.989495422	0.977120652
0.5	0.950067631	0.951626032	0.947258343
1	0.872902422	0.876892868	0.880000325
1.5	0.813415749	0.829011726	0.829745036
2	0.769519451	0.766653766	0.797299199
2.5	0.734510939	0.732665659	0.773919065
3	0.70214125	0.710984045	0.755178139
4	0.64520017	0.654620441	0.726585075
5	0.594370483	0.603919792	0.700989627
10	0.398249415	0.405265073	0.588586992
%N	RMSE 1	RMSE 2	RMSE 3
0.1	0.1756	0.5701	3.4044
0.2	0.5557	0.7138	3.4372
0.5	1.2177	1.2145	3.5289
1	2.4049	2.3020	3.8591
1.5	3.5947	3.4333	4.3386
2	4.7911	4.5811	4.9014
2.5	6.0144	5.7683	5.5142
3	7.1903	6.9181	6.1615
4	9.6413	9.3265	7.5952
5	11.9639	11.6183	9.0032
10	24.1792	23.7512	16.8592

The study performed on synthetic images shows that while hybrid model filter works well for T1 weighted Thorax MRI, the filtering of Brain MRI requires a robust filter which gives better results for a wider range of range of $\%N$.

Acknowledgements The authors acknowledges the MR image dataset providers (brainweb and NLM, USA) for providing open access to their datasets.

References

1. BrainWeb: Simulated brain database for synthetic MR images. http://www.bic.mni.mcgill.ca/brainweb/
2. National Library of Medicine, USA. The visible human data. https://www.nlm.nih.gov/research/visible/gettingdata.html

3. Sagheer SVM, George SN (2020) A review on medical image denoising algorithms. Biomed Signal Process Control 61:102036. ISSN 1746-8094. https://doi.org/10.1016/j.bspc.2020.102036
4. Chung KJ, Souza R, Frayne R (2020) Restoration of lossy JPEG-compressed brain MR images using cross-domain neural networks. IEEE Signal Process Lett 27:141–145. https://doi.org/10.1109/LSP.2019.2961072
5. Liu T, Xiong F, Qi XY, Xiao J, Villeneuve L, Abu-Taha I, Dobrev D, Huang C, Nattel S (2020) Altered calcium-handling produces reentry-promoting action potential alternans in atrial fibrillation-remodeled hearts. JCI Insight 5. https://doi.org/10.1172/jci.insight.133754
6. Yu J, Chen L, Zhou S, Wang L, Li H, Huang S (2020) Adaptive image denoising for speckle noise images based on fuzzy logic. Int J Imaging Syst Technol. https://doi.org/10.1002/ima.22442
7. Lai R, Mo Y, Liu Z, Guan J (2019) Local and nonlocal steering kernel weighted total variation model for image denoising. Symmetry 11:329
8. Chang H, Li C, Gallogly AH (2018) Brain MR image restoration using an automatic trilateral filter with GPU-based acceleration. IEEE Trans Biomed Eng 65(2):400–413. https://doi.org/10.1109/TBME.2017.2772853
9. Khmag A, Ramli AR, Al-haddad SAR et al (2018) Natural image noise level estimation based on local statistics for blind noise reduction. Vis Comput 34:575–587. https://doi.org/10.1007/s00371-017-1362-0
10. Khaparde A, Deshmukh V (2017) Optimized multi-focus image fusion using genetic algorithm. Adv Sci Technol Eng Syst J 2:51–56. https://doi.org/10.25046/aj020106
11. Saladi S, Nagarajan AP (2017) Analysis of denoising filters on MRI brain images. Int J Imaging Syst Technol 27:201–208. https://doi.org/10.1002/ima.22225
12. Liu XY, Lai C-H, Pericleous KA (2015) A fourth-order partial differential equation denoising model with an adaptive relaxation method. Int J Comput Math 92(3):608–622
13. Hossein KM, Hassanpour H (2014) Image denoising using anisotropic diffusion equations on reflection and illumination components of image. Int J Eng Trans C: Asp 27:1339–1348
14. Rajan J, Keizer K, Kaimal R (2008) An improved hybrid model for molecular image denoising. J Math Imaging Vision 31:73–79. https://doi.org/10.1007/s10851-008-0067-4
15. Ling H, Bovik AC (2002) Smoothing low-SNR molecular images via anisotropic median-diffusion. IEEE Trans Med Imaging 21(4):377–384
16. You Y-L, Xu W, Tannenbaum A, Kaveh M (1996) Behavioral analysis of anisotropic diffusion in image processing. IEEE Trans Image Process 5(11):1539–1553
17. Whitaker RT, Prizer SM (1993) A multi scale approach to nonuniform diffusion. CVGIP: Image Underst 57(1):99–110

Flight Delay Prediction Using Random Forest Classifier

R. Rahul, S. Kameshwari, and R. Pradip Kumar

Abstract Airlines focus on minimizing cost while ensuring on-time arrivals in their operations to avoid revenue loss. Especially network carriers with hub connections ensure that the incoming flights are on time for passenger, crew, and aircraft transfer by avoiding delays. Delay in time sums up billions of dollars in the aviation sector, predicting delay time helps in re-planning flight plans in a way to avoid delay. The existing deterministic models and real-time prediction system for delay time calculation lacks accuracy. The paper mainly focuses on using the airline arrival data and building a machine learning model (Random Forest Classifier) to predict delay time and probability. As random forest, in general, is robust, more flexible, and makes effective estimates this model will help in improving the overall performance of the system.

Keywords Random forest classifier · Turn around time · Supervised learning

1 Introduction

A 2010 Federal Aviation Administration study found that flight delays cost the US economy \$US32.9 (\$A45.52) billion a year. India is one of the busiest aviation hubs in the world and holds top position in air traffic and setting new records every year. The Air traffic has grown enormously and expected to have a growth which would be above 25% in the travel segment. Accurate estimations of the time of flight's turnaround time prior are beneficial for most of the stakeholders in the air traffic management, control industry and airport management system, as they could lead to reductions in potential safety risks and improvements in resources allocation. Especially network carriers with hub connections need to ensure that the incoming flights are on time for passenger, crew and aircraft transfer. The existing deterministic models and real time prediction system for turnaround time calculation lacks accuracy. The paper mainly focuses on predicting the turnaround time delay using

R. Rahul (✉) · S. Kameshwari · R. Pradip Kumar
Sri Ramakrishna Engineering College, Coimbatore, India
e-mail: rahul.1701157@srec.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_7

random forest classifier. Libraries such as Scikit-Learn and Matplotlib have been used to visualize outputs.

Random forests are one of the trees based supervised learning algorithm. This is one among the most flexible and easy to use algorithm. A forest in general consists of many trees and it is said that the more trees it has, the stronger the forest is. The algorithm works by creating decision trees from randomly selected data samples, collecting predictions from each tree and selecting the best solution through the process of voting. Random forests have a wide range of applications, such as recommendation systems, image classifications and it also helps in feature selection by selecting the most decisive features for the classifier. It can also be used to classify loyal loan applicants, stock market analysis and can even predict diseases.

2 Related Work

Many efficient algorithms have been developed using various techniques to predict flight delay Young Jim Kim et al. in [1] has analysed air traffic pattern and build a robust prediction model, recurrent neural network is used and a greater accuracy has been obtained. Further the sequence of flight data such as arrival, departure time is used to model long short-term memory RNN. Alice Sternberg et al. in [2] has reviewed and summarized existing flight delay prediction models, computation methods, mainly they have focused on significant models which uses advanced machine learning techniques.

Guan Gui et al. in [3] designed prediction tasks which uses various regression and classification task, LSTM is used for predicting flight delay but with minimal data accuracy is not obtained. Bin Yu et al. in [4] uses data of commercial cargo flight data in Beijing international airport, multifactor approach has been used along with SVM. The developed system can handle large amount of data and can recognize key factor which influences delay.

2.1 Dataset

The data set for training the model is inferred from previous flight plans. The data set to build and train the models is sourced from Airportdata.com. It is a complete source of information about airport ground handling and related service company data. Data.world is an open source website which makes it easy to get clear, accurate datasets on aviation. The data set used here is On-time data for all flights that departed NYC (i.e. JFK, LGA or EWR) in 2013. The parameters taken into consideration are year, month, day, date, actual departure time, scheduled departure time, departure delay, actual arrival time, scheduled arrival time, arrival delay, carrier, flight number, flight tail number, origin of flight, destination, amount of time spend in air, distance between airports, hour of scheduled departure, minute of scheduled departure, scheduled date.

3 System Implementation and Results

Random forest classifier analysis the data in the tree and outputs the result of various decision tree, it is widely used in handling series of data and yields good accuracy when high amount of input data is used. Firstly, the data is inputted into the model Fig. 1 shows a sample dataset used to train the model. The data consists of various parameters such as date, time, aircraft details etc., but not all the values in the data are used only data specific to train the model effective is been used.

More than 10,000 entries of data are used in training the model, from the data set the parameters taken to train the model are: Month, day of the month, day of the week, origin, destination, Arrival delay, departure delay. We have cleansed the data by observing the pattern of data and removed data such as data of extreme values, constantly occurring data, incomplete data i.e. data that has only 4 or 5 parameters not all the parameters. *Isnull().sum()* function is used to count the null values in the data, these data are removed from the data set. Further the data is organized using *math.floor* function where the time is narrowed down to hours where the minutes are removed. Dummy values are added into the data to increase the model accuracy.

The data is the split into a ratio of 2:8 i.e. 20% of test data and 80% training data. We have also tried splitting the data into 30% for testing and 70% for training but after training the model the results of accuracy are in the lower side when compared to 20 and 80% split. Figure 2 shows the output of the shape of the test and train

	YEAR	QUARTER	MONTH	DAY_OF_MONTH	DAY_OF_WEEK	UNIQUE_CARRIER	TAIL_NUM	FL_NUM	ORIGIN_AIRPORT_ID	ORIGIN	...	CRS_ARR_TIME	ARR_TIME
0	2016	1	1	1	5	DL	N836DN	1399	10397	ATL	...	2143	2102.0
1	2016	1	1	1	5	DL	N964DN	1476	11433	DTW	...	1435	1439.0
2	2016	1	1	1	5	DL	N813DN	1597	10397	ATL	...	1215	1142.0
3	2016	1	1	1	5	DL	N587NW	1768	14747	SEA	...	1335	1345.0
4	2016	1	1	1	5	DL	N836DN	1823	14747	SEA	...	607	615.0

Fig. 1 Sample input data

Fig. 2 Output shape of the test and train data

```
train_x.shape
(8984, 14)
test_x.shape
(2247, 14)
```

Figure. 3 Prediction output

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

```
0.8602581219403649
```

data. The train data consists of 8984 values and train data consists of 2247 values summing up of total 11,231 values.

The split data is fitted into the model by importing the necessary library such as sklearn were the random forest classifier is available. the classifier is stored into a variable called model, the data is fitted into the model using fit () function. After this process the random forest classifier is trained and throws the output that contains min sample leaf, min sample split, random state etc. The accuracy of the prediction model is obtained, the model yields an accuracy of 86% as shown in Fig. 3. The prediction accuracy can further be increased by inputting lot of data specific to the airport.

Confusion matrix and precision score values of the model are calculated to find the efficiency of the model. Confusion matrix is used to describe the performance of the any classification model, it comprises of mainly 4 values such False positive, false negative, true positive and true negative. The performance calculated for the random forest classifier are 1882, 54, 260, 51 respectively and the output of the values is shown in Fig. 4. The precision ratio of the model is calculated and the precision value of the model is 0.99 which is a better score when compared to the output yielded from other classification models. The accuracy of the random forest classifier can be increased by fine-tuning the hyper parameters and testing various algorithms and choosing the best fit. Commonly the flight delay is calculated upon the pre-trained models which are trained with data of normal pattern, the model should also be re-trained upon an unprecedented situation such as natural calamities, pandemic's etc.

The delay time is predicted with 4 parameters such as date, time, origin and destination of the flight. Figure 5 shows the output of the probability of delay to flight plan from JFK to ATL and another on from ATL to SEA, the delay values

Fig. 4 Output of confusion matrix

```
from sklearn.metrics import confusion_matrix
confusion_matrix(test_y, predicted)
```

```
array([[1882,   54],
       [ 260,   51]])
```

```
from sklearn.metrics import precision_score

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

```
0.9972375690607734
```

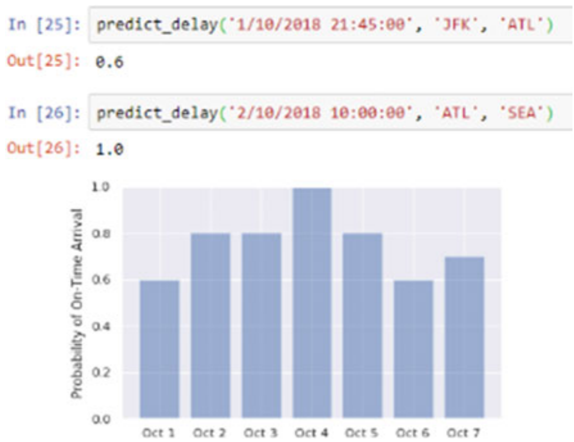


Fig. 5 Output of predicted values

are 9 min and 2 min respectively. The bar graph represents the output of on-arrival probability of a fight between JFK and LAX for the dates between October 1 and October 7. The values predicted may vary depending on the external factors which are not taken into account such as weather, passenger flow etc. Yet a substantial delay time and probability of delay can be predicted.

4 Conclusion

The output data of proposed delay time prediction system displays the output in a user-friendly GUI. The real-time input data such month, day of the month, day of the week, origin, destination should be entered, then the data will be processed using the prediction system and the predicted output i.e. the delay time and probability of delay will be displayed. Random forest classification is used and the prediction model yields an accuracy of 86% and the model can be used to predict the flight delay in American states.

5 Future Enhancement

Further the delay time prediction model can be improved by splitting the parameters and performing a multi-modal approach the parameters can be split into primary and secondary factors. The primary factor should include no. of passenger inbound and outbound, amount of fuel to be used during the flight, weather and temperature, Cargo weight, Catering loading time, Fuelling time, Cleaning time, Boarding time, De-icing time and the secondary factors to be considered are Type of the aircraft stand, Aircraft stand distance from terminal, Type of airport (linear, curvilinear or open apron), Aircraft ground inspection time (SPOM), Type of airliner (low cost or legacy), Boarding method, Passengers age, No. of workers involved in each department, Pushback time. The primary data along with the secondary data will provide accurate results, which can help the airport management team (turnaround manager) in situational awareness. These data will be helpful in identifying the conflicts beforehand and also during turnaround, which could be useful to reframe the flight's turnaround plan.

References

1. Kim YJ et al (2016) A deep learning approach to flight delay prediction. In: 2016 IEEE/AIAA 35th digital avionics systems conference (DASC). IEEE
2. Sternberg A et al (2017) A review on flight delay prediction. arXiv preprint arXiv:1703.06118
3. Gui G et al (2019) Flight delay prediction based on aviation big data and machine learning. IEEE Trans Veh Technol 69(1):140–150
4. Yu B et al (2019) Flight delay prediction for commercial air transport: a deep learning approach. Transp Res Part E: Logist Transp Rev 125:203–221
5. Cao W-D, Lin X-Y (2011) Flight turnaround time analysis and delay prediction based on Bayesian network. Comput Eng Des 5:1770–1772
6. Pal M (2005) Random forest classifier for remote sensing classification. Int J Remote Sens 26(1):217–222
7. Rodriguez-Galiano VF et al (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogramm Remote Sens 67:93–104

A Framework Using Markov-Bayes' Model for Intrusion Detection in Wireless Sensor Network

Gauri Kalnoor and S. Gowrishankar

Abstract The existing Intrusion Detection Systems which are mostly designed for detecting the particular form of Intrusion for Wireless Sensor Network (WSN) has many restrictions for different types of attacks and network structures. A novel strategy for intrusion detection based on knowledge has to be applied where the attacks are prevented from creating deviation of normal features and also from various other aggregated shapes. Multiple types of attacks should be detected over different structures of network. Every year, security is provided to the networks such that intrusions are prevented, spending around more than billions of dollars, all over the world. Among them, few disruptions that occur for vital systems are considered as the most serious type of threat mainly in areas like hospitals, military, banks and other critical applications. Firstly, clusters are discovered in the features of the network using mean shift unsupervised clustering algorithm, in the phase of training. In the next stage, the revealed clusters are generalized as an anomaly, if there is definite amount of deviation taken at the preliminary stage from normal cluster which are captured during the initial stage of training, with occurrence of no attacks. Thus, as samples are traced in different stages, the threats need to be averted, with many solutions possible, and one of the best solutions possible is to design a model of Intrusion Detection System (IDS), with the approach of Bayesian and Hidden Markov Network. In the proposed framework, the IDS are designed with different processing levels of training and testing based on connection records.

Keywords IDS · WSN · Training · Hidden Markov model (HMM) · Bayesian network · Knowledge based intrusion detection

G. Kalnoor (✉) · S. Gowrishankar
Department of Computer Science and Engineering, BMS College of Engineering, Bangalore, India
e-mail: kalnoor.gauri@ieee.org

S. Gowrishankar
e-mail: gowrishankar.cse@bmsce.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_8

1 Introduction

A Network comprised of various tiny nodes and are wireless in nature are distributed randomly with sensing data is represented as Wireless Sensor Network (WSN). These Nodes in the network are randomly scattered over the network area, to collect and transmit the sensed data in a synchronized way. The most advantageous feature of this network, as compared to that of conventional wireless networks is basically on low cost and more convenience in deployment.

Machine Learning (ML) is the learning process which learns or improves from the experience or study, and then acts without any programming explicitly [1, 2]. ML makes the process of computing more reliable, cost- effective and efficient. It analyses quickly, complex data and even more accurately which produces the models. ML is mainly categorized into unsupervised, supervised, reinforcement and semi-supervised learning. The ML has the ability to provide the most generalized solutions within an architecture which can learn and improves the performance. It plays the most pivotal role in many different fields, due to its interdisciplinary nature. Many different techniques and algorithms of ML are used to determine the solutions for issues in WSN like security, anomaly detection, fault detection, routing, and event detection and so on. In this work, the IDS model is built using HMM and Bayesian network model.

The work is carried out by investigating the model based on what are the capabilities of Bayesian Network and HMM to build IDS. The model based on Bayesian Network is constructed using the training data. The predicted Conditional Probabilities and the variable's dependencies can be found by building model with Bayesian Network. Based on the information obtained from the network, the matrices of state alteration probabilities and emission probability can be calculated and initialized, where the parameters are considered as HMM constraints for building model. Intrusions are detected using HMM-based classifiers, where some number of attempts made by the intruder on the network systems [3]. The IDS is designed based on the predictive model of HMM, having capability of discriminating the behaviors between abnormal and normal during network traffic [4].

2 Related Research

The technique of typical anomaly detection is used which identifies the behavior with some amount of change in normal behavior termed as an anomaly. The extensive survey has been done based on different techniques used to improve detection and save energy in WSN.

In [5], the author utilized the techniques of lightweight detection and tree classification to achieve the tradeoff between energy saving and high rate of detection. But, there is no discussion by the authors, about the unknown attacks which are not in the dataset referenced. The algorithm using wrapper based and feature selection is applied to develop a lightweight IDS. The IDS is used to remove features those are redundant and employs the neural network-based decision tree such that feature selection is optimized [6]. The detection paradigm is applied to increase the ability of generalization with incorporation of neural networks, but its capacity to detect unseen patterns was not complete due to the lack in updating the decision function.

An algorithm based on Watchdog Cloned Selection named as a bio-inspired approach was applied by the authors in [7]. The detection of Known attacks was successful, but unsuccessful in detecting unknown attacks. The methods used for intrusion detection were featured to minimize energy consumption and maximize the accuracy of intrusion detection which were featured to minimize energy consumption and maximize accuracy independent rate, failing to detect the unknown attacks.

In [8], the authors have proposed the method based on Synthetic Minority over-sampling technique (SMOTE), such that the dataset is balanced, using the algorithm called random forest where the classifier is trained well for intrusion detection. In the article, the authors have discussed about simulations that has been conducted based on the intrusion dataset and its benchmark. Also, it has been shown that the accuracy was reached up to 92.39%, when random forest algorithm was applied. It has been proved that; the accuracy is higher than that when other algorithms were applied. Once the minority samples were oversampled, and after combining the random forest algorithm with the SMOTE, the accuracy was increased to 92.57%. Thus, the authors have shown that an effective solution was provided based on the proposed algorithm, which in turn solved the class imbalance problem and also improves the performance of network for intrusion detection.

3 The Model

In the proposed work, the traffics of network is discretized by dividing the time into number of time slices denoted by Δt.

The model proposed is described in four steps below as shown in Fig. 1. This technique is performed at the node level of base station.

1. *Preprocessing of dataset*
2. *Training the dataset*
3. *Detection of trained data*
4. *Updating.*

4 Hidden Markov and Bayesian Network Model

A novel approach based Hidden Markov Model (HMM) is applied to various other applications such as Prediction of Gene, speech recognition, Speech synthesis, Crypt analysis and so on. But, in our proposed work, the novelty is that, HMM approach

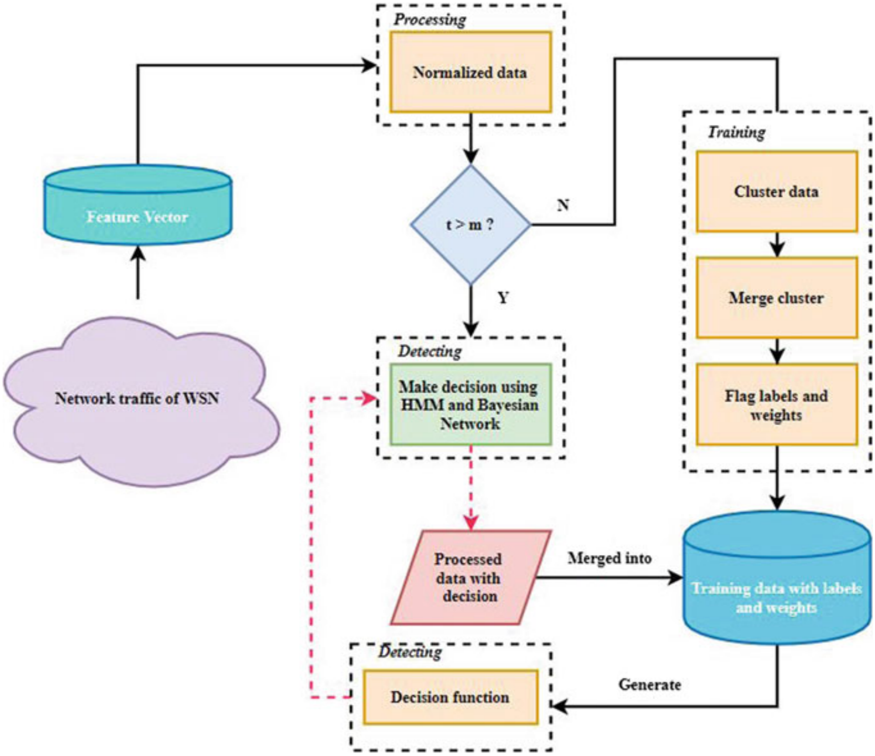


Fig. 1 Proposed system model

is applied in Intrusion detection system. Hidden Markov Models are the "generative models" where the hidden states and joint distribution of the observations are almost equivalent to the prior distribution of hidden states. This is called as transition probabilities. Also, these models are based on the observations with conditional distribution and the given present states. Thus, it is called as emission probabilities.

A. *Bayesian Network Model*

1. From the set of training data, the set of relevant variables are chosen, which the state variables are of HMM.
2. The ordering of these variables is chosen.
3. Assuming that the variables are denoted as X_1, X_2, X_m, where X_1 is first variable in the ordering, X_2 is the second and so on.
4. For i = 1 till m;
   (a) The X_i node is added to the network
   (b) Set Parent (X_i) becomes a minimal subset of {X_1, Xi_{-1}}, such that, the conditional independence of X_i and also all other members of {X_1, X_2, Xi_{-1}} is obtained, given Parent (X_i).

Table 1 The samples of data set after the discretization of attributes

Protocol type	Flag	Scr_Bytes	Dst_Bytes	Count
udp	SF	< = 162.155	<= 334	<= 377.616
icmp	SF	<= 695.122	<= 0	> 377.616
tcp	SF	> 695.122	> 334	<= 65.487
tcp	SF	<= 695.122	> 334	<= 65.487
udp	SF	<= 162.155	<= 0	<= 65.487
tcp	SF	<= 695.122	> 334	<= 65.487
tcp	SF	<= 695.122	> 334	<= 65.487
tcp	REJ	<= 162.155	<= 0	<= 184.033
tcp	S0	<= 162.155	<= 0	<= 184.033
tcp	SF	<= 695.122	> 334	<= 65.487
icmp	SF	<= 695.122	<= 0	> 377.616
tcp	SF	<= 695.122	> 334	<= 65.487
icmp	SF	> 695.122	<= 0	> 377.616
icmp	SF	> 695.122	<= 0	> 377.616
tcp	SF	<= 695.122	> 334	<= 65.487

(c) Define P (X_i = k | Assignments of Parents (X_i)), the Probability table.

IDS Framework Using HMM and Bayesian Network

A framework to build an IDS design based on HMM and Bayes network is presented. The IDS framework comprises of distinctive processing levels, in WSN: Read dataset, data pre-process, Bayes net, HMM parameters initialization, states and sequence generation, state transition estimation, matrix of emission probability and the model evaluation (Table 1).

Baum-Welch (Forward–backward) Algorithm

1. Apply BW re-estimation formula to find the most maximum likelihood parameters of HMM.
   For every HMM parameter matrix/ vector requiring the re-estimation, allocate storage for summations of numerator and denominator of the form:

$$\mu_j = \frac{\sum_{j=1}^{t} L_j(t) O_t}{\sum_{j=1}^{t} L_j(t)} \tag{1}$$

and

$$\sum j = \frac{\sum_{j=1}^{t} L_j(t)(o_t - \mu_j)(o_t - \mu_j)^{|}}{\sum_{j=1}^{t} L_j(t)} \tag{2}$$

The storage locations are known as accumulators.

2. For a model M with N states, let the forward probability, $\alpha_i(t) = P\big(o_1, o_2, ..o_j, x_t = j\big|M)$ called as joint probability for observing the first t vectors and remaining in state j at time t.
3. For states 1 < j < N,
   Initialize $\propto_1(1) = 1$ and $\propto_j(1) = \propto_{1j} b_j(o_1)$.
4. Calculate the forward probability which gives efficient value, by applying the following recursion:

$$\propto_1(t) = \left[\sum_{i=2}^{N-1} \propto_1(t-1)\propto_{ij}\right] b_1(o_1)$$

5. For each state j and time t, use the probability and the current observation vector, $L_j(t)$ and o_t respectively. Update the accumulator for that current state.
6. Calculate the new parameter values, using the final accumulator values.
7. For this iteration.

The HMM Parameters 'A' and 'B' which are estimated using the algorithm if the value of $P = P(O|M)$ is not greater than the previous iteration value, then stop.

Else, repeat the steps 1–6 using the new re-estimated values of parameter.

5 Simulation Experiments and Result Analysis

To assess the performance of the proposed methodology, the experiments have been conducted. This experiment scenario was simulated by using NS2 on a PC, with Inter (R) Core, 3.54 GHz and 8 GB RAM. Two types of network structure were considered: Flat network structure and Hierarchical network structure. The sensor nodes were deployed and scattered in the region with the dimension of 1000 (m) × 1000 (m) based on the type of network structures. The nodes are also deployed at the base station in the center of the deployed region. Among all the nodes, ten percent of the nodes were detected as malevolent nodes which performed different types of attacks. Comparing the flat network, the hierarchical network type of WSN, is more appropriate for networks with large scale nodes which reduces the energy consumption at a node and also bandwidth of the communication medium. Thus, the total number of nodes considered was 100 and the total nodes with attack type were 10. In the hierarchical protocol type, the relationship among the nodes is shown in Figs. 2, 3 and 4.

The simulation results of proposed work were analysed and compared with the methods of other mainstream intrusion detection, such as K-means, Decision Tress (DT), PCA-based Centralized intrusion detection (PCACID), Mean shift and also Logistic Regression (LR). The comparison is made to evaluate the performance of network based on the detection efficiency and it's adaptively to network structures. The simulated results showed that, among all the experiment type scenarios, the

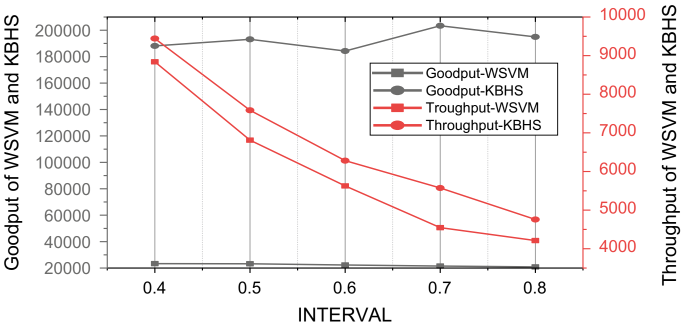


Fig. 2 Interval versus Good put and throughput data using KBHS and WSVM algorithm

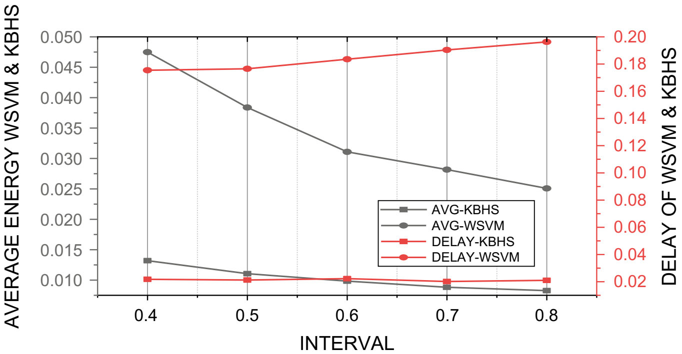


Fig. 3 Interval versus average energy and delay using KBHS and WSVM algorithm

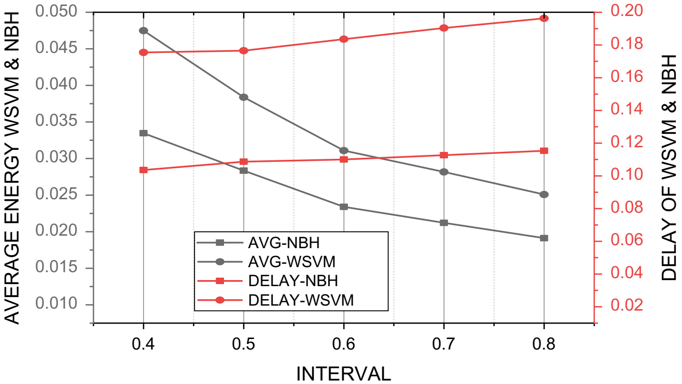


Fig. 4 Interval versus average energy and delay using NBH and WSVM algorithm

false alarm rate and average detection rate (DR) of proposed work, obtained were 1.76% and 97.54% respectively. However, the small deviation obtained were 1.054 and 0.920% for both FAR and DR.

Comparison of DR with different scenarios between proposed and other mainstream methods are analysed and then tabulated.

6 Conclusion

A model is designed for intrusion detection by applying the novel algorithm KBHS considering the network structure for WSN. The recommended method employs the knowledge-based strategy for detection, where the attacks can get slightly diverged from the normal features and thus obtaining different feature space. In the framework, the decision function is updated periodically where hidden Markov model and Bayesian network is used. The training and testing of data set is performed using the HMM as applicator. The HMM is trained for normal connection records using TCP. Thus, during training the model, the initialization of appropriate values was necessary, as the functioning of the decision function depends on these values. The parameter estimation and data training are done by applying powerful approach of HMM and Bayesian model in the decision function and create the IDS to find whether the traffic of the network is normal or intrusions have been detected. The model is run at the base station to detect the attack. Thus, the proposed model has better performance w.r.t detection rate, accuracy, FAR and other parameters when compared with the mainstream IDS techniques deployed in WSN.

References

1. http://www.intrusiondetectionsystem.org/
2. http://jabrauum.blogspot.com/
3. Alqahtani M (2019) A genetic-based extreme gradient boosting model for detecting intrusions in wireless sensor networks. Sensors 19:4383–4383
4. Ball MG, Qela B, Wesolkowski S (2015) A review of the use of computational intelligence in the design of military surveillance networks. Recent Adv Comput Intell Defense Secur 621:663–693
5. Can O, Sahingoz O (2015) A survey of intrusion detection systems in wireless sensor networks. In: 6th international conference on modeling, simulation, and applied optimization, pp 1–6
6. He D et al (2015) Robust anonymous authentication protocol for health-care applications using wireless medical sensor networks. Multimed Syst 21:49–60
7. Huang JY et al (2011) Shielding wireless sensor network using Markovian intrusion detection system with attack pattern mining. Inf Sci 231:32–44
8. Li M, Lin HJ (2015) Design and implementation of smart home control systems based on wireless sensor networks and power line communications. IEEE Trans Ind Electron 62:4430–4442

Effective Text Comment Classification Using Novel ML Algorithm—Modified Lazy Random Forest

Tejashri Ghodke and V. M. Khadse

Abstract Machine learning (ML) algorithms are methods used to classify data. The various patterns or classes can be classified with the help of these various ML algorithms. There are numerous areas where these algorithms can be used. One such area is to detect whether the comment, sms or text message is SPAM or Normal message. So the aim of this work is to identify the best machine learning algorithms to detect SPAM text message on two different dataset. The first dataset is collected from YouTube comment dataset and second is the SMS dataset. The Random Forest (RF) is the ensemble learning method for classification, regression and other tasks that operates by developing a multitude of decision trees at learning phase and outputting the class. Its one variant which performs well as compare to normal RF is Lazy RF, as the study shown and is the base for this research work. In this work, we have proposed one more novel variant of LRF and the different machine learning algorithms are compared in terms of accuracy with proposed Modified Lazy Random Forest. The results are compare with two techniques, first is the simple hold out, and second is the K-fold cross validation. For both cases the proposed algorithm performs well to detect the SPAM messages for the both datasets.

Keywords IDS · Machine learning · Deep learning · Performance matrices

1 Introduction

The machine learning algorithms are used in various areas such as classification problems, regression problems, Virtual Personal Assistants, Traffic Predictions, Videos Surveillance and much more. One such area is SPAM message detection. The hard task was to find the best ML algorithm for detection of SPAM in text dataset then achieving the improvement in respective algorithm. The RF is the best choice in terms of promising results which it produces. The thorough literature review showed that LRF is well performing algorithm and we proposed the novel MLRF algorithm. The comparative study showed that the proposed algorithm performs well as compared to

T. Ghodke (✉) · V. M. Khadse
College of Engineering Pune, Pune, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_9

current state-of-art algorithms. To cross verification we have found the results with hold out and K-Fold cross validation techniques. The L-RF and NB algorithms are analyzed on the YouTube API dataset. For the detailed insights the training and testing dataset split is considered from 10% Train–90% Test, 20% Train–80% Test, 30% Train–70% Test, 40% Train–60% Test, 50% Train–50% Test, 60% Train–40% Test and 70% Train–30% Test. The completed work is performed using python programming language with TensorFlow, NumPy etc. libraries. The dataset pre-processing is done with the python with the built-in hash function on string data. The processed dataset is then applied as an input to Modified Lazy-Random Forest and then to Naïve Bayes algorithm. The Modified Lazy-Random Forest is performing well on the given dataset as compared to other compared algorithm.

2 Literature Review

The work done by Aggarwal et al. in [1] has shown the classifier algorithmic rule for YouTube videos. The classifier is developed and tested with the sample dataset created by them. The YouTube contains tremendous variety of videos of assorted sorts as well as copyright profaned videos, business spam, hate and political theory promoting videos, vulgar and pornographic material and privacy incursive content. The obtained results established that accuracy of projected approach is quite 80%. They first type classifier approach to sight the privacy incursive harassment and offense videos having unwanted content on YouTube. The analysis results found that the accuracy of those classifiers i.e. VVD, VAVDS, VAVDP and RVDC is 83, 84, 90, and 97% severally. It indicates that known discriminatory options will be wont to exploit the harassment detection on YouTube unto an affordable accuracy.

The unstructured dataset creates problems in normal functioning of the popular machine learning algorithms in terms of classification, recognition and detection problems. The authors Phakhawat et al. in [2] performed experiment on YouTube dataset and balanced the usage of the SMOTE technique along with examination of the usage of ubiquitous algorithms such as multinomial Naïve Bayes (MNB), Decision tree (DT) and Support vector machines (SVM). After the experimentation SVM indicates promised results with an accuracy 93.30% on filtering task along with this 89.44% on classification problem. Their SMOTE approach could overcome the imbalanced data problem and provides an promised outcomes. Moreover, analyzing the results of SVM, the use of SMOTE provides an accuracy with 93.30% in considering to 76.41% with no re-sampling approach. The final result improved with 16.9% on Emotion Filtering dataset.

Video coding tutorials allow expert and novice programmers to visually examine actual builders write, debug, and execute code. Previous lookup in this area has focused on supporting programmers discover applicable content in coding tutorial movies as well as understanding the motivation and desires of content material creators [3]. A dataset of 6000 feedback sampled from 12 YouTube coding movies is used to behavior our analysis. The consequences also show that an extractive frequency-based summarization method with redundancy control, can sufficiently capture the important issues existing in viewers' comments.

The researchers in [4] strived on comparative findings of the usual filtering procedures used for YouTube comment spams is carried. The study extended datasets obtained from YouTube, the use of its Data API. According to the retrieved results, excessive filtering accuracy (more than 98%) can be achieved with low-complexity algorithms, mentioning the possibility of acquiring a suitable browser extension to alleviate comment spam on YouTube. The expansion in social media reputation is crucial and should be a positive impact with greater apparent in the final decade. This extends in recognition as well as a need to motivated malicious, scammers, and spammers to goal these platforms.

The research work done in paper [5], by S. Thiago et al. proposed a lazy model of the Random Forest (RF) classifier (named as LazyNN RF), primarily created for a noisy classification operations. The LazyNN RF "localized" coaching projection is composed via examples that higher resemble the examples to be classified, received via nearest neighborhood learning set projection. This gives marked evidence in favor of the exploring records regional in RF models. As future work, they deliberate to investigate distance metric learning. The theoretical analysis regarding the bias/variance trade-off of their proposal is also done.

The proposed model by Shreyas et al. [6] for the detection of spam comments on the video-sharing website—YouTube is done with message classification as in two types, the promotional purpose and irrelevant. The authors strived to classify these comments by utilizing ML algorithms such as RF, SVM, Naive Bayes with specific custom heuristics such as N-Grams are proven to be very efficient in detecting spam comments. The authors have presented a method for automated detection of spam comments on the YouTube platform.

The research worked completed by done by Tripathy et al. [7] used the IMDb movie review dataset. The primary goal of this paper classification of reviews on social platform websites into meaningful classes. For this purpose, they used many supervised ML algorithms such as NB, Maximum Entropy (ME), Stochastic Gradient Descent (SGD), and SVM. Different machine learning algorithms are proposed for the classification of movie reviews of IMDb dataset (IMDb 2011) using n-gram techniques viz., Unigram, Bigram, Trigram, a combination of unigram and bigram, bigram and trigram, and unigram and bigram and trigram. The proposed model gave 86.23% accuracy using Unigram + Bigram + Trigram (U + B + T) method for NB algorithm. The ME, SVM, and SGD gave accuracy 83.36%, 88.94%, 83.36% respectively using U + B + T method. The future research includes smiley comments which they not worked on.

The P. Sethi, V. Bhandary et al. have worked on emails and messages with legal, economic and technical issues [8]. A pin point is being pone with the help pf Bayesian filters for preventing the SPAM messages and emails issue. They also analyzed and studied the relative strengths of various ML algorithms in for the detection of spam messages which are communicated on mobile devices. They collected the data from on open public dataset and developed two datasets for their testing and validation cases. Accuracy in detecting spam messages is considered as the base point for the comparing the results. Our results clearly demonstrate that various ML algorithms with different features leads to work differently in detecting spam messages.

The work done in [9] by P. Kolari et al. mainly focuses on Emails and Communication. The huge usage of emails in digital world brought many problems such as SPAM mail or Normal Mail. Their research focused on email content analysis in Turkish language. For classification purpose they used RF classifier algorithm and Vector Space Model from ML methods. Both techniques are subjected to different performance matrices and their performances are compared.

In the proposed system [10] by C. Chen et al., a huge number of URL are scanned and analyzed with the help of many APIs for the identification of whether these URLs are malicious or not on time. They firstly carried the "Spam Drift" problem in statistical features based on Twitter spam detection. For solving this problem, they proposed a Lfun approach. Using this scheme, classifiers have been re-trained by the modified "changed spam" tweets that are learned from unlabelled patterns, which reduced the impact of "Spam Drift" significantly. They evaluated the performance of Lfun approach using performance matrices such as Detection Rate and F-measure. Their research concluded that both detection rate and F-measure are gained much when applying with their Lfun approach. They also compared Lfun to four traditional machine learning algorithms, and find that our Lfun outperforms all four algorithms in terms of overall accuracy, F-measure, and Detection Rate.

The email classification in two classes SPAM or Non-SPAM is carried by authors Olatunji [11], with the help of Extreme Learning machines (ELM) and SVM for classification. In this work, they attempted investigation on how SVM and ELM compared to the unique and important problem of Email spam detection, which is a classification problem. Empirical results from experiments conducted using a ubiquitous dataset resulted that both techniques outperformed than the best earlier published techniques on the same popular dataset employed in this study. However, SVM performed better than ELM on a comparison scale based on accuracy. But in terms of speed of operation, ELM outperformed SVM significantly.

The authors Vishaghini et al. [12] proposed the use of weighted SVM for spam classification using weight variables captured by the proposed KFCM algorithm. For the experiment UCI Repository, SMS Spam dataset is used with 4601 records having 57 attributes. Their research achieved a result that shows that WSVM with KFCM exhibits lower mis-classification rate than SVM. In the end, they also analyzed the results on performance matrices such as accuracy and Precision. Their improved WSVM (with KFCM) achieved 96.5% accuracy.

3 System Design

3.1 Data Pre-Processing

Dataset: The Datasets are YouTube SPAM dataset and SMS SPAM dataset. In which various messages and comments are present which are latest in nature. For the classification purpose we have pre-processed this dataset.

Numericalization: The dataset consists of string values for some of the fields such as multi language pattern, non numeric values, therefore these attribute values are converted into numeric values to play better on them. Python language plays vital role here, with hash value function. After converting the all values into numeric values the dataset is given to training model.

3.2 Training with the Model

Modified Lazy Random Forest: It is obvious that the MLRF model consists of two parts—KNN processed dataset feature values for records and the second Random Forest part. Basically RNN gives processed data to Random Forest and then Random Forest gives detection capability. Due to the introduction of KNN the Random Forest performs lately and takes more time to generate results which make it named as Lazy Random Forest and we have modified the Lazy RF and proposed the MLRF.

Training set: Providing training dataset. With 50% training and 50% test dataset. Further experimentation is done with 55% Train–45% Test subsequent up to 95% Train–5% Test. This technique of train and test is called “Hold-Out” method. K-Fold Cross validation technique is also applied for greater learning experience of the machine learning algorithm. The reason behind using this split for training and testing dataset is that it shows best accuracy and best performance in ML algorithms.

3.3 SPAM Classification

The last step of the architecture is the detection which classifies the record SPAM or Normal message. The proposed algorithm classifies both binary and multi-class classification. It is also possible to use this proposed algorithm anywhere where classification problem exist.

4 Proposed Algorithm

The proposed algorithm is shown in Fig. 1. The algorithm is self-explanatory and starts with the input dataset. The input dataset is then processed with KNN to identify the nearest neighbor and its distance, those values are then appended to base dataset record. The new processed dataset is then applied to Random Forest algorithm record

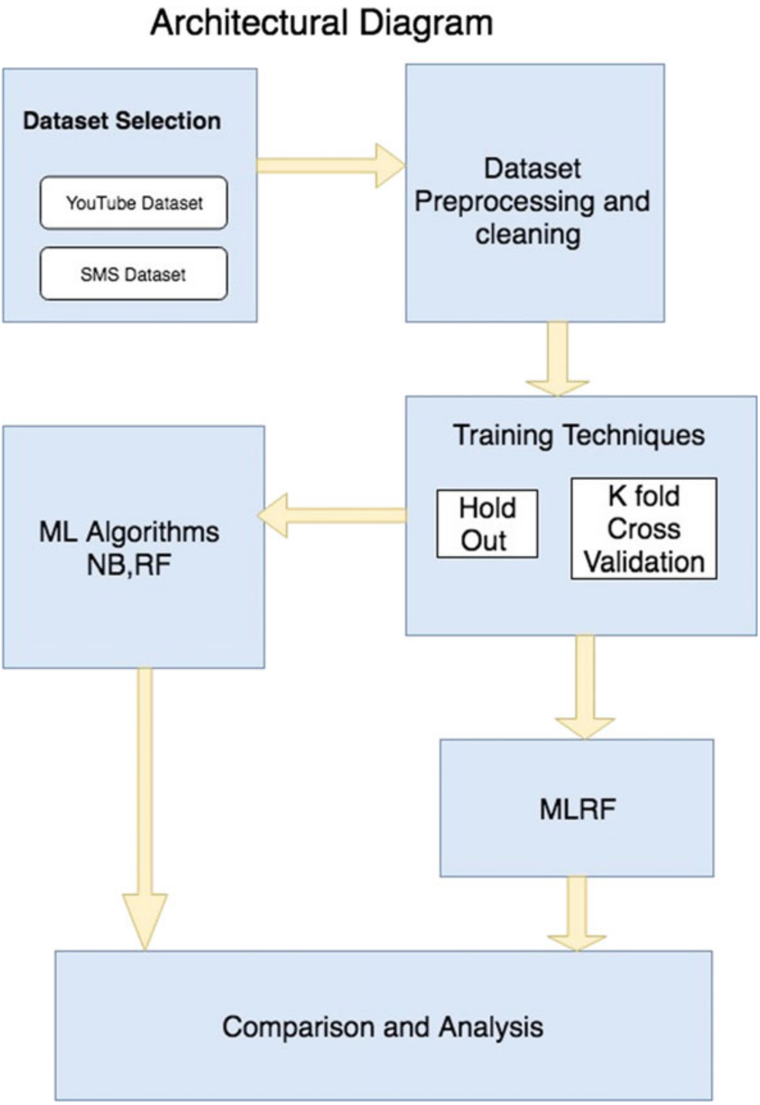


Fig. 1 Architectural diagram

Table 1 Hold-out method results on YouTube dataset

Testing % split	MLRF	RF	Nave Bayes
50	61.89175145	61.87175043	54.5445751221423
55	62.5352862	62.5049232	55.26593064
60	63.19597205	63.16359697	53.82534079
65	63.20912378	63.2	54.72081517
70	63.61755301	63.60259981	54.39428297
75	63.98035817	63.95147314	54.82993197
80	64.58103649	64.57204767	54.35992579
85	64.40829847	64.38949783	54.51109737
90	64.4198363	64.13095811	54.65883323

by record i.e. line by line. This act caused the greater accuracy than simple Random Forest algorithm.

5 Experiment and Results

This section describes the experiment that we have done and results which we have obtained. The complete experiment is done in python programming language and various libraries are used such as scikit-learn, numpy etc. The experiment is done using hold out technique and the results are shown in Table 1.

6 Results with HOLD-OUT Method: Youtube and SMS Dataset

6.1 *Hold-Out Method Results on Youtube Dataset*

The first accuracy in Table 1 is for 50% Train–50% Test, second 55% Train–45% Test subsequently last up to 95% Test–5% Test. The results are varying in nature because the program is taking the random records for its learning purpose. Same pattern of training-testing is spread across the following results. The same algorithms are applied on SMS dataset for the experimentation and the results are shown in Fig. 2.

Table 2 for SMS dataset with Hold out method. The results are compared and it is shown that the proposed Modified Lazy Random Forest is performing well over the compared algorithms.

Fig. 2 Proposed algorithm

```
1.function GetNN_NN_Class(Dataset,x',k)
2. for each x belongs to Dataset do
3.     add into new Dset=(nearestNDistance,PredClass)
4. end For
5. end function
6. function ClassifyRF()
7. Select RF(Dset)
8. Split Train-TEst
9. return RF_Results
10.End Function
```

Table 2 Hold-out method results on SMS dataset

Testing % split	MLRF	RF	Nave Bayes
50	81.69657423	79.3800978793	55.90111643
55	81.42387077	78.88124439	56.52759085
60	80.70127002	79.4036444	58.17529472
65	81.43589744	79.20512821	57.83492823
70	81.00023929	78.84661402	57.28643216
75	81.48979134	78.23648194	38.02690583
80	81.48979134	78.39949324	56.69856459
85	82.34797297	75.88751496	58.60215054
90	81.75109693	78.1976195	51.61290323

6.2 Hold-Out Method Results on SMS Dataset

The results in Table 2 clearly shows the MLRF gives promising results over other RF and Naive Bayes ML algorithm. The YouTube dataset contains 5 features and SMS dataset contains 2 features.

7 Results with K-Fold CV Method: YouTube and SMS Dataset

7.1 K-Fold CV Method Results on YouTube Dataset

The first accuracy in Table 3 is for YouTube dataset with first fold, second with accuracy of MLRF as 80.71% for second fold, third 82.22% for third fold and subsequently we have performed this task for five folds with K = 5 and the results are more interesting than hold out-method. The K-Fold cross validation has great learning

Table 3 5-Fold CV method results on YouTube dataset

Logistic regression	SVM	MLRF
0.505829596	0.802421525	0.807174888
0.507181329	0.836317774	0.822262118
0.517055655	0.800089767	0.833931777
0.488330341	0.830341113	0.816876122
0.509874327	0.802010772	0.829443447
Average	Average	Average
50.56542496	81.083619	82.17571713

capability for all the data records. The unseen data records are not covered in hold-out method but it is exactly apposite to K-Fold cross validation method. It learns all the data pattern records in the dataset for promising results.

The first accuracy in Table 4 is for SMS dataset with first fold, second with accuracy of MLRF as 80.78% for second fold, third 83.84% for third fold and subsequently we have performed this task for five folds with K = 10 and the results are more interesting than hold out-method. The K-Fold cross validation has great learning capability for all the data records. The unseen data records are not covered in hold-out method but it is exactly apposite to K-Fold cross validation method. It learns all the data pattern records in the dataset for promising results. The same results are compared with the results of other ML algorithms i.e. logistic regression and SVM.

Table 4 10-Fold CV method results on SMS dataset

Logistic regression	SVM	MLRF
0.498207885	0.829749104	0.806451613
0.513464991	0.818671454	0.807899461
0.499102334	0.849192101	0.838420108
0.515260323	0.816624776	0.825852783
0.52064632	0.807648115	0.807899461
0.513464991	0.85432675	0.867145422
0.491921005	0.836624776	0.824057451
0.484739677	0.825852783	0.795332136
0.515260323	0.838420108	0.825852783
0.50448833	0.859964093	0.836624776
Average	Average	Average
50.5655618	82.07074059	82.5528143

Fig. 3 LRF algorithm

```
Function GET-K-NEAREST-NEIGHBOURS(D_train,x',k)
{
        FOR-EACH (x from D_train)
{
maxPriorityQueue.insert (x,COMPUTEDISTANCE(x,x'))
}
RETURN (maxPriorityQueue.topK(k))

Function CLASSIFY (D_train,x',k)
{
        Dknn <= GETKNEARTNEIGHBOURS (D_train,x',k)
        Rfmodel <= RANDOMFOREST.train (Dknn)
        RETURN rfmodel.classify(x' )
}

}
```

7.2 *Lazy Random Forest*

More formally, assume Dtrain = (xi, yi) is the learning set of records. Actual learning task is started when this learning records are received by classified. The nearest x are identified for the purpose of projecting the training set to a subset of training records for finding the class of x. Assume that x is given, the set Dknn(x) contains k nearest neighbors of x which are computed and used for learning task of RF classifier. Many distance matrix techniques can be used here: (i) The inverse of cosine similarities, (ii) adaptive distance metrics that define non-isotropic neighborhood based on the observed characteristics of the input space, (iii) distance matrix learned from training records. They here considered the simple neighborhood definition based on cosine similarities. These procedures are executed for each test record for classification. The strategy is outlined in the algorithm in Fig. 3.

7.3 *Native Bayes*

Naive Bayes (NB) is being studied thoroughly since the 1960s. It was introduced (even not given that name) into the text retrieval community in the 1960s, and persists a popular technique for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines [2]. It also finds application in automatic medical diagnosis. Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in terms of the number of variables (features) in a learning task.

8 Conclusion

Initially the results for comment message detection of SPAM or normal is done with help of current state-of-art algorithms such as Random Forest, Naive Bayes, Lazy Random forest etc. and those results are compared with the proposed algorithm. The results are more promising than those current state-of-art algorithms, the results have proved that with accuracy, false positive rate, precision, recall and with the confusion matrix. To refine the experimentation and results in thoroughly, the two techniques have been implemented called hold out and K-Fold Cross validation. These two techniques have shown the promising results. The numericalization step has been introduced for values for the attributes. The proposed MLRF has best performing capacity on text comment classification. Moreover, the proposed algorithm can be applied to any of the classification problems such as in security field for attack detection, smart predictions and much more. The future work includes applying the various different dataset of different domain and then analyze how the proposed MLRF performs.

References

1. Aggarwal N, Agrawal S, Surekaa A (2014) Mining YouTube metadata for detecting privacy invading harassment and misdemeanor videos. In: Twelfth annual international conference on privacy, security and trust (PST), 2014. IEEE, pp 84–93
2. Sarakit P, Theeramunkong T, Haruechaiyasak C (2015) Improving emotion classification in imbalanced YouTube dataset using SMOTE algorithm. In: 2015 international conference on advanced informatics: concepts, theory and applications (ICAICTA). IEEE, pp 1–5
3. Poch E, Jha N, Williams G, Staten J, Vesper M, Mahmoud A (2017) Analyzing user comments on YouTube coding tutorial videos. In: Proceedings of the 25th international conference on program comprehension. IEEE Press, pp 196–206
4. Abdullah O, Ali MA, Karabatak M, Sengur A (2018) A comparative analysis of common YouTube comment spam filtering techniques. In: 2018 6th international symposium on digital forensic and security (ISDFS). IEEE, pp 1–5
5. Salles T, Gonalves M, Rodrigues V, Rocha L (2018) Improving random forests by neighborhood projection for effective text classification. Inf Syst
6. Aiyar S, Shetty NP (2018) N-gram assisted Youtube spam comment detection. Procedia Comput Sci 132:174–182
7. Tripathy A, Agrawal A, Rath SK (2016) Classification of sentiment reviews using N-gram machine learning approach. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2016.03.028
8. Sethi P, Bhandari V, Kohli B (2017) SMS spam detection and comparison of various machine learning algorithms. In International conference on computing and communication technologies for smart nation (IC3TSN). IEEE, pp 28–31
9. Kolari P, Java A, Finin T, Oates T, Joshi A (1999) Detecting spam blogs: a machine learning approach. In: Proceedings of the national conference on artificial intelligence, vol 21, no 2, 2006. AAAI Press; MIT Press, Menlo Park, CA; Cambridge, MA; London, p 1351
10. Kamble S, Sangve SM (2018) Real time detection of drifted Twitter spam based on statistical features. In International conference on information, communication, engineering and technology (ICICET). IEEE, pp 1–3

11. Olatunji SO (2017) Extreme learning machines and support vector machines models for email spam detection. In IEEE 30th Canadian conference on electrical and computer engineering (CCECE). IEEE, pp 1–6
12. Vishagini V, Rajan AK (2018) An improved spam detection method with weighted support vector machine. In International conference on data science and engineering (ICDSE). IEEE, pp 1–5

Unraveling Deep Learning Performance in Cross-Sensor Iris Recognition

Meenakshi Choudhary, Vivek Tiwari, and U. Venkanna

Abstract Iris recognition has shown dominant characteristics in the field of biometrics, however, iris acquisition from heterogeneous sensors is still a key barrier in its wide-scale deployment. In the past few years, convolutional neural networks achieved substantially improved performance in the domain of computer vision and outset to employ for iris applications. This paper analyzes the performances of various preeminent deep learning models while realizing cross-sensor iris recognition. Their quantitative performances are analyzed based on the experiments performed over two public iris datasets; ND-CrossSensor-Iris-2013 and ND-iris-0405. The experimental results depict that these models have enough potential to transfer their knowledge to the domain of iris recognition and to recognize the microstructures within the iris region to unveil improved feature perception. Furthermore, a comparative analysis is carried out among these models based on five manually designed experiments regarding cross-sensor iris recognition, where the best performance is reported by the DenseNet model.

Keywords Cross-sensor iris recognition · CNN customization · Knowledge transfer · Pre-trained models

1 Introduction

Iris Recognition (IR) is one of the most prominent techniques for inimitably identifying an individual based on the iris microstructures. Iris owns more stable and distinctive features compared to various other biometric traits. Over the last decade, researchers obtained excellent improvement in iris recognition and the state of the arts reported up to 99% accuracy [1]. However, IR encounters some challenges with image acquisition through heterogeneous cameras/sensors [2], which are manufactured by distinct companies with different technologies. Therefore, such sensors differ in the illumination, the underlying hardware, and wavelength range. Including dissimilar sensors during iris enrollment and authentication yields reduction in IR accuracy that

M. Choudhary (✉) · V. Tiwari · U. Venkanna
Department of Computer Science Engineering, DSPM IIIT Naya Raipur, Naya Raipur, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_10

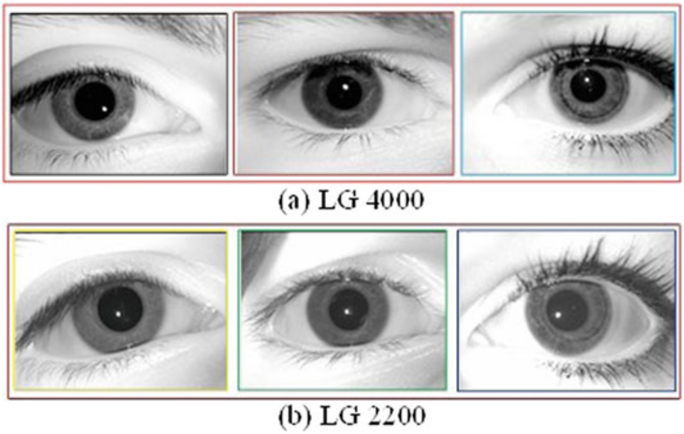


Fig. 1 Iris images captured by two distinct sensors, i.e. LG4000 and LG2200

is characterized as "Sensor mismatch" problem, primarily reported by Bowyer et al. [3]. The authors inspected that the iris recognition performance suffers if dissimilar sensors are used to capture the iris. Figure 1 illustrates iris images collected from two different sensors, i.e. LG4000 and LG2200, where the discrepancy in iris images can be clearly observed.

The first real-time IR system was prepared by Daugman [4] in 1993. Besides, some iris segmentation methods based on circular Hough transform [5], gabour filter [6] etc. for unconstraint iris segmentation were proposed. Additionally, for iris feature extraction, tunable filter bank [7], cross-spectral patterns [8], etc. are employed. Arora et al. [9] introduced an approach to enrich iris sensor interoperability through first perceiving the sort of sensor used, followed by selective image enhancement. However, such techniques employee handcrafted features, where parameters are empirically defined. More specifically, they do not support implicit parameter updations via learning from input iris images.

In past few years, Convolutional Neural Networks (CNNs) have reported innovative change in computer vision, through automatic weight updations. They have been widely employed for face [10] as well as iris applications [11–13]. Liu et al. [14] employed pair-wise convolutional filters to compare two iris images collected through distinct sensors, in parallel. However, the authors emphasized on iris verification rather than iris recognition. Besides, another deep-learning based multi-biometric system, i.e. IrisConvNet [15] was designed for person identification using the fusion of left and right iris features. However, the model was not examined for cross-sensor iris recognition. Further, DeepIrisNet [16] was designed with two distinct convolution and inception based deep CNN architectures for cross-sensor iris recognition. Though, the model's designing didn't emphasize the issues such as vanishing gradient and weight saturation that occurs while increasing the model depth. Furthermore, Nguyen et al. [17] explored some pre-trained CNNs in terms of the recognition accuracy at different layers. However, the work didn't emphasize cross-sensor iris recognition. Next, Wang et al. [18] presented 'MiCoReNet' for eye recognition, designed with the interleaved convolutional and residual layers. However, the shallow architecture makes the model insignificant to perform classification among huge number of iris subjects. In recent, a fusion-model [2] was introduced that performs cross-sensor iris recognition through fusion of CNN features with SIFT and BSIF counterparts. The deep learning framework was designed with alternative convolutional and residual blocks to cope with vanishing gradient and weight saturation.

One major obstacle in using deep learning techniques for iris recognition is the lack of a huge training data. Fortunately, the concept of transfer learning has been introduced in order to deal with the limited data size. Here, CNN models pretrained on other datasets (e.g. ImageNet [19]), can be unswervingly employed in the IR domain. Indeed, the pre-trained CNN models have been successfully utilized to accomplish several computer vision tasks. For instance, authors in [20] presented that in spite of the pre-training on the ImageNet dataset, the VGG model works substantially well for iris recognition. Nevertheless, after releasing VGG model, several other innovative models have been introduced in the literature. This paper explores such CNN frameworks for cross-sensor iris recognition, predominantly those succeeded in the ImageNet challenge.

We assume that instead of architecting and training a novel CNN to realize cross-sensor iris recognition, employing preeminent CNN architectures that have been successfully established in large-scale computer vision contests, may exhibit noble performance. A huge variety of state-of-the-art CNN models could be found in ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [19] conducted yearly to assess the forthcoming deep learning frameworks for comprehensive image classification and recognition. Such models are then released in the public domain to accomplish feature extraction from variety of images. Earlier research presented that these standard CNNs are effective for several computer vision tasks such as object localization and recognition, semantics recognition, action recognition etc., along with the image classification. This paper investigates the performance of various pre-trained CNN models that succeed the ILSVRC contest with highest recognition (or classification) accuracy.

The primary contributions of the paper are three-fold:

- First, it explores some standard deep learning based architectures that were succeeded in the ILSVRC challenge.
- Second, all standard CNNs are examined to realize cross sensor iris recognition through transfer learning and the corresponding results are analyzed.
- Third, these models are compared to find out the finest one for cross-sensor recognition.

The remaining portion of the paper is structured as: Sect. 2 explores the standard CNNs used in this work. Section 3 describes the dataset partitioning and experimental setup to assess various CNNs. Section 4 represents the experimental outcomes and performance comparison in the form of ROC curves. Section 5 summarizes the paper.

2 Proposed Work

2.1 Description of Standard CNN Models

The ILSVRC event has been proven as an imperative spot for the performance assessment of each novel CNN architecture, particularly after the contribution of technology hulks such as Microsoft, Google and Facebook. The size of the appealing CNNs has increasingly developed from 8 to 152 layers in years 2012–2015, meanwhile the classification error rate has substantially fallen from 16.4 to 3.57%. In this section, each architecture is analyzed in detail and their remarkable assets are emphasized.

AlexNet achieved an inventive progress in the ILSVRC challenge 2012 by introducing a CNN model that substantially surpassed the hand-crafted counterparts and contributed in top-5 lowest error rates (16.4%) [21]. Indeed, AlexNet is a twisted form of the LeNet [22] that exploits a massive train dataset (ImageNet) and GPU's computational power that facilitates $10\times$ speed-up in model training. Moreover, the hyper-parameter tuning of AlexNet aids to the increased performance, and thus won the ILSVCR 2013 contest. In our work, the output feature vectors from all convolutional (5) and fully connected layers (2) are examined and fed to the classifier for iris recognition.

VGG was the runner-up in ILSVRC 2014, the authors in [23] suggested that employing smaller filters (3×3) within the convolutional layer yields enriched performance. Indeed, multiple overlapped small filters may imitate the effect of bigger one and yield the improved generalization. In this vein, they developed another model, i.e. VGGNet. Although, several versions of VGG have been introduced, where VGG-16 and VGG-19 are the most common ones with 16 and 19 layers, respectively. In this work, the output feature maps from all convolutional (16) and fully connected layers (2) are examined.

GoogLeNet (or Inception) was the leading architecture in ILSVRC 2014. In particular, authors [24] from Google presented the Inception v1 framework, implemented as GoogLeNet that reported an error rate of 6.7%. The key invention was an inception unit that acts like a tiny model inside a larger one. Besides, the novel idea was to utilize (1×1) filters to combine and shrink the features quantity prior to appealing rich corresponding blocks. It aids in aggregating convolutional features in a superior manner, which is quite difficult by merely stacking additional convolutional layers. Further, authors suggested slight enhancements in terms of regularization mechanism such as batch normalization along with re-arrangements of filters within the inception unit to build Inception v2 and v3 [25]. In recent, researchers have provided identity links to increase the flow of gradients in Inception v4 [26]. In this work, the output feature maps from all (5) convolutional and inception (12) layers are examined to constitute CNN Features.

ResNet was the winner of ILSVRC 2015 challenge. Authors in [27] from Microsoft presented the concept of identity (or residual) connections to feed the outputs from two consecutive convolutional layers as an input to the succeeding layer. This identity connection increases gradient propagation within the network, enabling the framework to convert into very deep (152 layers) network. This architecture succeeded in the ILSVRC 2015 contest and contributed in top-5 with 3.57% error rate. Here, the outputs of all convolutional layers (1) and bottleneck layers (17) are examined.

DenseNet was proposed by researchers [28] from Facebook in 2016 that contains a dedicated connection between each pair of CNN layers in a feed-forward manner. Such densely connected model yields various benefits such as, eliminating the vanishing gradient problem, facilitating feature reuse, speeding up feature propagation, and extensively dropping the number of parameters. Since the DenseNet model is too huge (with approx. 400 layers), we consider only up to first two dense blocks to examine the layer's outputs for cross-sensor iris recognition.

YOLO (You Look Only Once) was introduced in 2016 [29] for object detection in images/videos. It predicts the dimensions of the bounding box with probability of containing the object. The YOLO v1 contains 26 layers and can detect 49 objects. However, it is unable to detect small objects. Next, the YOLO v2 [30] with 30 layers was introduced in 2017. It was able to detect 9000 classes at one time, thus named as YOLO9000. Afterward, YOLO v3 [31] with 126 layers, was proposed in 2018 to handle much smaller objects that can detect more than 80 different objects. Moreover, another version of YOLO, i.e. YOLO v4 [32] has also been released in 2020 to achieve even better accuracy and optimal speed in object detection. Since the proposed work emphasizes iris recognition instead of iris detection, we have not explored YOLO.

2.2 Image Augmentation

The deep CNN models need massive data/images per class for training that is currently unavailable for iris recognition. In this view, some additional samples are supplied to the model during training through ImageDataGenerator, an open source tool of Keras library [33]. It applies various transformations on input image matrix to generate similar images to enhance the training data. In our work, raw iris images (without iris segmentation and normalization) are used as input to remove pre-processing overhead.

2.3 Feature Extraction

In this work, the abovementioned CNNs, i.e. AlexNet, ResNet, GoogleNet, VGG, and DenseNet, are employed to extract features from the raw iris images. Since all the standard CNNs are pre-trained on 224×224 sized images, we down-sample the dataset into the same dimensions. Each model contains multiple layers, and each layer constitutes features with varying dimensions. More in detail, initial layers build generic features such as points, edges, blobs etc., whereas later layers form task specific features. Besides, we have also added some customizations over the models to obtain more impressive results. For instance, in case of ResNet and DenseNet, the excessive depth may cause slower training that can be mitigated through discarding some layers from the model. Finally, the features extracted by the models are arranged individually in the form of 1D feature vectors that are next fed to the classifier. In order to examine the generalization of each model for cross-sensor iris recognition, their respective features are fed to a certain classifier to compute the respective recognition accuracy.

2.4 Classification

Next, the classifier is utilized for final classification, which is trained on the feature vector. In our work, the softmax classifier is employed due to its popularity regarding multi-class image classification. Specifically, in our experiments, there are large number of iris subjects to discriminate, therefore employing multi-class SVM would not worth. This is due to its increased complexity in implementing one vs all strategy that is identical to joining M binary SVMs, where each distinguishes a given class against other classes. On the other hand, softmax outputs a vector with M terms, representing the probability of belongingness to each output class.

3 Experiments

To estimate the performances of aforementioned models, a group of experiments has been created on two publicly available datasets, ND-iris-0405 [34] and ND-CrossSensor-Iris-2013 [35]. The following subsections begin with the description of the datasets along with various experiments designed to realize cross sensor iris recognition.

Table 1 Dataset exploration

Sensor datasets	# iris subjects	# iris samples
ND-0405	712	64,980
LG4000	1352	29,986
LG2200	1352	116,564

3.1 Dataset Description

Table 1 explores the aforementioned datasets. ND-Iris-0405-Dataset contains total 64,980 images from 356 subjects. With right and left iris samples of 480×640 dimension from each subject, it contains total 712 unique iris instances, collected through LG2200 iris sensor. The ND-CrossSensor-Iris-2013 dataset comprises 29,986 images collected using LG4000 sensor, whereas 116,564 images are captured from LG2200 sensor. These images are from 676 unique iris subjects, therefore total 1352 unique instances are formed with left and right irises.

3.2 Experimental Design

The experimental design for cross-sensor iris recognition is demonstrated in this subsection. Since both datasets contain right and left iris samples of each subject, they are allocated different labels. Moreover, as the iris samples are unevenly distributed across iris subjects, each dataset is partitioned into three groups in order to balance the sample distribution within each part. In specific, Part-1 encompasses iris classes with maximum iris samples, while, Part-3 encompasses iris categories with least number of samples. Tables 2, 3, and 4 illustrate the partitioning of datasets into groups within ND-0405, LG2200, and LG4000 sensor datasets, respectively.

The same splitting is separately performed on ND-cross-sensor dataset.

Various experiments are performed on abovementioned datasets for cross-sensor iris recognition. Some experiments are designed for individual parts, while other pools images from similar parts of different sensors. Here, each part is partitioned into training and validation sets with 9:1 ratio. The training and validation sets are used for model training while images from probe set containing identical classes are utilized for testing.

Table 2 Split of ND-0405

	# iris subjects	# iris samples
Part-1	174	26,586
Part-2	280	28,753
Part-3	190	7786

Note: Some subjects with very less number of samples are not involved in the experiments

Table 3 Split of ND-cross-sensor (LG2200)

	# iris subjects	# iris samples
Part-1	196	26,586
Part-2	572	30,879
Part-3	269	9576

Table 4 Split of ND-cross-sensor (LG4000)

	# iris subjects	# iris samples
Part-1	626	14,106
Part-2	445	9920
Part-3	281	5960

Table 5 Cross-sensor experimental design

Exp	Training and validation				Testing		
	Datasets	# iris subjects	Train-set	Validation set	Datasets	Probe-set	# iris subjects
1	LG-2200	196	23,927	2659	LG-4000	1457	47
2	ND-0405	280	25,878	2875	LG-4000	2106	78
3	LG-4000	281	5364	596	LG-2200	1275	75
4	ND-0405 + LG-4000	280 + 281	25,878 + 5364	2875 + 596	LG-2200	9688	173
5	LG-2200 + LG-4000	196 + 281	23,927 + 5364	2659 + 596	ND-0405	5655	145

Cross-sensor Dataset Creation: Here, five different experiments are conducted to assess model's performance on cross-sensor iris recognition, where different pairs of parts are taken from heterogeneous sensors in order to produce training and probe sets as given in Table 5. The experiments Exp1, Exp2, and Exp3 are performed on couples of autonomous sensors for ND-0405 versus LG-4000, LG-2200 versus LG-4000, and LG-4000 versus LG2200, respectively. However, Exp4 and Exp5 are performed over different combination of heterogeneous sensors i.e. ND-0405 + LG-4000 versus LG-2200 and LG-2200 + LG-4000 versus ND-0405, respectively.

4 Experimental Outcomes

The performance assessment of all CNN models for cross-sensor recognition is carried out by executing all the experiments given in Table 5. The design of such experiments attempts to incorporate training and testing samples from different sensors. Besides, first three experiments consider single sensor, while remaining assimilate multiple sensors to create train and test samples. Figure 2a–e represent the results in terms of ROC curves corresponding to five experiments listed in Table 5. These curves represent the discrimination of all CNN models for all five experiments.

In all experiments AlexNet is entirely used without customization, whereas other models are altered for feature extraction. For instance, it is observed during execution of Exp1 that VGG constitute most dominant features at 12th convolutional layer. Therefore, we remove the subsequent layers and consider merely up to 12th layer. Besides, the 15th convolutional layer of GoogleNet is more constructive compared to

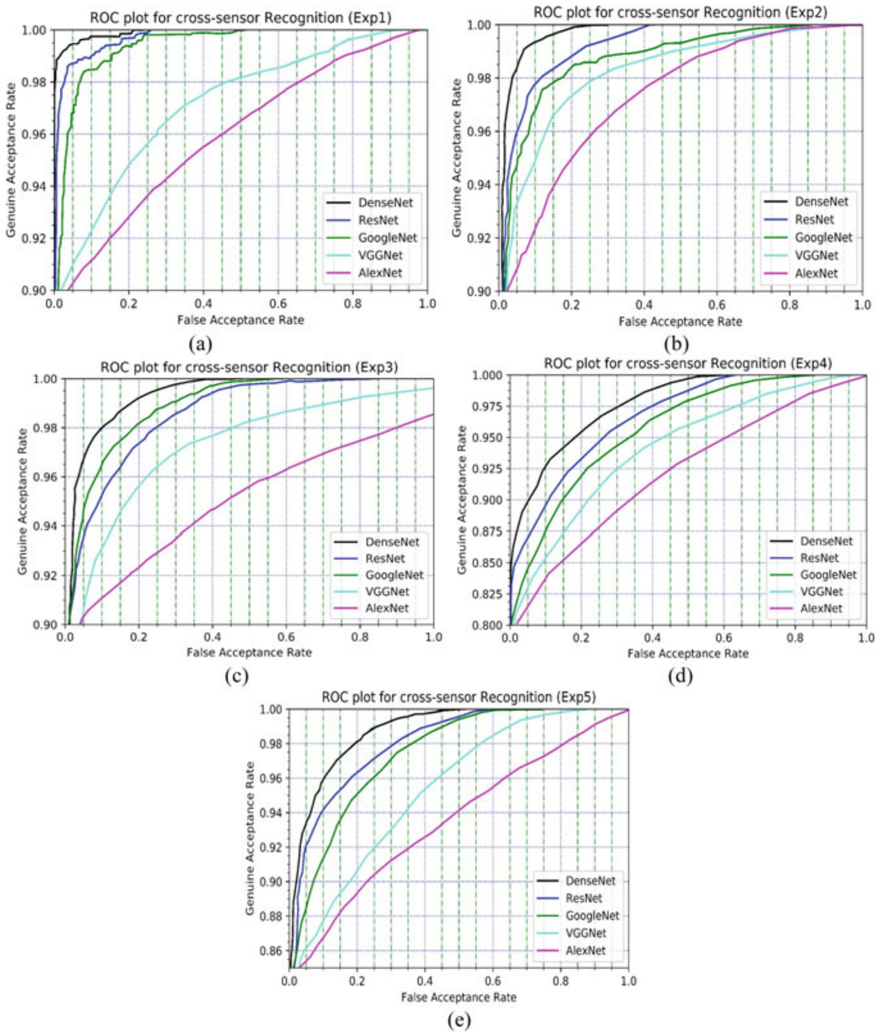


Fig. 2 ROC plots depicting performances of standard CNNs for Cross-sensor Iris Recognition

others. On the other hand, ResNet and DenseNet construct best features at layers 12 and 22, respectively. Therefore, we extract the iris features from the abovementioned layers corresponding to each model and fed them to Softmax classifier. The ROC curves resulted by all models for Exp1 are depicted in Fig. 2a. Similarly, For Exp2, the discrimination capability of all models is depicted in terms of ROC curves in Fig. 2b. Here, VGG extracts best discriminating features at 10th convolutional layer. Further, GoogleNet, ResNet and DenseNet exhibit finest outcome at 10th, 13th and 15th layers, respectively. Further, in Exp3 and Exp4, VGG gives peak performance at 12th layer then starts diminishing. Whereas, Inception, ResNet and DenseNet reach to highest accuracy at 10th, 13th and 22nd layer respectively. Figure 2c and d compares the discrimination performance exhibited by all CNN models. Finally, in Exp5, the VGGNet, Google Inception, ResNet and DenseNet achieve peak performance at 11th, 10th, 14th and 22nd layers respectively. The corresponding ROC curves are given in Fig. 2e.

The variance in the top layers can be clarified by the assets of respective CNNs. GoogleNet rapidly touches the peak compared to others due to complex inception layers, where each layer acts as a sub-unit within a larger model. As opposed to it, ResNet with skip links is efficient in enabling the quick gradient flow within the network. Due to this, the model performs well at the increased depth, and thus yields an advanced peak in the cross-sensor iris recognition accuracy. Besides, DenseNet having the resilient dense connections, enables neurons to collaborate effortlessly, resulting the top recognition amongst all CNN frameworks for cross-sensor iris recognition. It may be observed here that top results are not approaching the posterior layers of deep CNNs. It may be justified as the iris image is quite simple than the images within the ImageNet dataset, where huge structural differences exist in an extensive sort of objects. In contrast, we have designed the experiments with few iris subjects to discriminate. Therefore, it is meaningless to include layers in increased quantity in order to learn the iris feature representation. Thus, topmost accuracies are reported near the mid layers or before them.

In specific, we can observe from Fig. 2 that in all the experiments, DenseNet attains the best discrimination amongst the five CNN models. Next, ResNet and Inception reach similar performances near 12th to 15th layers. Besides, VGG with its simple design performs with comparatively less discrimination accuracy from 10th to 12th layers. Furthermore, in our experiments, the recognition accuracy of AlexNet linearly increases, which specifies that the layers exist in it might not entirely capture the discriminatory features within the given iris images.

5 Conclusion

This paper attempts to accomplish cross-sensor iris recognition using preeminent CNN models primarily trained on ImageNet dataset, through transfer learning and fine-tuning. The experimental results have delineated that even though the CNNs models are actually trained on animal features, they are adequate for iris recognition. The knowledge transfer can cope with the limited size dataset, which creates an obstacle in employing deep CNN models in the domain of iris recognition. Besides, some image augmentation strategies can further improve model training. The experimental results also depict that these CNN models can be customized according to the requirements of underlying application, i.e. they may be used entirely or some portion itself may be sufficient to achieve the best performance. These CNN models aid in automatic feature-production that can benefit various wide scale applications.

Although, the standard CNN models have the potential to perform well in any arbitrary application, their customization is significantly complicated. This can be understood by the fact that, in order to achieve best results, each layer should be examined based on the extracted features. There is no universal remark, where the best result would be found. Besides, selecting the network portion to freeze and fine-tune is also a challenging task, i.e. how many initial layers need to be freeze is completely a hit and run method.

References

1. Wang K, Kumar A (2019) Toward more accurate iris recognition using dilated residual features. IEEE Trans Inf Forensics Secur 14(12):3233–3245
2. Choudhary M, Tiwari V, Venkanna U (2019) Enhancing human iris recognition performance in unconstrained environment using ensemble of convolutional and residual deep neural network models. Soft Comput 1–15
3. Bowyer K, Baker S, Hentz A et al (2009) Factors that degrade the match distribution in iris biometrics. Identity Inf Soc 2(3):327–343
4. Daugman J (1993) High confidence visual recognition of persons by a test of statistical independence. IEEE Trans Pattern Anal Mach Intell 15(11):1148–1161
5. Okokpujie K, Noma-Osaghae E, John S et al (2017) An improved iris segmentation technique using circular Hough transform. Lecture notes in electrical engineering, vol 450. Springer, pp 203–211
6. Radman A, Jumari K, Zainal N (2013) Fast and reliable iris segmentation algorithm. IET Image Proc 7(1):42–49
7. Barpanda S, Sa P, Marques O et al (2018) Iris recognition with tunable filter bank based feature. Multimed Tool Appl 77(6):7637–7674
8. Trokielewicz M, Bartuzi E (2018) Cross-spectral iris recognition for mobile applications using high-quality color images. J Telecommun Inf Technol 3:91–97
9. Arora S, Vatsa M, Singh R et al (2012) On iris camera interoperability. In Proceeding IEEE international conference on biometrics: theory, applications, and systems, Arlington, VA, USA, pp 346–352
10. Ding C, Tao D (2018) Trunk-branch ensemble convolutional neural networks for video-based face recognition. IEEE Trans Pattern Anal Mach Intell 40(4):1002–1014
11. Choudhary M, Tiwari V, Venkanna U (2020) Iris anti-spoofing through score-level fusion of handcrafted and data-driven features. Appl Soft Comput 106206
12. Choudhary M, Tiwari V, Venkanna U (2019) An approach for iris contact lens detection and classification using ensemble of customized DenseNet and SVM. Futur Gener Comput Syst 101:1259–1270
13. Choudhary M, Tiwari V, Venkanna U (2020) Iris liveness detection using fusion of domain-specific multiple BSIF and DenseNet features. IEEE Trans Cybern. https://doi.org/10.1109/TCYB.2020.3005089
14. Liu N, Zhang M, Li H et al (2015) Deepiris: learning pairwise filter bank for heterogeneous iris verification. Pattern Recogn Lett 82(2):154–161
15. Al-Waisy S, Qahwaji R, Ipson S et al (2018) A multi-biometric iris recognition system based on a deep learning approach. Pattern Anal Appl 21(3):783–802
16. Gangwar A, Joshi A (2016) DeepIrisNet: deep iris representation with applications in iris recognition and cross-sensor iris recognition. In: IEEE international conference on image processing (ICIP), Phoenix, AZ, pp 2301–2305
17. Nguyen K, Fookes C, Ross A et al (2018) Iris recognition with off-the-shelf CNN features: a deep learning perspective. IEEE Access 6:18848–18855
18. Wang Z, Li C, Shao H et al (2018) Eye recognition with mixed convolutional and residual network (MiCoRe-Net). IEEE Access 6:17905–17912

19. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis (IJCV) 115(3):211–252
20. Razavian AS, Azizpour H, Sullivan J, Carlsson S (2014) CNN features off-the-shelf: an astounding baseline for recognition. In: IEEE conference on computer vision and pattern recognition workshops, pp 512–519
21. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks, vol 2. US, pp 1097–1105
22. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE
23. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition, CoRR
24. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: IEEE conference on computer vision and pattern recognition
25. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 2818–2826
26. Szegedy C, Ioffe S, Vanhoucke V (2017) Inception-v4, inception-ResNet and the impact of residual connections on learning. AAAI Conference on Artificial Intelligence, pp 4278–4284
27. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778
28. Huang G, Liu Z, Weinberger KQ, van der Maaten L (2017) Densely connected convolutional networks. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 4700–4708
29. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
30. Shafiee MJ, Chywl B, Li F, Wong A (2017) Fast YOLO: a fast you only look once system for real-time embedded object detection in video. arXiv preprint arXiv:1709.05943
31. Redmon J, Farhadi A (2018) Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767
32. Bochkovskiy A, Wang CY, Liao HYM (2020) YOLOv4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934
33. Chollet F et al (2015) Keras. https://github.com/fchollet/keras
34. Bowyer K, Flynn P (2009) The ND-IRIS-0405 Iris Image Dataset. CoRR
35. ND-CrossSensor-Iris-2013. Available at: https://sites.google.com/a/nd.edu/public-cvrl/data-sets

Travelling Salesman Problem Using GA-ACO Hybrid Approach: A Review

Ankita

Abstract Travelling Salesman Problem is the most widely studied combinatorial optimization problem which has attracted many researchers in the last few years. It is recognized as one of popular NP-Hard problem that has broad search space. TSP is a difficult problem which can't be solved by conventional methods particularly when the number of cities increases. So, to solve this problem efficiently, heuristic approach is the most feasible solution. Genetic Algorithm and Ant Colony Optimization are commonly used metaheuristics to solve TSP successfully. To solve TSP, this paper considers the three eminent heuristic approaches namely ACO, GA and hybrid GA-ACO.

Keywords Ant colony optimization · Genetic algorithm · Hybridization · Metaheuristics · Travelling salesman problem

1 Introduction

Optimization is a very common process that forms an essential part of our day-to-day life. It can be seen in almost all phases of life such as financial, industrial, business, engineering, etc. All living beings use an optimization to make their life easier. It is a methodology of making something fully perfect, effective. Most of the engineering problems can be formulated as optimization problems. TSP is hardly solved by conventional methods particularly when the number of cities increases. To solve such optimization problems, many heuristics and approximation algorithms have been developed. However, they are not efficient enough to solve TSP. So, to solve this problem efficiently, we need metaheuristics. Metaheuristics are one of the most exciting techniques that are helpful to solve complex optimization problems in real time. They play a pivotal role in the situation where the traditional methods fail to deliver the satisfactory results. Metaheuristics are gaining a splendid success in solving various hard optimization problems which exist in real world. Nature has

Ankita (✉)
Department of Computer Science and Applications, Kurukshetra University, Kurukshetra, India
e-mail: ankita.30@kuk.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_11

been a great and an immense source of inspiration for solving hard and complex problems. Nature inspired algorithms such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), Simulated Annealing, Particle Swarm Optimization, etc. have been applied to many combinatorial optimization problems to obtain reasonable solutions.

TSP is a widely studied NP Hard problem in computational metaheuristics. It is a mathematical problem that can be represented in terms of a graph which describes the position of nodes. ACO and GA are two bionic optimization algorithms that can solve TSP successfully and efficiently. ACO can find the shortest distance between food source and the nest. Although ACO produces satisfactory solutions to solve TSP, finding optimal solutions has always been a challenge to it. This allows GA to solve TSP appropriately but sometimes it can have stuck in local optimum solutions. Thus we need a hybrid approach of ACO and GA that can overcome their individual problems and achieve the best solution.

The structure of paper is introduced as follows: Sect. 2 discusses the Travelling Salesman Problem in detail; in Sect. 3, approaches to solve TSP such as GA, ACO, hybridization is discussed; finally conclusion is given in Sect. 4.

2 Travelling Salesman Problem (TSP)

TSP is a well known classic algorithmic problem which belongs to the field of computer science. It is one of the best known NP-hard problems, which means that there is no exact algorithm to solve it in polynomial time [1]. The minimal expected time to obtain optimal solution is exponential [1]. In this approach, some cities are provided and the aim is that a shortest path has to be found between these cities. Each city has to be visited only once and finally returns to the first city. The challenge of the problem is to minimize the total length of the trip. The origins of the TSP in mathematics are not really known all we know for certain is that it happened around 1930 [2]. In 1930, it was defined by Karl Menger but its fame in research got hype during 1960s to 1980s. Although, TSP is very simple to describe but it is very difficult to solve. In terms of graph theory, it can be defined as finding the Hamiltonian cycle of minimum weight for the given graph. TSP can be symmetric or asymmetric. In symmetric approach, distance between city i and j is equidistant to that of city j and i. Whereas, in asymmetric TSP, edges might have different weights in both directions.

TSP can be represented as a complete weighted graph $G = (N, E)$ where N is the collection of all those cities that have to be traveled and E is the collection of edges which link all the cities. A cost c_{ij} is associated to each edge $(i, j) \in E$, where c_{ij} is the distance between cities i and j. c_{ij} is defined in the Euclidean space as follows:

$$c_{ij} = \sqrt{\left(\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2\right)}$$

TSP has achieved its fame in various engineering applications. It has been utilized to design electronic devices, hardware devices in communications, in networking, etc. Some of the practical problems which can be identified as TSP are Vehicle Routing Problem, Ship Scheduling, Frequency Assignment Problem, Flowshop Scheduling and many more.

3 Approaches to Solve TSP

3.1 Genetic Algorithm (GA)

Genetic Algorithm is an adaptive metaheuristic search algorithm that belongs to evolutionary algorithm. GA is one of the oldest and most successful optimization techniques. It can be used to generate extremely good solutions for optimization and search problems. It can be viewed as a class of local search method that is based on natural selection and population genetics. It mimics the process of natural selection which infers that only the fittest attributes in the environment will be able to survive and reproduce, thus forms a new population. Some of these attributes will pass down to new generations. Genetic material from the previous generation is given to the subsequent generation maintaining fitness strength [3]. It simulates "survival of the fittest" among individuals of consecutive generations for solving a problem.

In the process of GA, selection of the fittest individuals is done for reproduction to form a new generation. The GA comprises of many solutions known as chromosomes or individuals. Each chromosome in a GA includes many genes, which specify the traits of each individual. A form of population is created by a set of chromosomes. The attribute of each chromosome is analyzed by a fitness function. Fittest chromosomes are chosen to form a new set of chromosomes. Then, a pair of fit chromosomes is selected and crossover is applied on it to produce a new offspring. At last, mutation step is applied on the selected chromosomes to maintain the diversity in population and to reduce the possibility of getting stuck in local optimum. This whole process is repeated till the required result is obtained or the termination condition is met.

Operators of Genetic Algorithm: Once the initial generation is created, GA uses three operations to evolve from one generation to another.

(1) **Selection**: It is inspired by the principle of 'Survival of the Fittest'. Firstly, random population is generated to evolve over generations. Then, fitness function is used to evaluate the performance of each individual. Based on their fitness score, the best parents are selected to pass their genes to the successive generations. The most common approach for selection is Fitness Proportionate Selection or Tournament Selection.
(2) **Crossover**: It is inspired by mating between individuals. One or more crossover sites are chosen randomly, and then the genes of individuals are exchanged to

create a new offspring. However, to include all locations exactly once in TSP, ordered or partially mapped crossover is used (Fig. 1).

(3) **Mutation**: It causes local modifications to the new generations randomly. It can be performed by inserting new genes in offspring or flipping the existing ones. Its main aim is to maintain diversity in the population. Since, in TSP, cities can't be dropped or added to the existing ones, swap mutation can be used in which two cities swap their places in the route (Fig. 2).

Basic steps of GA are as follows:

Step 1: Start
Step 2: Generate initial random population
Step 3: Calculate fitness of each individuals
Step 4: Is termination condition satisfied? If yes goto end, if no goto step 5

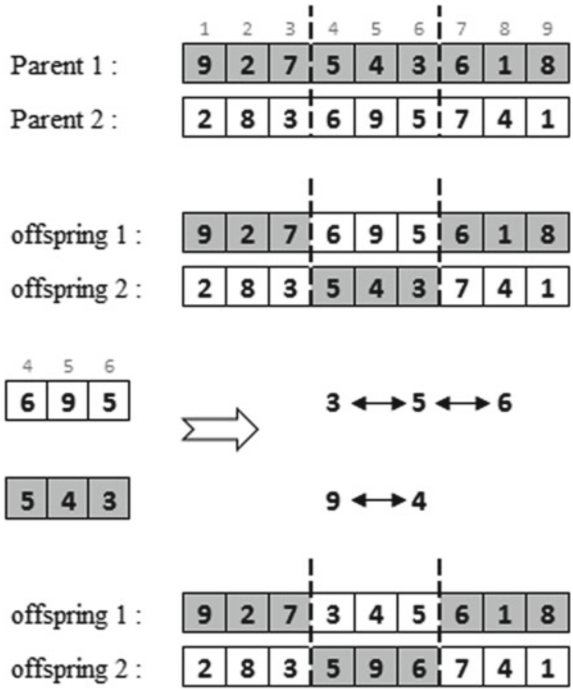


Fig. 1 Partially mapped crossover

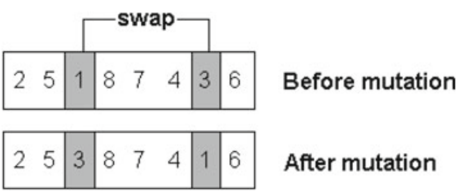


Fig. 2 Swap mutation

Step 5: Select the best individual based on fitness value
Step 6: Apply crossover operator on selected individual
Step 7: Apply mutation operator by flipping genes
Step 8: Generate a new population
Step 9: End

GA has wide range of applications to solve optimization problems such as Travelling Salesman Problem, Scheduling Problems, Image Processing, Vehicle Routing Problem, Machine Learning, etc.

3.2 Ant Colony Optimization (ACO)

ACO is one of the oldest and widely used optimization algorithm which is inspired by foraging behaviour of real ants. It was introduced by Dorigo in 1992 and was formalized as a new metaheuristic in 1999 [4]. It is a nature based search technique used to find the solution of combinatorial optimization problems.

The basic concept of ACO is to mimic the real behaviour of ants to determine the shortest path from the nest to food. When ants move from nest to food, they release a highly volatile chemical substance known as pheromone on the ground. They leave pheromone to attract and guide other ants on the path that they have traveled. This mechanism is called ‘stigmergy’ which means indirect communication among the self-organizing agents. Other ants of the colony sense that pheromone and choose the path that has more pheromone deposited. The path, of which pheromone intensity is more, will have more chance to be selected by the ants. The ants which could not smell the pheromone walk along a random route. Finally, the ants choose the shortest path to transport food to their nest (Fig. 3).

ACO algorithm mainly consists of three procedures that constitute the major part of this algorithm.

(1) **Construct Ant Solutions**: In this step, “ants” use heuristic information to construct paths stochastically i.e. to build the solution of optimization problem.
(2) **Update Pheromones**: In this process, pheromone trails are modified. The value of pheromone trails increase when ants traverse the path and decrease when the pheromone evaporation takes place.
(3) **Daemon Actions**: After construction of solution and before updating the pheromone value, there is a requirement of some problem specific actions. These are called daemon actions which refer to the decisions taken based on the global information.

Mathematical model of ACO for TSP

ACO was first introduced and applied to TSP by Marco Dorigo et al. [5]. In this, some ants are randomly located to given n cities which are at the distance d_{ij}, where d_{ij} is the distance between city i and j. The choice of next city to be visited depends on the amount of pheromone trail left on the path. The probability of kth ant at node

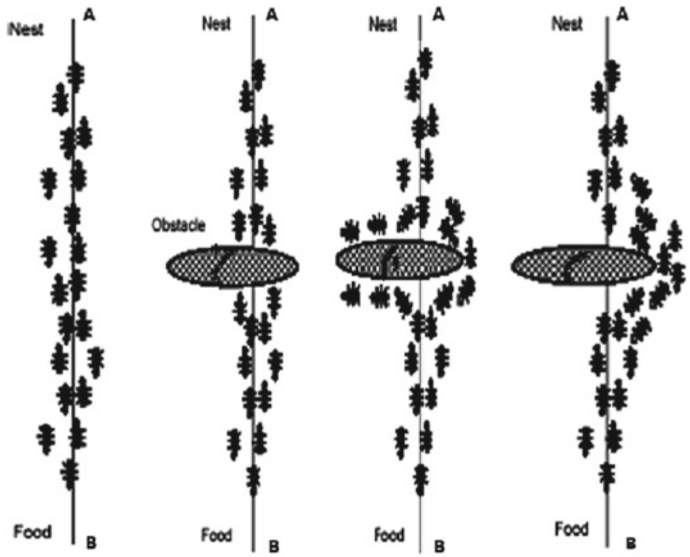


Fig. 3 Ant colony optimization

i choosing node j using pheromone trail t_{ij} is written as:

$$p_{ij}^{k} = ([t_{ij}]^{\alpha}[nij]^{\beta})/(\sum_{l\in \text{allowed k}} [t_{ij}]^{\alpha}[nij]^{\beta}) \text{ if } j \in \text{allowed}_k$$

where, t_{ij} = intensity of pheromone trail between cities i and j,
α = parameter to regulate the control of t_{ij},
n_{ij} = visibility of city j from city I,
β = parameter to regulate the control of heuristic visibility,
k = set of cities that have not been visited yet.

After each ant has chosen a city, amount of pheromone trials are updated by given equation:

$$t_{ij}(t+1) = (1-\rho)\, t_{ij}(t) + \rho \cdot t_0$$

where ρ = decay parameter, $0 < \rho < 1$;
$t_0 = 1/n$;
n = number of cities in TSP.

Basic steps of ACO are given below:
Step 1: Start
Step 2: Create ants
Step 3: Put ants on an entry state

Step 4: Select next state for each ant
Step 5: Is it a final state? If yes goto step 6, if no goto step 4
Step 6: Deposit Pheromone
Step 7: Perform Daemon Activities
Step 8: Evaporate Pheromone
Step 9: Is termination condition met? If yes goto step 10, if no goto step 2
Step 10: End

In the last few years, ACO has proved to be a successful technique for hard combinatorial optimization problems. It has powerful capability over other metaheuristics approach, when graph may change dynamically. ACO is more applicable for problems where predefined source and destination are given. It has been extensively applied to various combinatorial problems such as Travelling Salesman Problem, Vehicle Routing Problem, Job Shop Scheduling Problem, Quadratic Assignment Problem, Knapsack Problem, etc. Nowadays, this approach enjoys a rapidly growing popularity for combinatorial optimization problems and can be extended to continuous search domains as well.

3.3 Hybridization

Hybrid is formed by combining two or more entities. Two or more components having distinguished genotypes and perform essentially the same function are merged together. Hybridization is defined as the process of reproduction of genetically distinct populations. Hybrid algorithms are two or more algorithms that run together and complement each other to produce a profitable synergy from their integration [6]. Nowadays, hybrid optimization approaches are gaining a widespread popularity for solving hard optimization problems. For complex real world scenarios, one specific metaheuristic model fails to provide satisfactory results. So, there is a need to combine different algorithmic techniques. Different metaheuristic concepts are hybridized with each other. The main motivation behind the hybridization of different algorithms is to exploit the complementary character of different optimization strategies, that is, hybrids are believed to benefit from synergy [7]. Such well designed hybrids often perform better than their counterparts.

Hybrid GA-ACO for solving TSP

The hybrid algorithm takes the advantages of both Genetic Algorithm and Ant Colony Optimization. ACO has some drawbacks such as it can easily fall to local optimal solution. Though, it solves many optimization problems successfully, still its convergence couldn't be proved. GA can generate global optimal solution if well defined problem with sufficient time is provided to it. This makes it well suited to solve TSP. Still some limitations are there i.e. when fitness function is not properly defined, GA may converge towards local optima.

So, due to various limitations of GA, ACO individually, hybrid approach is preferred over them. Based on the views of different researchers, many articles are published which give detailed idea about hybridization of GA and ACO.

A hybrid optimization algorithm is proposed which is based on GA and ACO for solving TSP approach. TSP is assessed with some parameters like random data, sample data. Instincts of ant colony are combined with evolutional process of GA to find the shortest route to seek food. Gong et al. [8] proposed a hybrid approach in which every chromosome of GA acts as an ant of ACO. The pheromones values of ACO are used to find the building blocks of GA and to guide the selection of genetic operation points. Another hybrid approach is used to shorten the route chosen by ants. Initial answers of ACO are selected by GA from the obtained data routes for the mutations. Then, for ants search, GA answers have been used.

Sometimes, ACO parameters are optimized by GA in accordance with the output of ACO. At the same time, distance parameter is the input value of GA.

A hybrid pseudo-parallel genetic algorithm with ACO is discussed by Li et al. [9]. During different sub populations of GA, work of information exchange is done by ACO.

While comparing performance of GA and ACO of solving TSP, it has been observed that GA can be implemented easily, quickly as well as cost efficiently. When the search space is large and complex, it has rapid global search capability. Though ACO is greedier than GA, still it provides better results, particularly with large problems. It is suitable for finding high quality solutions and can easily accommodate with other algorithms. When experimental results of GA and ACO for solving TSP are compared, it has been proved that ACO provides far better results than GA.

4 Conclusion

Over the last few years, a lot of research has been conducted to solve TSP using different metaheuristics. GA and ACO are one of the most widely studied nature inspired metaheuristic algorithms for solving TSP. There are many limitations of GA and ACO individually. In order to overcome respective shortcomings, hybrid approach of GA and ACO is used to solve TSP. In this paper, various approaches to solve TSP such as GA, ACO and hybrid GA-ACO are discussed in detail. After different papers related to hybridization of GA-ACO reviewed, it has been concluded that the hybrid approach is far better than the individual approach. In future, this hybrid approach can be used to solve other combinatorial optimization problems.

Bibliography

1. Brezina Jr I, Čičková Z (2011) Solving the travelling salesman problem using the ant colony optimization. Manag Inf Syst 6(4):010–014

2. Rai K, Madan L, Anand K (2014) Research paper on travelling salesman problem and it's solution using genetic algorithm. Int J Innov Res Technol 1(11):103–114
3. Singp S, Kumari A (2013) Analysis of travelling salesman problem using ant colony optimization, genetic algorithm and hybrid of ant colony optimization and genetic algorithm. Int J Eng Res Technol 2(7):1896–1900
4. Duan H, Yu X (2007) Hybrid ant colony using memetic algorithm for travelling salesman problem. In: Proceedings of the 2007 IEEE symposium on approximate dynamic programming and reinforcement learning, pp 92–95
5. Dorigo M (1996) The ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern 26(1):1–13
6. Ting TO, Yang XS, Cheng S, Huang K (2015) Hybrid metaheuristic algorithms: past, present, and future. Recent advances in swarm intelligence and evolutionary computation, vol 585, pp 71–83
7. Blum C, Puchinger J, Raidl GR, Roli A (2011) Hybrid metaheuristics in combinatorial optimization: a survey. Appl Soft Comput 11(6):4135–4151
8. Gong D, Ruan X (2004) A hybrid approach of GA and ACO for TSP. In: Proceedings of 5th world congress on intelligent control and automation, vol 3, pp 2068–2072
9. Li S, Chen H, Tang Z (2011) Study of pseudo-parallel genetic algorithm with ant colony optimization to solve the TSP. Int J Comput Sci Netw Secur 11(3)

Efficient and Robust Indian Number Plate Recognition Through Modified and Tuned LPRNet

Rahul Singh Dangi, Ashish Kuvelkar, Samrit Kumar Maity, and Sanjay Wandhekar

Abstract This paper presents a novel approach for Indian Number Plate Recognition using the Convolution Neural Network (CNN). We generated and augmented synthetic data in Indian standards with variant background and font. We also collected and labelled real data of Indian License plates to create a genuine Indian oriented dataset. We compared existing models like SVM, KNN and CRNN in terms of accuracy and inference time, and selected LPRNet. We first trained the existing LPRNet model, which is designed for Chinese License plates on synthetic data and secured 85% accuracy. We then modified the LPRNet model and secured 93% accuracy on the same synthetic data. We applied the post-processing technique of pattern recognition rules for Indian standards of License plates to further improve accuracy. The Proposed system achieved 95% accuracy after fine-tuning with real data. Our model is very lightweight and can easily be deployed on Nvidia-Jetson nano, Intel Neural Compute Stick.

Keywords Deep learning · CNN · LPRNet · Indian license plate recognition · Indian number plate recognition

R. S. Dangi (✉) · A. Kuvelkar · S. K. Maity · S. Wandhekar
High-Performance Computing Technologies Group, Centre for Development of Advanced Computing (C-DAC), Pune, India
e-mail: drahul@cdac.in

A. Kuvelkar
e-mail: ashishk@cdac.in

S. K. Maity
e-mail: samritm@cdac.in

S. Wandhekar
e-mail: sanjayw@cdac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_12

1 Introduction

Research on Deep Learning (DL) and Machine Learning (ML) is trying to solve challenges that were considered un-solvable or extremely difficult to solve. Technological advancement in hardware, mathematics, and related fields is pushing the boundaries of possibilities every day. ML/DL technology became a major force behind the techno-social revolution.

One ML application that has a wider social impact is Vehicle Number Plate recognition or License Plate Recognition (LPR). It has high importance in ensuring road safety, controlling crimes, managing crowds etc. Currently, Number plate recognition is done in a limited, semi-automatic way within a few major Indian cities wherein capturing number plate images is automatic and machine supported but recognition is mostly manual. Research works exploring a fully automatic way of number plate recognition [4, 12–14] achieves accuracy up to only considerable level. Also, there is limited information about the performance of inference under the deployable scenario. There are mainly 2 aspects in ANPR namely license plate detection and recognition. The work done in [24, 25] have achieved 100% accuracy in license plate detection but license plate recognition is still a promising research area seeking strong improvement.

To perform license plate recognition, two methods are being used. The first method is to segment individual character image and perform character recognition on those segmented images by using machine learning models like Support Vector Machines (SVM) [1] or K-Nearest Neighbours (KNN) [2] where the models are pre-trained on EMNIST [3] handwritten dataset. The issue with this method is it degrades the overall accuracy as separating character images is not 100% accurate. The other method is using a neural network on a complete license plate input image and training it for its corresponding license plate number. Unlike the first method, the second approach eliminates the chances of failures caused by the separation of individual characters from the full image. In this work, we adopted the second method of treating input by our target model i.e. LPRNet. In a relevant work [4] Recurrent Convolutional Neural Network (RCNN) [5] model is used achieving accuracy up to 76%. LPRNet [8] is based on deep neural networks that use only Convolution Neural Network (CNN) [6] and does not use any Recurrent Neural Networks (RNN) [7] resulting in lesser inference time.

In the present work, LPRNet is modified to make it suitable for Indian Number Plates. For training LPRNet, 10K samples of synthetic number plates were randomly generated as per Indian standards [11]. Data augmentation techniques like scaling, rotation, and transformation were applied to impart variation in data. Our modified LPRNet was fine-tuned further by transfer learning on real data of 4000 images collected at the Centre for development of Advanced Computing (C-DAC), Pune premises. The license plates were segmented from the captured vehicle images using the License plate detection model YOLOV3 [9, 10]. Later, labelling of the images was achieved manually through the web portal designed by us for this purpose.

The results from LPRNet are post-processed with the pattern matching technique to filter out invalid License plate numbers in accordance with Indian standards [11].

This paper is organized as below. First, we will discuss existing models for license plate recognition to understand and compare their accuracy and inference time. Next, we will discuss our approach which uses a modified LPRNet. Then we will discuss the experimental setup we built to test and draw performance of our model.

2 Related Work

There are several algorithms available for License plate recognition starting from machine learning algorithms like SVM [1], KNN [2] to Recurrent Neural Networks [7].

The approach followed in [12] is a two-step process wherein in the first step, single character input images are generated using intensity projection and contouring techniques. In the second step, machine learning algorithms are applied to those uni-character images for character identification. For determining character in the uni-character image, machine learning algorithms like KNN or SVM are trained for 26 Alphabets (A–Z) and 10 digits (0–9). It has achieved best-case accuracy of 89.5% with letters and 97.7% with digits on EMNIST dataset using SVM model as given in [13].

The problem with character segmentation approach is:

1. For tilted image intensity, projection approach to separate individual character images fails
2. Segmentation approach using contouring fails if two or more character appears as connected in the input image.

Apart from the above issues, another issue is with the Machine Learning model used. Let's say there are around 10 characters in a License plate. The probability of the best fit ML model is 95%. Then the probability of the License plate number detected correctly is:

$$Probability(License\ Number\ recognition) = 0.95^{10} = 0.59$$

The algorithm in [14] uses a 16-layer CNN model for License plate recognition. The spatial transformer module is used to make CNN invariant to large variations. The maximum length of letters on the input license plate is assumed to be 11. So, CNN has 11 parallel classifiers in the final layer which uses softmax to classify each of 11 characters to one of 37 classes (A–Z, 0–9, null). It is trained on a synthetic dataset. The complete number plate recognition accuracy is 57% and the character level accuracy is 92%.

The algorithm in [4] is based on CNN combined with RNN. In CNN, VGG-Net network is used, which does the task of character images segmentation and RNN models dependencies in between segmented images and does the task of classifying the character images into its corresponding character class. The license plate recognition accuracy is 76% and the character level accuracy is 95.1%.

3 Proposed Approach: Modified LPRNet

In this section, we propose a modified LPRNet model architecture as per Indian standards.

3.1 Existing LPRNet Model

LPRNet was designed originally for Chinese License plates and it gave 96% accuracy with Chinese data but the same model was not giving equally accurate results on Indian standard data when trained from scratch. Hence, we customized the model in order to improve accuracy.

LPRNet design consists of

1. a location network with Spatial Transformer Layer (optional)
2. light-weight convolutional neural network (backbone)
3. per-position character classification head
4. character probabilities for further sequence decoding.

The basic building block is taken from SqueezeNet Fire blocks [15] and Google Inception blocks [16–18]. The small basic block structure is shown in Table 1. It takes a 3-dimensional input and returns output with the same dimensions and the same size as in Table 1. It has 4 levels of convolution to extract depth features which will finally accumulate to decide which character it is.

The backbone network architecture is in Table 2. The backbone takes a raw RGB image as input and calculates spatially distributed rich features. Wide convolution (with 1×13 kernels) utilizes the local character context instead of using LSTM-based RNN. The backbone subnetwork output can be interpreted as a sequence of character probabilities whose length corresponds to the input image pixel width.

Table 1 Small basic block

Layer type	Parameter/dimensions
Input	Cin*H*W
Convolution	Cout/4 1*1 stride 1
Convolution	Cout/4 3*1 stride = 1, padh = 1
Convolution	Cout/4 1*3 stride = 1, padw = 1
Convolution	Cout 1*1 stride = 1
Output	Cout*H*W

Table 2 Modified LPRNet architecture

#	Layer	LPRNet	Modified LPRNet
1	Input	94*24 RGB image	**128***24 RGB image
2	Convolution	64 3*3 stride 1	**96** 3*3 stride 1
3	Max pooling	64 3*3 stride 1	**96** 3*3 stride 1
4	SBB1	128 3*3 stride 1	**192** 3*3 stride 1
5	max pooling	64 3*3 strides (2,1)	**96** 3*3 strides (2,1)
6	SBB2	256	**Removed**
7	SBB3	256	Same
8	Max pooling	64 3*3 strides (2, 1)	Same
9	Dropout	0.5	Same
10	Convolution	256 4*1 stride 1	**128** 4*1 stride 1
11	Dropout	0.5	Same
12	Convolution	68 1*13 stride 1	**37** 1*13 stride 1

Since the decoder output and the target character sequence lengths are of different lengths, we apply the method of CTC loss [19] for segmentation-free end-to-end training. CTC loss is a well-known approach for situations where input and output sequences are misaligned and have variable lengths. Moreover, CTC provides an efficient way to go from probabilities at each time step to the probability of an output sequence.

3.2 Modified LPRNet Model

The reason for LPRNet for not being equally good under Indian context can be the differences in the Indian Number Plates and the Chinese Number Plates as follows:

- Indian License plates have spacing in between digits and alphabets.
- Indian License plates have around 10 characters, unlike Chinese license plates which have around 7 characters as in Fig. 1.
- Each character in the Indian License plate can only belong to one of 36 classes (A–Z or 0–9), whereas there are 68 classes in the case of Chinese License plates.

Fig. 1 Chinese number plate different from Indian number plate

In the Design of Experiments (DOE) for attaining better accuracy for Indian scenario, the following changes were incorporated and results were tested experimentally:

- Input Image size is changed from 96 to 128 as Indian License plates have 10 characters, unlike Chinese License Plates along with spacing between alphabets and digits.
- Convolution Filters in layer 2 is changed from 64 to 96 and layer 4 from 128 to 192.
- Layer 6 is removed.
- Convolution filters in layer 10 are reduced to 128 as they will be mapped to 37 classes instead of 68 classes.
- Convolution layer 12 will have 37 filters instead of 68 which is as per Indian standards [11].

3.3 Post Processing

It was observed that the result from the LPRNet model had few deterministic errors like detecting zero (0) as the letter 'O' and one (1) as the letter 'I' and vice versa. To eliminate such errors, we applied pattern matching based on the Indian standard rules.

P1 P2 P3 P4

- Where P1 is matched to the list of State or Union territory
- P2 should be a 2-digit number, specific to a Regional Transport Office (RTO)
- P4 should be a 4-digit number
- P3 was kept optional. As it is one or two characters added when P4 goes out.

Special cases of Delhi, Gujarat, and Bihar were also considered in pattern matching as per [11].

4 Experiment Setup

In this section, we will explain about dataset creation, labelling and training hyperparameters. We prepared labelled data as per Indian standards.

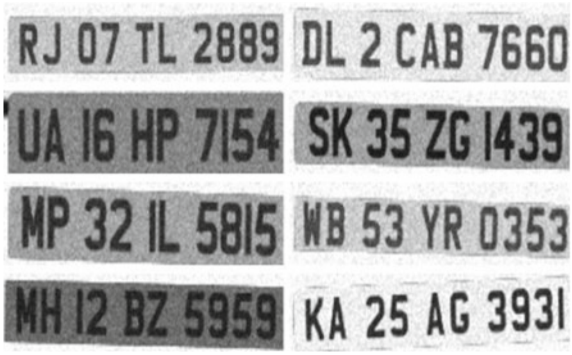


Fig. 2 Synthetic data generated as per Indian Standards

4.1 Dataset

Due to the unavailability of real training data and save labelling time we generated synthetic data with characteristics similar to real number plates and complying Indian standards.

4.1.1 Synthetic dataset

We created a synthetic dataset as in Fig. 2 of 10,000 samples as per Indian standards. The font we have chosen is UKNumberplate.ttf. The pixel size of the plate is kept 200 * 40 pixels. Special cases of Delhi, Gujarat and Bihar are also considered as per RTO guidelines [11] while producing synthetic dataset. A brief set of guidelines that was followed while creating the synthetic dataset is as follows:

- First two characters of a license plate represent a specific Indian state. So, they were randomly chosen from a list where each item is of two-character length and corresponds to 28 states and 9 union territories.
- Next two-digit number is randomly generated. In the special case of Delhi, Gujarat, and Bihar, if there is a zero in the first position then it is omitted. Next string of characters (variable length from 0 to 2) was randomly generated and in the special case of Delhi, Gujarat and Bihar a character is randomly selected out of (C, S, P, Y, V, T, R) and prepended to the said string as per RTO rules [11].
- The last four digits were randomly generated.

  We then augmented the data by using random rotation, scaling, and translation.

4.1.2 Real dataset

Real data is collected by placing a DSLR camera at C-DAC, Pune. Three-hour video data is captured at 30 frames per second (FPS). Variations in terms of angle, distance, of the camera, are incorporated. 4,000 License plates images were detected which belonged to 100 vehicles. Letter distribution of our License plate data is in Fig. 4 (Fig. 3).

License plates were segmented using YOLO V3 [9, 10] which is trained on the open dataset by Archaf Khazri [20] on NVIDIA P5000 accelerator card. The web-based application is designed to label the segmented license plate images in parallel by multiple users. The sample dataset is in Fig. 3.

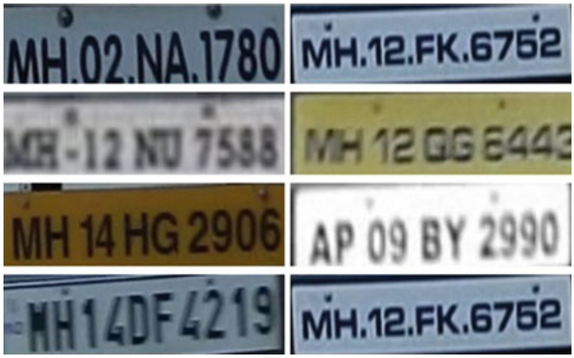


Fig. 3 Real data collected from C-DAC premise

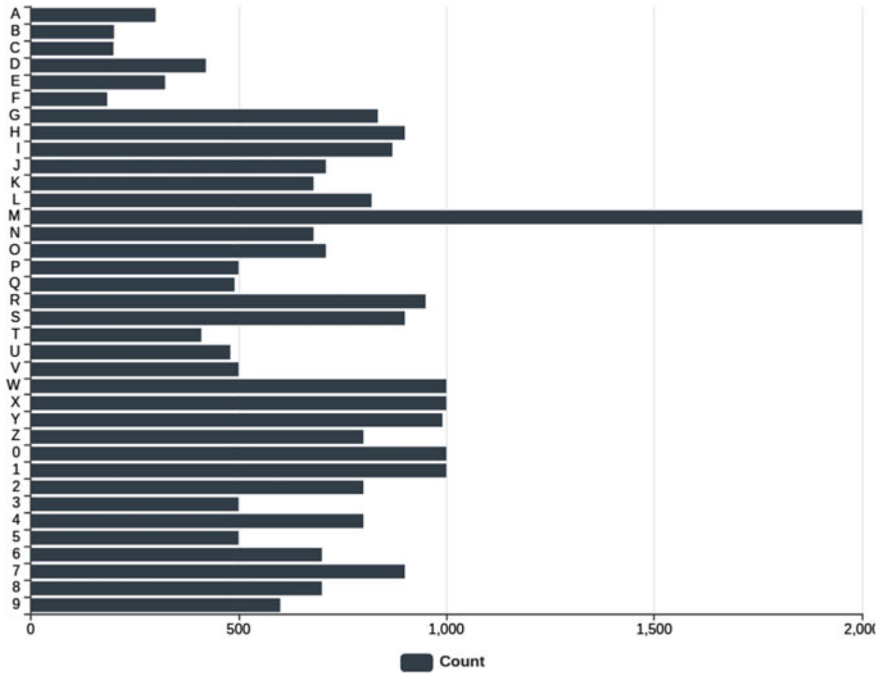


Fig. 4 Letter distribution of real data

Table 3 Hyperparameters for training

Parameters	Synthetic data	Transfer learning (real data)
Data	Training 10K, validation, 2k, testing 2k	Training 3.2K, validation 0.4k, testing 4k
Learning rate	0.003, decay rate = 0.9 after 5 k iterations	0.003, decay rate = 0.9 after 5k iterations
Batch size	30	4
Iterations per epochs	300	200
Total epochs	200	50

4.2 Training Details

To train our model we used C-DAC PARAM Shavak DL GPU System [23, 26], which is equipped with the latest × 86 based Intel Xeon server series processor along with one Nvidia Quadro P5000 accelerator card. It has 64 GB of RAM and 8 TB of secondary storage.

Hyperparameters: Hyperparameter selection is based on multiple experiments performed over GPU in TensorFlow [21] framework and plotting the results in Tensorboard [22]. The hyperparameters are in Table 3.

First, the modified LPRNet was trained on Synthetic data and then we performed transfer learning with real data.

5 Results

In this section, we present the results of the experiments carried out. All the plots are built during runtime using Tensorboard [22].

We used synthetic data to modify the LPRNet model and achieved 93% accuracy on Synthetic number plates with our modified network.

5.1 Synthetic Data Results

We performed training on synthetic data as per Hyperparameters in Table 3. We achieved 93% number plate accuracy on the synthetic dataset with modified LPRNet as compared to 85% accuracy with the original LPRNet as in Fig. 5. Also, the character level accuracy is above 99% with modified LPRNet as in Fig. 6.

From Fig. 8 it can be seen that there is a dip in 30K iteration, which is an indicator of the best fit model. Figure 7 indicates that if we go beyond 30K iteration we will be

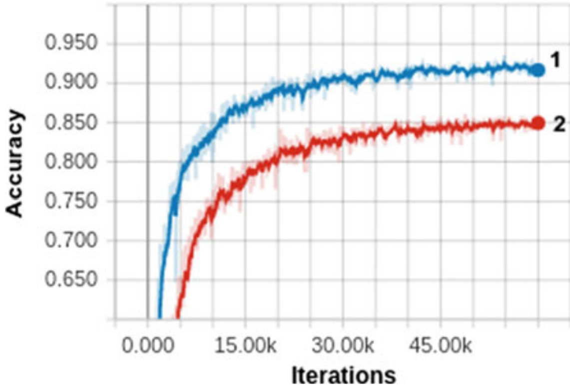


Fig. 5 Test plate accuracy on synthetic data: #1 is modified LPRNet whereas #2 is original LPRNet

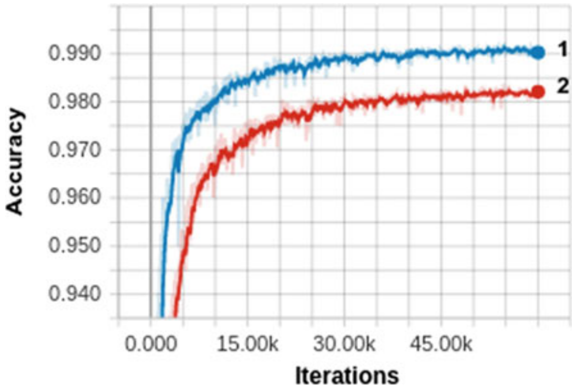


Fig. 6 Test character accuracy on synthetic data: #1 is modified LPRNet whereas #2 is original LPRNet

overfitting our model. We saved this model on iteration 30K and further performed transfer learning with real data as in the next section.

5.2 *Transfer Learning on Real Data*

We achieved 95% number plate accuracy on the real dataset with modified LPRNet as in Fig. 9. Also, the character level accuracy is above 99% with modified LPRNet in Fig. 10. Figures 11 and 12 helped us in selecting the best fit model.

5.3 *Confusion Matrix and Test Accuracy*

Test accuracy on real data is summarized in Table 4 for real and synthetic data with best fit models. The post-processing technique of pattern matching reduced the False

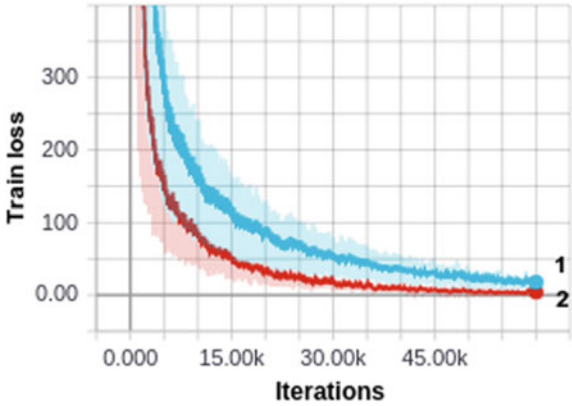


Fig. 7 Train loss on synthetic data: #2 is modified LPRNet

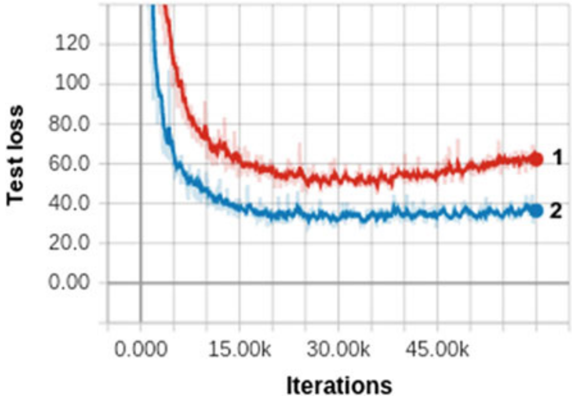


Fig. 8 Test loss on synthetic data: #2 is modified LPRNet, #1 is original LPRNet

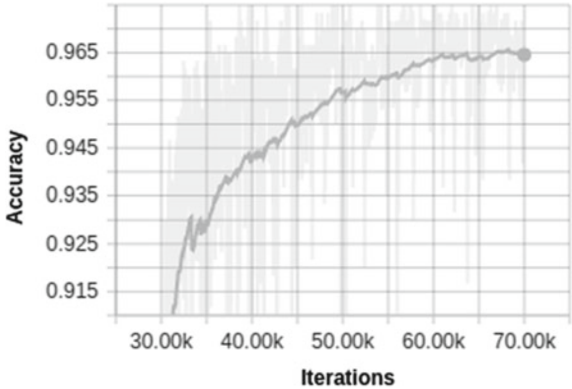


Fig. 9 Test plate accuracy on real data

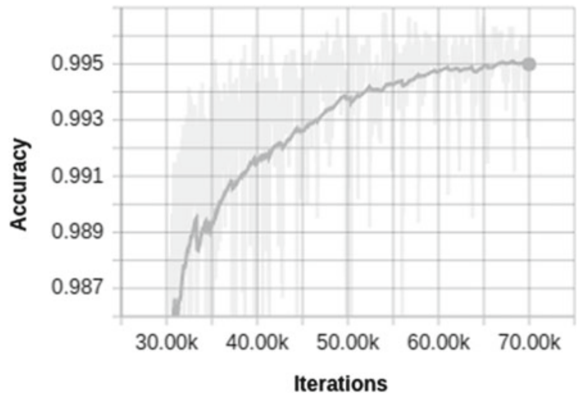


Fig. 10 Test character accuracy on real data

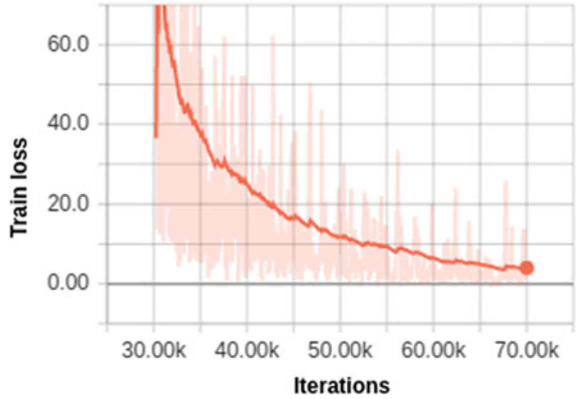


Fig. 11 Train loss on real data

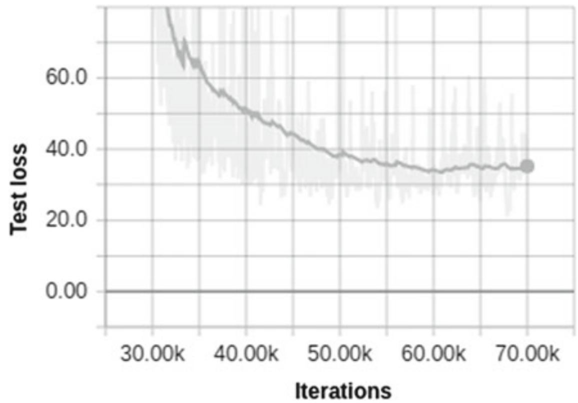


Fig. 12 Test loss on real data

Table 4 Test accuracy on synthetic and real data

	Accuracy synthetic data	Accuracy on real data
Character level	99	99
Number plate	93	95

Table 5 Confusion matrix for character-level accuracy for similar characters on real data

	I	1 (one)	O	0 (zero)
I	43	9	0	0
1 (one)	23	87	0	0
O	0	0	21	13
0 (zero)	0	0	43	84

positives and False negatives with similar-looking characters, the confusion matrix for the same is in Table 5.

5.4 Sample Results

Though our model achieved the state-of-the-art results even with blurred images as in Fig. 13 there were cases when our model predicted some incorrect results also as in Fig. 14.

6 Conclusion and Future Work

In this paper, we have customized LPRNet based on experiments to achieve the state of the art 95% accuracy for Indian Number Plate data. For future work, we can work on an end to end pipeline of number plate detection as well as recognition using only CNN networks, which will use our modified LPRNet for recognition. This complete pipeline will be time-efficient and have better accuracy.

Fig. 13 Sample results on number plates where it is correctly recognized

Fig. 14 Sample results on number plates where it is incorrectly recognized. The ground truth is shown in parenthesis

References

1. Srivastava D, Bhambhu L (2010) Data classification using support vector machine. J Theor Appl Inf Technol 12:1–7
2. Cunningham P, Delany S (2007) k-Nearest neighbour classifiers. Mult Classif Syst
3. Cohen G, Afshar S, Tapson J, van Schaik A (2017) EMNIST: an extension of MNIST to handwritten letters

4. Cheang T, Chong YS, Tay YH (2017). Segmentation-free vehicle license plate recognition using ConvNet-RNN
5. Li W, Cao L, Zhao D, Cui X, Yang J (2013) CRNN: integrating classification rules into neural network. In: The 2013 international joint conference on neural networks (IJCNN), Dallas, TX, 2013, pp 1–8
6. O'Shea K, Nash R (2015) An introduction to convolutional neural networks. ArXiv e-prints
7. Sherstinsky A (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Phys D: Nonlinear Phenom 404:132306. Crossref. Web
8. Zherzdev S, Gruzdev A (2018) LPRNet: license plate recognition via deep neural networks
9. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection, pp 779–788. https://doi.org/10.1109/CVPR.2016.91
10. Redmon J, Farhadi A (2018) YOLOv3: an incremental improvement
11. http://www.rto.org.in/vehicle-registration-plates-of-india/format-of-number-plates.htm
12. Anagnostopoulos C-N, Anagnostopoulos I, Ioannis P, Loumos V, Kayafas E (2008) License plate recognition from still images and video sequences: a survey. IEEE Trans Intell Transp Syst 9:377–391. https://doi.org/10.1109/TITS.2008.922938
13. Baldominos A, Sáez Y, Isasi P (2019) A survey of handwritten character recognition with MNIST and EMNIST. Appl Sci 2019:3169. https://doi.org/10.3390/app9153169
14. Jain V, Sasindran Z, Rajagopal A, Biswas S, Bharadwaj H, Ramakrishnan K (2016) Deep automatic license plate recognition system, pp 1–8. https://doi.org/10.1145/3009977.3010052
15. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5mb model size. arXiv:1602.07360 [cs]
16. Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2016) Inception-v4, Inception-ResNet and the impact of residual connections on learning. arXiv:1602.07261 [cs]
17. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2014) Going deeper with convolutions. arXiv:1409.4842 [cs]
18. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2015) Rethinking the inception architecture for computer vision. arXiv:1512.00567 [cs]
19. Hannun A (2017) Sequence modelling with ctc. Distill. https://distill.pub/2017/ctc
20. https://www.kaggle.com/achrafkhazri/labeled-licence-plates-dataset
21. Poots J (2019) Tensorflow: past the basics
22. Jensen C (2020) HACK 1-Basic TensorBoard Visualization
23. Agrawal S, Das S, Valmiki M, Wandhekar S, Moona R (2017) A case for PARAM Shavak: ready-to-use and affordable supercomputing solution. In: 2017 international conference on high performance computing & simulation (HPCS), Genoa, 2017, pp 396–401
24. Laroca R, Severo E, Zanlorensi L, Soares de Oliveira L, Gonçalves G, Schwartz W, Menotti D (2018) A robust real-time automatic license plate recognition based on the YOLO detector. https://doi.org/10.1109/IJCNN.2018.8489629
25. Laroca R, Zanlorensi L, Gonçalves G, Todt E, Schwartz W, Menotti D (2019) An efficient and layout-independent automatic license plate recognition system based on the YOLO detector
26. https://www.cdac.in/index.aspx?id=lu_param_shavakDL_launch

An Improved Machine Learning Prediction Model for Diabetes

Aayushi Bansal and Anita Singhrova

Abstract Diabetes is a chronic disease which if not treated on time may cause serious complications. Machine learning has been very useful in the diagnosis of diabetes thus saving lives of many. The PIMA diabetes dataset used in the research work consisted of many missing and outlier values which need to be treated. According to the medical practitioners the values of insulin, skin thickness, BMI can never be zero and both diabetic and non diabetic will have different range of values for various parameters therefore a technique to replace the missing and outlier values with respective medians is proposed. After preprocessing the dataset, PCA algorithm is applied for dimensionality reduction and thereafter two similar clusters are formed using the kmeans clustering algorithms. The predicted output of the various classifiers is then given as input to the voting classifier which uses the soft voting technique to make the final prediction. The accuracy of the proposed technique comes out to be 98.70% and also an improvement in precision, recall and f1 score has been noted. The mean squared error of the proposed algorithm is very less as compared to the previous classifiers.

Keywords Machine learning · Diabetes · PCA · Kmeans · Preprocessing · Missing values · Outliers · Accuracy

1 Introduction

Machine learning is having its applications in various sectors these days. Machine learning techniques along with the data mining techniques have proved to be very successful in predicting various diseases. Data mining helps in finding useful patterns from the data while using those patterns machine learning is able to draw conclusions and make deductions. The data should be in understandable format for carrying out analysis [1]. The machine learning techniques are used for prediction, classification and also in giving suggestions. It helps in the early prediction of cancer, diabetes, eye disorders, brain tumors and thus helping in the effective and timely treatment. If

A. Bansal (✉) · A. Singhrova
Deenbandhu Chhotu Ram University of Science and Technology, Murthal, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_13

the diseases are diagnosed at the early stages then its chances of killing the human being gets less. Early diagnoses lead to early treatment. ML is generally of two types i.e. supervised and unsupervised. In supervised a labeled dataset is given as input to the model while in unsupervised the dataset is not labeled. Supervised technique is used for diabetes prediction since the diabetes dataset is a labeled dataset. The model can be trained on the dataset and then that trained model can be used to make the prediction on the new data [2].

The main aim of the paper is to make a good model which can accurately predict the diabetes disease given a dataset as the input. Diabetes is a fatal disease which can damage various organs of human body. It cannot be transmitted from one person to other through touch but it can be passed genetically. Some of its medical indications are more thirst, repeated urination, high blood sugar levels etc. Thus it is focused to be diagnosed as early as possible and in accurate way so that person can start treatment in early stages and cure it [3]. It is not possible to cure the disease with oral medicines only [4]. Due to insufficient insulin production the blood sugar levels of a diabetic person reach very high. Available researches on blood sugar levels in diabetic patients rely on a constant glucose screening. Insulin, a hormone, produced by the pancreas, is responsible for controlling the blood sugar levels in human beings. Hence the decrease in the insulin in the body leads to fluctuations in the level of blood glucose which in turn causes diabetes [5]. Diabetes is of mainly of two types i.e. Type 1 and Type 2. Type 1 also known as juvenile diabetes begins in the childhood as the pancreas is not able to produce sufficient insulin. Type 2 is also known as gestational diabetes as it can also happen after pregnancy in women [6]. The sugar and carbohydrate produced by food is to be used by the body for energy creation but due to malfunctioning of pancreas it is not able to convert food to useful energy.

Pima Indian diabetes dataset is used for the research work. The dataset consists of records of women of age above 21 years. The dataset is labeled and consists of 768 records showing people having diabetes as 1 and not having as 0. The dataset consisted of many missing values such as insulin = 0, skin thickness = 0 etc. which is not practically possible hence data preprocessing is done to improve the prediction accuracy. The missing values and the outlier values in the dataset were statistically analyzed and were replaced with new values. The dataset consists of many dimensions hence, PCA (Principle component Analysis) was applied on the dataset to reduce the dimensions. The dataset having many dimensions is reduced to a small dataset having fewer dimensions. Kmeans algorithm was used to cluster the similar type of data before giving it to the model for training and the value predicted by kmeans is used as the target vector in the model training phase. Kmeans algorithm created two clusters of similar data. A voting classifier is further used to make the prediction on the given information. The voting classifier uses a soft voting technique to make the final prediction and finally good accuracy is achieved by the model. The paper is organized as follows: Sect. 2 discusses the literature review, Sect. 3 discusses the methodology, Sect. 4 discusses the results, and Sect. 5 is the future scope and conclusion.

2 Literature Review

Deeraj Shetty et al. (2017) developed an Intelligent Diabetes Disease Prediction model with Bayesian and KNN (K-Nearest Neighbor) algorithms to extract features from dataset and analyzed performance on those [7].

Santosh Rani et al. (2018) used ML algorithms to extract and process data and achieved features that were good at predicting diabetes in a person. The information could be used for future analysis as well [8]. Vrushali R. Balpande et al. (2017) designed a data mining model to predict the risk level of diabetes on various body organs [9]. Girdhar Gopal Ladha et al. (2018) performed a computation analysis that provided the parametric study along with the information of the features and the technique to implement the classification model [10]. The main aim of almost all research works was to detect diabetes disease in early stage. This work could be used to develop an efficient diabetes prediction model in nearby future. Messan Komi et al. (2017) stated that improper rise in the level of blood sugar was the main cause of diabetes disease [11]. Five data mining algorithms were used for predicting diabetes disease in early stage. The tested results showed that ANN (Artificial Neural Network) approach showed highest accuracy than the other ones. Xu et al. (2019) [12] recommended model that was based on Weighted Feature Selection of Random Forest (RF-WFS) and XGBoost (extreme gradient boosting) Ensemble Classification model. The recommended model used RF-WFS and XGBoost (extreme gradient boosting) classifiers for selecting best features. B. V. Baiju et al. (2019) provided a comprehensive review of data mining based disease prediction algorithms [13]. The recommended technique carried out the preprocessing of input data set and eliminated the noisy features and also estimated disease influence measure (DIM) on the basis of input data point attributes in the next step. The recommended technique predicted diabetes disease on the basis of DIM value. Wenqian Chen et al. (2017) stated that the growth in the popularity of data mining algorithms was noticed in healthcare domain because of their classification potential [14]. A hybrid prediction model was designed by use of K-means clustering algorithm for reducing data. Also, a classification model called J48 was used for carrying out classification. A higher accuracy was obtained. Girdhar Gopal Ladha et al. (2017) stated that one among eleven people all over the world was diagnosed with diabetes disease according to a survey performed in the year 2015 [15]. The major purpose of this work was to discover appropriate techniques for diagnosing this disease in early stage. This work also reviewed different data mining algorithms for detecting diabetes disease on the basis of different signs. Han Wu et al. (2018) suggested the use of a new prediction model type 2 diabetes mellitus (T2DM) [16]. This model was divided into two segments on the basis of a sequence of preprocessing events. These segments were identified as improved K-means algorithm and the logistic regression algorithm. The achieved results depicted that the recommended model showed 3.04% more prediction accuracy than other existing models. The recommended model performed well in both cases. Hence, this model was very efficient in the real time handling of diabetes disease. Changsheng Zhu et al. (2019) recommended a new model based on data mining to diagnose and predict diabetes disease in early stage [17]. K-means was a simple clustering algorithm. Improving accuracy rate of k-means clustering and logistic regression was the main aim of this work. The recommended model included three algorithms. These algorithms were identified as PCA (principal component analysis), k- means and logistic regression. The tested outcomes depicted that PCA algorithm made improvements in the k- means approach. This k-means algorithm showed 25% improvement in accuracy rate while logistic regression showed 1.98% more accuracy rate.

3 Methodology

This section describes the methodology adopted for the research work. The work is carried out in two phases, step 1 being preprocessing and step 2 being the application of PCA, Kmeans and voting classifier. The explanation of following is provided in this section: The dataset, data preprocessing which involves replacing of missing and outlier values in the dataset, dimensionality reduction using PCA, clustering of similar data using KMeans algorithm and finally using a voting classifier for making the prediction. Preprocessing is done on the data set to deal with the missing and outlier values. Since the dataset was small so it was not feasible to remove the rows having outliers and missing values thus a method to replace the missing and outlier values was proposed. The PCA algorithm was applied to reduce the dimensions so that the algorithm is clearly able to find the relations by reducing the highly correlated variables. A new set of variables consisting of the principle components is created which has the highest variation among all the other variables [18].

3.1 Toolkit

Python programming language is used for implementation of the algorithm as it is the best programming language to be used for machine learning after R language. Anaconda distribution is a free open source distribution which provides the Jupyter Notebook IDE for running the python scripts. Anaconda is generally used because it has many libraries for machine learning preinstalled in it. Hence it becomes easier for the user to use it. Using this distribution it was easy to design (code) the proposed algorithm. It consists of all the packages for data cleaning and plotting on the graphs.

3.2 Dataset Description

The diabetes dataset used was the Pima Indian Diabetes dataset which was downloaded from the Kaggle website and is also freely available on UCI Repository. It is an open source dataset. It consists of records of female patients who are living in Arizona (USA). It consists of 768 records out of which 500 records were non diabetic patient and 268 were diabetic patients. It is comprised of 8 features and 1 column is for the target which represent the diabetic patient as 1 and non diabetic patient as 0.

3.3 Proposed Approach

The proposed approach consisting of two steps is explained. The preprocessing of dataset was performed and then the pca, kmeans and voting classifier was applied to get a model with enhanced accuracy.

3.3.1 Data Preprocessing

Preprocessing is the essential step of data analysis process. It is the technique of converting raw data into clean and usable form. It is an important step as it involves removal of irrelevant data. The raw data consists of many outliers and missing values that need to be treated because it leads to inaccurate results. Data preprocessing enhances the data efficiency. It generally involves cleaning, transformation, integration and reduction of data [19]. For making the dataset more productive preprocessing of the dataset was done using the Anaconda libraries. When the dataset was closely analyzed using box plot it was found that it consists of many missing values in zero form and outlier values. As consulted by the doctors and medical practitioners it was known that insulin value, skin thickness, blood pressure of a person can never be zero. Also according to the professional dieticians the people showing positive for diabetes have different values for insulin, glucose, diabetes pedigree function than the ones who show negative results. Since the dataset consisted of many rows showing these values to be zero, and replacing the missing values with the mean of the whole dataset is not a suitable task. Thus a technique to replace these zero values with some value was proposed so that the model becomes efficient.

The proposed approach is to divide the whole dataset into two parts i.e. one part consisting of people having diabetes and one part of people not having diabetes. The division of the dataset and its division are as shown in Fig. 1. Thereafter, for the two different parts of the dataset, box plots for different features were plotted. The box plot also known as whisker plot is plotted by using the python libraries. The outlier values were spotted in the box plots. The minimum and maximum values in the dataset were found and using the upper and lower quartile values the interquartile range was calculated. The interquartile range thus helped in finding the outliers. The data below (lower range − 1.5 * interquartile range) and the data above (upper range + 1.5 * interquartile range) are considered to be outliers [20]. After replacing the missing and outliers of both the dataset the two datasets are merged to form a single dataset which could be used for further processing.

Fig. 1 Dataset preprocessing

3.3.2 Step 2. Algorithm Design

After preprocessing of the dataset, PCA is applied on the dataset which gives a reduced feature space which less highly correlated variables and a huge variability among variables. The reduced feature space dataset is provided as an input to the kmeans clustering algorithm. All the data having similar features is clustered into groups. The number of clusters formed is two since we have two outcomes i.e. 1 and 0. This clustered data when given to the voting classifier gives the final prediction of the testing data. The proposed algorithm is shown in Fig. 2. The brief description of the used techniques is given further.

PCA Algorithm

PCA is a linear technique. New features are created by linear combination of existing features. PCA is a technique used for reducing the dimensions of the data [21]. The datasets have many dimensions which are correlated to each other which is difficult to understand. So PCA help in reducing the correlation by reducing the feature space. As an output it gives a dataset with fewer features and with high variability. The huge dimensional data consists of many features of negligible importance. This is converted into dataset with fewer features. The extracted features are called as principle components of the data.

PCA helps in presenting the information in a few plots [22]. It is also called projection method as it converts data from p to k dimensions where (k < p). It helps in finding out the hidden patterns from data by removing irrelevant detail. The many correlated variables are converted to uncorrelated values. The first principle component shows the highest variability in data [23]. Then with the help of these the eigen vectors and eigen values are calculated which are helpful in getting the principle components. Hence this data is further used to be clustered by the Kmeans algorithm.

3.4 Kmeans Clustering Algorithm

Kmeans is a clustering algorithm and is used to cluster large sets of data. In this all the data with same features are clustered. It is a standard algorithm for creating the clusters. Squared error function is the foundation of this algorithm. This simple algorithm was proposed by MacQueen for clustering purpose. The clusters generated by this algorithm are independent. All the data in one cluster is similar. The data points in the clustered are grouped by using the Euclidean distance. The algorithm starts by choosing the initial cluster centers. The algorithm consists of two steps. In the first step randomly some cluster centers are chosen. The points are clustered using these cluster centers. This algorithm solely depends on the initial cluster centers formed. In the next phase, new cluster centers are chosen by using the mean value of the previous cluster and the points are then reassigned the new clusters. The same continues until and unless the whole data is properly clustered. This algorithm provides good efficiency (Fig. 2).

The clusters so formed after applying the pca and kmeans clustering algorithm are shown in Fig. 3. The data in cluster with blue dots represent the people not having diabetes and orange dot represent the people having diabetes. The two principle components are shown along the x and y axis. The data is properly clustered by applying the above techniques. Now this data is further given to the different classifiers which further by the use of voting technique make the final prediction. A soft voting technique is used by the voting classifier where the weights are assigned to the outputs of various classifiers on the basis of the probability vales they give.

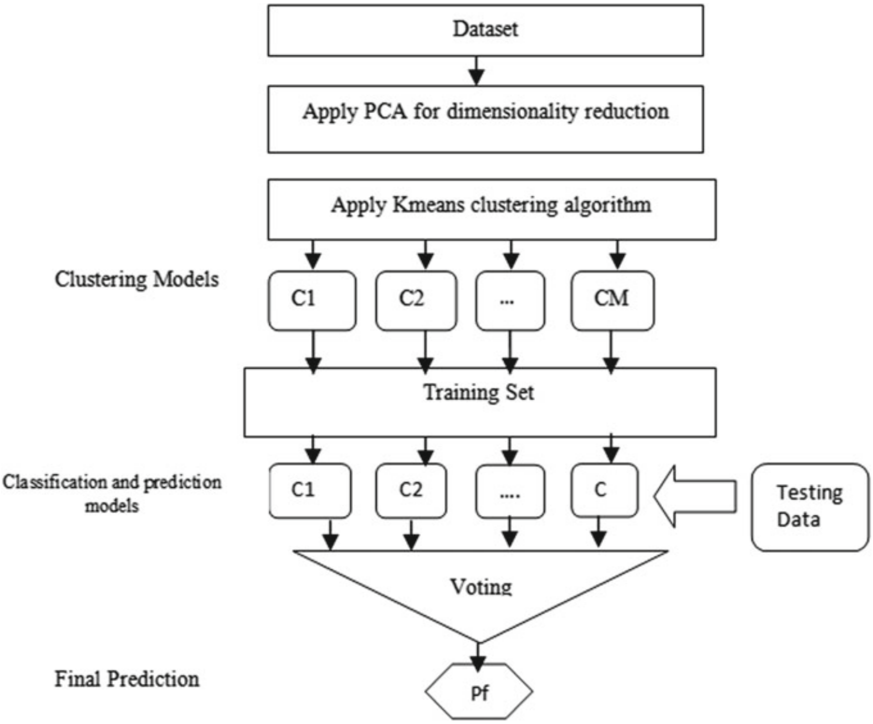
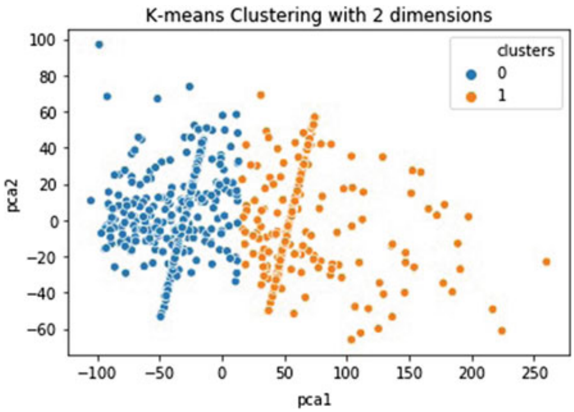


Fig. 2 Prediction model designed for accuracy

Fig. 3 Separation of the datasets by kmeans clustering algorithm



3.5 Voting Classifier

A voting classifier is a machine learning model that takes as input a number of models and gives an output based on the highest probability of selected outcome. The output of many classifiers is aggregated into the voting classifier and then the final output is chosen by taking the majority voting of all the classifiers. Voting classifier works on the basis of two voting mechanisms hard voting and soft voting. In hard voting, the output predicted class is the class with the highest predicted vote. While in the soft voting technique, the output class is the prediction based on the average of probability given to that class. Voting classifier is said to perform much better than the individual classifiers since it aggregates the result of various classifiers and then gives its output which is higher as compared to a single classifier. Soft voting is done by assigning weights to the classifiers on the basis of their accuracy. The classifiers having higher accuracy are given higher weights as compared to the ones with lower accuracy.

4 Results

The paper focused on improving the accuracy of diabetes prediction as it is of immense use these days to design models which are much more accurate than the previous models. It is very important to diagnose a person positive for diabetes as accurately and quickly as possible. This designed model has proven to be good in the prediction of diabetic patients accurately. It gives an accuracy of 98% in predicting diabetes. Earlier much attention was not given to replace the missing values and outliers in the dataset. Dividing the dataset into two parts, one with having diabetes and the other of people of not having diabetes. Then the two datasets were closely analyzed using statistical tools and it was found out that both kinds of people have different readings for various parameters. That is why it was much more feasible to find different medians and then replace the outliers and missing values with respective median values of both datasets. This proved to be quiet useful in clustering the data in an organized way into clusters.

PCA technique was applied to reduce the dimensions of the data as irrelevant features were of no use and also with so many features it becomes impossible to understand the data. The reduced dimensions of the data had only two dimensions which were uncorrelated and had high variability.

The dataset after applying the PCA were fed to the kmeans clustering algorithm. Many different clustering models were created and theses models had data separated into clusters consisting of two clusters. The features were given to the clustering algorithm and the predicted output of this model was used to form the target vector as after clustering it was accurately predicted to which class the tuple belongs to. The model further produced the training dataset which was further given to the other classification models and the output of those models was given to a voting classifier which by using the soft voting technique predicted the output.

Table 1 shows the comparison of the accuracy of the proposed model with other models. With the help of proposed algorithm an accuracy of 98.70% was achieved. The accuracy so achieved is because of the good clustering due to kmeans and also preprocessing. The accuracy is good and the model can now very efficiently predict a diabetic patient. Figure 4 shows the graph representing the comparison in accuracy of the various classifiers with the proposed algorithm.

The comparison of all other parameters is shown in Table 2. The table also shows the execution time of the algorithms. As it is very clearly seen in Fig. 5 graph that the execution time of the proposed algorithm is the highest. The algorithm takes so long time due to many processing steps included in the algorithm. Thus this algorithm is lacking at this parameter. It was seen that the precision rate, recall rate, and f1 score had shown a significant amount of improvement due to the proposed algorithm. A graphical representation of values of these parameters is seen in Fig. 6. And also the model has the least mean squared error among all other models. The error has reduced to large amount due to which our model becomes much more efficient and there are very few chances of wrong prediction by this model. This model gives such low error and high efficiency as it makes the use of a voting classifier which

Table 1 Comparison of accuracy of the various machine learning algorithms and the proposed approach

MLA name	Accuracy (%)
K neighbors classifier (KNC)	75.32
Random forest classifier (RFC)	69.48
Bagging classifier (BC)	71.43
Logistic regression (LoR)	78.03
Gaussian NB (GNB)	70.13
Support vector classifier (SVC)	65.58
Gradient boosting classifier (GBC)	68.83
Linear regression (LR)	73.38
Decision tree classifier (DTC)	70.13
Ada boost classifier (ABC)	68.83
Proposed algorithm (PA)	98.70

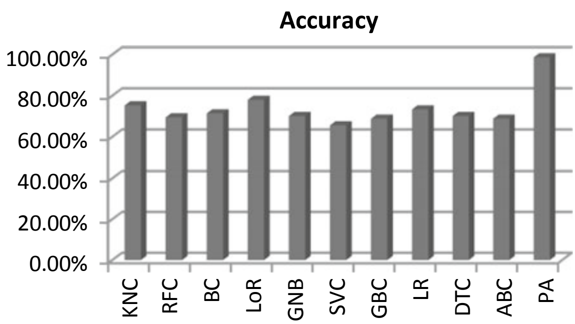
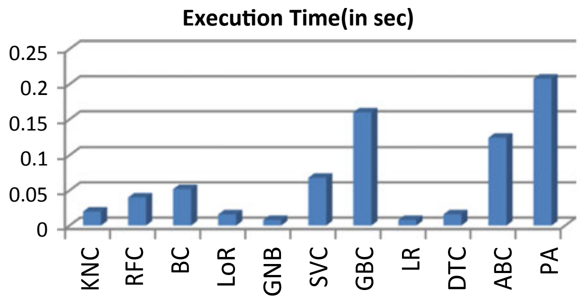


Fig. 4 Comparison of the accuracies of various machine learning algorithms with the proposed algorithm

Table 2 A comparison of the precision, recall, f1 score, mean squared error and execution time of other algorithms with the proposed approach

MLA	KNC	RFC	BC	LoR	GNB	SVC	GBC	LR	DTC	ABC	PA
Precision	0.6923	0.5789	0.6153	0.6585	0.5777	0	0.5555	0.6428	0.5813	0.5510	0.9608
Recall	0.5094	0.4150	0.4528	0.5094	0.4905	0	0.4716	0.5094	0.4716	0.5094	0.9985
F1 score	0.5869	0.4835	0.5217	0.5744	0.5306	0	0.5102	0.5684	0.5050	0.5294	0.9888
MSE	0.2467	0.3051	0.2857	0.2597	0.2987	0.3441	0.3116	0.2662	0.3181	0.3116	0.0129
Execution time (in seconds)	0.0199	0.0399	0.0519	0.0159	0.0079	0.0679	0.1599	0.0079	0.0159	0.1239	0.2079

Fig. 5 Comparison of the execution time of the machine learning algorithms with the execution time of the newly proposed approach



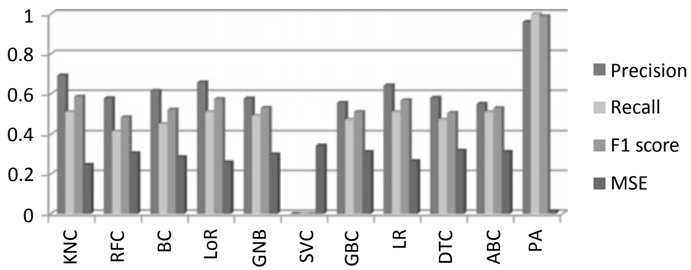


Fig. 6 Comparison of precision, recall, f1 score, mean squared error of the various machine learning algorithms with the proposed algorithm

takes as input the predictions from various classifiers and makes the final decision by combining the results of all.

5 Conclusions and Future Scope

The main motive of the research was to improve the accuracy of diabetes prediction. It was meant to design an algorithm that can very efficiently predict the disease. After carrying out a huge survey on the current topic, a novel algorithm was proposed by the combined application of preprocessing, PCA and kmeans clustering. A productive model giving better prediction accuracy was thus designed. To enhance the efficiency the data was preprocessed and also clustering was done to make it easy for the classification models to predict the correct output. The model was tested for other parameters such as precision, recall, f1 score, mean squared error and it was deduced that an increase in the values of precision, recall and f1 score was seen. The mean squared error had reduced to a much lower value and it is an indication of a much efficient model. The model lacked behind in one parameter which is the execution time. The proposed model took much longer to implement than the other models. And thus it is considered to a limitation of the proposed algorithm. But since the main motive was to improve the accuracy the accuracy had been improved without taking execution time in consideration. The voting classifier produced much accurately than the other models designed till now.

The future scope may include the application of the model in the diagnosis of new and fresh data of patients from the hospital. Also this model can be applied in prediction of other diseases. The model can be further improved by using some other dimensionality reduction techniques and also any other clustering algorithm may be used. Since much focus was not given on the execution time of the model, this parameter can be improved as a future work. A plus point of this model is that it can easily adjust on some datasets easily though their preprocessing technique may be different.

References

1. Amalarethinam DIG, Ph D (2018) Prediction of diabetes mellitus using data mining techniques: a survey
2. Alehegn M, Joshi R, Mulay P (2018) Analysis and prediction of diabetes mellitus using machine learning algorithm. Int J Pure Appl Math 118(9):167–175
3. Mirshahvalad R (2017) Diabetes prediction using ensemble perceptron algorithm, pp 190–194
4. Mercaldo F, Nardone V, Santone A, Nardone V, Santone A (2017) Diabetes mellitus affected Patients classification and diagnosis through machine learning techniques. Procedia Comput Sci 112:2519–2528
5. Zolfaghari R (2012) Diagnosis of diabetes in female population of Pima Indian heritage with ensemble of BP neural network and SVM. IJCEM Int J Comput Eng Manag 15(4):115–121
6. Joshi TN, Chawan PPM (2018) Diabetes prediction using machine learning techniques. IJERA 8(1):9–13
7. Patil N (2017) Diabetes disease prediction using data mining
8. Rani S (2018) Association clustering and time series based data mining in continuous data for diabetes prediction. In: 2018 second international conference on intelligent computing and control systems (ICICCS), pp 1209–1214
9. Balpande VR, Wajgi RD (2017) Diabetes using data mining technique. ICIMIA, pp 576–580
10. Ladha GG (2018) A computation analysis to predict diabetes based on data mining : a review. In: 2018 3rd international conference on communication and electronics systems (ICCES), pp 6–10
11. Komi M, Zhang X (2017) Lf = l SiN (Ili Li), no S Ix, pp 1006–1010
12. Xu Z, Wang Z (2019) A risk prediction model for type 2 diabetes based on weighted feature selection of random forest and XGBoost ensemble classifier. In: 2019 eleventh international conference on advanced computational intelligence (ICACI), pp 278–283
13. Baiju BV, John Aravindhar D (2019) Disease influence measure based diabetic prediction with medical data set using data mining. In: Proceedings—2019 1st international conference on innovations in information and communication technology (ICIICT), pp 1–6
14. Chen W, Chen S, Zhang H (2017) A hybrid prediction model for type 2 diabetes using K-means and decision tree, no 61272399
15. Ladha GG, Pippal RKS (2018) A review and analysis on data mining methods to predict diabetes. In: Proceedings—2017 7th international conference on communication systems and network technologies (CSNT), pp 334–337
16. Wu H, Yang S, Huang Z, He J, Wang X (2018) Type 2 diabetes mellitus prediction model based on data mining. Inform Med Unlocked
17. Zhu C, Uwa C, Feng W (2019) Informatics in medicine unlocked improved logistic regression model for diabetes prediction by integrating PCA and K-means techniques. Inform Med Unlocked 100179

18. Wang Q, Gao Q, Gao X, Nie F (2017) Angle principal component analysis. In: IJCAI international joint conference on artificial intelligence (IJCAI-17), August 2018, pp 2936–2942
19. Alasadi SA, Bhaya WS (2017) Review of data preprocessing techniques in data mining. J Eng Appl Sci 12(16):4102–4107
20. Thirumalai C, Vignesh M, Balaji R (2017) Data analysis using box and whisker plot for lung cancer. In: 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), March 2017, pp 1–6
21. Sumathi D, Selvaraj P, Burugari VK (2019) Principal component analysis for dimensionality reduction for animal classification based on LR. Int J Innov Technol Explor Eng 8(10):1118–1123
22. Mathur B, Kaushik M (2018) Data analysis utilizing principal component analysis. Int J Eng Res Technol 11(2):333–348
23. Paul LC, Al A Suman, Sultan N (2013) Methodological analysis of principal component analysis (PCA) method. IJCEM Int J Comput Eng Manag 16(2):2230–7893

University Recommendation System for Higher Studies in USA

Aishwarya Nalawade and Bhavana Tiple

Abstract With the increasing number of graduates seeking to pursue higher education, getting a student's admission into the university becomes more challenging. New graduate students are typically unaware of the postgraduate admission criteria and procedures and may spend a substantial amount of time obtaining guidance from consulting firms to help them define their admission opportunities. But this methodology can be biased and misleading considering the small number of universities that a human consultant may consider. Thus, a machine learning approach was built in this paper to automatically predict the prospects for postgraduate admissions, enabling graduates to recognise and target the most appropriate universities for their profile. The University Recommendation System for Higher Studies is a system framework that prescribes universities for students or users of the system who recommend universities to students looking for education for their higher studies. The main parameters used to recommend university are CGPA percentages, GRE Score, TOFEL Score, university rank, etc. This system uses machine learning algorithms and compares the accuracy of the algorithms.

Keywords Recommendation system · Higher study abroad · Data mining · K Nearest Neighbor

1 Introduction

The system collects data about the elements. They combine curves and profiles to encourage the user to dwell on the right choices regarding elements, personalities, policies and governance. For a long time, electronic and page availability has been expanding rapidly, and analysts have relied heavily on material to separate the inevitable data for better teaching. Therefore, the recommendation procedures are well known for assisting different dynamic settings. Because there are so many

A. Nalawade (✉) · B. Tiple
School of Computer Engineering and Technology, MIT World Peace University, Pune, India

B. Tiple
e-mail: bhavana.tiple@mitpune.edu.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_14

alternatives to universities, graduates have to invest a ton of energy to investigate subtleties, and they don't do it appropriately. Graduates need a system that recognizes their priorities and proposes the right university. The most confusing factor for graduates is the decision of the university. This system helps undergraduate students choose which universities to consider.

1.1 Problem Statement

The college recommendation policy plays an important role, as university studies require a lot of study of the magnitude of the variables to accommodate future perspectives and Looking for a good university. It is a difficult task for a graduate to pursue his/her higher studies. Graduates are searching at various angles, including university campuses, faculty, university cultural exercises and university foundations, and university surveys also seek further validation of subtleties. Depending on the field, best assessments, area and past rate, the college recommendation system recommends graduating to the best universities. This is important because it reduces manual work and can be robotized with the help of software.

1.2 Scope of the Project

This system is recommended by colleges, which evaluate students' chances of enrolling in their higher education, their chances of progress, or factors such as GRE, TOPEL or grade point. Other factors are seal, which should be considered as a research paper or internship. This system is useful for final year students.

The system is used by the dataset we use and we try to get as much accuracy as possible by comparing the number of algorithms such as K-Means, Knowledge Base, KNN, Regression and SVM. Check the algorithms and which dataset is best. Available systems do not consider many parameters, so colleges that recommend this system can easily help students and check where they are. By using different machine learning algorithms with good recommendation.

2 Literature Survey

Many students who wish to graduate want to complete their studies, preparing for the next grade, which is the master's degree. Many of them may wonder about the basic requirements for admission to universities and universities where they can be accepted according to their need [1]. The books contain a number of studies that conduct statistical research on adoption decisions. For example, the authors in [2] introduced a specialist program called PASS, in which psychological recording is used to predict the ability of high school students in Greece to enter national examinations. The authors in [3] used a hypothetical model to assess admission policies and standards based on factors such as GPA school, ACT school, residency race. The limitations of this study include not looking at other important factors such as past work experience, technical documentation. The authors of this paper in [4] analyzed the current state of acceptance by predicting student consent behavior using ML data analysis and strategies. They have used the aphorism process to analyze the behavior of students who wish to enroll in a particular college. They also use the Näıve Bayes algorithm, which helps students choose a curriculum and practice entry. In their program, they have a test for students who want to be accepted, and depending on their performance, they recommend a course that uses the Naive Bayes algorithm for students. But human intervention was required to make a final decision on the situation.

3 Our University Recommendation System

We represent our graduate school recommendation system Universities/colleges for applicants Those are the ones. The profile-based system displays a list of universities based on students' GRE verbal score, GRE quants score TOEFL score and CGPA. It will calculate the average and that average will be converted into the range. So first we will calculate average for each score that is GRE verbal, GRE quants and TOEFL score. So the formula for GRE verbal and GRE quants is (1),

$$\text{GREV}/\text{GREQ} = ((\text{score} - 130) * 100)/40 + 0 \tag{1}$$

So 130 is the minimum score of GRE verbal or GRE quants, 100 is the maximum range, 0 is the minimum range, score is the value of GRE verbal or GRE quants and 40 is the difference between the maximum and the minimum score that is 170 and 130 respectively. Same goes for TOEFL score is (2),

$$\text{TOEFL} = ((\text{score} - 0) * 100)/120 + 20 \tag{2}$$

So here 0 is the minimum score, 100 is the maximum range, 20 is the minimum range, score is the TOEFL. Score of the student is 120 and the difference between maximum score and minimum score that is 120 and 0 respectively. Now after calculating average of each score we can now obtain the range by formula given below is (3),

$$\text{Range} = (\text{GREQ} * 2 + \text{GREV} * 1.5 + \text{TOEFL} * 0.4 + \text{prcnt} * 1)/3.9 \tag{3}$$

So here 2, 1.5, 0.4 and 1 are the weights which are assumed. More weightage is given to GRE quants because normally universities gives first priority to GRE quants score. 3.9 is the addition of all the weights assumed. Now this range will be compared

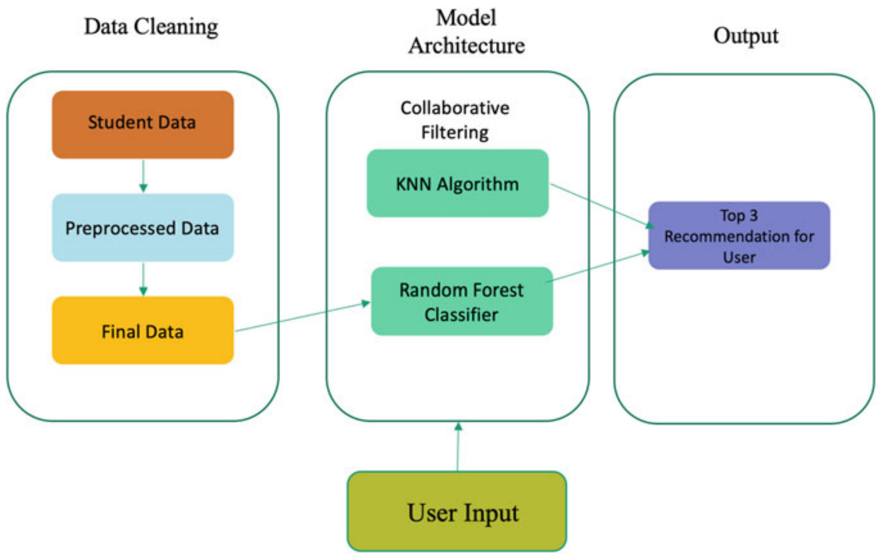


Fig. 1 System architecture

to the ranges of the universities. On the basis of that the list of universities will be listed with the user and system ratings.

4 System Architecture

4.1 DataSet

Initially the list of 173 universities had been narrowed up, removing the data they wanted. Universities still do have skewed results. Crawler is designed to include a list of students from Edulix, and access their profiles. Using the Python Library's Beautiful Soup model, the data from each profile are scrapped and extracted from the HTML after selecting a particular student class (Fig. 1).

4.2 Data Cleaning

Roughly 4500 raw samples of data were collected as a result of scraping. Each style suits to the student's profile. GPA, Undergraduate School, GRE Verbal Rating, GRE Quantitative Quality, GRE Analytical Writing Score, Number of Journal Articles, Number of Conference Articles, Industry Experience, Educational Experience, and Internship are the data points collected. This field is only a text file, so details from university undergraduates need to be cleaned up since it is not a preferred area. Input errors were therefore generated from numerous studios, which were resolved by cutting the string and removing the spaces inside them. GRE scores (Verbal, Quantitative, and AWA) have been refined as scores of both the old and new test versions have been included. Similarly, the GPA scores available are based on multiple point scales, and all GPA scores are weighted to a standard 5-point scale. Furthermore, other attributes, such as the institution's undergraduate students and the school to which they belong, are deemed to be unique features. In total 173 separate undergraduate universities and 23 different courses were found after filtration, each of which was used as binary characteristics.

5 Algorithms Used

5.1 Collaborative Filtering

More weight-age is given to GRE quants because normally universities gives first priority to GRE quants score. On the basis of that the list of universities will be listed with the user and system ratings. Steps:

- Users similarity calculation: It uses Pearson correlation, Euclidean distance and cosine similarity
- Selection of top N nearest neighbors and
- Prediction.

So, in collaborative based the list of universities will be displayed on the basis of student's GRE total score and TOEFL score. It will compare the cut-off of the universities and on the basis of that the list of universities will be listed with the user and system ratings.

5.2 Content Based Filtering

Recommendation to user is given only by the users individual behaviour and data. First it analyses the description of items preferred by user to decide the preferences that can be utilized to describe these items. Based on users' choices user profile is created, next each item attribute is compared with user profile so that only related items are recommended to the user. So in collaborative based the list of universities will be displayed on the basis of student's GRE verbal score, GRE quants score and TOEFL. It will compare the cut off of the universities and on the basis of that the list of universities will be listed with the user and system ratings.

5.3 Similarity and Distance

(1) *Pearson Correlation*: Pearson relationship quantifies the straight relationship between persistent factors (alluded to as). The unique formula for correlation, created by Pearson himself, utilizes unanalysed

$$\rho(X, Y) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{4}$$

Formula for two factors, X and Y is in Eq. (4) Raw inspections are focused by subtracting their means and re-scaled by an estimate of standard deviations. An alternate Method to show a similar amount is as far as presume values, means X, Y, and standard deviations X, Y in Eq. (5) Notice that

$$\rho(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{5}$$

the numerator of this division is much the same as above meaning of covariance, since mean and desire are frequently utilized reciprocally. Separating the covariance between two factors by the result of standard deviations guarantees that correlation will consistently fall between -1 and 1. This makes deciphering the correlation coefficient a lot simpler. Let the set of items evaluated by the two users and be meant by I, at that point similarity coefficient between them is determined as in Eq. (6) Here i signifies the rating of user u for item i,

$$\text{sim}(u, v) = \frac{\sum_{i=I_u \cap I_v}(r_u, j - \overline{r}_u)(r_v, j - \overline{r}_v)}{\sqrt{\sum_{i=I_u \cap I_v}(r_u, j - \overline{r}_u)^2}\sqrt{\sum_{i=I_u \cap I_v}(r_v, j - \overline{r}_v)^2}} \tag{6}$$

and r_u is the mean of all items given by user u. Additionally, $r_{v,\,i}$ signifies the rating of user v for item i, and r_u is the mean of items given by user v. Let the set of items judged by the two users and be meant by I, at that point closeness coefficient between them is decided as in Eq. (7)

$$\text{sim}(u, v) = \frac{\sum_{i=I_u \cap I_v}(r_u, j - \overline{r}_u)(r_v, j - \overline{r}_v)}{\sqrt{\sum_{i=I_u \cap I_v}(r_u, j - \overline{r}_u)^2}\sqrt{\sum_{i=I_u \cap I_v}(r_v, j - \overline{r}_v)^2}} \tag{7}$$

(2) *Cosine Similarity*: The cosine distance between two points is the edge that the vectors to those points make. This point will be inside the range 0°–180°, regardless of what rate measurements the space has. The similarity sim(u, v) between user u and v is calculated as shown in Eq. (8)

$$\mathrm{sim}(u, v) = \frac{u \cdot v}{|u||v|} = \frac{\sum_{i=1}^{N} u_i v_i}{\sqrt{\left(\sum_{i=1}^{N} u_i^2\right)\left(\sum_{i=1}^{N} v_i^2\right)}} \tag{8}$$

5.4 K Nearest Neighbor

K Nearest Neighbors algorithm could be a non-parametric method used for classification and regression. Within the K-NN classification, the output may be a class member. An object is characterized by the bulk of its neighbors, and its k could be assigned to the foremost common class among its neighbors (k is a positive integer, usually small). If $k = 1$, the article is just assigned to it single nearest neighbor class. The K-NN model was tested by adjusting the quantity of neighbors used and noticed that the next precision of 87.1% was obtained when the amount of neighbors was cherish 6. Exact variance with the amount of trees planted.

5.5 Random Forest Classifier

Random Forest may be a set of trees for making decisions. Like single-variety trees, they can be influenced by high volatility or high dependencies (depending on how they are adjusted). Typical for random forests are the quest for a natural balance between the two extremes. Since they require only a few music parameters, they can be used more effectively with the default parameter settings (i.e. they're effective).

6 Result and Analysis

By using tkinter and list box Top 6 universities displayed. It recommends top 6 universities to the students. Based on this student will choose one university for admission. This is very convenient tool for students which help students for choosing right university based on their requirements (Fig. 2).

7 Conclusion and Future Scope

We have applied KNN, SVM and linear regression on the attributes such as CGPA, GRE, and TOEFL. KNN gives preference to GRE score, SVM gives preference to CGPA and Linear Regression gives preference to TOEFL score. Hence the conclusion is that in this recommendation system KNN gives the highest accuracy that is

University Ranking	University	Acceptance Percentage
1	Carnegie Mellon University	66 %
9	Georgia Institiute of Technology	66 %
22	New York University	61 %
10	University of Maryland College Park	57 %
30	University of North Carolina at Charlotte	57 %
25	Rutgers University New Brunswick	56 %

Fig. 2 University list

87% as compare to SVM and Random forest which is 78% and 81% respectively by comparing the graphs of each model. We have studied new features of python language as list box, tables, graphs, TKINTER and animation. In future we will use sematic analysis on the reviews of the other users for the recommendation system and This project can be hosted as a weblink in any college (Fig. 3).

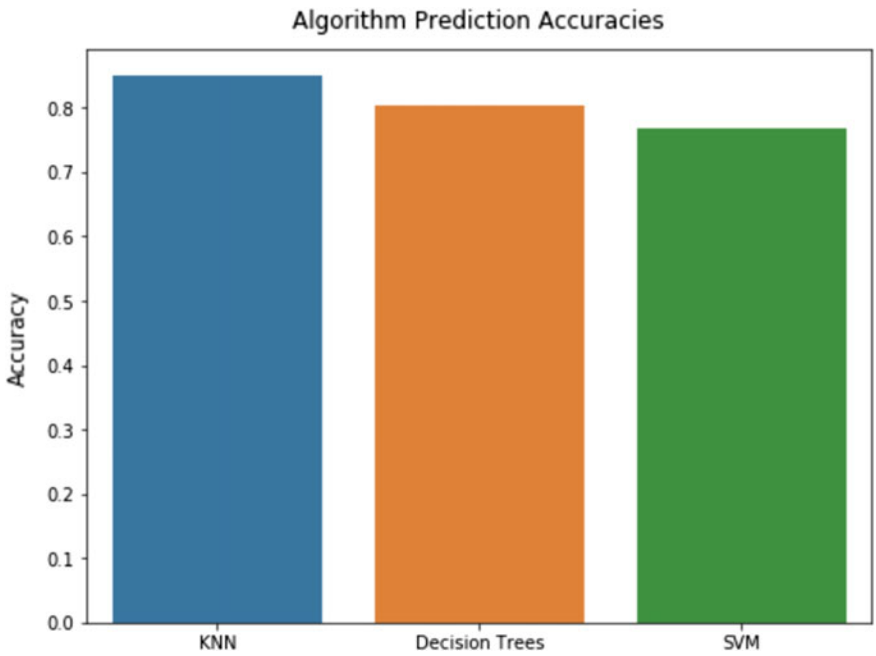


Fig. 3 Algorithm prediction accuracies

Acknowledgements Sincere thanks to my guide Bhavana Tiple for providing immense support in this research. Also, I thank Siddhivinayak Kulkarni Sir for helping me the best way possible.

References

1. Ricci F, Rokach L, Shapira B, Kantor PB (2017) Recommender systems handbook, 1st edn, Chap. 2. Springer US
2. Martinez SG, Lhadj AH (2018) Educational recommender systems: a pedagogical-focused perspective, vol 25. Springer International Publishing, pp 113–124
3. Vialardi C, Bravo J, Shafti L, Ortigosa A (2016) Recommendation in higher education using data mining techniques. In: International conference on educational conference, Cordoba, Spain, pp 190–198 (2016)
4. Sarwar B, Karypis G, Konstan J, Riedl J, Castro F, Vellido A, Nebot A, Mugica F (2015) Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th international conference on World Wide Web (WWW '01). ACM, New York, NY, pp 285–295
5. Linden G, Smith B, York J (2013) Amazon.com recommendations: item-to-item collaborative filtering. Amazon.com Inc., Seattle, Washington, United States, IEEE Computer Society
6. Schafer JB, Frankowski D, Herlocker J, Sen S (2017) Collaborative filtering recommender systems. In: The adaptive web. Springer, Berlin Heidelberg, pp 291–324
7. Melville P, Mooney RJ, Nagarajan R (2012) Content-boosted collaborative filtering for improved recommendations. In: Proceedings of the eighteenth national conference on artificial intelligence (AAAI-2012), Edmonton, Canada, pp 187–192, July 2012
8. Ekstrand MD, Riedl JT, Konstan JA (2016) Collaborative filtering recommender systems. Found Trends Hum-Comput Interaction 4(2):81–173
9. Herlocker JL, Konstan JA, Terveen LG, Riedl JT (2014) evaluating collaborative filtering recommender systems. ACM Trans Inf Syst 22(1):5–53
10. Santos OC, Boticario JG (2016) Requirements for semantic educational recommender systems in formal e-learning scenarios. Algorithms 4(2):131–154 (2016)
11. Bobadilla J, Serradilla F, Hernando (2019) Educational and scientific recommender systems: designing the information channels of the virtual university. International Encyclopedia of Education, 12th edn, vol 7. Elsevier, Oxford, UK, pp 112–118
12. Manouselis N, Drachsler H, Verbert K, Santos OC (2016) Recommender systems for technology enhanced learning, 1st edn. Springer-Verlag, New York
13. Schulz AG, Hahsler M, Jahn M (2010) Educational and scientific recommender systems: designing the information channels of the virtual university. Int J Eng Ed 17(2):153–163
14. Deloach S, Saliba L, Smith V, Tiemann T (2013) Developing a global mindset through short-term study abroad: a group discussion approach. J Teach Int Bus 15(1):37–59
15. Han J, Kamber M, Pei J (2011) Data mining: concepts and techniques, 3rd edn. Morgan Kaufmann Publishers, San Francisco

Comparison of Performances of Regression Model-Based Prediction of Meteorological Conditions

Ramya Sree Vejendla, Bhagya Lakshmi Pavuluri, Nandini Venigandla, Deepika Tinnavalli, and Shahana Bano

Abstract This paper clearly illustrates the working of different machine learning algorithms to determine the weather conditions. This involves the prediction of temperature by training with the pre-existing dataset of weather conditions on each day for around 40 years. This trained data is tested to evaluate the temperature of a certain day in the upcoming calendar by date or year. This illustration describes the comparison between different algorithm results and determines the most efficient algorithm. The algorithm involved were Linear Regression, Logistic Regression, and Clustering. These three algorithms involve different mechanisms such as predicting based on mean, probability, and grouping based on similar constraints. The model helps to select the most efficient algorithm which gives the approximate values nearer to accurate values. Though all the techniques involved in previous analysis are mostly based on mean analysis the result is almost approximate but under logistic regression, it either gives almost the accurate result or the wrong result. Here we introduce clustering since the date or year could be grouped under a certain condition where either based on the temperature of a certain year or the season.

Keywords Unsupervised data · Linear regression · Logistic regression · My Logit function

1 Introduction

Here in this work, the algorithms considered were Regression analysis implementing both linear regression and logistic regression and also clustering technique i.e. K-means clustering. Weather prediction [5] is all a different concept as we in the first place could predict that usually due to global warming and all the temperature rise from year to year increases. But the thing matters is at what rate does it affect and the range where it could lie and also the seasonal differences. Apart from this if

R. S. Vejendla (✉) · B. L. Pavuluri · N. Venigandla · D. Tinnavalli · S. Bano
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

S. Bano
e-mail: shahanabano@icloud.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_15

sudden calamities occur how it could affect the weather i.e., in terms of temperature, humidity, precipitation, etc. To predict and analyse certain adverse effects we implemented three different algorithms that were linear regression, logistic regression, and clustering.

1.1 Linear Regression

Linear regression is the linear representation of two different scalar values of which one is a dependent variable while the other is an independent variable. Here we predict [1] the unknown variable with respect to the other known variables. Here in linear regression, the linear procedure helps to predict the rate of change of a dependent variable and the regression describes the point values or the approximate values. Moreover, the linear regression is linear [3] not because of the rate of change is linear that is x is constantly dependent on y but because the theta or the relation is linear. There are different regression analysis [8] besides linear based on the number of dependent variables and independent variables such as multiple linear regression, ordinal regression, or multinomial regression.

1.2 Logistic Regression

Logistic Regression is used to predict [12, 13] the data and explain the relationship between one variable with one or more ordinal and nominal variables. It is a predictive analysis [2] and also a static model that uses a logistic function to determine a binary dependent variable, in logistic regression, the outcome is like continuous with many possible values but the outcome has only a limited number of possible values as results. Logistic regression implements the Logit function that is used in the method of classification.

1.3 K-Means Clustering

Clustering is like the most popular partitioning algorithms in clustering is the K-means cluster analysis. It tries to cluster data based on their similarity. Also, there exist specific clusters and the data is grouped into the same clusters. There are a lot of clustering techniques like k-means clustering, hierarchical clustering, and fuzzy clustering [11]. Of all these k-means is most popular as it is easy to implement it over unknown groups of data from complex data sets. And from these unlabelled data the centroids of the k clusters as to label new data.

2 Procedure

2.1 Dataset Extraction

Here the dataset is obtained from the real-world weather conditions and this dataset consists of about 19 variables and thousands of records. The dataset consists of both dependent and independent values to implement three different algorithms as each algorithm is applied to the different number of independent variables. Hence the dataset is accomplished satisfying all these factors.

2.2 Linear Regression

In linear regression analysis the complete working depends on a linear curve that deals with one dependent variable and one independent variable. Here in this apart from fitting the linear line it also performs [10] different steps namely analysing the correlation between the variables and the directionality of data, It estimates the model that is fitting the line for the values as per the axial dimensions and finally, it validates the usefulness of the fitting model. Linear regression also utilizes a linear formula.

$$Y = a + bx$$

where Y is the dependent variable as it depends on X,
X is the independent variable.

2.3 Logistic Regression

Logistic regression determines the relation between a dependent variable and one or more independent variables. These independent variables could be of any value like nominal value, ordinal value, or sometimes a ratio value. This is also similar to the multiple linear regression [7] model. This is obtained based on probability values that are unlike in the linear regression here it uses an exponential function. We implement a maximum likelihood estimation function to obtain the best fit.

$$\ln \subset \frac{p}{1-p} \supset = a + bx$$

$$\frac{p}{1-p} = e^{a+bx}$$

$$p = \frac{e^{a+bx}}{1 + e^{a+bx}}$$

2.4 Clustering

Clustering as mentioned divides the unlabelled data into clusters. Initially, we specify the number of clusters and now shuffle the dataset by deriving the centroid of each cluster. Now shuffle the dataset again and again until there is no change observed in the clusters ad data points. And finally, the goal of the k-means clustering is to find groups of data under a label i.e., unlabelled to labelled data.

$$objective\ functiona\ J = \sum_{J=1}^{k} \sum_{i=1}^{n} ||x_i^{(j)} - c_j||^2$$

where k = Number of clusters, n = Number of cases, xi = case i, cj = Centroid for cluster j.

3 Flow Chart

3.1 Linear Regression

To perform certain linear regression analysis initially view the dataset and pre-process the weather dataset. To perform regression analysis [9] the data needs to have both dependent and independent variables and ensure that there is no relation between the two dependent and independent variables. Train the dataset with selected records of data that could be well equipped between different sets of values. Now to get into working fit the data model to linear model that is in the linear pattern. Now to display and check the certain values with a summary function. To predict the head, implement confidence and prediction intervals. Now terminate the function (Fig. 1).

3.2 Logistic Regression

For the working of logistic regression analysis extract the weather dataset [4] and now obtain and analyse the dates and analyse the dates and temperature conditions. As a part of data pre-processing apply data visualization techniques and remove the misclassified data now similar to linear regression utilize the mylogit function to apply the regression analysis and determine the date and temperature analysis. Also, unlike linear regression which uses the simple linear term, it utilizes exponential coefficients with confidence interval fitting. Now to interpret the results by plotting the graph with results (Fig. 2).

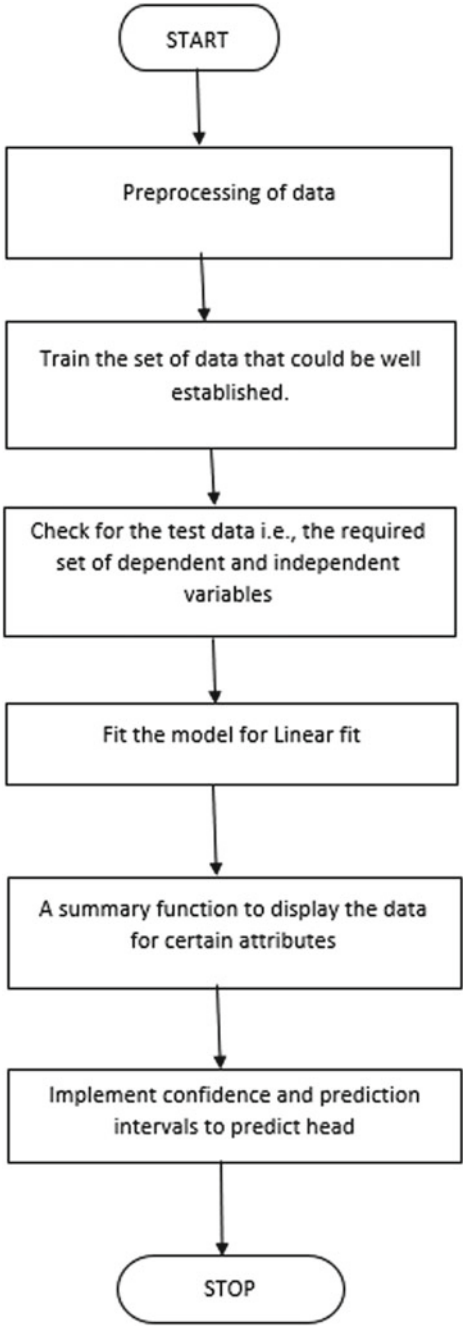


Fig. 1 Linear regression flowchart

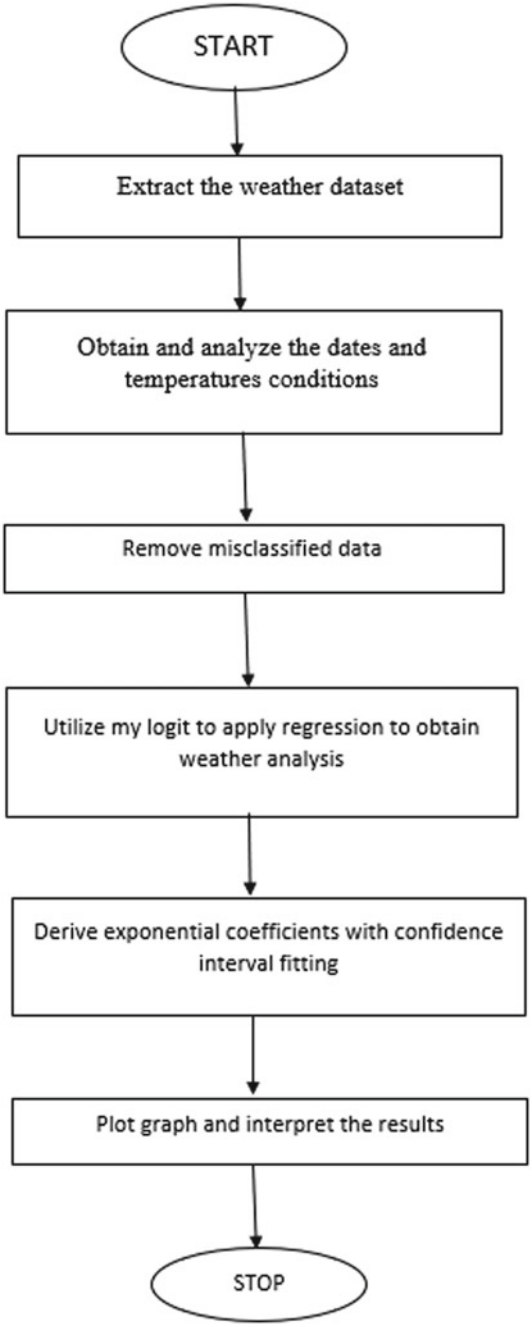


Fig. 2 Logistic regression flowchart

3.3 Clustering

To implement the clustering techniques whichever the algorithm was involved initially install the libraries and view the dataset. Now pre-process the dataset and visualize the data to remove the null variables if any exists. Apply K-means and select any of the well-established algorithms such as Hartigan, Lloyd…. Now if the algorithm exists to fit the methodology proceed with next step otherwise, repeat the application of the algorithm in otherwise conditions. Now obtain the column which you want to select. And now view the obtained clusters or plot the graph and analyse different temperatures and various constraints behind and now terminate the code (Fig. 3).

4 Results

4.1 Linear Regression

The main essence of the regression technique is the data set and here the dataset is a real-time dataset. I considered this dataset because it has a vast number of tuples predicting all the possibilities of a climate. In execution the dataset is completely read and displays the tuple data Now a set of data is set to be trained and the left is to be tested once the data is trained one can fix the data apart from which is trained is to text dataset. Now apply linear regression analysis i.e., lm to the tuples of a dataset whose relation is to be described and now here we get the summary as a quadrant representation with 4 quartiles Describing the minimum, maximum, first quartile, third quartile, and median values also with coefficient error values that is the deviation from expected results end the linear regression analysis to view this aa a plot based on the above-described analysis we plot them with respect to both the tuples selected earlier to fit. A linear fit is obtained describing the relation in a linear pattern (Figs. 4 and 5).

4.2 Logistic Regression

The dataset is the same as that used in linear regression analysis initially the dataset is being pre-processed and being checked for any missing values or out of boundary values. Consider the data and summarize it where u find the quartile values of each attribute involved. Besides I chose x for months from the dataset and then to proceed with the function that returns all the possible combinations of fog or fog temperature and more. And then a set of tuples that is years, months, and days of a year altogether. To interpret the logistic regression we consider the above tuples for a generalized linear model that on summarization gives us deviance values with respect to all

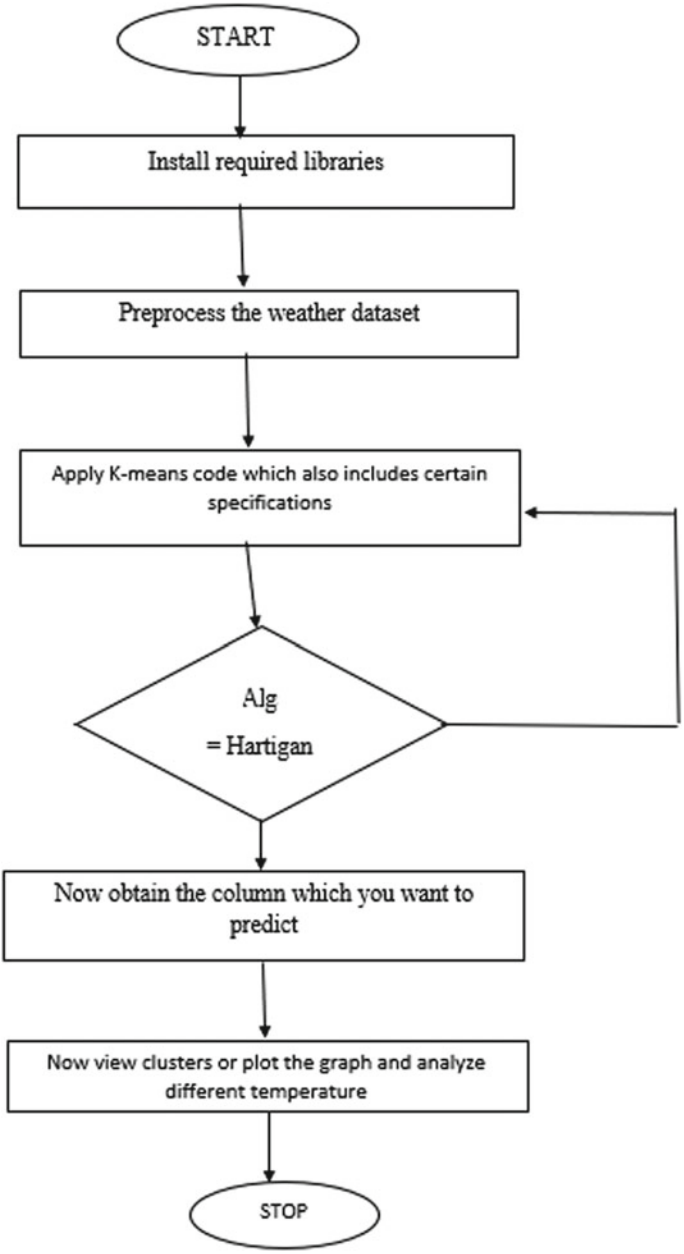


Fig. 3 Clustering flowchart

```
Call:
lm(formula = year ~ high_temp, data = train)

Residuals:
    Min      1Q  Median      3Q     Max
-0.5046 -0.4935 -0.4743  0.5065  0.5344

Coefficients:
             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 2.016e+03  4.096e-02 49225.456   <2e-16 ***
high_temp   4.097e-04  5.572e-04     0.735    0.462
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5001 on 2739 degrees of freedom
Multiple R-squared:  0.0001974, Adjusted R-squared:  -0.0001677
F-statistic: 0.5407 on 1 and 2739 DF,  p-value: 0.4622

> head(predict(lm.fit, test, interval = "confidence"), 5)
        fit      lwr      upr
1  2016.491 2016.472 2016.511
5  2016.492 2016.473 2016.511
8  2016.495 2016.475 2016.515
10 2016.493 2016.474 2016.511
14 2016.494 2016.475 2016.513
> head(predict(lm.fit, test, interval = "prediction"), 5)
        fit      lwr      upr
1  2016.491 2015.511 2017.472
5  2016.492 2015.511 2017.473
8  2016.495 2015.514 2017.476
10 2016.493 2015.512 2017.473
14 2016.494 2015.514 2017.475
```

Fig. 4 Linear regression analysis

different months and also the quartile values for data interpretation. A confint() function gives various data with normality and co-efficiency. Then it describes the exponential result analysis and views the plot of the values in a logistic pattern which might be an s-shaped graph including all the deviations of errors on the analysis (Figs. 6, 7 and 8).

4.3 Clustering

Clustering involves different algorithms of which k-means is most widely used due to its well approximate clustering. Also, in k-means, the initial phenomenon is a bit similar to linear regression where the data is trained. Initially, the data is pre-processed here the data is viewed and the data is viewed and several random sets are selected as a part of set data. We could select the number of clusters required

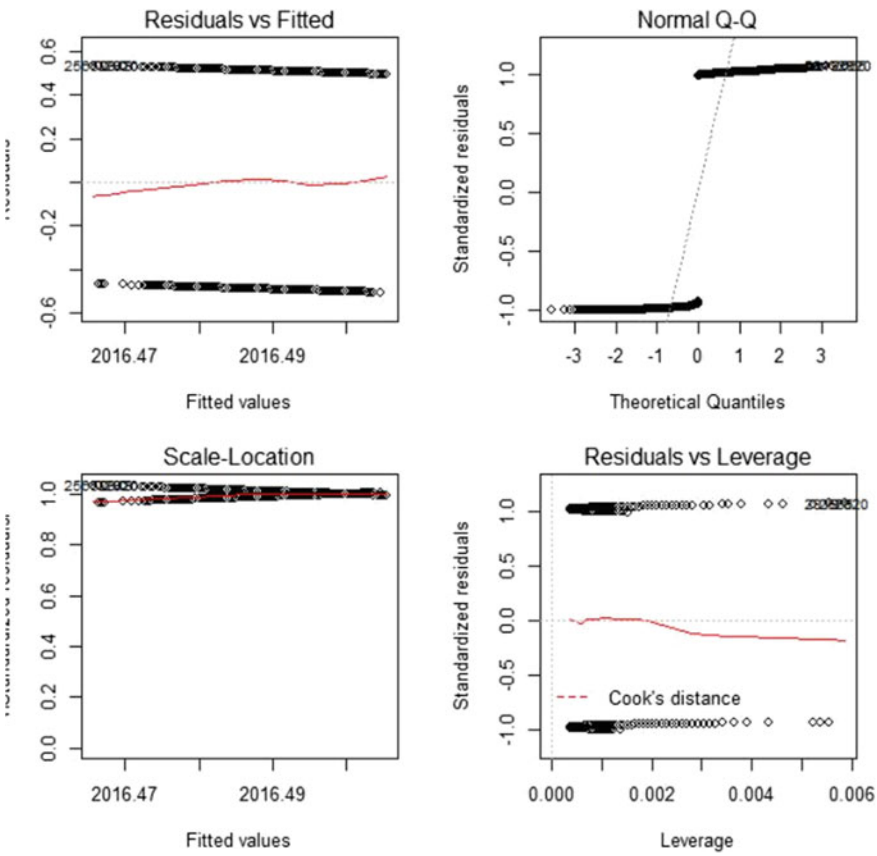


Fig. 5 Plots based on linear regression analysis years versus maximum temperature

in the asset of instruction along with certain k-means methodologies that are Lloyd, Hartigan, and others. Now based on a different technique that cluster interpretation differs and the cluster levels and values could be chosen with regard to certain iterations and all. The cluster records could be selected again and be evaluated with respect to several interpreted methodologies and the clusters are classified (Fig. 9; Table 1).

5 Conclusion

I would like to conclude the whole work in terms of accuracy. So, I performed all the different algorithm analysis on the same dataset with the same number of tuples and records. We implemented Linear regression, Logistic Regression, and K-means Clustering to predict the different range of temperatures i.e., high, medium, and

```
> confint.default(mylogit)
                    2.5 %       97.5 %
(Intercept) -837.55419860 724.72467881
month2        -1.23884332   0.66283679
month3        -0.96564877   0.98233463
month4        -0.83254204   1.18441854
month5        -0.82441257   1.00921423
month6        -0.58905039   1.23751525
month7        -0.37142737   1.49713804
month8        -0.11532125   1.96934994
month9        -0.21936199   1.86878585
month10       -0.98706440   0.74517039
month11       -1.38724105   0.43437541
month12       -1.72704225  -0.02951531
day           -0.01164503   0.03250412
year          -0.35831670   0.41642864
> plot(mylogit)
> exp(coef(mylogit))
 (Intercept)       month2       month3       month4       month5       month6       month7       month8 
3.157774e-25 7.497591e-01 1.008378e+00 1.192364e+00 1.096804e+00 1.382969e+00 1.755678e+00 2.526953e+00 
      month9      month10      month11      month12          day         year 
2.281224e+00 8.860809e-01 6.209946e-01 4.154975e-01 1.010484e+00 1.029482e+00 
> exp(cbind(OR = coef(mylogit), confint(mylogit)))
Waiting for profiling to be done...
                      OR     2.5 %   97.5 %
(Intercept) 3.157774e-25 0.0000000      Inf
month2      7.497591e-01 0.2877213 1.978369
month3      1.008378e+00 0.3807192 2.751325
month4      1.192364e+00 0.4387980 3.411961
month5      1.096804e+00 0.4330783 2.777864
month6      1.382969e+00 0.5479568 3.490974
month7      1.755678e+00 0.6849673 4.562578
month8      2.526953e+00 0.9096642 7.627880
month9      2.281224e+00 0.8198710 6.895693
month10     8.860809e-01 0.3638583 2.099856
month11     6.209946e-01 0.2456355 1.552503
month12     4.154975e-01 0.1724799 0.958320
day         1.010484e+00 0.9884654 1.033163
year        1.029482e+00 0.6987137 1.518992
> |
```

Fig. 6 Logistic regression analysis

low. Hence the analysis and its results on invariant data determine the error rate and accuracy where logistic regression is most accurate. Here we consider logistics as a good way to optimize and predict the weather or temperature conditions mostly from unlabelled as well as labelled also sometimes.

```
> mydata$month <- factor(mydata$month)
> mylogit <- glm( events~ month + day+ year, data = mydata, family = "binomial")
> summary(mylogit)

Call:
glm(formula = events ~ month + day + year, family = "binomial",
    data = mydata)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6355   0.2939   0.3792   0.4402   0.6961

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -56.414760 398.547853  -0.142   0.8874
month2       -0.288003   0.485131  -0.594   0.5527
month3        0.008343   0.496944   0.017   0.9866
month4        0.175938   0.514540   0.342   0.7324
month5        0.092401   0.467771   0.198   0.8434
month6        0.324232   0.465969   0.696   0.4865
month7        0.562855   0.476684   1.181   0.2377
month8        0.927014   0.531814   1.743   0.0813 .
month9        0.824712   0.532701   1.548   0.1216
month10      -0.120947   0.441905  -0.274   0.7843
month11      -0.476433   0.464707  -1.025   0.3053
month12      -0.878279   0.433051  -2.028   0.0425 *
day           0.010430   0.011263   0.926   0.3544
year          0.029056   0.197643   0.147   0.8831
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 801.55  on 1437  degrees of freedom
Residual deviance: 774.47  on 1424  degrees of freedom
  (2217 observations deleted due to missingness)
AIC: 802.47

Number of Fisher Scoring iterations: 5
```

Fig.7 Deviances based on regression analysis

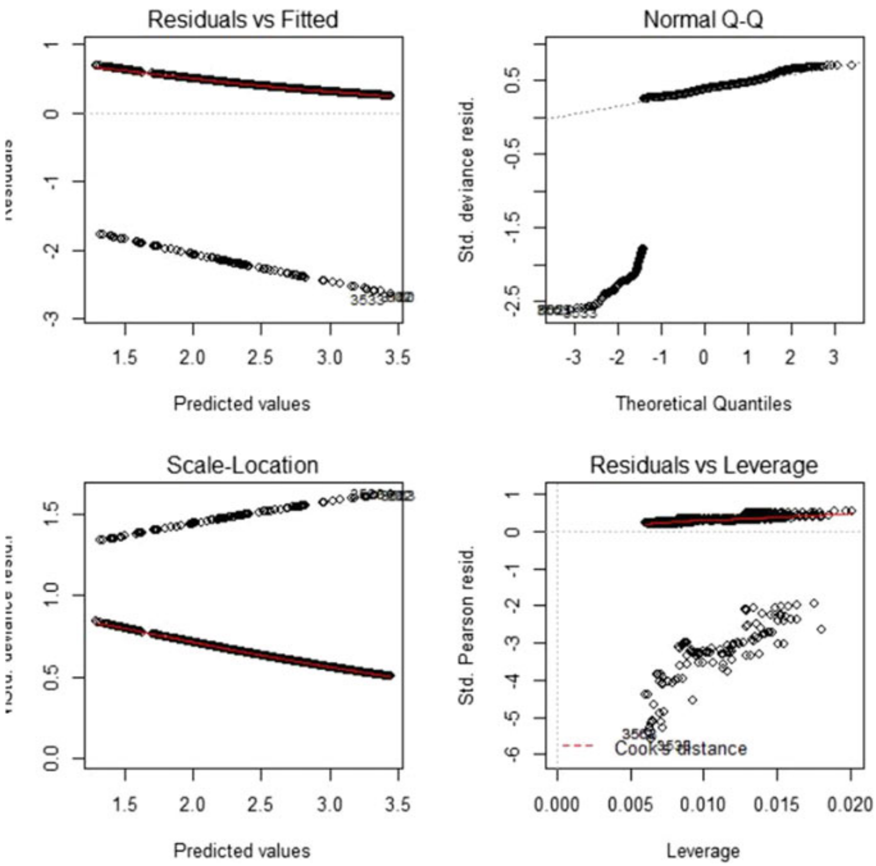


Fig. 8 Plots based logistic regression years versus maximum temperature

```
> kmeans(x,2,iter.max = 10,nstart=1,algorithm=c("Hartigan-Wong","Lloyd","Forgy","MacQueen"),trace=FALSE)
K-means clustering with 2 clusters of sizes 5, 5

Cluster means:
       [,1]       [,2]
1 0.3972067 0.01419765
2 1.1891598 1.18307225

Clustering vector:
 [1] 1 1 1 1 1 2 2 2 2 2

Within cluster sum of squares by cluster:
[1] 0.8198241 0.4264660
 (between_SS / total_SS =  80.0 %)

Available components:

[1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss" "betweenss"    "size"
 "iter"
[9] "ifault"
> km=kmeans(iris[1:4],3)
> colnames(x)<-c("x","y")
> x<-rbind(matrix(rnorm(10,sd=0.3),ncol=2),matrix(rnorm(10,mean=1,sd=0.3),ncol=2))
> cl=kmeans(x,2)
> m=kmeans(x,5,nstart=25)
> m=kmeans(x,5,iter.max = 10)
> cl1=kmeans(x,5,algorithm=c("Lloyd"))
> cl1=kmeans(x,5,algorithm=c("MacQueen"))
> cl1
K-means clustering with 5 clusters of sizes 1, 1, 4, 1, 3

Cluster means:
        [,1]       [,2]
1  1.33630701  0.7055838
2  0.47538945  0.3155323
3 -0.03964517 -0.2634254
4  0.99155648  1.0973310
5  0.75428740  0.7864129

Clustering vector:
 [1] 2 3 3 3 3 4 5 5 5 1
```

Fig. 9 Cluster-based on K-means algorithm

Table 1 Comparison of data distribution [6] over different quartiles in all three algorithms

S. No.	Results	Decision tree	Random forest	K-NN
1	Median values	-0.4743	-0.3792	-0.4094
2	Residual error or deviance	55%	80%	61.20219%
3	Accuracy error	50%	80%	0
4	Data handling	Handles simple linear data	Handles nonlinear [3] and unlabelled data	Handles unsupervised and unlabelled data
5	Benefits	Easy to understand	Best results and deals with multiple data	Deals with any group of data i.e., ungrouped or unlabelled data

References

1. Comparative Analysis of Temperature Prediction using Regression Methods and Back Propagation Neural Networks Survey on Weather Forecasting Using Data Mining, Proc. IEEE Conference on Emerging Devices and Smart Systems (ICEDSS 2018) 2–3 March 2018, 978–1–5386–3479–0/18.

2. Weather Analysis to predict rice cultivation time using multiple linear regression to escalate farmers exchange rate, 2017 International Conference on Advanced Informatics, Concepts, Theory and Applications.
3. Cloud based flight delay prediction using logistic regression, 2017 International Conference on Intelligent Sustainable Systems.
4. Numerical weather prediction using nonlinear auto regressive network for the Manaus region Brazil 2017 Innovations in power and Advanced Computing Technologies.
5. Weather monitoring using Artificial Intelligence 2016 2nd International Conference on Computational Intelligence and Networks.
6. Weather Visibility Prediction Based on Multimodal Fusion 2019 IEEE access year.
7. Prediction of Climate Variable using Multiple Linear Regression 2018 4th International Conference on Computing.
8. Rainfall Prediction using Regression Model by IJRTE 12016.
9. Comparative study of different weather forecasting models 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS).
10. Haze weather recognition based on multiple features and Random forest, 2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC).
11. Analysis of Weather Prediction using Machine Learning and Big DATA 2018 International Conference on Advances in Computing and Communicational Engineering.
12. Weather Forecasting Using Artificial Neural Network, Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018) IEEE, 978–1–5386–1974–2/18.
13. Rainfall Prediction based on Deep Neural Network: A Review, Proceedings of the Second International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020) IEEE.
14. Rainfall Forecasting in Bandung Regency using C4.5 Algorithm, 2018 6th International Conference on Information and Communication Technology (ICOICT).
15. Dynamic Line Rating Using Numerical Weather Prediction and Machine Learning, Year: 2017 | Volume: 32, Issue: 1 | Journal Article | Publisher: IEEE.

Vocal Source Builds Divergence in Gender Recognition

Guru Sree Ram Tholeti, Deepika Ghanta, N. V. S. Guru Sai Sarma Chilukuri, and Shahana Bano

Abstract Gender Classification is one of the crucial problems in fields of AI. Can a machine be able to classify Gender (Male/Female)? Yes, the above problem is done using various image classification techniques which uses feature extraction from given set of images. As human being's we can know the gender of the person when we talk to them, can this be achieved by a machine? This is what we are going to work in this research. This research is going to examine the performances of different Machine Learning algorithms and Deep learning algorithms on Gender Classification based on voice. We have used Multi Layer Perceptron (MLP), Random Forest, Decision Tree and Logistic Regression models and compare their performance to find out the best classifier for our data set. We got 96.84% accuracy using MLP, 96.42% using Random Forest, 96.21% using Decision Tree and 89.37% using Logistic Regression. Multi Layer Perceptron stands high with a modest difference in accuracy.

Keywords Multi layered perceptron (MLP) · Random forest · Decision tree · Logistic regression · Gender · Voice

1 Introduction

Performing different tasks on speech is one of the valuable and interesting work now a days, mostly in "**Natural Language Processing**" (NLP) and other branches of AI. Gender identification is considered to be one of the major problems in the field of signal processing [1]. Gender Differences occur in Voice based vocal folds, length of vocal folds, pitch of vocal box and use age of articulators. A close scrutiny, some of the human vocal features, reveals that classifying gender goes way beyond just the frequency and the pitch of a person [1]. The Data Set "Voice.csv" is used in our work taken from Kaggle. This is a classification work in which we have 2

G. S. R. Tholeti (✉) · D. Ghanta · N. V. S. G. S. S. Chilukuri · S. Bano
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

S. Bano
e-mail: shahanabano@icloud.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_16

classes and 21 features in our dataset, Male represented as 1 and Female represented as 0. In crime detection, it will be helpful. People commit different types of crimes through phone calls or voice messages. We have used Multi Layer Perceptron (MLP), Random Forest, Decision Tree and Logistic Regression models. Among all the models Multi Layer Perceptron gave the highest accuracy and evaluation metrics such are AUC-ROC curve, confusion matrix, Recall score, accuracy. MLP solves the problem stochastically so it is helpful fitness approximation. Gender classification has applications like, it is able to improve the intelligence of a surveillance system analyse the customer's demands for store management, and allow the robots to perceive gender etc. [2].

2 Methodology

2.1 Multi Layered Perceptron (MLP)

The imported data is then split into k folds where k = 2. An MLP model is built. The weights and biases are initialized with random values. In the initial iteration, an input pattern is presented and output values are calculated [3]. This process is continued for until the loss value reach the global minima by updating the weights and biases. Once the above process is done, results will be displayed and the process is stopped (Fig. 1).

2.2 RandomForest

The imported data is splitted into training and testing data. It creates multiple number of decision trees. The outputs are collected from each decision tree. The decision of the model is obtained by decision given by maximum number of trees in the forest is taken as the final decision. Then the new data point is classified into the corresponding class (Fig. 2).

2.3 DecisionTree

The imported data is splitted into training and testing data with a split ratio. The system entropy is calculated in the beginning. The information gain of each input feature is calculated. The features with maximum information gain will be selected, this attribute will become a node in the decision tree. Continue to expand the decision tree with combination of the attributes from root to that level. Once the tree is completely derived with given features then display results and stop (Fig. 3).

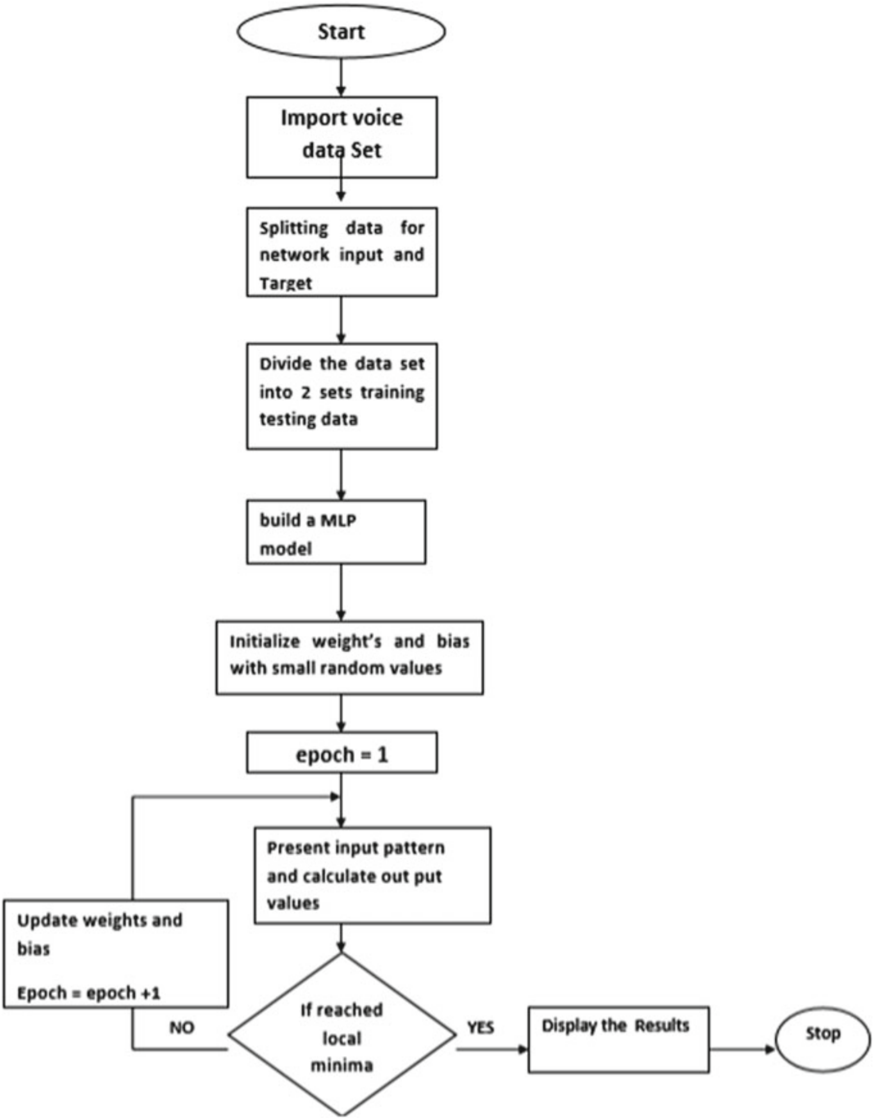


Fig. 1 Multi layered perceptron flow chart

2.4 LogisticRegression

Initially the Voice data is imported into our environment. The data is splitted into train and test set with a split ratio of 70:30. Then the equation of the linear model is sent to the sigmoid function. The sigmoid function generates a number which is between 0 and 1 which will help us in classifying the given point i.e. the sigmoid

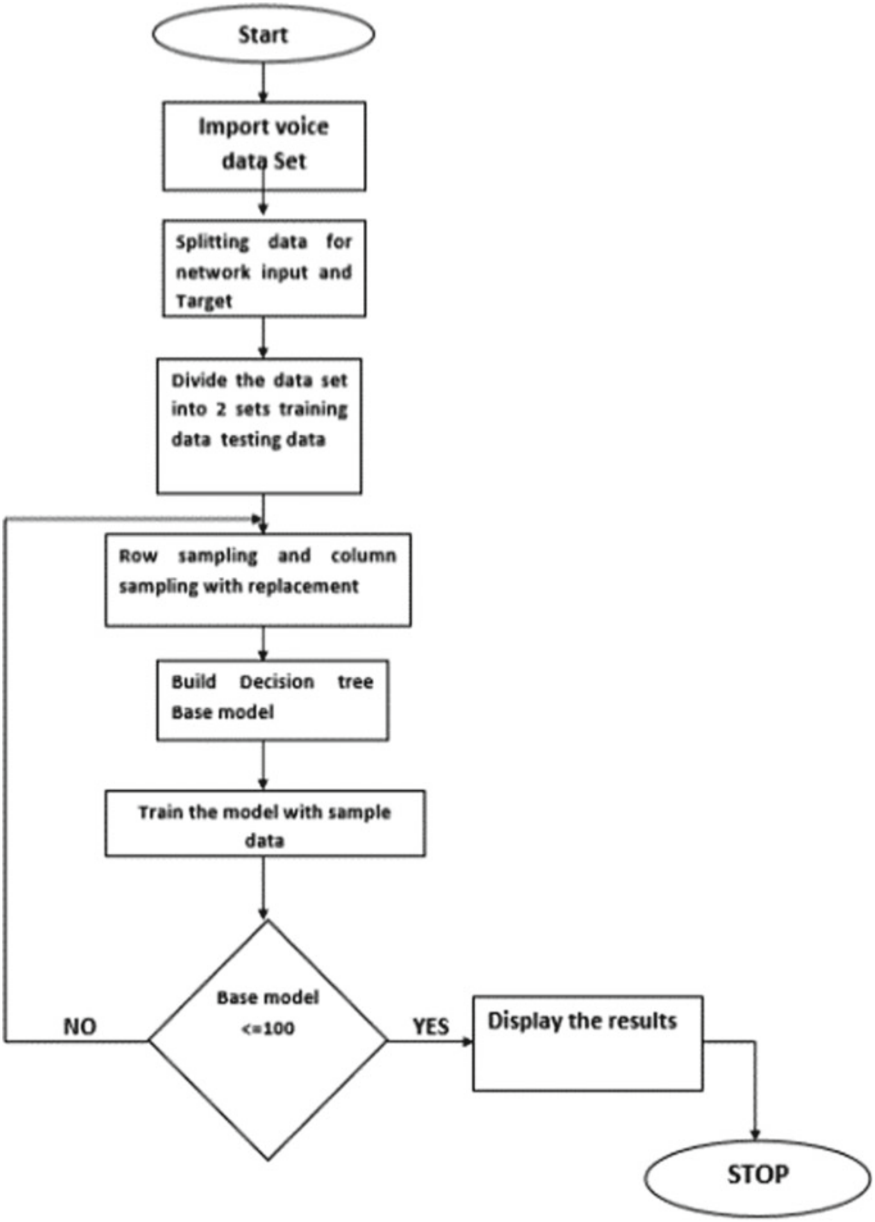


Fig. 2 Random forest flow chart

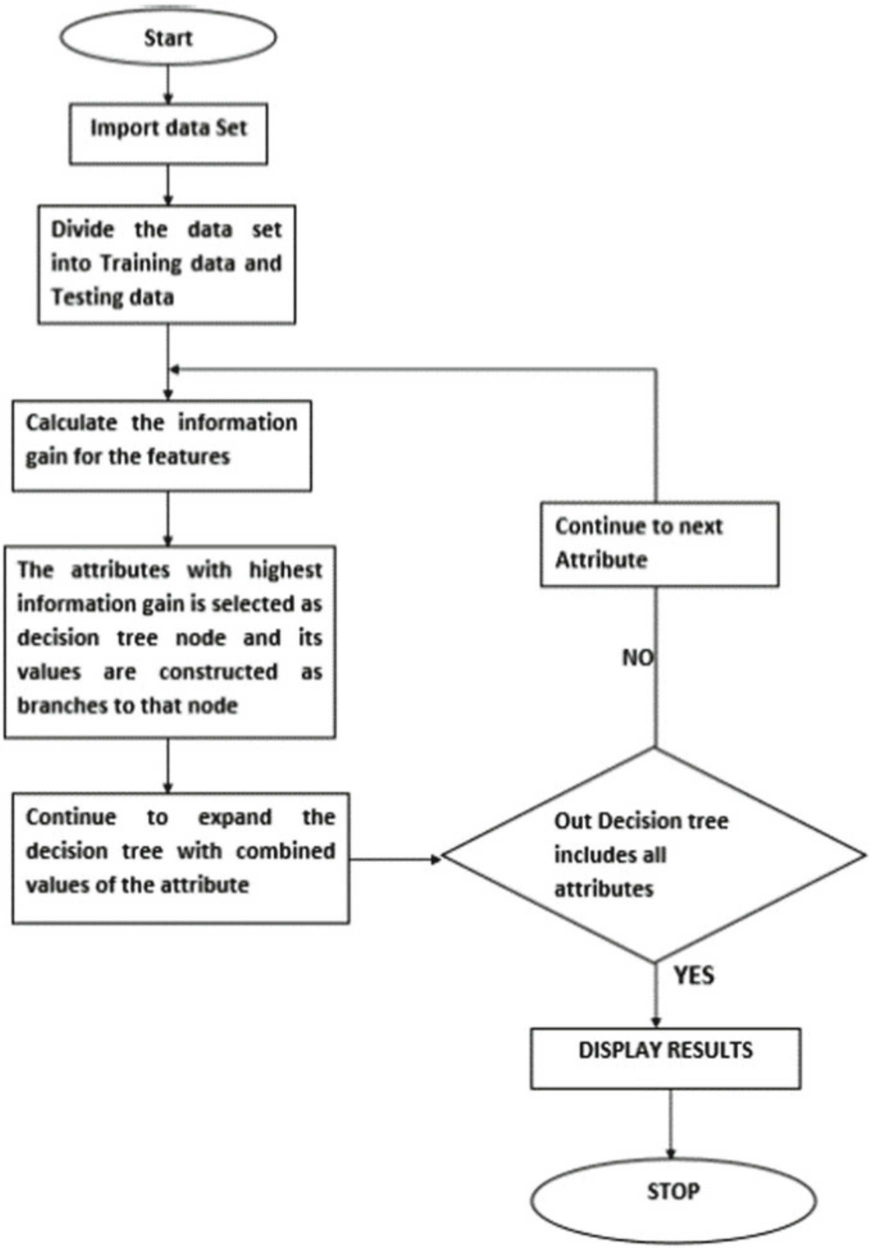


Fig. 3 Decision tree flow chart

will generate the probability of the person to be placed in male and female. The optimized parameters for theta are calculated by using gradient descent optimizer. Then the classified results are displayed (Fig. 4).

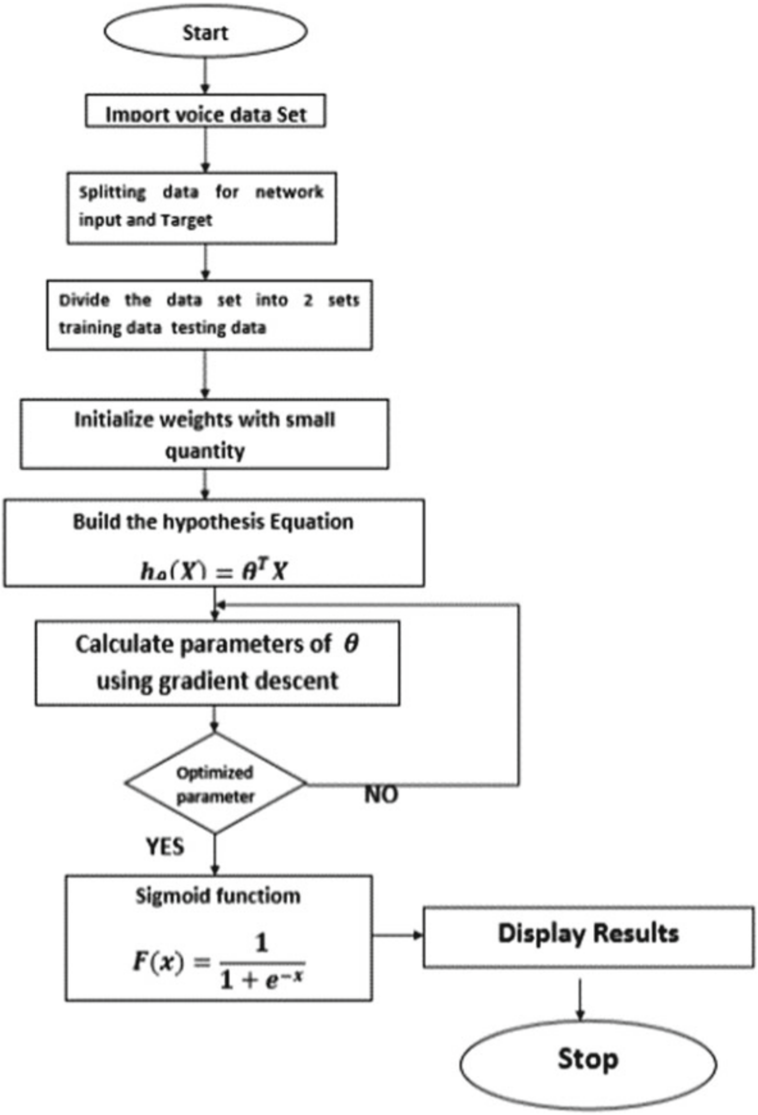


Fig. 4 Logistic regression flow chart

3 Procedure

Data Set

Voice Data set is taken from Kaggle website which is source for several datasets. There are 21 features and 3168 records present in dataset. Since data set that we are considering having 2 classes (Male, Female) our work is confined to binary classification.

Some features in dataset:

- Meanfreq (mean frequency)
- Sd (standard deviation)
- Median
- Mode
- centroid
- Skew
- Gender.

3.1 Multi Layered Perceptron (MLP)

Multi Layer Perceptron is used for data which is not linearly separable [4]. MLP model used in my work consists of 1 input layer, 2 hidden layers and 1 output layers, we used "ReLU" Activation function (Eq. 2) in hidden layers and "SIGMOID" Activation function (Eq. 1) for output Layer [5].

Sigmoid Activation Function

$$F(s) = \frac{1}{1 + e^{-s}} \tag{1}$$

ReLU Activation Function

$$\mathrm{F(s)} = \max(0, \mathrm{s}) \tag{2}$$

Input layer of our model consists of 20 neurons each one indicates individual features in our dataset the neurons present in input layers are linear i.e. the input and outputs of the input neurons are same. The outputs of input layer are connected to the hidden layers nodes which are computational neurons which will compute the output using below equation (Eq. 3).

$$output = \sum_{i=1}^{n} (w_i x_i) + bias \tag{3}$$

W weight of connection;
x input value.

Finally the output of output layer is compared with original output in training data, which is known as loss function, here our main goal is to minimize this loss to global minima, loss is caliculated by Eq. (4).

$$loss = \sum_{i=1}^{n} (y - \hat{y})^2 \tag{4}$$

In the above formula

Y o/p of the record in training data;
ŷ o/p obtained by the network.

initially we got a loss of: 0.0391.
To reduce the loss function, we use gradient descent optimizer (Eq. 5)

$$W(new) = W(old) - \eta * \frac{\partial\, loss}{\partial\, weight} \tag{5}$$

η Learning rate of the network.

Finally we got a loss of: 0.0167; Accuracy: 96.84.

3.2 Random Forest

Random Forest classifier is an boosting technique of ensemble model, in random forest the base model used is decision tree where the and the final decision is based on majority vote technique, my model creates 100 decision trees.

The complete training data is random sample with replacement and given as input to all the decision trees, while testing the testing data is given to all the trees and the all the decision that is given by majority of the trees in the forest is taken as the final decision of the tree, the main advantage of Random forest is it overcomes the high variance problem faced by decision tree, my random forest model gives an accuracy of 96.42% and F1 score of 96.42%.

3.3 Decision Tree

A decision tree is a decision building or a decision support algorithm with the help of tree like structure in which the nodes indicates the attributes and the branches indicates the decision, decision tree is build on 2 things mainly, Information gain and Gini index, our decision tree model use information gain (Eq. 6) as a criteria and for calculating information gain we required entropy given by Eq. (7).

$$\text{Information Gain(T, X)} = \text{entropy(X)} - \text{Entropy(X, T)} \tag{6}$$

T Target variable;
X Individual attribute

$$\text{Entropy(X)} = -\text{p} \log_2 \text{p} - \text{q} \log_2 \text{q} \tag{7}$$

X individual attribute;
p probability of class 1;
q probability of class 2.

By using above to formulas we got "meanfun" attribute as our main root with 0.5 information gain.

Accuracy of the model on the given data set is 96.21 and F1 Score of 96.27.

3.4 Logistic Regression

Logistic regression is mainly used for data which is linearly non separable, Logistic regression hypothesis function (Eq. 8) can be given by

$$H_\theta(x) = \theta^T * x \tag{8}$$

θ^T weight vector;
x attributes

and applies an activation function namely sigmoid function to it. The sigmoid function take the predicted value from linear model and converts into a value between 0 and 1.

The sigmoid function is given by Eq. (9),

$$Sigmoid\left(\theta^T * x\right) = \frac{1}{1 + e^{-\theta^T * x}} \tag{9}$$

According to probability value of sigmoid function, we classify our input data into classes present in our dataset. We use a cost function (Eq. 10), which will be optimized to find the best parameters of theta.

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^i \log h_\theta\left(x^i\right) + \left(1 - y^i\right)\log\left(1 - h_\theta\left(x^i\right)\right)\right] \tag{10}$$

To minimize the cost function we used a gradient descent optimizer and it can be computed by using Eq. (11)

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^i\right) - y^i\right)x_J^i \tag{11}$$

for logistic regression model over voice data set we got an accuracy of 88.53% and F1 score of 89.06% and Cost is 0.325.

4 Results

4.1 Multi Layered Perceptron(MLP)

In the result of MLP classifier we have Accuracy of testing data as 96.84% (from Fig. 5) and an accuracy of 96.16% on training data (from Fig. 6), from confusion metrics we have that 162 persons are classified correctly into female class and 145 persons are classified correctly into male class and there are 0 persons who are male and classified into female and 10 persons who are classified into male but actually female. The Roc Score of our model is 97.09% which indicates the trust level of the model.

```
print("\nAccuracy score: %f" %(accuracy_score(testY,predictions_MLP) * 100))
print("Recall score : %f" %(recall_score(testY,predictions_MLP) * 100))
print("ROC score : %f\n" %(roc_auc_score(testY,predictions_MLP) * 100))
print(confusion_matrix(testY,predictions_MLP))
print("F1 score : %f\n" %(f1_score(testY,predictions_MLP, zero_division=1)*100))
```

```
Accuracy score: 96.845426
Recall score : 100.000000
ROC score : 97.093023

[[162  10]
 [  0 145]]
F1 score : 96.666667
```

Fig. 5 Result of MLP model

```
mlp = MLPClassifier(hidden_layer_sizes=(50,3),max_iter=301)

y_t = y_t.reshape(-1, 1)
print(x_t.shape,y_t.shape)
mlp.fit(x_t,y_t)
print("Training set score: %f" % mlp.score(x_t,y_t))
print("Test set score: %f" % mlp.score(x_tes,y_tes))
```

```
(2217, 20) (2217, 1)
```

```
E:\Anaconda\lib\site-packages\sklearn\neural_network\_multilayer_perceptron.py:934: DataConversionWarning: A column-
s passed when a 1d array was expected. Please change the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, warn=True)
```

```
Training set score: 0.961660
Test set score: 0.968454
```

Fig. 6 Training and testing set scores of MLP model

```
print("\nAccuracy score: %f" %(accuracy_score(y_test,RF_ypre) * 100))
print("Recall score : %f" %(recall_score(y_test, RF_ypre) * 100))
print("ROC score : %f\n" %(roc_auc_score(y_test, RF_ypre) * 100))
print(confusion_matrix(y_test, RF_ypre))
print("F1 score : %f\n" %(f1_score(y_test, RF_ypre, zero_division=1)*100))
```

```
Accuracy score: 96.424816
Recall score : 96.624473
ROC score : 96.425444

[[459  18]
 [ 16 458]]
F1 score : 96.421053
```

Fig. 7 Result of random forest model

4.2 Random Forest

In the result of Random Forest classifier we have Accuracy of testing data as 96.42% (from Fig. 7), from confusion metrics we have that 459 persons are classified correctly into female class and 458 persons are classified correctly into male class and there are 16 persons who are male and classified into female and 18 persons who are classified into male but actually female. The Roc Score of our model is 96.42% which indicates the trust level of the model.

4.3 Decision Tree

In the result of Decision Tree classifier we have accuracy of testing data as 96.21% (from Fig. 8), from confusion metrics we have that 450 persons are classified correctly into female class and 465 persons are classified correctly into male class and there are 19 persons who are male and classified into female and 17 persons who are classified into male but actually female. The Roc Score of our model is 96.42% which indicates the trust level of the model.

```
print("\nAccuracy score: %f" %(accuracy_score(y_test,dt_ypre) * 100))
print("Recall score : %f" %(recall_score(y_test, dt_ypre) * 100))
print("ROC score : %f\n" %(roc_auc_score(y_test, dt_ypre) * 100))
print(confusion_matrix(y_test, dt_ypre))
print("F1 score : %f\n" %(f1_score(y_test, dt_ypre, zero_division=1)*100))
```

```
Accuracy score: 96.214511
Recall score : 96.074380
ROC score : 96.217062

[[450  17]
 [ 19 465]]
F1 score : 96.273292
```

Fig. 8 Result of decision tree model

```
print("\nAccuracy score: %f" %(accuracy_score(y_test,lr_ypre) * 100))
print("Recall score : %f" %(recall_score(y_test, lr_ypre) * 100))
print("ROC score : %f\n" %(roc_auc_score(y_test, lr_ypre) * 100))
print(confusion_matrix(y_test, lr_ypre))
print("F1 score : %f\n" %(f1_score(y_test, lr_ypre, zero_division=1)*100))
print("Cost : %f\n" %np.sqrt(((lr_ypre - y_test) ** 2).mean()))
```

```
Accuracy score: 89.379600
Recall score : 94.939271
ROC score : 89.154537

[[381  76]
 [ 25 469]]
F1 score : 90.279115

Cost : 0.325890
```

Fig. 9 Result of logistic regression model

4.4 Logistic Regression

In The result of Logistic regression classifier we have Accuracy of testing data as 89.37% (from Fig. 9), From confusion metrics we have that 381 persons are classified correctly into female class and 469 persons are classified correctly into male class and there are 25 persons who are male and classified into female and 76 persons who are classified into male but actually female. The Roc Score of our model is 89.15% which indicates the trust level of the model.

We perform classification on same dataset using four model their Accuracy and F1 Score is given by “Table 1”.

In “Table 2” we can find the different metrics in “Multi Layer Perceptron” model such as accuracy, Recall, Roc Score and F1 score.

Table 1 Accuracy and F1 score of all models used

S. No.	Model name	Accuracy	F1 score
1	MLP	96.84	96.66
2	Random forest	96.42	96.42
3	Decision tree	96.21	96.27
4	Logistic regression	89.37	90.27

Table 2 Accuracy, recall, roc-score, F1-score of MLP model

S. No.	Model	Accuracy	Recall	ROC score	F1 score
1	MLP	96.84%	100%	97.093%	96.66%

5 Conclusions

In this work we tried to classify gender based on voice, we performed this work only on 2 types of genders, the 3rd type of gender is not considered in our work. Now a days we can handle many apps based on voice in which gender classification is done first, in our work we took 1584 male and 1584 female records and achieved an accuracy of 96.84 and error of 0.03 and 0.94 True positive rate and 1.00 true negative rate.

References

1. Jadav S (2018) Voice-based gender identification using machine learning. In: 2018 4th international conference on computing communication and automation (ICCCA). Greater Noida, India, pp 1–4. https://doi.org/10.1109/CCAA.2018.8777582
2. Gupta P, Goel S, Purwar A (2018) A stacked technique for gender recognition through voice. In: 2018 eleventh international conference on contemporary computing (IC3) Noida, pp 1–3. https://doi.org/10.1109/IC3.2018.8530520
3. Badhon SM, Rahaman MH, Rupon FR (2019) A machine learning approach to automating Bengali voice based gender classification. In: 2019, 8th international conference on system modeling and advancement in research trends (SMART). https://doi.org/10.1109/smart46866.2019.9117385
4. Hassan B, Ahmed R, Li B, Hassan O, Hassan T (2019) Autonomous framework for person identification by analyzing vocal sounds and speech patterns. In: 2019 5th international conference on control, automation and robotics (ICCAR), Beijing, China, pp 649–653. https://doi.org/10.1109/ICCAR.2019.8813463
5. Chaudhary S, Sharma DK (2018) Gender identification based on voice signal characteristics. In: 2018 international conference on advances in computing, communication control and networking (ICACCCN), Greater Noida (UP), India, pp 869–874. https://doi.org/10.1109/ICACCCN.2018.8748676

An Analytical Prediction of Breast Cancer Using Machine Learning

N. V. S. Guru Sai Sarma Chilukuri, Shahana Bano, Guru Sree Ram Tholeti, Sai Pavan Kamma, and Gorsa Lakshmi Niharika

Abstract Breast Cancer is one of the most occurring cancer among women affecting about 2 million people. There is 98% chance of 5 years survival rate if detected at early stage. The data about Breast cancer used in this paper is the Wisconsin dataset which is taken from Kaggle. This is a classification problem, there are two classes (0 representing a non-malignant tumor, 1 representing malignancy). Min Max scalar is used for preprocessing of data to limit data within certain range (known as scaling). The algorithms used for classification are Support Vector Classifier, Random Forest, Naïve Bayes, Decision Tree, K-Nearest Neighbours. Support Vector Classifier and Random forest gave the highest accuracy, Evaluation metrics such are Area Under Curve-Rectified Operational Characteristics curve, confusion matrix, Recall score, accuracy. To avoid overfitting cross validation is used where k fold value is 3.

Keywords Artificial neural network · Random forest · Decision tree · Cross validation

1 Introduction

Being the most frequently occurring cancer in women, breast [1] cancer affects around 10% of women at some point in their life. It is the second leading contributor to women's death after lung cancer. 25% of all cancers in women including 12% of all new cases are caused by breast cancer [2]. Topics like medial science rise rapidly when certain approaches like data mining is applied due to better possibility of prediction of diseases [3–7], reducing medicine costs, improving health of patient by revamping the quality of healthcare along with value by saving people's lives through real time decisions. The paper provides you with an analysis of performance and comparison of accuracy in classification between the algorithms [8–13] such as: Logistic Regression, Suport Vector Machine [14], Random Forest and Naïve Bayes,

N. V. S. Guru Sai Sarma Chilukuri · S. Bano · G. S. R. Tholeti · S. P. Kamma · G. L. Niharika (✉)
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

S. Bano
e-mail: shahanabano@icloud.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_17

being the major influential algorithms of data mining used in the research community [15].

KNN: Assumes that similar things exit with in proximity.

Algorithm

1. Load the data.
2. Initialize K value to choose number of neighbours.
3. For each sample in the data
   a. Calculate the distance between the query and the current sample from the data.
   b. Add the distance and the index of the example to an ordered collection.
4. Sort the ordered collection of distances and indices in ascending order by the distances.
5. Pick the first K entries from the collection.

In general distance between the samples is calculated using Euclidean distance (Eq. 1).

$$\mathrm{d(x, y)} = \sqrt{\sum_{i=0}(x_i - y_i)^2} \tag{1}$$

Different types of distances are

1. Manhattan distance: $\mathrm{D(x, y)} = \sum_{i=0}^{n}|x_i - y_i|$
2. Chebyshev distance: $\mathrm{D(x, y)} = \max(|x_i - y_i|)$

Advantages

KNN needs less training period because of instance base leaning. New data can be added seamlessly which will not impact the accuracy because KNN requires no training for classification. KNN is easy to implement. Less number of parameters.

Disadvantages

Does not work well with larger datasets because the cost of calculating the distance between the data points is very high. Does not work well with higher dimensions. Sensitive to noisy data, missing values, and outliers.

Naïve Bayes

Navie Bayes [16] (Eq. 2) is probabilistic classification algorithm whose crux is based on Bayes theorem.

$$P(h|D) = \frac{P(D|h) * P(h)}{P(D)} \tag{2}$$

P (h | D) is posterior probability
P (D | h) is likelihood
P (h) is prior probability
h is hypothesis
D is Data.

Assumptions
Predictors/features are in independent (Eq. 3) (i.e. one feature does not affect another feature).

$$P(A \cap B) = P(A) * P(B) \tag{3}$$

Algorithm

- Calculate the probabilities of membership of each class label (i.e. probability of data points associated to a class).
- The class having the highest probability is the most suitable class.
- The above statement refers to calculating MAP (Maximum A Posteriori).
- MAP for hypothesis is:
  - MAP (h) = max $P(h|D)$
  - MAP (h) = max $\frac{P(D|h)*P(h)}{P(D)}$
  - MAP (h) = max $P(D|h) * P(h)$

  Gaussian Naïve Bayes classifier (Eq. 4):

$$P(x_i|y) = \frac{1}{\sqrt{2 * \pi * \sigma_y^2}} * \exp\left(-\frac{\left(x_i - \mu_y\right)^2}{2 * \sigma_y^2}\right) \tag{4}$$

Advantages
The convergence is quicker given that Naïve Bayes condition of Independence holds. As Naïve Bayes is generative model it easy to deal with missing values. Needs less training data.

Disadvantages
Naïve Bayes assumes that the features are independent. There is zero frequency problem where Naïve Bayes algorithm assigns zero to class that model has never seen during training. To solve zero frequency, we need to use smoothing techniques.

Random Forest
Random forest is voting based supervised discriminative classification algorithm.

Algorithm

- Randomly select "Q" features from total "N" features where Q $\ll$ N
- Build the decision trees associated with "Q" features from selected data points
- Find the best tree within the forest
- Take the test features and use the randomly created decision trees for classification and store the predicted outcomes
- Calculate the votes for each predicted target
- Highly voted predicted target would be the final answer.

Advantages
It can be used for both classification and regression problem. There is no problem of overfitting given that there are a greater number of trees present in the forest. Can handle missing values.

Disadvantages
Needs a lot of computational power and memory storage to calculate a greater number of decision trees**.** Predictions are slow. Needs longer training period.

Logistic Regression
Logistic Regression is supervised discriminative classification algorithm mostly used when target feature is dichotomous**.** Logistic regression fits a S shaped curve (Eq. 5) known as sigmoid curve to the target variable.

Sigmoid Function:

$$s(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

Output of sigmoid curve will always be within range of [0, 1].

Assumptions
Observations to be independent of each other, assumption of linearity.

Advantages
Easy to train, Less number of computations, Easy to Implement and Interpret and Features need not to be scaled.

Disadvantages
For Multinomial classification we need to use one versus all or one versus one Classification, Prediction is not possible and Target variable should be discrete.

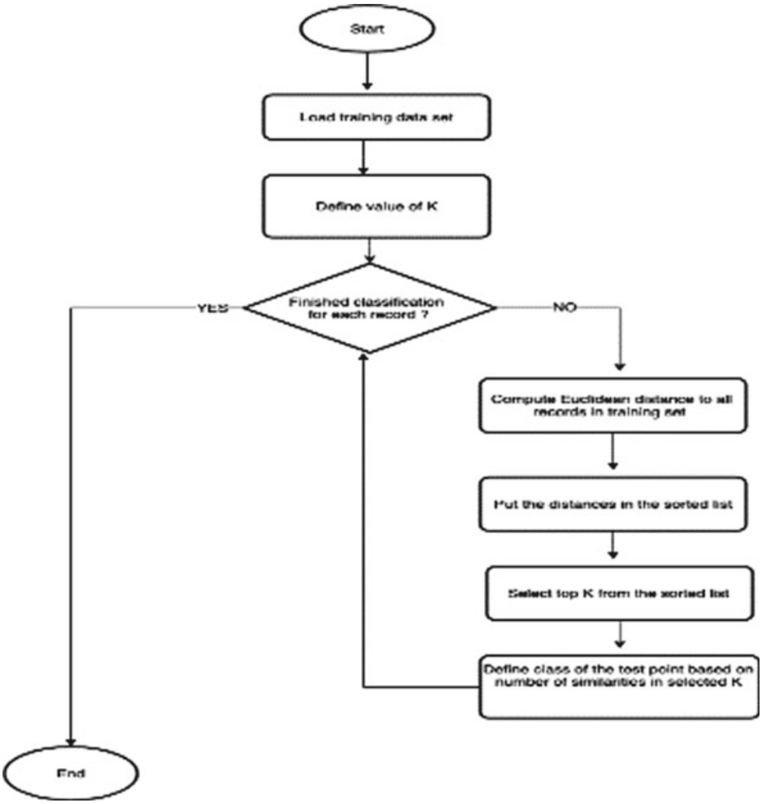


Fig.1 KNN Flow chart

2 Methodology

2.1 K-Nearest Neighbour

K-NN is very simple in the implementation. K-Nearest Neighbour is highly efficient regarding the search space; non linear separability can be achieved with K-Nearest Neighbour. Few parameters to tune distance metric and k-value [17] (Fig. 1).

2.2 Random Forest

Random Forest is ensemble classification algorithm in which features selected are done at random. Among all the available classification methods, random forests provide the highest accuracy. The random forest algorithm can also handle big data

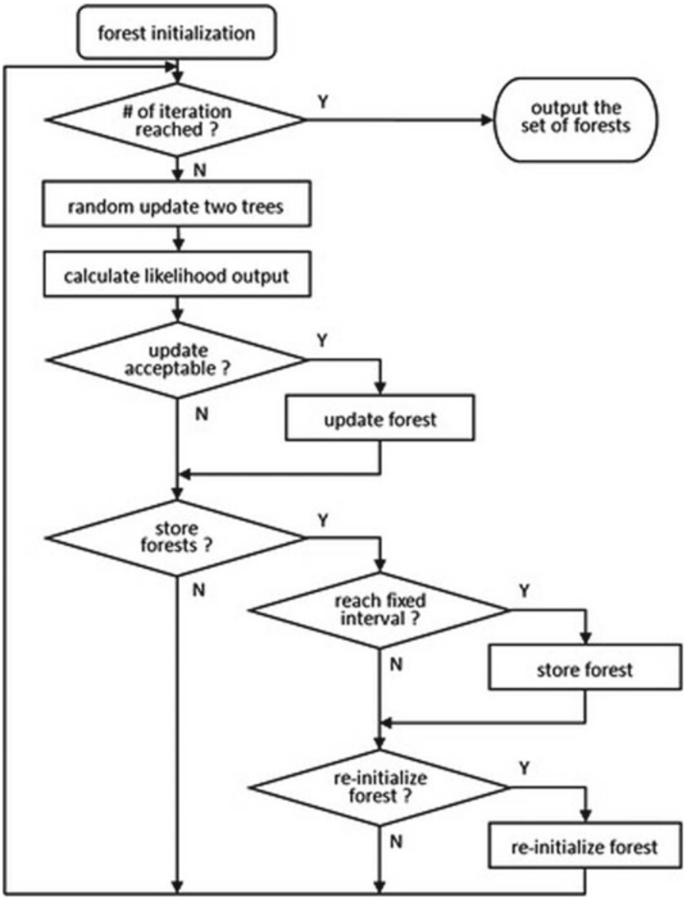


Fig. 2 Random forest Flow chart

with huge number of variables running into thousands. When the data is imbalenced it automatically balance data sets. Random forest also handles variables fast, making it suitable for complicated tasks (Fig. 2).

2.3 Artificial Neural Network

Artificial Neural Networks will identify the patterns in the and has a capability to learn the hidden patterns by them selfes. Input of neural network is stored in the networks instead of a database; hence there is no loss of data and does not affect its working. Learning in neural network is change in weights of neural by backpropagation which uses optimization algorithms such as gradient descent, adam etc. Neural networks can be implemented in parallel by using multiple cores of a processor without affecting the performance of the system (Fig. 3).

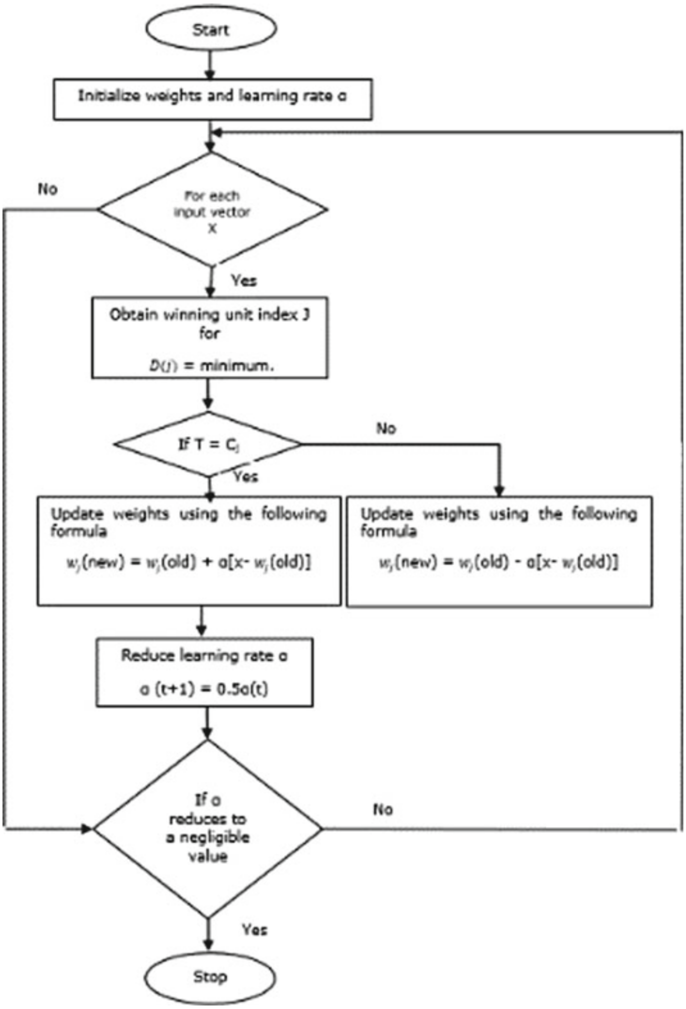


Fig. 3 ANN flow chart

3 Procedure

3.1 Data Set and Pre-processing

Breast cancer Wisconsin Data set [18] is taken from Kaggle website which is source for several datasets. There are 32 parameters and 570 rows present in dataset. Some of parameters present in dataset: Diagnosis of Breast cancer, Radius Mean, Texture Mean, Perimeter Mean, Area Mean. Id number of patients is dropped from data set. Diagnosis is the target variable which consists of two classes (Malignant = 1 or Benign = 0). Malignant means critical, Benign means not harmful. 63% of the diagnosis feature is Benign, 37% is Malignant. The data set has no missing values. Diagnosis feature is categorical in nature. As target feature (Diagnosis) is dichotomous (Malignant and Benign classes) we use Point Biserial Correlation is used to measure the strength of associion between the Independent features and Dependent variable.

- $$r_{pb} = \frac{M_0 - M_1}{s_y}\sqrt{\frac{n_0}{n} * \frac{n_1}{n}} \tag{6}$$
  - $M_0 = Mean\,of\,data\,group\,1$
  - $M_1 = Mean\,of\,data\,group\,2$
  - Sy = Standard deviation of continuous data
  - n_0, n_1 = number of items in respective groups
  - n = total number of elements in two groups.

Diagnosis feature is bimodal distribution. Which means distribution consists of two peaks. Min Max scalar (Eq. 7) is applied to Independent features to limit the values between 0 and 1 (i.e. Scaling of independent features is done using Min Max Scalar).

$$X_{sc} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{7}$$

80% of data is used for training and 20% of data is used for testing. To stop overfitting of data during training cross validation is used with K-fold value as 3.

3.2 Decision Tree

Decision Tree [19] is constructed based on Information Gain and Gini Index. We need to calculate the entropy to obtain Information Gain. Entropy tells us about appropriate measure of the randomness of a system. Lower the value of Entropy (Eq. 8) higher information is obtained by the model.

$$E(T) = \sum_{i=1}^{c} -p_i * \log_2(p_i) \tag{8}$$

$$E(T, X) = \sum_{x \in c} P(C) * E(C) \tag{9}$$

- T refers to Target variable
- X refers to individual attribute.

Target variable in my data set is Diagnosis which consists of two classes Malignant (1) and Benign (0)

$$\begin{aligned} \text{E(Diagnosis)} &= \text{E(Malignant, Benign)} \\ &= \text{E}(357, 212) \\ &= \text{E}(0.62, 0.38) \\ &= -(0.62) * \log_2(0.62) - (0.38) * \log_2(0.38) \\ \text{E(T)} = \ \text{E(Diagnosis)} &= 0.96 \end{aligned}$$

Information Gain (I)

$$I(T, X) = E(T) - E(T, X) \tag{10}$$

Max I (T, X) (Eq. 10) is selected as root node and attribute values are taken as branches to that node [20].

3.3 Gaussian Naïve Bayes

Gaussian NB (Eq. 11) is a variant of Naïve Bayes algorithm. Gaussian NB is generative model which means that the model will learn the join probability distribution [20].

$$p(x_i|y) = \frac{1}{\sqrt{2 * \pi * \sigma_y^2}} * e^{-\frac{(x_i - \mu_y)^2}{2 * \sigma_y^2}} \tag{11}$$

- σ is standard deviation
- μ is mean

From the above formula we could calculate the Gaussian distribution of data to obtain likely hood (L).

Assumptions
Variance is independent of y and x. All features are independent. As dependent feature consists of two classes M and B.

Prior probability of cancer being Malignant (M):

$$\mathrm{P(cancer = M)} = \frac{357}{(357 + 212)} = 0.62$$

Prior probability of cancer being Benign (B):

$$\mathrm{P(cancer = B)} = \frac{212}{(357 + 212)} = 0.38$$

Given a record of a patient we could determine the type of cancer based on the features by calculating the score for M and B

$$\begin{aligned} \text{data} = (&\text{radius-mean} = 17.99,\ \text{texture-mean} = 10.38,\ \text{perimeter-mean} = 122.8, \\ &\text{area-mean} = 1001.0, \ldots,\ \text{fractal-dimension-worst} = 0.1189). \end{aligned}$$

Probability for calculating the given record is Malignant:

$$\begin{aligned} &\mathrm{P\,(cancer = M\,|\,data) = p\,(cancer = M)} \\ &\quad * \mathrm{L\,(radius\text{-}mean = 17.99\,|\,cancer = M)} \\ &\quad * \mathrm{L\,(texture\text{-}mean = 10.38\,|\,cancer = M)} \\ &\quad * \mathrm{L\,(perimeter\text{-}mean = 122.8\,|\,cancer = M)} \\ &\quad * \ldots * \mathrm{L\,(fractal\text{-}dimension\text{-}worst = 0.1189\,|\,cancer = M)} \\ &\mathrm{P\,(cancer = M\,|\,data) = 0.78} \end{aligned}$$

Probability for calculating the given record is Benign:

$$\begin{aligned} &\mathrm{P\,(cancer = B\,|\,data) = P\,(cancer} = B) \\ &\quad * \mathrm{L\,(radius\text{-}mean = 17.99\,|\,cancer = B)} \\ &\quad * \mathrm{L\,(texture\text{-}mean = 10.38\,|\,cancer = B)} \\ &\quad * \mathrm{L\,(perimeter\text{-}mean = 122.8\,|\,cancer = B)} \\ &\quad * \ldots * \mathrm{L\,(fractal\text{-}dimension\text{-}worst = 0.1189\,|\,cancer = B)} \\ &\mathrm{P\,(cancer = B\,|\,data) = 0.22} \\ &\max(P(cancer = M|data), P(cancer = B|data)) = \max(0.78, 0.22) \end{aligned}$$

So, given record of a patient’s cancer is Malignant.

3.4 Artificial Neural Network

ANN [21] is a supervised learning algorithm used for finding certain patterns in the given features. It is also used for classification, regression.

$$\mathrm{y} = \sum_{i=1}^{30} w_i * x_i + b \tag{12}$$

- w is weights
- x is inputs
- b is bias

$$y_{out} = activation(y) \tag{13}$$

Randomly initialize 30 weights.
Consider the following record:

$$\begin{aligned}
\text{record} = (&\text{radius-mean} = 17.99,\ \text{texture-mean} = 10.38,\ \text{perimeter-mean} = 122.8,\\
&\text{area-mean} = 1001.0, \ldots,\ \text{fractal-dimension-worst} = 0.1189).\\
y &= w_1 * \text{radius-mean} + w_2 * \text{perimeter-mean}\\
&\quad + \ldots w_{30} * \text{fractal-dimension-worst}\\
y &= 0.125 * 17.99 + 0.256 * 10.38 + \cdots + 0.584 * 0.1189\\
y &= 10.8755\\
y_{out} &= \frac{1}{1 + e^{-y}}\\
y_{out} &= \frac{1}{1 + e^{-10.8755}}\\
y_{out} &= 0.3215.
\end{aligned}$$

Now by applying backpropagation we update the weights

$$\begin{aligned}
Error(e) &= \sum 0.5 * (target - y)^2\\
\frac{d(Error)}{dw} &= -2(target - y)
\end{aligned}$$

4 Results

See Table 1.

Table 1 Results

Algorithm	Accuracy	AUC	Recall	TP out of 79 positives	TN out of 35 Negatives
K-NN	94.73	0.98	85.71	78	30
LR	95.61	0.99	92.30	73	36
Naïve Bayes	93.85	0.99	92.30	71	36
SVC	92.98	0.99	82.05	74	32
Random forest	95.61	0.99	92.30	73	36

4.1 K-Nearest Neighbor Evaluation

See Figs. 4 and 5.

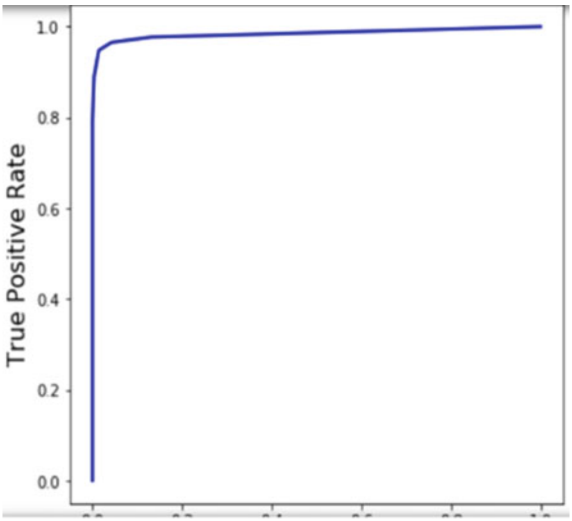
Area under the curve (AUC score): 0.98.

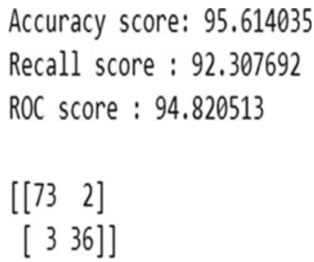
Fig. 4 Evaluation of KNN

```
Accuracy score: 94.736842
Recall score : 85.714286
ROC score : 92.224231

[[78  1]
 [ 5 30]]
```

Fig. 5 KNN ROC curve





```
Accuracy score: 95.614035
Recall score : 92.307692
ROC score : 94.820513

[[73  2]
 [ 3 36]]
```

Fig. 6 Evaluation of logistic regression

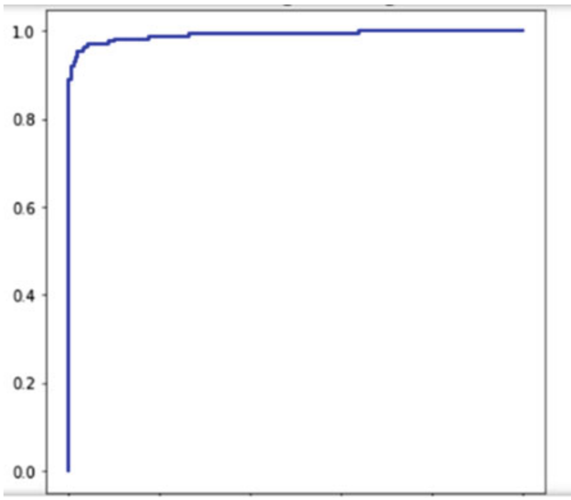


Fig. 7 LR ROC curve

4.2 Logistic Regression Evaluation

See Figs. 6 and 7.

AUC score: 0.99.

4.3 Naïve Bayes Evaluation

See Figs. 8 and 9.

AUC score: 0.99.

```
Accuracy score: 93.859649
Recall score : 92.307692
ROC score : 93.487179

[[71  4]
 [ 3 36]]
```

Fig. 8 Evaluation of Naïve Bayes

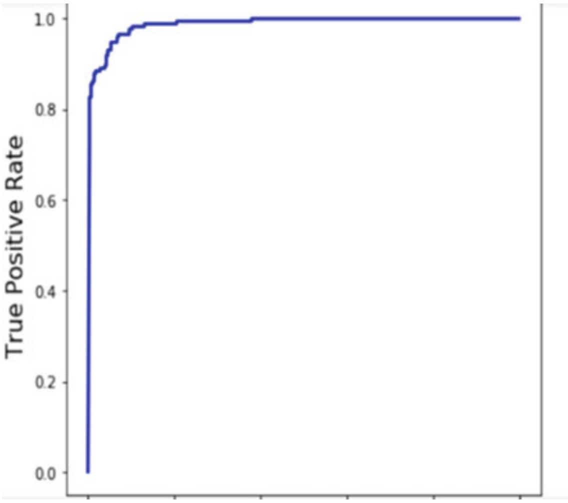


Fig. 9 Naïve Bayes ROC curve

4.4 *SVC Evaluation*

See Figs. 10 and 11.

```
Accuracy score: 92.982456
Recall score : 82.051282
ROC score : 90.358974

[[74  1]
 [ 7 32]]
```

Fig. 10 Evaluation of SVC

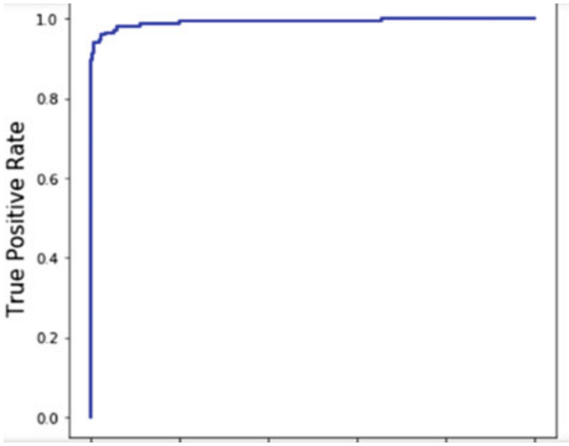


Fig. 11 SVC ROC curve

Fig. 12 Evaluation of decision tree

```
Accuracy score: 94.736842
Recall score : 94.871795
ROC score : 94.769231

[[71  4]
 [ 2 37]]
```

AUC score: 0.99.

4.5 Decision Tree Evaluation

See Figs. 12 and 13.

AUC score: 0.92.

4.6 Random Forest Evaluatıon

See Figs. 14 and 15.

AUC score: 0.99.

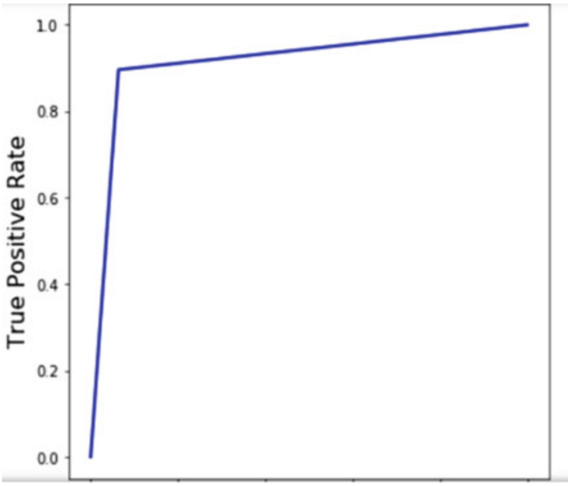


Fig. 13 Decision tree ROC curve

Fig. 14 Evaluation of random forest

```
Accuracy score: 95.614035
Recall score : 92.307692
ROC score : 94.820513

[[73  2]
 [ 3 36]]
```

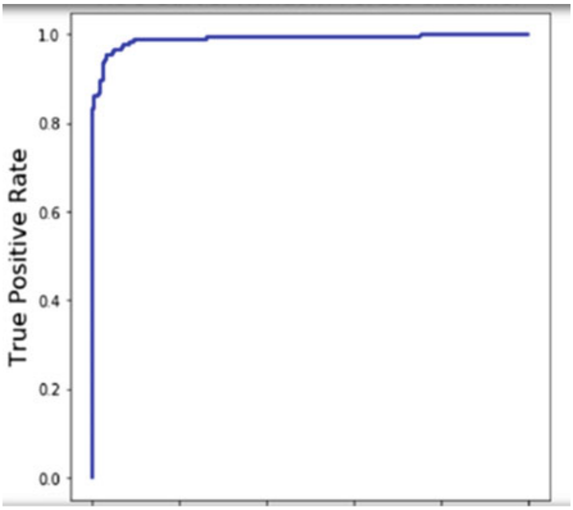


Fig. 15 Random forest ROC curve

5 Conclusion

K-NN, S.V.M, logistic regression [22], Naïve Bayes, and Decision Tree were used in the project. Random forest and Logistic regression have the highest accuracy, recall score, and AUC. Logistic regression and Random forest have 95% accuracy (as mentioned in Figs. 6 and 15) and has a smaller number of false positives and false negative compared to remaining classification algorithms.

References

1. Chelladurai R, Selvakumar R, Poonguzhali S (2018) Automatic segmentation of multiple lesions in ultrasound breast image. Int J Eng Technol (UAE) 7:665–670
2. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. CA Cancer J Clin 66(1):7–30. https://doi.org/10.3322/caac.21332. Epub 2016 Jan 7. PMID: 26742998
3. Sunny J, Rane N, Kanade R, Devi S (2020) Breast cancer classification and prediction using machine learning, Int J Eng Res Technol (IJERT) 09(02):576–580
4. Varsha ML, Kashyap MHP, Bodhith E, Prasad MSR (2020) Prediction of heart diesease using machine learning techniques. Int J Sci Technol Res 9(02):4389–4392
5. Siva Kumar P, Sarvani V, Prudhvi Raj P, Suma K, Nandu D (2017) Prediction of heart disease using multiple regression analysis and support vector machines. J Adv Res Dynam Control Syst 9(18):675–682
6. Srinivas V, Aditya K, Prasanth G, Babukarthik RG, Satheeshkumar S, Sambasivam G (2018) A novel approach for prediction of heart disease: machine learning techniques. Int J Eng Technol (UAE), 7(2.32):108–110
7. Anisha PR, Vijaya Babu B (2018) EBPS: effective method for early breast cancer prediction using Wisconsin breast cancer dataset. Int J Innov Technol Explor Eng 8(2S):205–211
8. Jerez JM, Molina I, García-Laencina PJ, Alba E, Ribelles N, Martín M, Franco L (2010) Missing data imputation using statistical and machine learning methods in a real breast cancer problem. Artif Intell Med 50(2):105–115
9. Ghasem Ahmad L, Eshlaghy A, Pourebrahimi A, Ebrahimi M, Razavi A (2013) Using three machine learning techniques for predicting breast cancer recurrence. J Health Med Inform 4:124–130.
10. Asri H, Mousannif H, Al Moatassime H, Thomas N (2016) Using machine learning algorithms for breast cancer risk prediction and diagnosis. Proc Comput Sci 83:1064–1069
11. Amrane M, Oukid S, Gagaoua I, Ensarİ T (2018) Breast cancer classification using machine learning. In: 2018 electric electronics, computer science, biomedical engineerings' meeting (EBBT), Istanbul, pp 1–4. https://doi.org/10.1109/EBBT.2018.8391453
12. Jadhav M, Thakkar Z, Chawan P (2019) Breast cancer prediction using supervised machine learning algorithms. Int Res J Eng Technol (IRJET) 06(10): 851–854
13. Shravya Ch, Pravalika K, Subhani S (2019) Prediction of breast cancer using supervised machine learning techniques. Int J Innov Technol Explor Eng (IJITEE) 8(6):1106–1110. ISSN: 2278-3075
14. Huang MW, Chen CW, Lin WC, Ke SW, Tsai CF (2017) SVM and SVM ensembles in breast cancer prediction. Plos One 12(1):e0161501
15. Probst P, Wright MN, Boulesteix AL (2019) Hyperparameters and tuning strategies for random forest. Wiley Interdisc Rev Data Mining Knowl Discov 9(3):e1301
16. Banchhor C, Srinivasu N (2016) CNB-MRF: adapting correlative Naïve Bayes classifier and MapReduce framework for big data classification. Int Rev Comput Softw 11(11):1007–1015. https://doi.org/10.15866/irecos.v11i11.10116

17. Rachapudi V, Venkata Suryanarayana S, Subha Mastan Rao T (2019) Auto-encoder based K-means clustering algorithm. Int J Innov Technol Explor Eng 8(5):1223–1226
18. RamyaSri R, IshaSanjida S, Parasa D, Bano S (2019) Food survey using exploratory data analysis. In: 2019 2nd international conference on intelligent communication and computational techniques (ICCT), pp 258–264
19. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu TY (2017) Lightgbm: a highly efficient gradient boosting decision tree. In: Advances in neural information processing systems, pp 3146–3154
20. Roshini T, Sireesha PV, Parasa D, Bano S (2019) Social media survey using decision tree and Naïve Bayes classification. In: 2nd ınternational conference on ıntelligent communication and computational techniques (ICCT), Jaipur, India, pp. 265–270 (2019)
21. Longo GA, Zilio C, Ortombina L, Zigliotto M (2017) Application of Artificial Neural Network (ANN) for modeling oxide-based nanofluids dynamic viscosity. Int Commun Heat Mass Transfer 83:8–14
22. Gurram D, Narasinga Rao MR (2017) A comparative study of support vector machine and logistic regression for the diagnosis of thyroid dysfunction. Int J Eng Technol (UAE) 7(1.1):326–328

A Synopsis of Monocular Depth Estimation

Shubham Chaudhari, Aaryamaan Rao, Rohit Vardam, and Mandar Sohani

Abstract Depth information is necessary for automated devices or software to accomplish tasks which require the knowledge of the surrounding environment without error. Monocular depth estimation uses a single 2D image to estimate depth, which makes it an exigent method. However, recent studies in monocular depth estimation based on convolutional neural networks indicate favorable performance in accuracy. Research on monocular depth training traverses through increasingly complex model architectures, and optimization of loss functions, all of these have recently helped to close the gap between the accuracy percentage of the traditional methods and trained models. In this review paper, we survey numerous existing literature on monocular depth estimation along with various datasets, as well as several supervised, unsupervised and semi-supervised algorithms. Furthermore, we figured out the drawbacks of the existing traditional methods and discussed contemporary methods using convolutional neural networks which may already have improved upon the said drawbacks. Lastly, we present our own findings on further scope in monocular depth estimation.

Keywords Monocular depth · Computer vision · Depth maps · SIDE · CNN · UNet

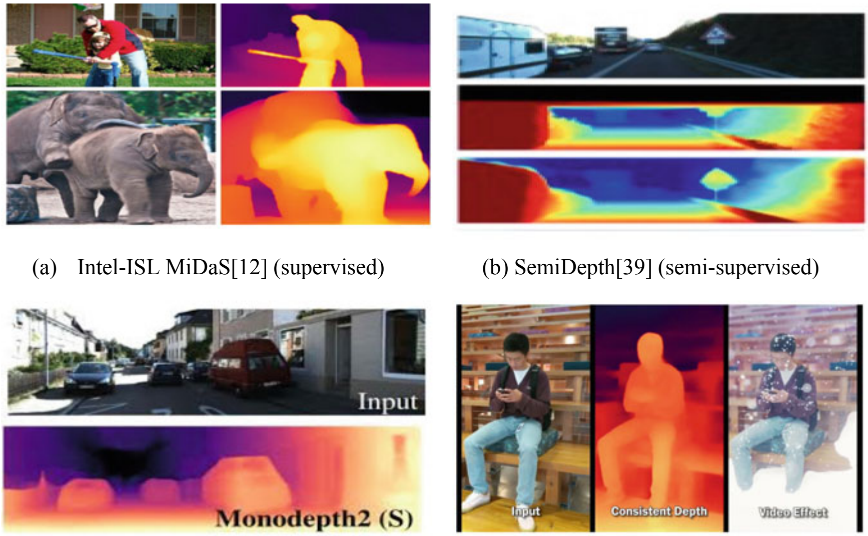
1 Introduction

Depth estimation is an essential requisite for several important applications such as scene interpretation and 3-dimensional scene reconstruction. This computer vision task also has far reaching applications in varied fields such as autonomous vehicles, robotics, as well as film production [1]. This paper aims to explore and review the current research and developments in the monocular depth estimation field. Depth information is essential for autonomous systems to perceive environments

S. Chaudhari (✉) · A. Rao · R. Vardam · M. Sohani
Mumbai University, Mumbai, India

M. Sohani
e-mail: mandar.sohani@vit.edu.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_18



(a) Intel-ISL MiDaS[12] (supervised) (b) SemiDepth[39] (semi-supervised)

(c) Monodepth2[16] (self-supervised) (d) Consistent depth[15] (supervised)

Fig. 1 Illustration of extracted depth maps from monocular images using various methods

and estimate their own state. Traditional depth estimation methods, such as identifying structure from motion and stereovision matching, are dependent on feature correspondences of multiple viewpoints. However, the predictions are sparse depth maps (Fig. 1).

Inferring depth information from a single image (monocular depth estimation) is often referred to as an ill-posed problem [2]. This is largely because a single two-dimensional image captured from a monocular camera system setup may be projected to an infinitely many number of three-dimensional scenes. As depth estimation from multiple viewpoint images started achieving significant breakthroughs, the focus also fell on estimation from a single viewpoint since it has reduced computational cost and constraints in comparison with multi-viewpoint depth estimation systems or stereoscopic depth estimation systems [1]. With the rapid development of deep neural networks, monocular depth estimation based on deep learning (using CNNs) has been widely studied recently and has achieved promising performance in accuracy [3]. A depth map is an image object that contains information about the distance between the surfaces of objects from a given viewpoint. A depth map is extracted from a source image and is typically maintained in grayscale format, but RGB-D maps are also common. Dense depth maps are predicted from monocular images by DCNNs in an end-to-end manner. In order to upgrade the accuracy of depth estimation, different kinds of network frameworks, loss functions, and training strategies are employed as per which fits the specific scenario the best. Convolutional neural networks (CNNs) represent an interesting method for self-learned image processing and form a link between general feed-forward neural networks and adaptive filters.

Many deep convolutional neural network-based approaches have achieved significantly encouraging results with the help of sufficiently large quanta of data, relying on prior knowledge to tackle unknown depth information. CNNs possess salient features of translation invariance and spatially local connections (receptive fields). Owing to these factors, CNNs represent a very convenient and modern approach for reconciling the Monocular Depth Estimation problem. Estimating depth from a single image is a difficult task even for humankind as we rely on a variety of factors to perceive depth including texture, lighting angles, occlusion, object scales, perspectives, etc. and also account for the entire scene by considering the total shapes and dimensions [1]. The most common approaches for the MDE problem so far have been supervised learning tasks. Recently we have also observed semi self-supervised or fully self-supervised approaches being proposed. The common CNN architecture consists of an encoder setup which is responsible for dense feature extraction, and a second decoder setup which up-samples the desired dense depth information based on the features to output a depth map. Some of the architectures which have given successful results in depth feature extraction and prediction include VGG, ResNet and DenseNet [1]. In these networks, repeated stride convolution and spatial pooling layers lower the spatial resolution of transitional outputs, which can be a bottleneck to obtain desired dense predictions. Some newer approaches combine traditional Convolutional Neural networks with Conditional Random Field i.e. CRF models [2]. This approach leads to better edge-conforming depth maps as output.

2 Datasets

While compiling this review paper, we researched work done by other researchers using different datasets. Their findings have been compiled below, subject to certain limitations. Each dataset is unique with different numbers of scenes, resolutions, settings, objects, illumination, etc. The standard widely used evaluation metrics are RSME, RSME log, Abs rel, and Sq. Rel, Accuracies [4].

2.1 NYU-Depth V2

The NYU-Depth V2 dataset by Silberman et al. [5] consists of video sequences which focus on indoor scenes and there are more than 464 scenes in this dataset taken from 3 different cities. Depth cameras from the Microsoft Kinect have been used to record RGB images and the ground truth labels for the same, producing 1449 densely labeled pairs of aligned RGB input images and targeted depth images. Each image is of resolution 640×480 which is rescaled in preprocessing while experimenting. This dataset as well consists of only RGB, and depth maps similar to Make3D.

2.2 KITTI Dataset

The KITTI dataset [6] has become a standard dataset used in innumerable machine vision problems. It is used for testing as well as evaluation, and the KITTI evaluation metrics have become commonly used benchmarks. This dataset can be used for optical flow, visual odometry, depth, object detection, semantic segmentation and tracking [4, 7–11]. Of the total 56 scenes, 28 scenes are used for training and the remaining for testing, by Eigen et al. [4]. The RGB images are classified into labels such as “city”, “residential” and “road”. Each scene consists of stereo image pairs with a resolution of 1242 × 375 pixels. A rotating LIDAR sensor is employed to perform sparse sampling on the depths of the images. This dataset was introduced in IJRR in 2013. Using the Eigen split, where KITTI dataset is split into two parts, consisting of 23,488 images for training, and 697 images for testing taken from 32 and 29 scenes respectively is the most commonly employed approach for this dataset.

2.3 Make3D

The Make3D dataset [12] contains 534 outdoor images, 400 for training and 134 for testing. The resolution of these images are 2272 × 1704 while the ground truth depth map resolutions are 55 × 305. Generally, input images are downsampled before feeding it to a neural network which helps to reduce heavy computation and train faster. This dataset only provides monocular RGB images and its respective depth maps. Since, there are no monocular sequences or stereo image pairs, it can not be used for semi-supervised and unsupervised learning methods for training. Instead, it is widely used as a testing set of unsupervised algorithms to evaluate the generalization ability of networks on different datasets [13]. Make3D dataset is suitable for supervised methods for training.

2.4 Gamehook Mod

Gamehook is a plugin [14] created by Philipp Krähenbüh which can extract optical flow, depth, and semantic segmentation from the game. Due to this, an unlimited amount of data is available for training as well as testing. The current available dataset is of 220 k images for training, 60 k images for testing and has been extracted from three different video games. It gives the advantage of getting the training images exactly as we need by changing the setting of the game using other mods which can control weather, texture, crowd density, lighting, season, etc. This plugin can extract state of the art optical flow, depth map, and segmentation. Since GTA V is a hyper realistic game, data extracted is visually closer to real world images, than other synthetic dataset. Data extracted from three different video games is as follows: 10 k

Table 1 Brief comparison between different datasets

Datasets	Statistics	Annotation
KITTI	94 k	Depth aligned with raw data
NYU depth V2	1449 + 470 k raw	Depth maps
SYNTHIA	220 k	Depth, Segmentation, 2D and 3D boxes for objects
Make3D	534	Depth maps
Gamehook	220 k + 60 k	Depth, optical flow, segmentation

images from FarCry Primal, 200 k images from GTA V, and 10 k from The Witcher 3.

2.5 Synthia Dataset

Synthia dataset [15] is one of the biggest dataset comprising more than 220 k sample images of RGB along with the targets such as depth, ego motion, and semantic segmentation. Synthia dataset is a collection of photo realistic frames rendered from a virtual environment having a variety of dynamic classes such as sky, building, poles, cars, pedestrian, cyclist, mist, road, sidewalk, fence, vegetation, sign, and lane markings enhancing the details and increasing the scope broadly. In addition to this, it has multiple dedicated seasons such as fall, spring, winter, and summer, dynamic lighting conditions, and weather. 'SYNTHIA-AL-TRAIN' dataset consists of more than 96 thousand sample images of resolution 640×480 (Table 1).

3 Methods

The methods to solve the MDE problem are numerous, but in order to simplify them for the sake of comprehension, they may be classified into three main groups—supervised, semi-supervised and unsupervised learning approaches.

3.1 Supervised

Supervised learning is a process to map the training dataset to target dataset via a model. Data constraints for supervised learning need to be structured data. Techniques such as Support Vector Machine (SVM), Linear regression, Logistic Regression, Classification Neural networks for example CNN, KNN, etcetera come under supervised learning.

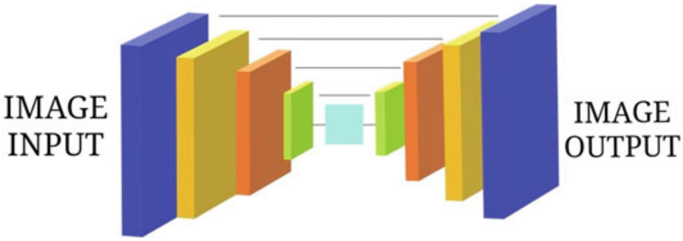


Fig. 2 UNet encoder decoder architecture

If X is the training dataset {x1, x2, x3, …, xi} consisting of i sample and for each sample there exist a target label or class {y1, y2, y3, …, yi} called Y; then supervised algorithm attains to learn G: X → Y, where X is the input space and Y is the output Space, and G is the hypothesis space. In supervised learning, image to image translation is attained using UNet architecture.

In UNet architecture, Fig. 2, there are two main components called encoder and decoder. Encoder has successive convolutional layers which yield low-resolution, high level features whereas decoders reconstruct the image using these features. Decoder has successive upsampling convolutional layers resulting in an increase of resolution of the output image. A successive convolutional layer can extract dense features and then learn to assemble a desired output precisely based on this information. Accuracy of the output images depends on the model architecture and loss function. Research has shown that to increase the accuracy of the supervised model, encoder and decoder must be optimized. Already pre-train feature extractor models such as EfficientNet, DenseNet, ResNet, Inception, MobileNet SSD, VGG etc. as depth encoders further integrating it with a decoder. Other methods such as Conditional Random Fields (CRFs), multi-planar local guidance, ordinal regression, probabilistic estimation, etc. can be used as training strategies.

3.2 Unsupervised

Unsupervised learning is a type of machine learning that looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision. In contrast to supervised learning that usually makes use of human-labeled data, unsupervised learning, also known as self-organization allows for modeling of probability densities over inputs (Fig. 3).

Recent research suggests that using an unsupervised pipeline for depth estimation is possible with the use of stereo image pairs using a photometric warp loss function, in order to replace a loss based on ground truth depth. Godard et al. [17] used binocular stereo pairs and trained a network to predict the depth that minimises the photometric difference between the true right image and one generated by warping the left image into the right's viewframe, using the predicted depth. Godard et al. [17]

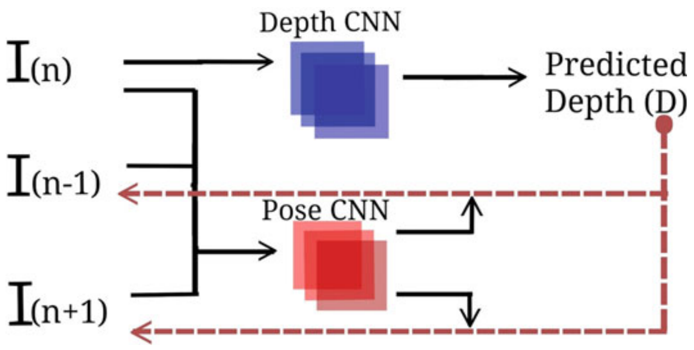


Fig. 3 Illustration of unsupervised learning of depth [16]

optimised the depth prediction by introducing a symmetric left–right consistency criterion and an improved stereo loss function. Their image generation model is not fully differentiable. To compensate this limitation, a Taylor approximation is employed to linearize the resultant loss in an objective that is more challenging to optimize viz. scaling ambiguity error.

Addressing the problem of unsupervised learning, Ranjan et al. [18], uses a framework in which networks learn to "collaborate" and "compete". In their paper, Competitive collaboration was framed as a three player game consisting of two players competing for a resource (datasets) that is regulated by the moderator. These two players compete for training data by analyzing static scene pixels and dynamic region pixels in a scene from an image sequence. This is moderated by a motion segmentation network that segments the static scene and dynamic regions, and distributes training data between the two players. The moderator needs to be properly trained to ensure a fair distribution. Therefore, the two players collaborate to train the moderator, such that it classifies static and moving regions correctly in alternating phases of the training cycles.

3.3 Semi-supervised

Unlabeled data, when used in conjunction with a small amount of labeled data, can produce considerable improvement in learning accuracy. The cost associated with the labeling process thus may render large, fully labeled training sets infeasible, whereas acquisition of unlabeled data is relatively inexpensive. In such situations, semi-supervised learning Fig. 4 can be of great practical value.

Luo et al. [19] demonstrates another semi-supervised learning approach. In their research they split the task of monocular depth estimation into two separate sub-tasks viz. view synthesis and stereo matching. Both procedures obey primary geometric principles and they could be trained even in the absence of expensive training data. The first stage is view synthesis. For a stereo pair, binocular views are rendered by two

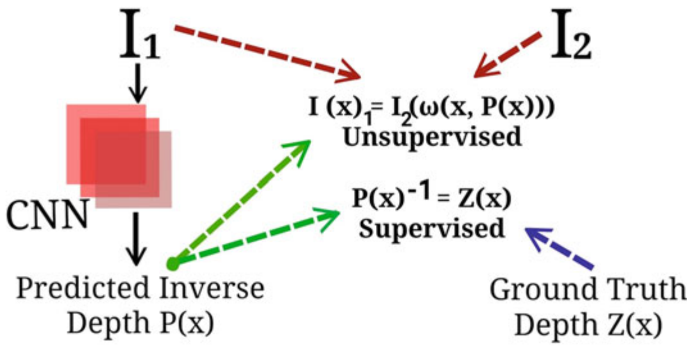


Fig. 4 Illustration of Semi-supervised general framework proposed in [11]

specifically synchronized and calibrated cameras, leading to a strong correspondence between pixels in the horizontal axis. They propose a new probabilistic scheme for transferring pixels from the original image. By this means, the transformation is directly generated from left image to right image through the use of a differentiable selection layer. After generating a high-quality view, their stereo matching network transforms the high-level scene analysis problem into a one dimensional matching problem, which significantly reduces their computational complexity.

3.4 Self-supervised

Self supervised training relies highly on the dataset that the model is being trained on because the data provides the supervision. Monocular training is significantly more accurate with diverse and large training sets [20]. However, collection of such varied and large datasets with accurate ground truth depth, necessary for supervised learning [4, 13] is a daunting challenge. Recent developments in various self-supervised approaches have displayed that it is instead possible to train monocular depth estimation models utilizing only synchronized stereo pairs [17, 21] or monocular video [22]. Between the two self-supervised approaches, monocular video is an appealing alternative to stereo-based supervision but it introduces it's own set of challenges. Adding to estimating depth, the model also needs to estimate ego motion between temporal image pairs during training. [23]. Godard et al. [23] introduced three architectural and loss innovations that when applied together would lead to significant improvements in monocular depth estimation training with monocular video, stereo pairs or both.

While recent techniques in self-supervised monocular depth estimation are almost equal to supervised methods in performance, they only pertain to low resolution images. Pillai et al. [24] spoke of a sub-pixel convolutional layer extension for depth super-resolution that accurately super-resolves the daunting problem of disparities from their low-resolution outputs. They also introduced a differentiable flip-augmentation layer that permitted the disparity model to learn an improved prior for disparities at image boundaries in an end-to-end fashion.

4 Prominent Work Done

4.1 Intel-ISL MiDaS

Monocular depth estimation improves with larger and more diverse training sets [14]. Intel Intelligent Systems Lab trained a robust monocular depth estimation model adopting five distinct datasets along with an enormous data source viz. 3D films [14]. Ranftl et al. [20] showcased the quantification of the value of a diverse set of existing datasets for monocular depth estimation and investigated optimal strategies for mixing datasets during training. Furthermore, it was shown that a principled approach based on multi-objective optimization [25] led to better results when compared to a simple mixing strategy [20]. Finally, Ranftl et al. [20] highlighted the significance of high-capacity encoders and the irrational effectiveness of pretraining the encoder on a large-scale auxiliary task, based on their experience.

Ranftl et al. [20] confirmed that mixing the data from complementary sources substantially enhanced monocular depth estimation. MiDaS outperformed competing methods across diverse datasets, setting a new milestone as seen in Fig. 1a for monocular depth estimation.

4.2 Consistent Video Depth Estimation

The traditional reconstruction of depth for all pixels in a monocular video used ad-hoc priors and was geometrically inconsistent. Luo et al. [26] showcased an algorithm that reconstructed dense and geometrically consistent depth for all the pixels in a monocular video. Luo et al. [26] trained for single image depth estimation using a convolutional neural network. They presented a new video-based reconstruction system that combined the strengths of classic and learning-based techniques. The geometric constraints were obtained using traditional methods where they were obtainable to maintain a consistent and accurate depth, while the learning based priors were used to fill the gap left by weakly constrained parts of the scene more reasonably than prior heuristics [23]. This was implemented by fine-tuning the weights of a single-image depth estimation network at test time. This test-time training strategy allowed them to use both short-term and long-term constraints and helped in the prevention of drifting over time.

In Fig. 1d shows sample screenshots taken from the videos along with the corresponding screenshots from the depth videos provided by Luo et al. [26]. The depth videos are fully dense and detailed, with sharp object boundaries. The reconstruction is flicker-free and geometrically consistent throughout the video. The improved quality and consistency of the depth videos by Luo et al. [26] show great potential for use in interesting new applications such as fully automatic video special effects that interact with the dense scene content.

5 Comparative Study

See Table 2.

6 Scope

MDE as it currently stands requires high-end hardware for its heavy computations. However, judging by how quickly GPU technologies are progressing, in the near future MDE algorithms and models are likely to be trained faster, and more experiments would be performed leading to optimization and faster estimation of depths. On achieving a detailed depth map output with bare minimum hardware requirements with less heavy computations, not only can Monocular Depth Estimation replace expensive sensors which are not always feasible, but will also provide more opportunities for this method to be implemented in various other fields.

6.1 3D Reconstruction and Point Cloud

A point cloud is a set of data points in three dimensional space having 3 freedom of degree. Point clouds form the initial base for 3D modelling as well as it helps to position 3D objects in the environmental space relative to the starting point into a scene. Reconstruction of the image with its respective depth map, a 3D perspective point cloud can be constructed [29].

6.2 Virtual Effects

The Depth-of-Field blur effect simulates the visual blurring of specific depth region pixels of objects located in front or back of the focus point in an image, while leaving other depths alone. This allows to refocus the image on certain objects or use artistic blurring. The focus point is associated with a focal distance. Any arbitrary focus

Table 2 A synopsis of diverse monocular depth estimation models

Model	Contribution/main focus	Year	Method
BTS [1]	Multi scale local planar guidance	2019	Supervised
DORN [27]	Ordinal regression	2018	Supervised
VNL [3]	Geometric constraints with DL	2019	Supervised
SOM [28]	Structure-attentioned memory network	2019	Supervised
DenseDepth [29]	Transfer learning + optimization	2018	Supervised
ACAN [30]	Deep-CNN-based method	2019	Supervised
SharpNet [31]	Multitask encoder-decoder network: occluding contours, depth, normal	2019	Supervised
PAP-depth [32]	Affinity layer learning	2019	Supervised
SARPN [33]	Multi scale depth map prediction using residual pyramid decoder	2019	Supervised
FastDenseNas [34]	Parameterisation of small networks	2018	Supervised
SENet-154 [35]	Deep-CNN-based method	2018	Supervised
FCRN_ROB [36]	Conditional random fields	2016	Supervised
RelativeDepth [2]	Novel algorithm for MDE using relative depth maps	2019	Supervised
Index network [37]	Index-guided encoder-decoder framework	2019	Supervised
RefineNet [38]	Optimized multitask encoder-decoder network: depth and segmentation	2018	Supervised
Eigen et al. [4]	CNNs	2014	Supervised
BESEG [39]	Modular two stage pipeline: depth, segmentation	2020	Supervised
AcED [40]	CNNs	2020	Supervised
GASDA [41]	Reducing the domain discrepancy using the bidirectional image style transfer	2019	Supervised
VDA [42]	Temporally consistent depth prediction	2019	Supervised
ProbMonoDepth [43]	Probabilistic monocular depth	2019	Supervised
SVS [19]	Stereo matching	2018	Semi-supervised
FIS-Nets [44]	CNN	2020	Semi-supervised
SemiDepth [45]	Left–right consistency	2019	Semi-supervised
3Net [46]	Trinocular assumption and network design	2018	Unsupervised
SIGNet [47]	Robust depth and flow perception without using geometric labels	2018	Unsupervised
Struts2depth [48]	Refinement: adapts learning and can transfer to new datasets or environments	2018	Unsupervised
CC [18]	Compete and cooperate network strategy	2018	Unsupervised
Monodepth [17]	Left–right disparity consistency	2017	Unsupervised

(continued)

Table 2 (continued)

Model	Contribution/main focus	Year	Method
SfMLearner [49]	Novel method to select and weakly rectify image pairs for better training	2020	Unsupervised
SDNet [50]	Multi task CNN architecture bases on Deeplabv3	2019	Unsupervised
SfMLearner [51]	Optimized CNNs architecture for unsupervised learning	2017	Unsupervised
LKVOLearner [52]	Differentiable implementation of DVO, a novel depth normalization strategy	2017	Unsupervised
DOPlearning [53]	Joint unsupervised training on depth, optical flow, and pose	2020	Unsupervised
SC-SfMLearner [54]	Geometry consistency	2019	Unsupervised
DeepLabV3+ [55]	Depth cues	2020	Unsupervised
Depth-VO-Feat [56]	Novel feature reconstruction loss	2018	Unsupervised
VOMonodepth [57]	Visual odometry	2019	Self-supervised
monoResMatch [58]	CNN, infusing stereo	2019	Self-supervised
SuperDepth [24]	Sub-pixel convolutions within a disparity estimation	2018	Self-supervised
SGDepth [59]	Novel semantic masking technique to improve depth for moving objects	2020	Self supervised
Depth hints [60]	Fused semi-global matching	2019	Self-supervised
SelfDepthNorm [61]	Multi task visual odometry network	2019	Self-supervised
RefinedMPL [62]	Accurate point clouds modelling for 3D detection	2019	Self-supervised
LSIM [63]	Siamese network for depth estimation	2019	Self-supervised
SemanticAware [16]	Semantics consistency between stereo pairs	2019	Self-supervised
PackNet-SfM [64]	Novel depth network	2019	Self-supervised
Monodepth2 [23]	Improved self-supervised depth estimation	2018	Self-supervised

point can be selected from the depth map to get the associated focal distance to create Depth-of-Field blur effect.

Different realistic simulations Fig. 1d can be simulated by superimposing it on an image using a depth map. For example, in a snow fall simulation it will take care of the object's structure. If there is a table object, snow will accumulate on the table while snow will accumulate on the floor, if not the table, giving the realistic meaning to the simulation.

6.3 Object Trajectory

The path of an object can be simulated realistically that it would take through a scene using the depth map. For example, if a ball were to roll through the scene at a certain depth, it would pass in front of or behind different objects, depending on their relative ordering in space.

6.4 Absolute Depth

Measuring the absolute depth of the objects just from 2D images has been a challenge. But, only relative depth of objects is extracted using depth estimation. Although, it is possible to estimate absolute depth as well with some prior information such as the sizes of some of the objects as well as calibrating the camera initially.

6.5 Parallax Simulation

Generating a new perspective of the scene which does not exist in the initial input images can be possible with the help of depth map. Though, this can be done for small variation of degrees only by extending the small known parallax to a larger one, as if a more shifted image were taken.

6.6 Drawbacks of Traditional Methods

LiDAR systems have been the most common and reliable way for various applications that require depth estimation. It works by first illuminating the target with a laser light and then measuring the reflection of it using an active sensor. The differences in laser return times are used for constructing a 3D representation. Although LiDAR has been cheap for large applications in recent times, it still remains very expensive for smaller application areas.

Another common method is depth estimation using stereo vision which involves the concept of triangulation and stereo matching which works well in many environments. However, the depth information obtained from stereo cues is dependent on the difference between the pair of images and is therefore limited by the distance between the two cameras. The accuracy tends to worsen as the distances considered are larger because even miniscule estimation errors in angles, translate into large errors in distances. Monocular cues on the other hand such as texture variations and color gradients contain useful and important depth information that can be extracted from a single image.

7 Conclusion

Monocular depth estimation has become a prominent field in research over the years. Our focus in this survey was to create a base for others to begin working on monocular depth estimation with an ease of access to numerous resources to aid an in depth exploration of the field. We overviewed and inspected several top open source datasets. In addition to this, we have also performed an exhaustive literature survey of existing published papers on MDE, classifying them based on their training method types as supervised, semi-supervised, self-supervised, or unsupervised. Significant analysis was performed on numerous published papers indicating their training method, type of approach and the main contribution of the paper. We also discussed the keen future scope of monocular depth estimation.

References

1. Han JH, Han M-K, Ko DW, Suh IH, From big to small: multi-scale local planar guidance for monocular depth estimation
2. Lee J-H, Kim C-S, Monocular depth estimation using relative depth maps. Korea University
3. Yin W, Liu Y, Shen C, Yan Y (2019) Enforcing geometric constraints of virtual normal for depth prediction. In: ICCV
4. Eigen D, Puhrsch C, Fergus R (2014) Depth map prediction from a single image using a multi-scale deep network. In: NeurIPS
5. Silberman N, Hoiem D, Kohli P, Fergus R (2012) Indoor segmentation and support inference from RGBD images. In: ECCV. Courant Institute, New York University
6. Geiger A, Lenz P, Urtasun R (2012) Are we ready for autonomous driving? the kitti vision benchmark suite. In: 2012 IEEE conference on computer vision and pattern recognition. IEEE, pp 3354–3361
7. Mayer N, Ilg E, Hausser P, Fischer P, Cremers D, Dosovitskiy A, Brox T (2016) A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4040–4048
8. Zhao C, Tang Y, Sun Q (2019) Deep direct visual odometry. arXiv preprint arXiv:1912.05101
9. Wang P, Chen P, Yuan Y, Liu D, Huang Z, Hou X, Cottrell G (2018) Understanding convolution for semantic segmentation. In: 2018 IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1451–1460
10. Chang M-F, Lambert J, Sangkloy P, Singh J, Bak S, Hartnett A, Wang D, Carr P, Lucey S, Ramanan D et al. (2019) Argoverse: 3d tracking and forecasting with rich maps. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 8748–8757
11. Kuznietsov Y, Stuckler J, Leibe B Semi-supervised deep learning for monocular depth map prediction
12. Saxena A, Chung SH, Ng AY NIPS (2005) Learning depth from single monocular images
13. Saxena A, Sun M, Ng AY, Make3D: learning 3D scene structure from a single still image. IEEE Trans Pattern Anal Mach Intell (PAMI)
14. Krahenbuhl P (2018) Free supervision from video games. In: 2018 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Salt Lake City, UT
15. Ros G, Sellart L, Materzynska J, Vazquez D, Lopez AM (2016) The SYNTHIA dataset: a large collection of synthetic images for semantic segmentation of urban scenes. In: IEEE CVPR
16. Chen P-Y, Liu AH, Liu Y-C, Wang Y-CF, Towards scene understanding: unsupervised monocular depth estimation with semantic-aware representation

17. Godard C, Mac Aodha OM, Brostow GJ (2017) Unsupervised monocular depth estimation with left-right consistency. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 270–279
18. Ranjan A, Jampani V, Balles L, Kim K, Sun D, Wulff J, Black MJ (2019) Competitive collaboration: joint unsupervised learning of depth, camera motion, optical flow and motion segmentation. In: CVPR
19. Luo Y, Ren J, Lin M, Pang J, Sun W, Li H, Lin L (2018) Single view stereo matching. In: CVPR
20. Ranftl R, Lasinger K, Hafner D, Schindler K, Koltun V, Towards robust monocular depth estimation: mixing datasets for zero-shot cross-dataset transfer
21. Garg R, Vijay Kumar BG, Reid I (2016) Unsupervised CNN for single view depth estimation: geometry to the rescue. In: ECCV
22. Zhou T, Brown M, Snavely N, Lowe D (2017) Unsupervised learning of depth and ego-motion from video. In: CVPR
23. Godard C, Mac Aodha O, Firman M, Brostow GJ (2019) Digging into self-supervised monocular depth estimation. In: Proceedings of the IEEE/CVF international conference on computer vision (ICCV), pp. 3828–3838
24. Pillai S, Ambrus R, Gaidon A, SuperDepth: self-supervised, super-resolved monocular depth estimation. Toyota Research Institute
25. Sener O, Koltun V (2018) Multi-task learning as multi-objective optimization. In: NeurIPS
26. Luo X, Huang J-B, Szeliski R, Matzen K, Kopf J (2020) Consistent video depth estimation
27. Fu H, Gong M, Wang C, Batmanghelich K, Tao D (2018) Deep ordinal regression network for monocular depth estimation. In: CVPR
28. Zhu J, Shi Y, Ren M, Fang Y, Lien K-C, Gu J, Structure-attentioned memory network for monocular depth estimation
29. Alhashim I, Wonka P, High quality monocular depth estimation via transfer learning
30. Chen Y, Zhao H, Hu Z, Attention-based context aggregation network for monocular depth estimation
31. Ramamonjisoa M, Lepetit V, SharpNet: fast and accurate recovery of occluding contours in monocular depth estimation
32. Zhang Z, Cui Z, Xu C, Yan Y, Sebe N, Yang J (2019) Pattern-affinitive propagation across depth, surface normal and semantic segmentation. In: CVPR
33. Chen X, Chen X, Zha Z-J, Structure-aware residual pyramid network for monocular depth estimation
34. Nekrasov V, Chen H, Shen C, Reid I (2019) Fast neural architecture search of compact semantic segmentation models via auxiliary cells. In: CVPR
35. Hu J, Ozay M, Zhang Y, Revisiting single image depth estimation: toward higher resolution maps with accurate object boundaries. Takayuki Okatani
36. Laina I, Rupprecht C, Belagiannis V, Tombari F, Navab N, Deeper depth prediction with fully convolutional residual networks
37. Lu H, Dai Y, Shen C, Xu S, Index network
38. Nekrasov V, Dharmasiri T, Spek A, Drummond T, Shen C, Reid I, Real-time joint semantic segmentation and depth estimation using asymmetric annotations
39. Ng MH, Radia K, Chen J, Wang D, Gog I, Gonzalez JE, BEV-seg: bird's eye view semantic segmentation using geometry and semantic point cloud. University of California, Berkeley
40. Swami K, Bondada PV, Bajpai PK, ACED: accurate and edge-consistent monocular depth estimation
41. Zhao S, Fu H, Gong M, Tao D (2019) Geometry-aware symmetric domain adaptation for monocular depth estimation. In: CVPR
42. Atapour-Abarghouei A, Breckon TP, Veritatem dies aperit-temporally consistent depth prediction enabled by a multi-task geometric and semantic scene understanding approach
43. Xia Z, Sullivan P, Chakrabarti A, Generating and exploiting probabilistic monocular depth estimates
44. Wang B, An J, FIS-nets: full-image supervised networks for monocular depth estimation

45. Amiri AJ, Loo SY, Zhang H, Semi-supervised monocular depth estimation with left-right consistency using deep neural network
46. Poggi M, Tosi F, Mattoccia S, Learning monocular depth estimation with unsupervised trinocular assumptions
47. Meng Y, Lu Y, Raj A, Sunarjo S, Guo R, Javidi T, Bansal G, Bharadia D, SIGNet: semantic instance aided unsupervised 3D geometry perception. In: CVPR
48. Casser V, Pirk S, Mahjourian R, Angelova A, Depth prediction without the sensors: leveraging structure for unsupervised learning from monocular videos
49. Bian J-W, Zhan H, Wang N, Chin T-J, Shen C, Reid I, Unsupervised depth learning in challenging indoor video: weak rectification to rescue
50. Ochs M, Kretz A, Mester R, SDNet: semantically guided depth estimation network
51. Zhou T, Brown M, Snavely N, Lowe DG Unsupervised learning of depth and ego-motion from video
52. Wang C, Buenaposada JM, Zhu R, Lucey S, Learning depth from monocular videos using direct methods. Carnegie Mellon University
53. Wang G, Zhang C, Wang H, Wang J, Wang Y, Wang X, Unsupervised learning of depth, optical flow and pose with occlusion from 3D geometry
54. Bian J-W, Li Z, Wang N, Zhan H, Shen C, Cheng M-M, Reid I (2019) Unsupervised scale-consistent depth and ego-motion learning from monocular video. In: NeurIPS
55. Tel SG, Wolf L, Single image depth estimation trained via depth from defocus cues
56. Zhan H, Garg R, Weerasekera CS, Li K, Agarwal H, Reid I (2018) Unsupervised learning of monocular depth estimation and visual odometry with deep feature reconstruction. In: CVPR
57. Andraghetti L, Myriokefalitakis P, Dovesi PL, Luque B, Poggi M, Pieropan A, Mattoccia S, Enhancing self-supervised monocular depth estimation with traditional visual odometry
58. Tosi F, Aleotti F, Poggi M, Mattoccia S, Learning monocular depth estimation infusing traditional stereo knowledge
59. Klingner M, Termöhlen J-A, Mikolajczyk J, Fingscheidt T, Self-supervised monocular depth estimation: solving the dynamic object problem by semantic guidance
60. Watson J, Firman M, Brostow GJ, Turmukhambetov D, Self-supervised monocular depth hints
61. Zhan H, Weerasekera CS, Garg R, Reid I, Self-supervised learning for single view depth and surface normal estimation
62. Vianney JMU, Aich S, Liu B, Refined monocular PseudoLiDAR for 3D object detection in autonomous driving
63. Goldman M, Hassner T, Avidan S, Learn stereo, infer mono: siamese networks for self-supervised, monocular, depth estimation
64. Guizilini V, Ambrus R, Pillai S, Raventos A, Gaidon A, 3D packing for self-supervised monocular depth estimation. Toyota Research Institute

Automated Car Parking System Using Deep Convolutional Neural Networks

Preeti Agrawal, Akash Patole, Yogesh Patil, and Parminder Kaur

Abstract Most of the people start their day by facing the same problem of finding vacant space in parking. Seeking a vacant parking slot during peak hours in areas like Colleges, Shopping Malls, Cinema theatres, Exhibitions and Convention Center has always been frustrating for many drivers. So here we are with the solution, in this paper we are addressing the vacant space with the help of deep learning method convolution neural network (CNN) and we are using PKLot Dataset which is already available. Vacant space is addressed with the help of red and green color boundaries. Red color indicates the car is present and the green color indicates there is vacant space. This system is efficient, effective and user friendly for finding the vacant space in parking.

Keywords PKLot dataset · Convolutional neural network (CNN) · Vacant slot · Occupied slot

1 Introduction

As the current time forces all of us to focus on the light where the situation is, as the ratio of population is increasing in the same way the use of vehicles is increasing day by day, so it is necessary to find the solution for people's problems. Surveys say that traffic generated by cars searching for vacancies in parking spaces is up to 40% of the total traffic. Now that is a serious issue to look after, and the smart parking system is the most effective way of solving the current and future traffic-related problems, including traffic congestion.

Object recognition is one of the important parts of computer vision, the same object of a similar class may have a different shape, size and color which make the recognition of objects difficult. So there are many traditional methods for object recognition. Handcrafted features are a Traditional method. Another way of object

P. Agrawal · A. Patole (✉) · Y. Patil · P. Kaur
MGM's Jawaharlal Nehru Engineering College, Aurangabad, India

P. Kaur
e-mail: parminderkaur@jnec.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_19

detection is deep learning concept, the famous algorithm of deep learning convolutional neural network (CNN) is used. The advantage of deep learning methods is that we can train our model with a large number of datasets. As we have used deep learning concepts so we have trained our model with a PKLot dataset which nearly consists of 695,899 images. The results are found to be satisfactory for all the test images. This system is efficient, effective and user friendly for finding the vacant space in parking.

2 Literature Survey

There are many systems already available for finding the vacant space. Most of the research is already done, but our motto is that the system should be cost-effective and friendly to use. Various technologies are used in different systems like [1]. In this research paper, they have used traditional object recognition approaches to apply feature extraction, part deformation handling, occlusion handling and classification sequentially, but the problem with this [1] is we can't use large dataset. We have overcome this problem by using deep learning concepts in [2]. In this research paper they train their dataset by using R-CNN after applying R-CNN deep learning method on a dataset, vehicle detection process has been successfully performed, but the problem with this system is they separate different object based on the region [2]. Another system [3], which is working on real-time images by using Deep learning Convolution Neural Networks (CNN) algorithm and a binary support vector Machine (SVM) classifier to detect the occupancy of outdoor parking space from images. This classifier is trained and tested by the features learned by the deep CNN from datasets (PKLot) [4]. In this paper, they propose an automatic method for detecting and counting cars in unmanned aerial vehicle (UAV) images. This proposed method begins with the segmentation of input image and then CNN is used combined with linear support vector machine (SVM) classifier to classify regions into "car" and "no-car" classes. Besides, small areas are analyzed using rectangular windows to locate cars more accurately and remove false positives [5]. In this research paper, the system includes a couple of sensors, microcontroller and microcomputer, the output of the system is placed at the entrance of the parking slot.

3 Methodology Used

See Fig. 1.

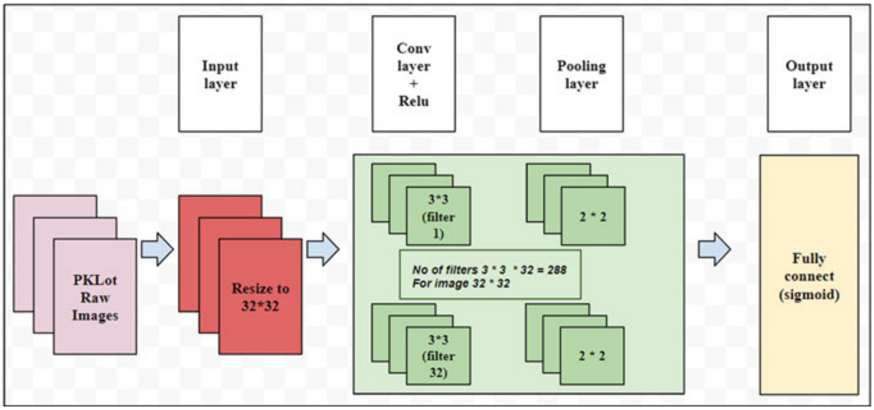


Fig. 1 Architectural building blocks of CNN model

3.1 Creation of CNN Model

Convolution Neural Networks have shown very good performance in object detection or classification purposes. The CNN model consists of various steps: first, we have to build the Convolution Neural Network architecture, second we have to create training and test labelled dataset. In Convolution Neural Networks there are multiple layers but for our project, we use CNN layer, pooling layer and fully connected layer. Any raw image of any size can be given as input to the algorithm designed which is first resized to the desired image patch size of 32 × 32. Resized image patch is fed as input to CNN. Image patch is represented as a 3D tensor of dimensions N × N × B, where N represents length and width of the image and B is the number of bands/channels. Therefore, all the factors discussed above play a significant role in making deep networks get trained. CNN architecture is as shown below (Fig. 1) as after performing CNN layer, polling is done and as shown below polling is done after CNN layer has performed. The output of the CNN layer is the input to the polling layer (Fig. 2). There are two approaches to polling layer average and max but we use max-pooling layer, the main use of this layer is to reduce the size of the original image. Even though pooling results in some amount of information loss, it still is found beneficial for the network as a reduction in size leads to less computational overhead for the upcoming layers of the network and it also works against over fitting. After performing a polling layer then we perform a fully connected layer which is the output layer of CNN (Fig. 2). The way this fully connected layer works is that it looks at the output of the previous layer and determines which features most correlate to a particular class. The output of the previous layer polling layer is compared with multiple predefined outputs then it calculates the probability of correctness of the output. If we define 80% probability for correct output. Then it calculates the probability if probability >80% then it displays 1 means there is a car and if probability <80% then it displays 0 means that slot is empty.

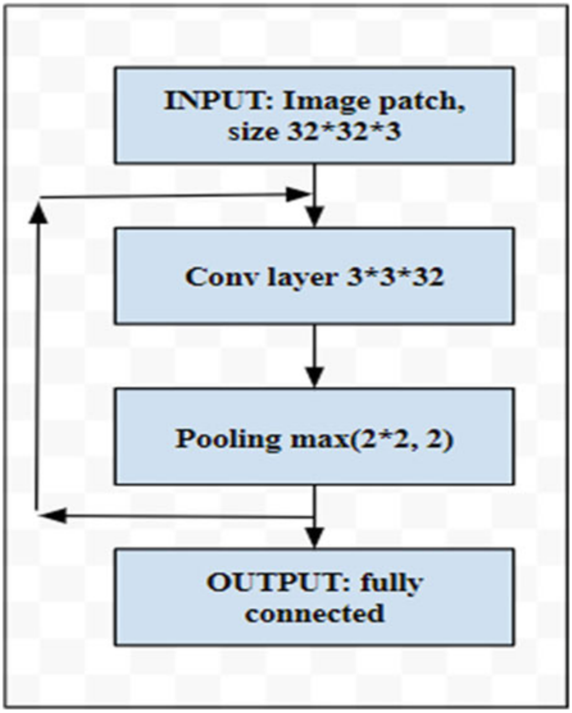


Fig. 2 Convolution neural network model

3.2 Creation of Training and Test Labelled Dataset

CNN's require a large number of images for training purposes. The image resolutions, size and scale of objects impact the training process, as task-relevant information varies with spatial resolution. How an object is recorded in an image depends on the objects location, angle of capture and its size. This is to be considered for data augmentation. CNN can deal with change in location as weights are shared in convolutional layers. Majority of the researchers have used hand-labelled dataset created for both training and testing and since labelling images is a very time-consuming process. In our project, we have used the PKLot dataset. The PKLot dataset [6] contains 12,417 images of three different parking areas (Fig. 3a–c) from which 6,95,899 segmented images (Fig. 4) of parking slots were generated and labelled in the dataset. This image acquisition was made by five minutes of the time-lapse interval over 30 days during the daytime on three different weather conditions namely Cloudy, Rainy, and Sunny weathers. And the accuracy of the training sets is available in an XML file for each image of the parking area. Following (Fig. 3a–c) are the examples of three different parking areas and (Fig. 4) is the example of segmented images of both empty slots and occupied slots.

Before been fed to CNN model, the segmented images are all scaled into 32 * 32 pixels and normalized. As we are using the sigmoid function in the output layer in the

a. PUCPR parking area

b. UFPR04 parking area

c. UFPR05 parking area

Fig. 3 **a**–**c** Three different parking areas

Fig. 4 Segmented images

CNN so we will have binary classification i.e. (0 or 1). Hence we create the dataset into two classifiers Empty and Occupied.

3.3 Experimentation

We have implemented the Deep Learning Concept using Convolutional Neural Network (CNN). This CNN model needed a huge amount of data to get trained for successful implementation, the data is taken from PKLots dataset which is freely available. Keras documentation is an open-source framework which provides the support to build the convolution, pooling and fully connected (Dense) layers. Our model was implemented in the Spyder platform with Tensor flow backend on Intel i5-9300H CPU @ 2.40 GHz, NVIDIA GeForce GTX 1650 graphic device with 4G byte graphic memory having Windows 10 operating system installed.

After the creation of the CNN model with Keras framework, they are trained by providing more than thousands of images from PKLots dataset. The number of epochs was 5 which has given 99.9% of the model accuracy. The model was trained with three different weather conditions i.e. Cloudy, Rainy, Sunny.

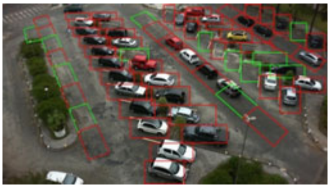
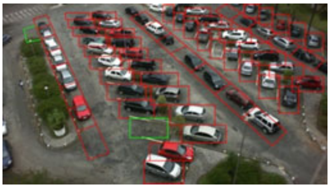
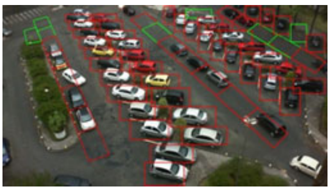
Table 1 shows some of the three test cases of the three different weather conditions. The table shows the output image so that we can see the accuracy by counting the correct parking slot spotted with the help of the model, and then based on that, we can see in which weather condition the model gives the best performance. Each image contains the same parking area where we manually identified 65 parking slots. The correct spotted slot is counted from those slots and then their percentage is shown in output accuracy from those three test cases of each weather, the total average accuracy is calculated to know in which weather condition the model works with the best accuracy. By observing the table, we can calculate the average accuracy of the weather. Cloudy weather has 75.3%, Rainy weather has 85.2% and Sunny weather has 90.7%. So we can conclude here that in sunny weather the model shows the best performance, wherein in cloudy weather it will give poor performance.

4 Result and Discussions

The cropped image is given as an input to the CNN model, the final output of the model is the image that helps the driver to identify that particular slot is vacant or occupied (Fig. 5).

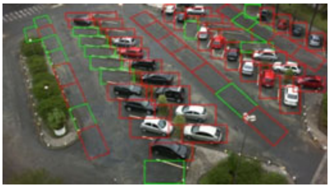
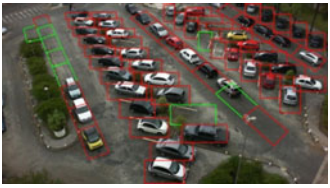
Here is the example of an output image of one slot that shows whether a car is present or not with the help of a rectangle box, as the red rectangular box shows that there is a car (occupied) and green rectangular box shows that slot is empty (no car). To evaluate the accuracy of the CNN model, the confusion matrix is generated. To increase accuracy in the CNN model we use multiple filters or different size of a filter such as 3 * 3, 5 * 5, 7 * 7, 11 * 11 and many others then change in accuracy value

Table 1 Test cases

Test case Id	Weather condition	Output image	Correctly spotted slot	Output accuracy (%)	Total weather average accuracy (%)
TC_01	Cloudy weather		48 out of 65	73.8	75.3
TC_02			37 out of 65	56.9	
TC_03			62 out of 65	95.3	
TC_04	Rainy weather		58 out of 65	89.3	85.2
TC_05			62 out of 65	95.3	

(continued)

Table 1 (continued)

Test case Id	Weather condition	Output image	Correctly spotted slot	Output accuracy (%)	Total weather average accuracy (%)
TC_06			46 out of 65	70.7	
TC_07	Sunny weather		58 out of 65	89.3	90.7
TC_08			55 out of 65	84.6	
TC_09			64 out of 65	98.4	

is recorded. We calculate loss, accuracy during the evaluation of the CNN model. In our project, we used the PK-LOT dataset. In this dataset, there are thousands of images of parking lots. In these parking lots, we consider one particular parking lot for confusion matrix generation. For example, if there are 64 slots in one image of the parking slot.

- There are two predicted possible classes: "Yes" and "No". In our project yes means that slot is occupied, and No means that slot is empty.
- The classifier made a total of 64 predictions in that particular image.
- Out of 64 cases, it predicted "Yes" 54 times, and "No" 10 times.
- In reality, 49 slots are already occupied by another car, and 15 slots are empty (Table 2).

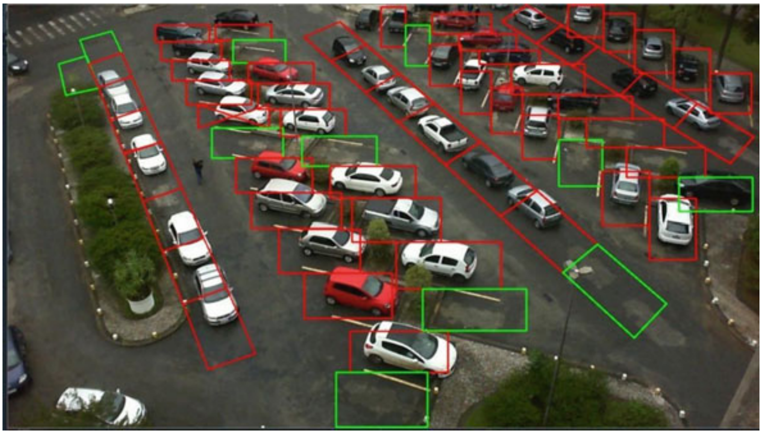


Fig. 5 Output image

Table 2 Classifier and actual value prediction

	Predicted: No	Predicted: Yes
Actual: No	9	6
Actual: Yes	1	48

We analyze the result of the model after generating the confusion matrix. From the confusion matrix, we calculate accuracy, Misclassification Rate, True positive rate, true negative rate (Table 3).

Table 3 Analysis of accuracy, misclassification rate

	Predicted: No	Predicted: Yes	
Actual: No	TN = 9	FP = 6	15
Actual: Yes	FN = 1	TP = 48	49
	10	54	

TP (True Positive) These are cases in which we predicted yes (that slot already occupied by another car)
TN (True Negative) These are cases in which we predicted No (that slot is empty)
FP (False Positive) We predicted yes, but that slot is empty
FN (False Negative) We predicted No, but actually that slot is not empty

Now we calculate the rates from the confusion matrix.

- **Accuracy**:
  (TP + TN)/Total = (48 + 9)/64 = 0.89.
- **Misclassification Rate**:
  How frequently is it wrong?

(FP + FN)/Total = (6 + 1)/165 = 0.04.

- **True Positive Rate**:
  When it is actually yes, how frequently does it predict yes?
  TP/actual yes = 48/49 = 0.97.
- **True Negative Rate**:
  When it is no, how frequently does it predict No?
  TN/actual No = 9/15 = 0.6.

The CNN architecture is built with 3 convolution layers. We used these CNN convolution layers to get more exact features from the recognition process. The model trained for 5 epochs with 194,229 images from training data per steps. The accuracy obtained from training is 99.90%. The graph shows the relation between accuracy and validation accuracy (Fig. 6), loss and validation loss (Fig. 7).

Fig. 6 Accuracy

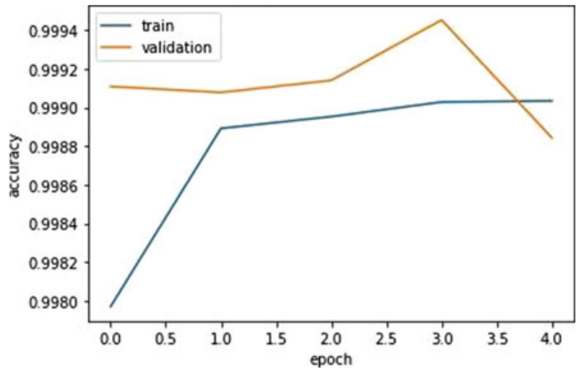
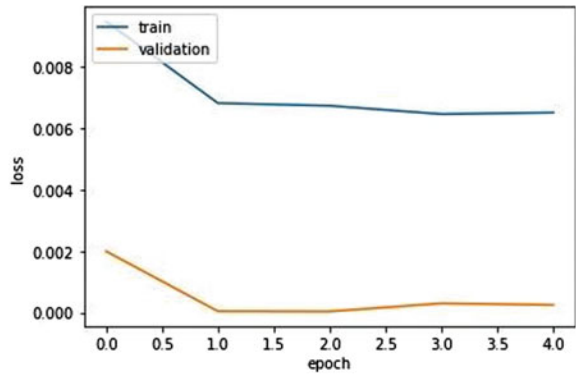


Fig. 7 Loss



5 Conclusion and Future Scope

In this paper, we have used the PKLot database which consists of thousands of images as the input to the model. Different season images are considered like rainy, sunny and winter and accuracy according to them is calculated by using confusion matrix and accuracy is different for different season images. The result shows that the model designed can identify whether the slot is occupied or empty. As we have used CNN model in that we used three-layer (a) convolution layer (b) max-pooling layer (c) fully connected layer. By using CNN algorithms we build the model with 99.9% accuracy. The CNN model provides a cheap and reliable solution for detection of empty slots in the parking area. We trained our model by using images of PKLot database then we classify them into various slot manually like rainy, sunny and winter then we crop the images and send as input to the CNN model and GUI is designed as the output screen for the model where red rectangular box indicates that it is occupied and green color rectangular box shows that place is vacant.

Future work will extend the system to allow the user to know the specific slot number which is vacant so that it will save time and the driver can directly go on that slot number and park the car and the system will also show the navigation for vacant slots. The system also provides users to book the slot for a particular time interval and if a user won't arrive in a given time then the slot will automatically become unbook and available for other users.

References

1. Feizabadi S (2014) Joint deep learning for car detection. https://www.researchgate.net/publication/270222817_Joint_Deep_Learning_for_Car_Detection
2. Yilmaz AA, Güzel MS, Askerbeyli İ, Bostanci E (2014) A vehicle detection approach using deep learning methodologies. Ankara University
3. Acharya D, Yan W, Khoshelham K (2019) Real-time image-based parking occupancy detection using deep learning. In: 5th annual conference of Research@Locate, CEUR workshop proceedings, vol 2087, Adelaide, Australia, pp 33–40
4. Ammour N, Alhichri H, Bazi Y, Benjdira B, Alajlan N, Zuair M (2017) Deep learning approach for car detection in UAV imagery. www.mdpi.com/journal/remotesensing
5. Pimple S, Uttarwar G, Patil N, Bhagunde N (2019) Cost effective stand-alone parking assistance system. Government College of Engineering, Aurangabad
6. de Almeida PR, Oliveira LS, Britto AS, Silva EJ, Koerich AL (2015) Pklot a robust dataset for parking lot classification

Currency Exchange Rate Prediction Using Multi-layer Perceptron

Appala Raju Middi and Venkata Sai Rishita Middi

Abstract Financial forecasting is an estimate of a future financial outcome and this outcome is related to some kind of value. We can measure this outcome for a company to predict its future stock or to detect the viability of a human for the sanction of a loan. In all these cases, we want to estimate the future outcome based on historical data. Various methods have been developed lately, to make time series predictions. In this work, we have used Multi-layer perceptron algorithm to predict the Currency Exchange rate between US dollar and EURO. The training network has been compiled using TensorFlow.

Keywords Perceptron · Neural network · Currency · Fluctuations · Prediction

1 Introduction

The value of a firm's future contractual transactions in foreign currencies is affected by exchange rate movements. The sensitivity of the firms' contractual transactions in foreign currencies to exchange rate movements is because of transaction exposure. It is very essential to consider techniques to reduce exchange rate risk when a firm is exposed to exchange rate fluctuations due to international transactions. Multinational corporations are exposed to exchange rate fluctuations in multiple ways like transaction exposure, economic exposure and translation exposure. A perceptron is a binary classifier that classifies linearly separable data. The output label is assigned as a function of the input and weights. A single layer perceptron has limitations because it cannot be used for data that is not linearly separable.

The system proposed in [1] uses a model of neural network that is used to predict the exchange rate of Rupiah (Indonesian Currency) against US Dollar. The root mean square error (RMSE) is used as an indicator in the success of the proposed model. In

A. R. Middi (✉)
Christ University, Bengaluru, India
e-mail: middi.raju@christuniversity.in

V. S. R. Middi
Brown University, Providence, RI, USA

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_20

[2], AUD dollar has been measured against the US Dollar using time-series models in collaboration with news sentimental analysis. The numerical results showed that adding sentimental analysis has significantly reduced the prediction error. Prediction of Thai Baht by using Hidden Markov Models has been performed in [3], that is using four different factors like dollar index, interest rate etc. The forecast of the currency exchange rates between US Dollar with GBP by using ANN and Meiosis Genetic algorithms [4]. The research performed in [5] focusses on currencies of ASEAN countries w.r.t. US Dollar. A group of differential equations are employed in Grey forecasting. Ranjit et al. [6] applies Long short-term architecture for predicting the value of Nepalese Rupee against USD, GBP and EURO using their historical exchange rates. A modified fuzzy relational model [7] has been used to forecast the exchange rate between US dollar and Switzerland Franc.

The accuracy of convolutional neural networks in the prediction of Hungarian Forint's direction of changes with daily forecasting has been determined in [8]. A comparison between various models namely, SVR, ANN, STM has been used to anticipate the exchange rate between USD/PKR [9]. A prediction scheme is proposed to efficiently predict the long-term exchange rate using Radial Basis Function Neural Network (RBFNN) [10] between USD and Japanese Yen.

This work is based on a system that predicts the currency Exchange rate using Multi-Layer Perceptron. The paper is divided into five sections: Sect. 1 is Introduction, Sect. 2 Preparing Data, Sect. 3 Building the model, Sect. 4 is Training and Testing Network, Sect. 5 Results and Conclusion.

2 Acquiring and Preparing Data

The data has been acquired from investing.com and it has been downloaded as a csv file. By default, the dataset starts from the most recent data points. The data has to be reframed/reversed to start with the oldest data point first. This is important to trace our problem as a supervised learning problem. We need to give our network some idea about how to figure out the dependencies between the dataset. The raw data consists of the date and the value of price on the specified date (see Fig. 1).

The raw data has to be converted to a form, understandable by the neural network. A logical assumption is made that there is a correlation between the price on a current date and the exchange price on previous dates. The neural network predicts the patterns out of these experiences. Two arrays are created X and Y representing Cause and Effect respectively. Effect is the price that we want to predict, and Cause is the value that we base our prediction on (see Figs. 2 and 3).

Fig. 1 Raw data acquired over a duration of time

```
'Raw values'
[('Jan 18, 2001', 1.0589),
 ('Jan 19, 2001', 1.0709),
 ('Jan 22, 2001', 1.0644),
 ('Jan 23, 2001', 1.0664),
 ('Jan 24, 2001', 1.0857),
 ('Jan 25, 2001', 1.0816),
 ('Jan 26, 2001', 1.0821),
 ('Jan 29, 2001', 1.0906),
 ('Jan 30, 2001', 1.0791),
 ('Jan 31, 2001', 1.0677)]
```

Fig. 2 The price on a previous date and the price on current date

```
'X/Y values (wihout empty ones)'
[(1.0589, 1.0709),
 (1.0709, 1.0644),
 (1.0644, 1.0664),
 (1.0664, 1.0857),
 (1.0857, 1.0816),
 (1.0816, 1.0821),
 (1.0821, 1.0906),
 (1.0906, 1.0791),
 (1.0791, 1.0677),
 (1.0677, 1.0651)]
```

```
'X'
[1.0589, 1.0709, 1.0644, 1.0664, 1.0857, 1.0816, 1.0821, 1.0906, 1.0791, 1.0677]
'Y'
[1.0709, 1.0644, 1.0664, 1.0857, 1.0816, 1.0821, 1.0906, 1.0791, 1.0677, 1.0651]
```

Fig. 3 The X and Y arrays

3 Building the MLP Model with Keras

The next task is to build the network for training. The training network is created out of pieces of processing units called layers. Initially, a container is created to hold the layers that process the data. The input to the layer 'x' is the exchange rate on a previous date. The output is a single dense layer. The trained network will be used

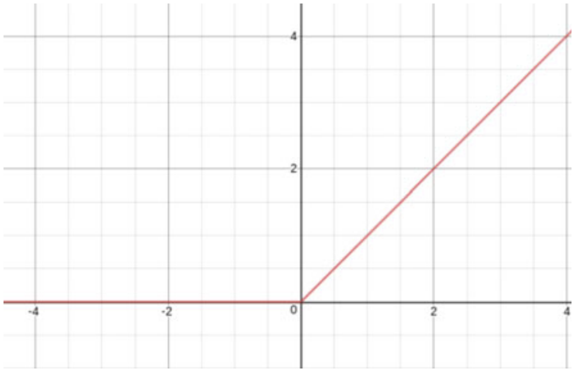


Fig. 4 Graph of ReLU function

for prediction. The network that we have used has one hidden layer i.e. the capacity of our network is 1.

Activation functions determine the output of a neural network. They aid in the ability of the neural network to converge, thus improving the computational efficiency. The activation function used in this work is ReLU (Rectified Linear Unit) (see Fig. 4). The model is compiled using TensorFlow before using it for prediction. Mean Square Error is used as the loss function. The mean square error between the input and output has to be minimized during training. The number of training sessions (epochs) will be gradually increased or decreased based on the metrics we get.

4 Training and Testing of the Model

Tuning the training parameters such as epochs and batch_size is very important for the network to converge to a minimum error value. The number of batches that will go through the model during training, updates the internal state of the model. As we increase the number of epochs, the value of mean square error decreases and reaches a minima. (see Fig. 5 for epoch = 10 batch_size = 10).

The dataset is divided into training set and testing set (80% training set and 20% test set). It is done to ensure that the training is done on one dataset and the verification is done on the test set. For the value of epoch = 46, an optimal solution is obtained.

Baseline prediction is one of those aspects of building time-series when we have a reference point. The closer the prediction is to the baseline, the stronger our model is. The model is not very strong when the number of epochs = 10 (see Fig. 6). The prediction graphs are quite deviated from the baseline.

```
Epoch 1/10
 - 0s - loss: 0.1039 - mse: 0.1039 - mape: 35.9473
Epoch 2/10
 - 0s - loss: 0.0064 - mse: 0.0064 - mape: 6.8045
Epoch 3/10
 - 0s - loss: 0.0018 - mse: 0.0018 - mape: 3.5748
Epoch 4/10
 - 0s - loss: 0.0017 - mse: 0.0017 - mape: 3.7466
Epoch 5/10
 - 0s - loss: 0.0016 - mse: 0.0016 - mape: 3.6855
Epoch 6/10
 - 0s - loss: 0.0016 - mse: 0.0016 - mape: 3.6093
Epoch 7/10
 - 0s - loss: 0.0015 - mse: 0.0015 - mape: 3.5112
Epoch 8/10
 - 0s - loss: 0.0014 - mse: 0.0014 - mape: 3.4308
Epoch 9/10
 - 0s - loss: 0.0013 - mse: 0.0013 - mape: 3.3429
Epoch 10/10
 - 0s - loss: 0.0012 - mse: 0.0012 - mape: 3.2234
```

Fig. 5 The value of mean square error reduces with each session

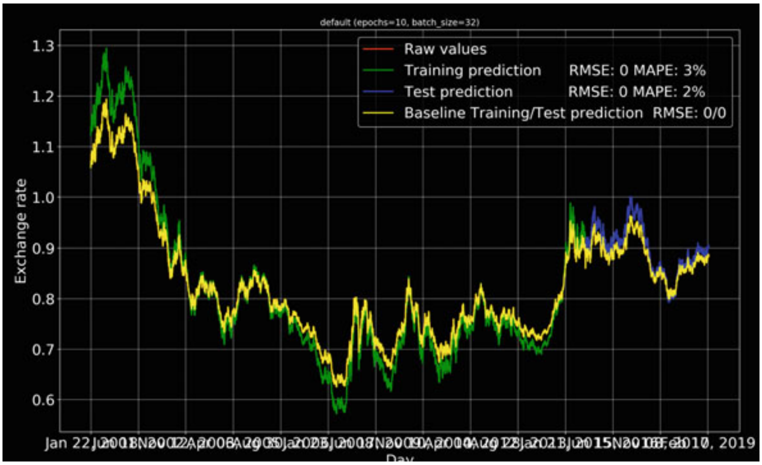


Fig. 6 Exchange rate versus date for epoch = 10

5 Results and Conclusion

The optimal solution is obtained at the value of epoch = 46 and batch_size = 32. The prediction was done for the subsequent days (see Fig. 7). There has been a significant improvement in the accuracy compared to [1, 2].

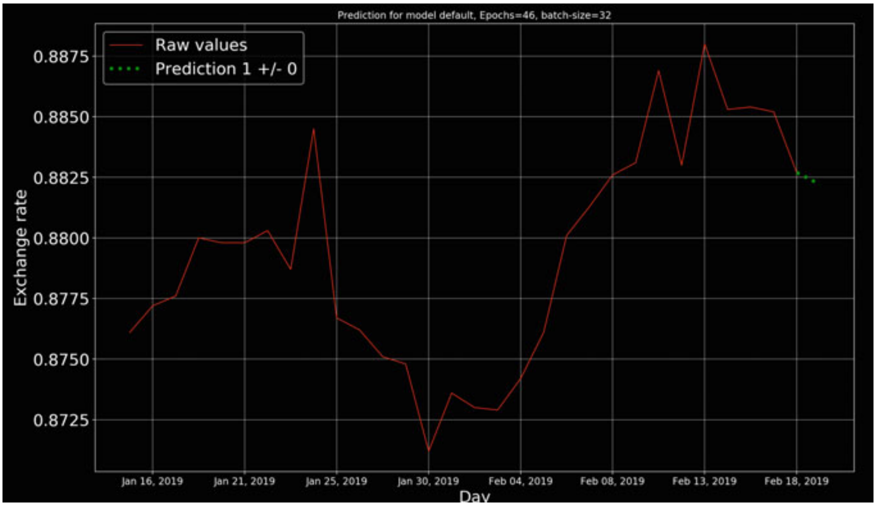


Fig. 7 Prediction graph

References

1. Arsi P, Somantri O, Hasanah U, Astuti T, Imron M, Waluyo R (2018) A proposed model of neural network for rupiah exchange rate's prediction. In: 2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE), pp 90–93. https://doi.org/10.1109/ICITISEE.2018.8720962
2. Lee C, Chang C, Hwang F (2019) Currency exchange rate prediction with long short-term memory networks based on attention and news sentiment analysis. In: 2019 international conference on technologies and applications of artificial intelligence (TAAI), Kaohsiung, Taiwan, pp 1–6. https://doi.org/10.1109/TAAI48200.2019.8959884
3. Nootyaskool S, Choengtong W (2014) Hidden Markov models predict foreign exchange rate. In: 2014 14th international symposium on communications and information technologies (ISCIT). Incheon, pp 99–101. https://doi.org/10.1109/ISCIT.2014.7011878
4. Nayakovit S, Khantanapoka K, Jaritngam U (2010) Prediction exchange rate of USD/GBP with intelligence cyberspace experimental. In: 2010 international conference on electronics and information engineering, Kyoto, pp V2-15–V2-19. https://doi.org/10.1109/ICEIE.2010.5559706
5. Wang T, Kuo S, Chen H (2011) Forecasting the exchange rate between ASEAN currencies and USD. In: 2011 IEEE international conference on industrial engineering and engineering management, Singapore, pp 699–703. https://doi.org/10.1109/IEEM.2011.6118007
6. Ranjit S, Shrestha S, Subedi S, Shakya S (2018) Comparison of algorithms in foreign exchange rate prediction. In: 2018 IEEE 3rd international conference on computing, communication and security (ICCCS), Kathmandu, pp 9–13. https://doi.org/10.1109/CCCS.2018.8586826
7. Askari MA, Menhaj MB (2016) A modified fuzzy relational model approach to prediction of foreign exchange rates. In: 2016 4th international conference on control, instrumentation, and automation (ICCIA), Qazvin, pp 457–461. https://doi.org/10.1109/ICCIAutom.2016.7483206
8. Galeshchuk S, Demazeau Y (2017) Forecasting Hungarian forint exchange rate with convolutional neural networks. In: 2017 international conference on behavioral, economic, sociocultural computing (BESC), Krakow, pp 1–3. https://doi.org/10.1109/BESC.2017.8256358

9. Memon IZ, Talpur S, Narejo S, Junejo AZ, Hassan EF (2020) Short-term prediction model for multi-currency exchange using artificial neural network. In: 2020 3rd international conference on information and computer technologies (ICICT), San Jose, CA, USA, pp 102–106. https://doi.org/10.1109/ICICT50521.2020.00024
10. Rout M, Majhi B, Mohapatra UM (2012) Efficient long range prediction of exchange rates using radial basis function neural network models. In: IEEE-international conference on advances in engineering, science and management (ICAESM-2012), Nagapattinam, Tamil Nadu, pp 530–535

Analysis of Web Application Firewalls, Challenges, and Research Opportunities

Subhash V. Pingale and Sanjay R. Sutar

Abstract According to survey in January 2020, over internet almost 1,295,973,827 web sites are hosted. Among them 72% websites are vulnerable to different attack like SQL Injection, Cross Site Scripting, Brute forcing attack, Phishing attack, password attack, birth day attack, malware attack and man in middle attack now days many hackers/attackers are trying to bypass the security mechanism by using new techniques so it's challenging task to respond newly and unknown attack with an effective solution. In this paper i am trying to find out the techniques, tools, and solutions used to detect attacks, such as intrusion detection systems (IDS), Web Application Firewall (WAF), machine learning (ML) techniques, this tools and techniques is used to discuss analysis of traditional technologies, drawback and produce more effective solutions. In this paper we compared the different web application firewall correspond to its policy control and one proposed a new web application firewall. In the proposed web application firewall we have discussed three approaches.

Keywords Web application firewall · Datasets · Security · Intrusion detection system

1 Introduction

In the era of social media, the internet users are increased drastically. Many users use many internet applications such as World Wide Web, File Transfer Protocol, Web media, messaging and peer to peer based software etc. due to this internet traffic is increased Internet traffic is classified into three main classes to the system administrator, internet service provider, and governments: First, the packet classification is

S. V. Pingale (✉)
Research Scholar, Dr. Babasaheb Ambedkar Technological University, Lonere, India
e-mail: subhash.pingale@sknscoe.ac.in

Assistant Professor, SKN Sinhgad College of Engineering, Pandharpur, India

S. R. Sutar
Dr. Babasaheb Ambedkar Technological University, Lonere, India
e-mail: srsutar@dbatu.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_21

used in web application firewall and Intrusion detection system to detect the patterns of denial of service or other malicious attacks. Increased Internet traffic is one of the challenge over the last five years. It requires in details understanding of the network because there are various types of network traffic for the internet service providers and a very large volumes of stream flows. With the increased bandwidth as well as number of services users can perform more and complex activities than before. Two types of signal are available baseband and broadband. Due to broadband signal user can easily perform tasks such as Voice over Internet Protocol, online shopping, file transfer, peer-to-peer-network and video sharing among users, the more complexity is increased when we are using different wireless technologies.

Web Application Firewall is used to protect web application by filtering the HTTP traffic between internet and web applications. Web Application firewall is used to protect web application from different types of attack such as Injection, cross site scripting etc.

In web application firewalls different policies are defined this policy is used to protect web application against the vulnerabilities in the application web application firewall is used to place between the web application and internet.

1.1 Web Application Firewall

Web application firewall is most useful solution used in application layer to detect and prevent the attacks which is passed through HTTP (Fig. 1).

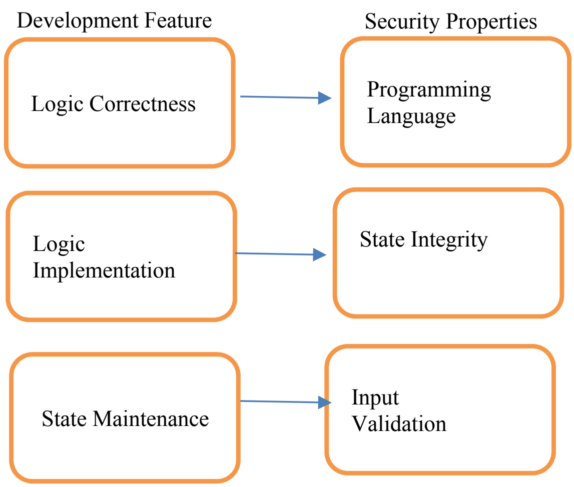


Fig. 1 Web application security properties

1.2 *Web Application Security Risk*

1.2.1 Injection

Injection attack can take place when the data which is not trusted and its sent to the engine as a query or through command in a web application. The attacker's query can trick the engine and execute the commands for hacking purpose [1].

1.2.2 Broken Authentication and Session Management

In some web application authentication as well as session management are not implemented correctly in such cases it allows attackers to compromise session tokens, passwords, keys to assume other users identities [1].

1.2.3 Cross Site Scripting (XSS)

Attackers uses the XSS script to execute in the browsers and hijack user sessions in the website or users can be redirected to malicious sites this type of attack happens when an application takes as input an untrusted data and pass it to a web browser without proper validation [1].

1.2.4 Cross Site Request Forgery (CSRF)

It is an attack vector which tricks web browser to execute an unwanted action in application to which user is logged in. CSRF attack sends a HTTP request by including the victimes session cookies. This type of attacker force to victime brawers to generate the request vulnerable applications are legitimate [1].

1.2.5 Insecure Direct Object Reference

Attackers can access the unauthorised data through the references. References is generated when a developer exposes the object of internal reference such any filename, directory, or database key without protection or access control check [1].

1.2.6 Security Misconfiguration

Security of application system, framework, web server, database server, application server, and platform is depending on the secure configuration. All these setting should be defined, implemented, and maintained with secure default [1].

1.2.7 Insecure Cryptographic Storage

Attacker may attack or modify the data which is not protected this data will get attacker from many web applications which are not properly protected sensitive data such as credit card SSNs, and authentication credentials with appropriate security algorithm [1].

1.2.8 Failure to Restrict URL Access

Many web applications are accessing Unified Resource Language (URL) without rendering protected link and buttons, actually applications need to perform similar access control each times these pages are accessed otherwise attacker can forge URLs to access this hidden pages anyway I. Insufficient Transport Layer Protection Transport layer fails to protect some application due to week algorithm or invalid certification when it is required to protect sensitive data [1].

1.2.9 Invalidated Redirects and Forwards

Some web application frequently redirect and forward to users to other websites which are not validated. Such websites uses untrusted data to determine the required pages without proper validation [1].

2 Related Work

2.1 Critcal Analysis of Web Application Firewall

Ryohei et al. [2] has proposed techniques to detect malicious code using machine learning Algorithm. The input is classified into two parts using learning process and classification process. Classifier is used to extract the features from non-malicious code and malicious code. It has described a feature vector such as (w1, w2, w3, …, wx) where wx is the weight of each term. All this term is extracted using extraction method from documents. Weights of all the term is calculated by an evaluation method. It has proposed two extraction methods such as tokenizing method and Blank separation method [3].

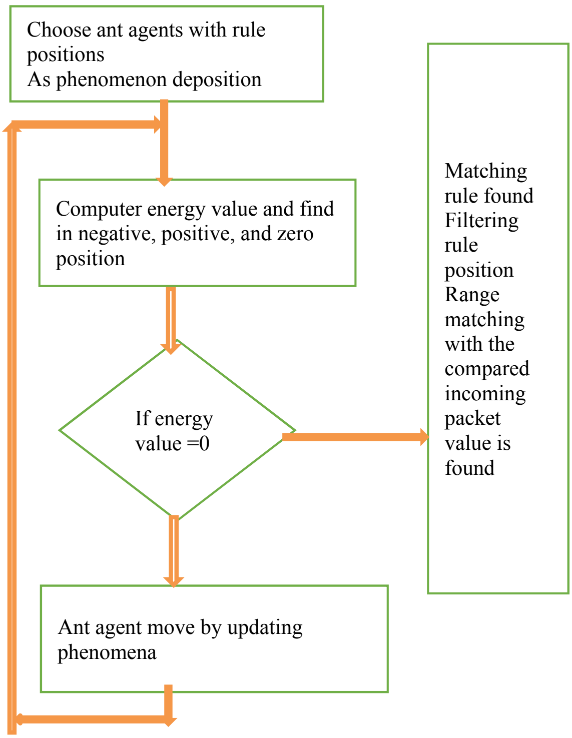


Fig. 2 Flow chart of an ant agent

2.2 APO Technique

Sreelaja et al. [4] has discussed an APO technique to filter the incoming packets in a network. In this technique it matches the rule with rules sets. Ant agent determines the rule by matching rule with the rule set. It describes the energy value of the ant agent. Figure 2 shows the model of the Ant system.

2.3 Dataset

In this paper I am going to do analysis of different dataset as shown in Table 1, it is observed that HTTP Dataset CSIC 2010 is more useful [5]. This dataset was developed in Spanish Research National Council. This dataset is analysis of Spanish e-commerce web application it includes 25,000 malicious traffic and 360,000 normal traffic in malicious traffics it include HTTP request such as injection, buffer overflow attack and information gathering.

Table 1 Analysis of datasets

Dataset	Public access	HTTP traffic	Labeled	Two classes	Modern attack	Not anonymized
UNB ISC		✓	✓	✓	✓	✓
PKDD		✓	✓	✓	✓	
LBLN	✓		✓	✓		
DEFCON	✓	✓	✓		✓	✓
DARPA 99	✓	✓	✓	✓		✓
CAPTURED Traffic	✓	✓			✓	✓
CSIC	✓	✓	✓	✓	✓	✓

We know that both the attacks static and dynamic are generated which include such as modern web attacks such as cross-site scripting, injection, information gathering, buffer overflow, CSRF, include parameter tampering.

In Table 1 blank cell is used to represent available feature is not present and tick symbol is used to represent specific feature is available in the dataset.

Followings are the key finding in dataset.

1. Many datasets do not contain HTTP Traffic
2. Many datasets are not publically available
3. Some datasets are not properly labeled
4. Traffic is anonymized
5. It is not balanced between normal traffic and attack.

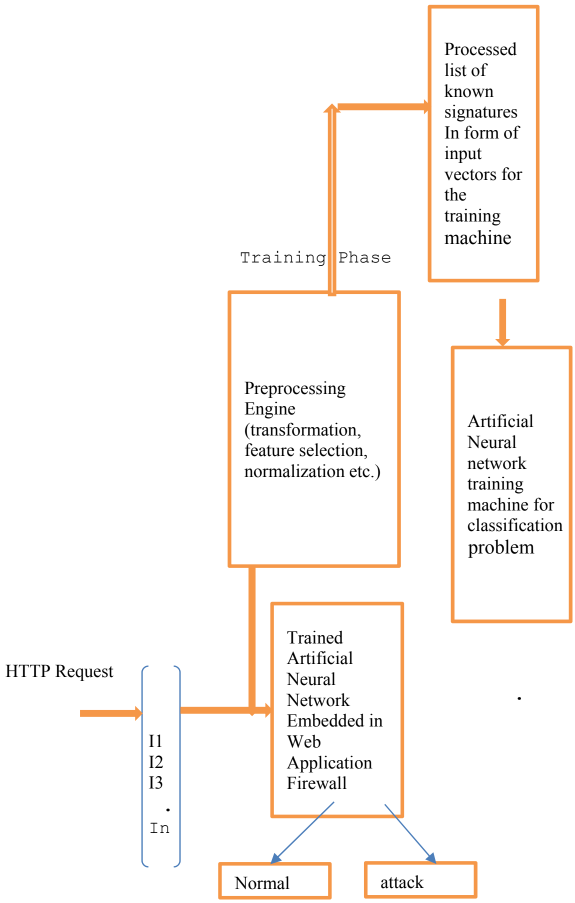
2.4 ANN based Web Application Firewall

Moosa [6] has stated a solution to the most Popular Vulnerabilities Structure Query Language of web application he has proposed this solution by using Artificial Neural Network with the vast number of attacks defined signature are increasing so need to be develop a good system that can work successfully with the scalable network hence he has developed a system using ANN whose basic component is neuron it is shown in Fig. 3.

3 Comparison of Web Application Firewalls

Web application Firewall is used to protect the web applications from attacker in this paper we have selected 10 web application firewall which include Snipper, mod security, F5, forti-web, citrix, Secure Internet Information Server, Easy Guard, Secure Sphere, Citrix, Web App secure, proposed web application firewall. All this firewalls

Fig.3 Components in neuron



are compared using following feature such as deep inspection, policy control, monitoring, blocking, response filtering attack prevention, session protection, authentication, and overall security performance. In Table 2 feature defense mechanism verses security policy control is shown. Values are represented by symbol. Tick symbol represents the concerned security policy control is present and blank cell represent concerned security policy control is not present.

Table 2 Feature analysis of web application firewall

Feature defense mechanism/security policy control	Web snipper	F5	I-security	Secure IIS	Easy guard	Secure sphere	Forti-web	Citrix	Web secure	Mod security	Proposed firewall
Time efficient	✓	✓		✓		✓	✓	✓		✓	✓
Well organized	✓	✓		✓		✓	✓	✓	✓	✓	✓
effective	✓	✓		✓	✓		✓				✓
Monitoring	✓			✓	✓	✓	✓		✓	✓	✓
blocking	✓		✓	✓		✓	✓	✓	✓	✓	✓
Deep inspection	✓		✓	✓		✓	✓			✓	✓
Attack prevention	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Website blocking	✓	✓	✓	✓	✓	✓	✓			✓	✓
Session protection	✓	✓		✓		✓	✓			✓	✓
Response filtering		✓		✓	✓		✓			✓	✓
Authentication and web SSO	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
Overall security	✓	✓		✓		✓	✓				✓

4 Proposed Approaches

4.1 Multimodal Networks

There are multiple types of features present in the dataset: categorical, ordinal, and string based values It is desired that all these values are not passed to the same model but rather to individual sub-models Multimodal network is a network that accepts multiple inputs sent to multiple sub networks for the selected features, we pass them to their respective channels and then combine back all the intermediate representations to get the final output.

Advantage: Better handling of feature types, context-specific training.

Scope for research: Feature engineering, deciding the sub networks that need to be used for the respective feature types.

4.2 Attention-Based Mechanism

There is a URL pattern field, which consists of a series of tokens. Attention is a sequential model to which we pass this URL and other strings present in the test case vector Helps for better understanding and checking if there is any spam or toxic words present in the URL for attack possibility RNNs or LSTMs are sequential models but they are more complex, with increased number of trainable parameters and cost. Attention mechanism requires lesser memory and time requirement while still getting the same results.

Advantages: Better focus on keywords in URLs, lesser complexity.

Scope for research: Custom architecture design, chance for improved performance.

4.3 Heuristic Based Selection

Genetic evolutionary algorithm that focuses on feature selection. There are multiple variables present in the dataset, not all of which contribute to deciding whether there is an attack or not.

We come up with a heuristic based approach where after every iteration only the fittest i.e. most contributing attributes survive and the redundant ones are removed.

Doing so helps gain convergence faster by deciding the ideal weights to be given to each attribute when checking for attack possibility.

This is a traditional AI approach wherein the engine aims to go to the global ideal state.

Advantages:

(1) Evolutionary, so does not get stuck with redundant columns for long.
(2) Heuristic function can be worked upon for faster convergence.

Scope for research: Novelty in heuristic function, feature engineering methods.

5 Followings Are the Key Finding from Above Survey

1. Large number of websites are hosted on the internet so risk of vulnerabilities is increased
2. need to develop automated web application firewall
3. need to work for faster convergence
4. due to highly increased web application there is need to provide security
5. WAF is highly customized for each environment.

6 Conclusion

Most of the web applications are vulnerable, web application firewall and traditional methods are not efficient. In this survey we have highlighted Analysis of datasets as well as analysis of web application firewall which is helpful to understand the working of WAF and need to decide the appropriate. So there is need to develop automated web application firewall which can blocks both the attacks signature based as well as anomaly based. In this paper I have proposed three approaches multimodal network, Self-attention, and Heuristic based selection.

References

1. Top10 OWASP, Top 10-2010 (2010) The ten most critical web application security risks. The Open Web Application Security Project
2. Komiya R, Paik I, Hisada M (2013) Classification of malicious web code by machine learning. In: 3rd international conference on awareness science and technology. Dalian, China
3. Razzaq A et al (2013) Critical analysis on web application firewall solutions. In: IEEE eleventh international symposium on autonomous decentralized systems (ISADS)
4. Sreelaja NK, Vijayaakshmi Pai GA (2010) Ant colony optimization based approach for efficient packet filtering in firewall. Elsevier
5. HTTP DATASET CSIC (2010) [Online]. Available: http://www.isi.csic.es/dataset
6. Moosa A (2016) Artificial neural network based web application firewall for SQL injection. In: 2016 annual conference on world academy of science, engineering and technology, vol 4. IEEE, 23 Mar 2010, pp 181–186

A Comparative Study for Predicting Burned Areas of a Forest Fire Using Soft Computing Techniques

Ibrahim Al-Shourbaji, Mohammed Alhameed, Anwer Katrawi, Fathe Jeribi, and Sophia Alim

Abstract Forest fires is an important environmental disasters that have many consequences in our life. As a result, early forecasting and rapid action of fires are needed to control such a phenomenon and saving lives. In this paper, Montesinho Natural Park (MNP) dataset is used to determine the burned areas of a forest fire, three prediction techniques were analyzed; namely Classification and Regression Tree (CART), Multivariate Adaptive Regression Splines (MARS) and Artificial Neural Network (ANN). These algorithms are implemented to specify the technique that would provide best prediction results. The performance of these algorithms was assessed based on three statistical measures, the Mean Absolute Error (MAE), the Mean Squared Error (MSE) and the Root Mean Squared Error (RMSE). From the Statistical measures, it is inferred that MARS algorithm provides the best results in term of performance compared with other methods. The obtained results confirm that MARS improves the forecasting accuracy and has the capability to forecast forest fires effectively with a comparable computational cost.

I. Al-Shourbaji (✉)
Department of Computer and Network Engineering, Computer Science and Information Technology College, Jazan University, Jazan 82822-6649, Kingdom of Saudi Arabia
e-mail: ialshourbaji@jazanu.edu.sa

M. Alhameed
Department of Computer Science, Computer Science and Information Technology College, Jazan University, Jazan 82822-6649, Kingdom of Saudi Arabia
e-mail: malhameed@jazanu.edu.sa

A. Katrawi
Faculty Member, Academic Development Deanship, Jazan University, Jazan 82822-6649, Kingdom of Saudi Arabia
e-mail: akatrawi@jazanu.edu.sa

F. Jeribi
Department of Information Technology and Security, Computer Science and Information Technology College, Jazan University, Jazan 82822-6649, Kingdom of Saudi Arabia
e-mail: fjeribi@jazanu.edu.sa

S. Alim
Independent Researcher, Bradford, UK

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_22

Keywords Soft computing · Forest fires · Forecasting · Burned areas · CART · MARS · ANN

1 Introduction

Forest fires can be considered as a serious environmental disaster that has many negative consequences in different aspect of life. However, an early detection can offer an essential tool for managing natural resources and preventing disasters [1]. Special attention must be paid to forest fires because they represent a serious threat to humans and nature that may have very negative consequences to our lives.

Methods of soft computing (SC) can be highly helpful in the identification and assessment of forest fires at an early level and can reliably determine the environmental causes and impacts [2–4]. The SC approaches are used to reach an appropriate response to a problem [5]. In several different applications, this increases the potential of the SC strategies to solve complicated problems [6]. Over recent years, several methods, including the Artificial Neural Network (ANN) strategies, knowledge-based systems, data mining and fugitive rule-based systems and genetic algorithms, have taken one or more SC-buzzing techniques for enhancing forest fire identification and evolution [7–12]. Several latest publications focused on forest fire prediction soft computational methods [13–15].

CART has been used to analyze the key variables impacting the frequency and risk of long-term fires in some areas in Europe. For the calculation of fire danger in southeast Italy, the CART approach was used [16]. The critical influence relationship between environmental variables that could influence the incidence of fires in northwestern Spain was analyzed in another study [17]. A novel approach was recently developed to classify the key factors affecting the frequency of fires in five European countries, including Spain, Portugal, Italy, Greece and France [18]. The CART approach is used to extract valuable details about the relationships between variables and is able to accurately predict the frequency of forest fires [17, 19–21].

Several research have used the MARS approach for forecasting possible forest fire events and for estimating burnt areas [22–24]. Compared to other techniques such as CART and ANN, which are mostly used to predict mapping forest characteristics, MARS has been verified to be accurate and efficient [25]. Moreover, it deals with dynamic data connections which can be used to forecast burned-areas under potential climatic conditions [26, 27]. In forecasting forest fires, ANN has been used and its accuracy has proved to be accurate and reliable [28–30]. This approach involves the detection of a collection of learning parameters: the learning rate, the number of hidden layers, and the number of nodes on the hidden layer [31]. Even, as there are further training iterations, the average performance of the prediction model will be impaired [32].

A variety of sensors are needed to gather such meteorological data for the identification of forest fires, including humidity, temperature and pressure, for example. The machine operator will later use these details to take reasonable responses; however, extracting useful knowledge from the data gathered presents a challenge to the system operators. If the system could accurately forecast the condition of the system based on recent or previously acquired unseen values, this task may have been simpler. Three techniques have been analyzed in this work to establish the best, precise and efficient methodology to predict the burning areas of a forest fire. CART, MARS and ANN are used in these approaches. The MNP data is used in these strategies as inputs. Based on three statistical tests, which involve MAE, MSE and RMSE, the efficiency and output of these strategies have been measured. This document would support to offer extra help for device managers to better prepare for emergency scenarios.

The remainder of this article is formulated the following way. Section 2 introduces four machine-learning algorithms to be used for burned-area prediction: CART, MARS and ANN. The methodological indicators used for the assessment of the methodology are presented in Sect. 3. Conclusions and future directions are provided in Sect. 4.

2 Materials and Methods

2.1 Description of Dataset

In this analysis, the MNP dataset situated in the northeast of Portugal is being used. The dataset, which has a mean temperature of 6–12 °C, was obtained from January 2000 to December 2003. Figure 1 displays the MNP's regional diagram.

The properties of the input consist of 12 attributes and 1 output attribute. Table 1 includes the characteristics and definitions. More material on the MNP dataset is available in [33].

Fig. 1 Montesinho natural park location

Table 1 Description of the attributes of the MNP dataset

No.	Attribute	Description
1	X	Coordinate—X, values ranging from 1 to 9
2	Y	Coordinate—Y, values ranging from 1 to 9
3	Month	Month of the year (January–December)
4	Day	Days of the week (Monday–Sunday)
5	FMMC	Fine fuel Moisture code described by Fire Weather Index System
6	DMC	Duff Moisture code described by Fire Weather Index System
7	DC	Drought code
8	ISI	Initial Spread Index
9	Temperature	In degree celsius
10	RH	Relative humidity (%)
11	Wind	Wind speed (km/h)
12	Rain	Rain (mm/m^2)
13	Area	Total burned area (ha)

2.2 Methods

2.2.1 Cart

The CART algorithm was first proposed in [34]. Its main goal is to classify the data in an easy and understandable way. In order to classify a problem, the value of the target variable (Y) is found by the use of some interest variable (X). The CART method is based on binary recursive partitioning that splits a group of data into two sub-groups to build a tree. In the splitting phase, if a feature value satisfies $\left(X_j^i \leq C\right)$, then the sample is placed to the right, otherwise, it is placed on the left. It uses one variable (X) in every splitting step. Thus, this method requires deciding the variable that is the best choice based on the split criterion. CART not only has the potential to dynamically align groups and manage missed values, but also has various merits, such as cost-sensitive instruction, complex construction of functions and calculation of likelihood tree [35]. Moreover, this method can easily handle outliers with no assumptions (e.g. number of analysis or pre-processing time), which makes it more effective and fast.

2.2.2 Mars

MARS was proposed in [36], as a method to organize relationships between variables in a high dimensional dataset. It builds a model by using divide and conquer strategy. Its main goal is to split the training datasets into separate number segments (basis functions) of differing gradients regression. This gives a greater flexibility for the model to handle linear and nonlinear behavior. The end-points of segments are called knots, where the primary purpose is to assess the end of one data area and the beginning of another. To generate base functions, it uses a stepwise strategy and uses an efficient regression algorithm to select the position of the knot among all variables. The MARS algorithm employs two main phases: forward and backward phases. Forward phase selects the potential knots for the basis function to improve the performance, while backward phase is responsible for removing the basis-functions that produce the smallest increase in the overall model function. Unlike CART, this method uses a set of equations to perform regression tasks and mathematical functions are used to seek optimal solutions for a problem. MARS has many advantages that make it more suitable for optimization and statistical goals. First, it has the capacity to capture complex data in high dimensional data and create an easy to understand model. Second, it has the capability to perform analysis on parameter relative importance, and therefore, it reduces computational time. The MARS model equations can be written as:

$$f(x) = a_0 + \sum_{k_m=1} f_i(x_i) + \sum_{k_m=2} f_{ij}\left(x_i, x_j\right) + \sum_{k_m=3} f_{ijk}\left(x_i, x_j, x_k\right) + \cdots \quad (1)$$

The first sum involves only one variable, the second sum includes the contributions from only two variable interactions, while the third sum represents the contributions from three variables over all the basis-functions and so on. In the first sum, each function can be written as:

$$f_i(x_i) = \sum_{\substack{k_m = 1 \\ i \in \boldsymbol{V}(\boldsymbol{m})}} a_m \boldsymbol{B_m}(x_i) \quad (2)$$

where x is the predictor and a_m is the coefficient of the basis functions $B_m(x)$.

While in the second sum, each bivariate function can be obtained as:

$$f_i\left(x_i, x_j\right) = \sum_{\substack{k_m = 2 \\ i, j \in \boldsymbol{V}(\boldsymbol{m})}} a_m \boldsymbol{B_m}\left(x_i, x_j\right) \quad (3)$$

Similarly, each function in the third sum can be represented as:

$$f_i\left(x_i, x_j, x_k\right) = \sum_{\substack{k_m = 3 \\ i, j, k \in \boldsymbol{V}(\boldsymbol{m})}} a_m \boldsymbol{B_m}\left(x_i, x_j, x_k\right) \quad (4)$$

where V(m) is the variable set linked with mth basis-function and B_m survives backward selection process [37].

2.2.3 ANN

ANN is a widely used classification method in different applications such as medical, industrial, financial, electrical transformers and energy [38]. Neural Network (NN) is divided into two main categories: single-layer perceptron and Multilayer Perceptron (MLP). The single-layer perceptron has a single layer of output layers, while the MLP consists of multi layers of simple taste, neurons that interact by using weighted connections. NN typically includes three main layers: As seen in Fig. 2, the input, output and secret layers are shown. The input and output layers are specified by datasets of input and output. The input number reflects the number of input nodes, while the output number represents the number of output nodes. Hidden layers are intermediate layers between the layers of input and output, and their key purpose is to multiply by weights each value reaches the hidden layer, showing the predetermined numbers, and the lines between the nodes display the movement of knowledge from one node to the next.

The following equations can be used to calculate the outputs of each neuron activity in the hidden and the output layers

$$\boldsymbol{h}_{\boldsymbol{j}} = \boldsymbol{f}\left(\sum_{i=1}^{n} w_{ij} x_i\right) \tag{5}$$

$$\boldsymbol{y}_{\boldsymbol{k}} = \boldsymbol{f}\left(\sum_{j=1}^{i} w_{jk} h_j\right) \tag{6}$$

where h_j is the number of a hidden node output,y_k is the output of output node, the function f represents the rule for mapping the neuron's total input to its output it is called activation function.

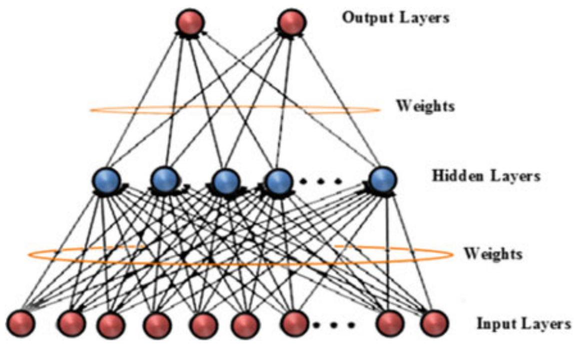


Fig. 2 Multilayer neural network

2.2.4 Performance Evaluation Metrics

The following statistical measures of prediction performance were used to decide how well CART, MARS, and ANN algorithms are performed.

These measures were used in several works on the MNP dataset for calculating error [37, 39, 40] due to their suitability for forest fire forecasting.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i'\right| \tag{7}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i'\right)^2 \tag{8}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_i'\right)^2}{n}} \tag{9}$$

where, the total number of data is n,y_i are the target values of record i and the predicted-values of record i are y_i', respectively.

3 Results and Discussion

The models were built using python 3.0 language libraries which is an environment for machine learning and it is widely used for research and education. All the experiments of this work were conducted on a machine with Win 7, Intel R core (TM) i3 CPU at 3.16 GHz PC and 3.25 GB RAM.

3.1 Data Preparation

The goal from this phase is to normalize numerical fields in the MNP dataset. The normalized e_i for variable E in the *i*th row is computed as:

$$Normalized\,(e_i) = \frac{e_i - E_{\min}}{c - E_{\min}} \tag{10}$$

where $E_{\min}$ and $E_{\min}$ represent the minimum and maximum values for variable E respectively.

For the data fields that are categorical in nature such as, day and month, they are encoded into a binary array using one hot encoding method where each category value is mapped to a binary vector instead of normalizing them between the [0, 1] range. For example, the day field is broken into each day of the week starting from

Sunday. This means that of the day is Monday, then it will be represented as a binary array [0, 1, 0, 0, 0, 0, 0] and Wednesday would be [0, 0, 0, 1, 0, 0, 0]. Similar fashion was also employed for the month field, starting from January.

As confirmed by the work of [33], and in order to improve prediction accuracy and speed-up the models generation process, all attribute except for the RH are used as inputs in the forecasting methods and they are split into training and testing dataset. Tenfold cross validation was used to avoid picking particular parts that are for training and testing [41]; the resulting data was split into 10 parts; the procedure starts by dividing the dataset into 90% for training and 10% for testing. In order to finalize the process, the procedure was repeated 10 times to allow each part of data being as a testing data for each algorithm were used in this work. The default values of the algorithms parameters are used to find best accuracy achieved by each algorithm were used in this work.

3.2 Performance Analysis

The CART, MARS, and ANN techniques are evaluated by the defined statistical measures. The performance results for the forecasting techniques are provided in Table 3 for the testing data.

The results show that MARS predicts burned area probability quite better. The results reveal that MARS outperforms CART and ANN methods. In Table 3 it can be seen that MARS algorithm has the smallest MAE of 1.215, MSE of 2.148 and RMSE of 1.399 in the test stage, while the CART has the largest results MAE of 1.585, MSE of 4.566 and RMSE of 2.112. In general, MARS has the best forecasting capability for forest fires.

MARS algorithm has achieved the best results compared to other forecasting techniques were used in this work because of its fast computing rate. This can be easily confirmed and checked in Table 3. Similar results were obtained by Moisen and Frescino [25] and Amatulli et al. [26]. The best model results for foresting burned areas of a forest fire are shown in Fig. 3. The difference between actual and expected values for the used dataset generated by MARS technique in the form of scatter-plots is provided. The predicted values are in blue color and the actual values are in red color.

Table 3 Comparison among different techniques based on the statistical measures

Perdition technique	Statistical measures		
	MAE	MSE	RMSE
CART	1.585	4.566	2.112
MARS	***1.215***	***2.148***	***1.399***
ANN	1.583	3.499	1.831

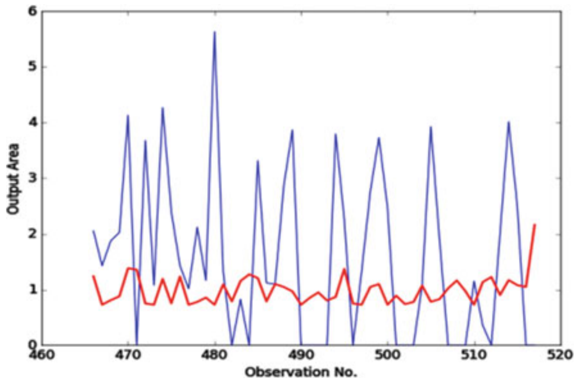


Fig. 3 Forest fires predicted values based on the best algorithm (MARS)

Fig. 4 Display of the training and testing time for different methods

3.3 Computational Execution Time

The required average computational time in seconds for training and testing the model used in this work is illustrated in Fig. 4. It can be seen in the figure that that MARS model required less execution time, while CART and ANN needed more execution time.

4 Conclusion

This work aimed to evaluate four machine-learning algorithms: CART, MARS and ANN using MNP dataset. A set of statistical measures are used to evaluate the used models and their performance: MAE, MSE and RMSE. According to statistical measures results, we can conclude that MARS would be efficient to obtain high level of accuracy in predicting and estimating of burned areas of a forest fire with a comparable computational cost compared to the methods were used in this work. The results from this work not only will assist systems operators to monitor their system efficiently but it will also provide useful information for long-term planning and estimating forests burned areas.

References

1. Mandallaz D, Ye R (1997) Prediction of forest fires with Poisson models. Can J For Res 27(10):1685–1694
2. Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira JM (2012) Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. Ecol Manag 275:117–129
3. Mahdipour E, Dadkhah C (2014) Automatic fire detection based on soft computing techniques: review from 2000 to 2010. Artif Intell Rev 42(4):895–934
4. Angelis DE, Ricotta C, Conedera M, Pezzatti GB (2015) Modelling the meteorological forest fire niche in heterogeneous pyrologic conditions. PLoS ONE 10(2):1–17
5. Zadeh LA (1994) Soft computing and fuzzy logic. IEEE Softw 11(6):48
6. Goel N, Singh S, Aseri TC (2013) A comparative analysis of soft computing techniques for gene prediction. Anal Biochem 438(1):14–21
7. Satir O, Berberoglu S, Donmez, C (2015) Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem. Geoms Nat Hazards Risk 7(2):1645–1658
8. Fowler A, Teredesai AM, Cock DE (2009) An evolved fuzzy logic system for fire size prediction. In: Proceedings of the 28th international conference on fuzzy information processing society (NAFIPS). Annual Meeting of the North American, pp 1–6
9. Yuan C, Zhang Y, Liu Z (2015) A survey on technologies for automatic forest fire monitoring detection and fighting using unmanned aerial vehicles and remote sensing techniques. Can J For Res 45(7):783–792
10. Radovanović M, Vyklyuk Y, Milenković M, Vuković DB, Matsiuk N (2015) Application of adaptive neuro-fuzzy interference system models for prediction of forest fires in the USA on the basis of solar activity. Therm Sci 19(5):1649–1661
11. Artés T, Cencerrado A, Cortés A, Margalef T (2016) Time aware genetic algorithm for forest fire propagation prediction: exploiting multi-core platforms. Concurr Comput Pract Exp 29(9):1–18
12. Davide A, Giovanni B, Giorgio V (2014) Calibrating Rothermel fuel models by genetic algorithms. In: Proceedings of the 7th international conference on forest fire research, pp 102–106
13. Holsten A, Dominic R, Costa L, Kropp JP (2013) Evaluation of the performance of meteorological forest fire indices for German federal states. For Ecol Manage 287:123–131
14. Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151
15. Siljander M (2009) Predictive fire occurrence modeling to improve burned area estimation at a regional scale: a case study in East Caprivi, Namibia. Int J Appl Earth Obs Geoinf 11(6):380–393
16. Amatulli G, Rodrigues MJ, Trombetti M, Lovreglio R (2006) Assessing long-term fire risk at local scale by means of decision tree technique. J Geophys Res Biogeosci 111:1–15
17. Lozano FJ, Suárez-Seoane S, Kelly M, Luis E (2008) A multi-scale approach for modeling fire occurrence probability using satellite data and classification trees: a case study in a mountainous Mediterranean region. Remote Sens Environ 112(3):708–719
18. Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira JM (2012) Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. Ecol Manag 275:117–129
19. McKenzie MJ, Peterson K (2000) Agee fire frequency in the interior Columbia River Basin: building regional models from fire history data. Ecol Appl 10(5):1497–1516
20. Aertsen W, Kint, V, Van Orshoven OK, Muys B (2010) Performance of modelling techniques for the prediction of forest site index: a case study for pine and cedar in the Taurus mountains. In: Turkey XIII world forestry congress, pp 18–23
21. Amatulli G, Camia A (2007) Exploring the relationships of fire occurrence variables by means of CART and MARS models. In: Proceedings of the international conference on IV wildfire, pp 13–17

22. Terrier A, Girardin MP, Périé C, Legendre P, Bergeron Y (2013) Potential changes in forest composition could reduce impacts of climate change on boreal wildfires. Ecol Appl 23(1):21–35
23. Balshi MS, McGuire AD, Duffy P, Flannigan MD, Walsh JE, Kicklighter DW, Melillo J (2008) Assessing the response of area burned to changing climate in western boreal North America using multivariate adaptive regression splines (MARS) approach. Glob Change Biol 15(3):578–600
24. Boulanger Y, Gauthier S, Burton PJ (2014) A refinement of models projecting future Canadian fire regimes using homogeneous fire regime zones. Can J For Res 44(4):345–376
25. Moisen G, Frescino TG (2002) Comparing five modelling techniques for predicting forest characteristics. Ecol Model 157(2):209–225
26. Amatulli G, Camia A, San-Miguel-Ayanz J (2013) Estimating future burned areas under changing climate in the EU-Mediterranean countries. Sci Total Environ 450:209–222
27. West AM, Kumar S, Jarnevich CS (2015) Regional modeling of large wildfires under current and potential future climates in Colorado and Wyoming USA. Clim Change 1–13
28. Maeda EE, Formaggio AR, Shimabukuro YE, Arcoverde GFP, Hansen MC (2009) Predicting forest fire in the Brazilian Amazon using MODIS imagery and artificial neural networks. Int J Appl Earth Obs Geoinf 11(4):265–272
29. Karapilafis FG, Iliadis L, S Spartalis, Katsavounis S, Pimenidis E (2013) Modeling spatiotemporal wild fire data with support vector machines and artificial neural networks. In: Proceedings of the international conference on engineering applications of neural networks, pp 132–143
30. Goldarag YJ, Mohammadzadeh A, Ardakani AS (2016) Fire risk assessment using neural network and logistic regression. J Indian Soc Rem Sens 1–10
31. Safi Y, Bouroumi A (2013) Prediction of forest fires using artificial neural networks. Appl Math Sci 7(6):271–286
32. Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals computing design and application. J Microbiol Methods 43(1):3–31
33. Al-Janabi S, Al-Shourbaji I, Salman MA (2017) Assessing the suitability of soft computing approaches for forest fires prediction. Appl Comput Inf. https://doi.org/10.1016/j.aci.2017.09.006
34. Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24(6):417
35. Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression and regression trees. CRC Press
36. Kaghed NH, Abbas TA, Ali SH (2006) Design and implementation of classification system for satellite images based on soft computing techniques. In: Proceedings of the 2nd international conference on information and communication technologies, Damascus, pp 430–436
37. Friedman JH (1991) Multivariate adaptive regression splines. Ann Stat 19(1):1–141
38. Al-Janabi S, Rawat S, Patel A, Al-Shourbaji I (2015) Design and evaluation of a hybrid system for detection and prediction of faults in electrical transformers. Int J Electr Power Energy Syst 67:324–335
39. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York, USA
40. Castelli M, Vanneschi L, Popovič A (2015) Predicting burned areas of forest fires: an artificial intelligence approach. Fire Ecol 11(1):106–118
41. Salzberg SL (1997) On comparing classifiers: pitfalls to avoid and a recommended approach. Data Min Knowl Disc 1(3):317–328
42. Chen SH, Jakeman AJ, Norton JP (2008) Artificial intelligence techniques: an introduction to their use for modeling environmental systems. Math Comput Simul 78(2):379–400
43. Zhang JH, Yao FM, Liu C, Yang LM, Boken VK (2011) Detection emission estimation and risk prediction of forest fires in China using satellite sensors and simulation models in the past three decades—an overview. Int J Environ Res Public Health 8(8):3156–3178
44. Saghri JA, Radjabi R, Jacobs JT (2011) Early forest fire detection using principal component analysis of infrared video. In; Application of digital image processing. Proceedings of SPIE, vol 8135, p 81351A

45. Davenport TM (2012) Early forest fire detection using texture analysis of principal components from multispectral video. Doctoral dissertation. California Polytechnic State University, San Luis Obispo
46. Xu D, Shao G, Dai L, Hao Z, Tang L, Wang H (2006) Mapping forest fire risk zones with spatial data and principal component analysis. Sci China Ser E Technol Sci 49(1):140–149
47. Garges DC (2015) Early forest fire detection via principal component analysis of spectral and temporal smoke Signature. Doctoral dissertation. California Polytechnic State University, San Luis Obispo
48. Van Wagner CE, Forest P (1987) Development and structure of the Canadian forest fire weather index system. Petawawa National Forestry Institute

Shill Bidding Detection in Online Auction

Anuja Vijay Pawar, Aishwarya Menon, Vedika Vishwanath Painjane, and Rashmi Dhumal

Abstract The acceptance of online auction systems has been increasing gradually over the years and this has led to a surge in many deceitful activities performed during transactions and functions in an online auction system (e.g. shill bidding, bid shielding, etc.). Exploitation of vulnerabilities in such systems for personal benefit in the past two decades has made winning the trust of customers a very difficult task. In this paper, we focus on the problem of shill bidding. To counteract shill bidding, we have used decision tree algorithm incorporating certain parameters which combine to point out who the shill bidder is in the online auction. At each stage we calculate the gain values, choosing the attribute with maximum value as the decisive factor for that particular pass. Each time a particular parameter for shill activity is detected it adds on to shill score and once this score exceeds our predefined boundary value that particular user is nabbed as a shill bidder. The admin can take appropriate steps to block this user in real-time to prevent him or her from participating in further auction process. Real time detection of the shill bidder as opposed to detection post auction proceedings is an aspect that has been looked into and worked upon.

Keywords Online auction · Decision tree · Shill bidding · E-commerce · Detection

1 Introduction

Auction is the process of selling a variety of items which include man-made things, industrial products, attractive antique collection and many more. In the traditional bidding format the bidder has to be present physically on the decided geographical location. But as the time changed and technology advanced, online auctions have gained popularity. In online action, the bidder need not be physically present; he/she can bid for a particular item through online process. Online auction provides equal opportunity to all the global bidders given the online platform, which is not possible

A. V. Pawar (✉) · A. Menon · V. V. Painjane · R. Dhumal
Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India

R. Dhumal
e-mail: rashmi.dhumal@rait.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_23

in the traditional auctions. The platform is user friendly and everything can be done on only one mouse click. The buyer gets an invitation to join the auction where he can decide what to bid on. The buyer can bid on items and will receive notification about whether he won the auction or not by the seller. Online auction has many advantages as stated above but there are related disadvantages too, one of the most recorded cyber crimes in online auction is the real time fraud that takes place. The Internet Crime and Complaint Center (IC3), which is monitoring for internet related crime reporting, stated that in 2007 more than 200,000 complaints were recorded and the number doubled up to 400,000 in just five years. The major disadvantage in many shill bidding detection methods is that the systems are not capable enough to nab the shill bidders in active time of auction. This means that fraud has already occurred and caused loss to both the auctioneer and other bidders in the auction. Many auctions place a limit for the maximum price but in this too; the detection takes place after the auction has ended. Shill bidding refers to fake bidding where the shill bidder inflates, i.e., increase or decrease the bid of the item suddenly. It is one of the major threats in cyber fraud and the auction hosts and users have lost over 250 million dollars in 2008. Shill bidding is not easy to detect in real-time scenario.

2 Literature Survey

Anowar et al. [1] emphasizes on moneymakers who use different patterns to conduct the shill bidding activity in the online auction. The paper defines various characteristics of the shill bidder. The detection in the bidding process can be done by using the classification models which are implemented in this paper. Various stages are explained in this paper like for the partitioning of the data into clusters hierarchical clustering is used. For labeling the bidder into normal and suspicious semi-automated approach is done. Multiple sampling methods are used to resolve the imbalance problems. Randomized Search Cross Validation which has different categories of fold like fivefold and tenfold is used to compare the efficiency of the models. Lastly, testing performance is done using the shill bidding classifiers. It focuses on various models for the fake bidder detection process but the only drawback could be seen as the shortage or inefficiency of the real-time training data-set.

Ford et al. [2] uses hierarchical clustering technique to separate the users into normal and anomalous groups. The anomalous group deviates from the norm and hence viewed as suspicious. Further analysis is done on the anomalous group to find if there are any potential shill bidders. Attributes like user, stage and auction attributes are used to identify the suspicious bidders. After identifying and labeling the clusters are formed. The formed clusters are used as training sets in decision tree algorithm. The decision tree creates its own classification and if this classification matches with the hierarchical clustering classification, it is marked as correct. The results are validated with three-fold cross-validation process. The result is mapped into three classes: normal, suspicious and very suspicious. However, there is an over fitting concern which results in fixing the maximum depth of the tree to 3. The result is generated after the auction, which makes it difficult to take action against the Shill bidder.

Majadi et al. [3] talk about identifying shill bidders by applying various techniques in a way that it prevents the fraudsters from looting the victim. The paper shows the process working in real time with the fake bidder. It shows classifications of shill bidding, reserve price shilling and competitive shilling. It consists of forming bidding patterns dynamically based on neural network-based classification using training data. This paper pays particular attention to the study of behavioral patterns of shill bidders for developing a tactic used to detect shill bidding based on previous research systems. The strategy consists of attacking frauds conducted by shill bidders by applying effective techniques like Bayesian network, neural network and statistical methods. The only future scope found in this paper was about the improvement of accuracy in bidding process and techniques used to nab the shill-bidder.

Mamun and Sadaoui [4] propose software architecture. This architecture consists of following layers, application layer, data layer and business layer. The first layer shows that the information about the user, present and past bids and product information. The second layer is used to store the credential information of the participating users. The third layer performs the main function as this layer processes the information of the user. The status module in business layer assigns statuses based on metrics like Using Period, Using Manner, and Shill Attempt. The security module in business layer is the most significant, as it performs Bidding Behavior Tracking and IP Tracking. It is used to block the fake bidders. The main disadvantage of this system is that it is not able to detect collusive shill bidders.

Wang [5] proposed classifying the fraudulent behavior patterns into three main groups on the basis of the underlying agents of such patterns, namely, those caused by a tenderer, a corresponding tenderer and also those by third parties like auction site. It detects suspicious bids using features like premium amplitude that records amplitude increases in prices of items, bid distribution that measures overall distributivity of prices and items across the particular auction. Here an interesting parameter is a similar goods which takes into account the measure of similarity or difference between goods being sold. The data set is trained as per the above mentioned parameters to nab shill bidders if a particular sequence of events that are suspicious may occur. Decision tree technique is used here to detect patterns of potential shill bidding with bid frequency being the decisive parameter, thus getting the status of the root attribute. Entropy values and gain are calculated in each iteration and each time the element with highest gain becomes the node of the decision tree for that particular pass of the algorithm. The advantage is its improved efficiency in nabbing shill bidders as it accurately determines who is a normal bidder and who is dishonest. It cannot completely avoid shill bidding but can be used to counteract its harmful effects to some extent.

3 Proposal

We have inferred a proposal on the basis of obstacles faced in the existing system for Fake Bidding Detection. The main focus of this system is to detect real-time shill bidding. For achieving the objective of forbidding fraudulent bidders.

3.1 Patterns Observed in Real Time Online Auctions

We propose to detect patterns of shill bidding in online auctions so as to apprehend those who are guilty of this crime. Some of the suspicious patterns that have been observed in real auction data include:

- Pattern 1: To steeply fluctuate prices the bidder resorts to continuously outbidding his own bids despite being the top bidder.
- Pattern 2: Short intervals between successive bids of the same person.
- Pattern 3: Huge price fluctuations from original price range.
- Pattern 4: More numerous bids in the starting half of auction.
- Pattern 5: A shill bidder has the most number of bids as compared to others.

3.2 Security Perspective

Security is managed by decision tree algorithm. The parameters that we input to the system for each user are:

- SelfOutBidding: Here this parameter will record how many times a particular user will outbid the bids that he has placed on an item in the website. The threshold of maximum number of consecutive bids to 2 successive bids. Any user exceeding this limit will be leading to increment his shill score. SelfOutBidding ≥ 3. Means shill score incremented by one.
- AvgOutBidInterval: In this case this parameter will record the average time interval between the successive bids of an individual. A shill bidder will ensure that there is a very less time gap between the bids he places on an item. The threshold for the time gap between successive bids as less than 5 min. Any person who bids in less than five minutes for the next bid is placed under the scanner. AvgOutBidInterval ≤ 5 min. Means shill score incremented by one.
- AvgBidIncr: In this case this parameters will record the percentage increase in the bid placed from what was previously placed by the same user for that particular item. A shill bidder will try to increment the price as much as possible to see a maximum deviation from original price. AvgBidIncr $< 10\%$. Means shill score incremented by one.
- FirstHalfFreq: This parameter records the sum total of bids made by users in the first half of the auction.

- SecondHalfFreq: This parameter records the sum total of bids made by users in the second half of the auction. The comparison of above two parameters is: if FirstHalfFreq SecondHalfFreq then the user is put under scanner for shill bidding because most shill bidders bid aggressively in the start of auction so as to get more time for their activity.
- BidOccurence: This parameter records bidding frequency of users.
- BidTot: Total number of bids submitted in the online auction that is processed by our system. The above two parameters are compared as: If BidOccurence > BidTot/2 then the user is put under scanner for shill bidding because the bids placed from a single user exceeds more than half of the total bids placed in the system.
- Shill Score: The parameter which records the instances of shilling for particular user and finally aids to nab a fraudulent bidder.

Further, it will be seen how these above mentioned parameters play their individual roles in identifying any suspicious patterns of use during an auction process.

Before discussing the working, the flow of the system is explored. The admin or auctioneer will be able to login into the website as the admin. The products will be separated according to the categories. After successful login, the admin will be able to add or remove categories of the products. He will be the person who would decide whether a product should be up for auction or not. The system is designed such that the auctioneer can review the shill activity and once he detects a fraudulent bidding activity he can nab the fraudulent bidder and prevent him from participating in any more auctions on the website by blocking his particular account due to which the shill bidder is no more able to place any more bids. He can log in with his credentials but as soon as he presses the "place bid" button he gets a notification that he has been blocked by the platform.

The basic flow for our methodology focuses on how the decision tree algorithm works with the parameters that we have assigned to nab the fake bidders as shown in Fig. 1. In our system, a few variables are assigned to every user at the time of joining the website to keep a record of all their interaction. Every time a new activity comes up be it a new bid or a new purchase, it is added to this record. This is an important

Fig. 1 Decision tree

part in our system because the algorithm takes these records as input for deciding if the user is a fraud. In a nutshell, we quantify the behavior of every user i.e., generate a value based on the behavior and evaluate exactly how suspicious that user is.

Whenever a user makes a bid on any item on the website, the system first finds out whether this user is bidding on this item for the first time or if he has done it before. If this is for the first time then it directly updates the previous value of the item. That is provided the new value exceeds the previous one. On the other hand, if this is not a first time bid then it takes Bid Frequency as input. The algorithm checks if the Bid Frequency is high and if so it assigns a shill score of 1. If not then it checks if the Bid Frequency of user on this item is low. If it is found to belong to the low frequency range, it can still be that the bidder is a shill bidder with a low number of bids and so it checks the average bid time parameter. A shill score of one is assigned to the user if his average bid time parameter is found to be high. However if the Average Bid Time is not high then algorithm does not proceed with further checking, having confirmed that this user is not a shill bidder.

On the other hand if the bid frequency parameter is neither low nor high then it goes to check if the outbid frequency is high or not. If the Average Outbid Frequency is found to be high then. It increments the shill score by one. If OutBidFrequency parameter is not high then it checks if it is low and if low the algorithm terminates further checking. If this Outbid Frequency parameter is neither high nor low this indicates that the algorithm has not as yet determined with full surety that this user is shill bidder or not. Hence it progresses to the next stage where the Average Bid Frequency parameter is checked. If this parameter lies in the low range, the algorithm stops further checking and determines that this user is a legitimate one. However if it is in high or medium range we assign the shill score one as the algorithm has, through several systematic checks has determined the user to be a shill bidder. Thus the flowchart given here explains the basic flow of the algorithm used in the system and aids to comprehend its working.

The main crux of the working is based on the classic decision tree approach where calculation of the entropy values and the corresponding gain is done. It is done for each iteration so that after the comparison of gain values of every attribute we can choose the attribute with maximum gain value. In the first iteration this value becomes the root of the decision tree. In the following iterations we get subsequent attributes as a part of the tree to ultimately the entire tree. In our given example of 8 users as shown in Table 1 we calculate the gain for each of the attributes mentioned above and BidOccurence becomes the root as it has maximum value as compared to the other parameters. BidOccurence becomes the starting point of our tree which is now trained to check out whether BidOccurence lies in the low, medium or high range.

A high BidOccurence will lead to increment of shill score whereas the values ranging in medium or low ranges are checked further. If BidOccurence is low AvgBidInterval is also checked where a low count of this parameter will signal to shill score not being increased and high will lead to an increment in shill score. If BidOccurrence falls in the medium ranges then we check the next parameter SelfOutBidding which if falls in the low ranges would not increase the shill score and high will increment shill score whereas a medium range score for this parameter

Table 1 Scenarios

SelfOutBidding	FirstHalfFreq	BidOccurrence	AvgOutBidInterval	AvgBidIncr	Shill status
1	p	N	4	12	Yes
1	q	R	8	15	Yes
5	b	Q	3	4	Yes
2	a	A	7	3	No
4	r	R	3	3	Yes
3	c	D	2	2	No
8	d	E	2	8	No

will lead us to the next parameter AvgBidIncr whose low ranges point to no suspicious activity but high value ranges point to suspicious activity and thus an increment in the shill score.

Each time the fraudulent activity is detected by the above mentioned parameters the shill score parameter is incremented by one. We have fixed a threshold of 3 for this based on our study of works by various others who have worked to combat instances of fraud in online auctions. Any user who records a shill score of three or exceeding three is labeled as a shill bidder by the system. The decision tree approach used thus leads us to nabbing the shill bidder with just mentioned parameters serving as inputs to it. It is trained to compute these values at every stage and compare and arrive at a proper conclusion by appropriate comparisons of various high, low or mid ranges of the values and thresholds set in the system. While implementing the system we were faced with a scenario which can be described as in Table 1.

In Table 1, the alphabets A, B Z correspond to their individual values in the alphabetic sequence and a…z show whether it occurs in the first or second half with m = 13 as the midpoint and the values preceding it indicating bids in the first half and those following it indicating bids in the second half.

Formulae used:

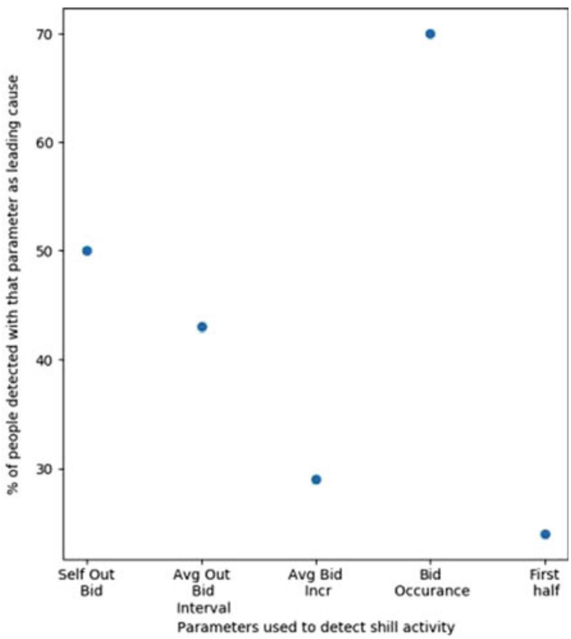
- P(YES) = COUNT(YES)/COUNT(TOTAL_SHILL_STATUS)
- P(NO) = COUNT(NO)/COUNT(TOTAL_SHILL_STATUS)
- ENTROPY = −P/(P + N).LOG2(P/P + N) − N/(P + N).LOG2(N/P + N)
- INFORMATION GAIN = CLASS ENTROPY − ENTROPY_SPECIFIC_ATTRIBUTE.

For the particular Table 1 we found maximum gain of the attribute “BidOccurence” in the first iteration which makes it the root of ID3 tree, following which we found that the attribute having the highest gain from amongst the others was AvgOutBidInterval, and so on it continues in the way described above to give you the complete tree with decision of who is the shill bidder and who are most likely to be shill bidders after which the auctioneer can use his discretion to declare the winner of the auction. The data is generated at run time and the decision tree will help the auctioneer decide to arrive at a proper decision.

4 Result Analysis

The results of the above mentioned work delve into various parameters that are taken into account to calculate the shill score of buyers participating in auction. A total of 80 people were monitored for any fraudulent activities on e-commerce website of which 15 turned out to be shill bidders. The shill score is used to nab potential shill bidders in online auction. As Fig. 2 shows, the major percentage of shill bidders nabbed was through BidOccurrence parameter. Frequently placed bids prove to be a decisive factor in nailing the fraudulent bidders and a whopping 70% of those caught was through this parameters. Second parameter that helps trace signs of shill bidding is SelfOutBid which stands at 50%. Here a bidder tries to overcome his own last placed bid with the next bid multiple times, intended to spur the auction in the direction of his malicious interest. AvgOutBidInterval measuring time interval between successive bids is the next factor (43%), while the last two, namely AvgBidIncr and FirstHalf do not contribute much to the shill nabbing proceedings with the total percentage of caught offenders standing at 29% and 24% respectively. We use the scenario given in Table 1 given above to demonstrate the result. The bidders whose shill score is greater than 3 will be nabbed and displayed to the auctioneer who can then take necessary action against the potential shill bidder.

Fig. 2 Analysis of fraudulent bidder detection



5 Conclusion

A fraudulent activity like shill bidding is a severe problem in electronic commerce. It breaks the trust of the honest user and also challenges the security. This paper presents a system to secure online auctions from shill bidding by detecting potential shill bidders on run-time which can be blocked by the auctioneer. The detection of the shill bidder in online auction is done by the combination of parameters mentioned above and the score generated is used to resolve if the bidder is a shill bidder or not. We have incorporated the given method into an e-commerce website.

References

1. Anowar F, Sadaoui S, Mouhoub M (2018) Auction fraud classification based on clustering and sampling techniques. In: 2018 17th IEEE international conference on machine learning and applications (ICMLA). IEEE, pp 366–371
2. Ford BJ, Xu H, Valova I (2010) Identifying suspicious bidders utilizing hierarchical clustering and decision trees. In: IC-AI, pp 195–201
3. Majadi N, Trevathan J, Bergmann N (2016) Analysis on bidding behaviours for detecting shill bidders in online auctions. In: 2016 IEEE international conference on computer and information technology (CIT). IEEE, pp 383–390
4. Mamun K, Sadaoui S (2013) Combating shill bidding in online auctions. In: International conference on information society (i-society 2013). IEEE, pp 170–176
5. Wang W-J (2016) Decision tree induct-based shill bidding detection. DEStech Trans Eng Technol Res (SSTE)

An Optimal Region Growing Segmentation Algorithm with Decision Tree Tumor Classifier

V. Sivakumar and N. Janakiraman

Abstract In the present day, to enhance the survival rate of the patients and the treatment procedure the detection of brain tumor segmentation plays an important role in the medical field. A large range of MRI images are generated at a time we cannot easily diagnose the brain tumor and it will increase the execution time and not easy to fragment in a short duration of time. Within a short time, we can easily segment and categorize the brain tumor, the automated brain tumor segmentation is used. In the proposed system, the modified region growing algorithm with a modified whale optimization algorithm used to segment the brain tumor segmentation and the MRI image will be classified by decision tree classifier technique. For preprocessing the MRI images in the brain tumor segmentation, the least mean square (LMS) technique is used. From the simulation, the proposed method gives the outcomes with a clear accuracy of 96.6%. MATLAB is used to implement this proposed strategy.

Keywords Least mean square · Modified region growing techniques · GLCM feature extraction · Whale optimization algorithm · Decision tree classifier

1 Introduction

The unusual cells or tissue from within the brain is known as a brain tumor. Tumor originates from some parts of the brain and expands across the entire part of the brain. Brain tumors can be divided into two types, one is Cancerous tumors and another one is benign tumors. The cancerous or malignant tumor is the most dangerous and fast-spreading tumor nowadays that spread across the brains and spines. It is also of two subdivisions. (1) Primary tumor initiates in the brain. (2) Secondary tumors spread across elsewhere. The benign or non-cancerous tumor is the group of cells or tissue that doesn't attack or not spread throughout the body. When it eliminated from the body it doesn't grow again whether the malignant tumors can be grown even if we removed. Strong headaches, balance loss, confusion, blurred vision are

V. Sivakumar (✉) · N. Janakiraman
Department of Information and Communication Engineering, KLN College of Engineering, Sivagangai, Tamil Nadu, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_24

the symptoms of brain tumors. But sometimes there is no symptom if that only risky. Treatments like chemotherapy, radiation, surgery are mostly used to cure.

The process of sorting out of different tumor tissue from normal brain cells with the aid of an MRI image of the brain. The MRI image gives the full information it will lead to better treatment and diagnosis. But yet it is still a risky task because of irregular forms and complex nature of tumors. The segmentation process is to simplify or editing the representation of an image into something more meaningful and easy to understand. Image segmentation used to pattern recognition, image analysis, low-level analysis and finally robotics.

Decision tree (DT) most used method for inductive inference in Data mining. Decision trees are used to split data into a set of branches (leaf nodes). The construction of the tree is taken in top/down recursive and divide-and-conquer techniques. Constructing a decision tree is simple and does not need previous knowledge. The decision tree initiates with the root node from the top of the tree. Feature with the maximum information gain should be in the root. The information characterized in the decision tree can be extracted and marked as IF–THEN rules. The DT classifier will analyze and classify MRI as abnormal or normal MRI images (Fig. 1).

In this paper, based on the use of modified region growing and Decision Tree. This proposed method consists of five phases viz., BRATS dataset as input MRI image, Pre-processing using Least-mean-square (LMS), Modified region growing to have normal region growing and orientation constraint, Feature extraction, Final classification using Decision tree (DT). Based on the wordings from the World Health Organization, brain tumors are divided into the following groups as Benign or Pilocytic. Glioblastoma. Anaplastic, Astrocytoma.

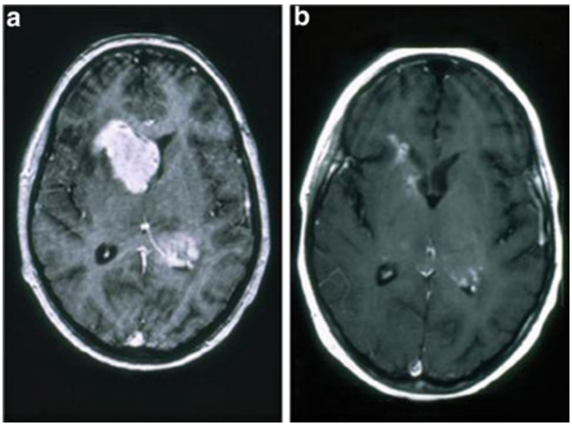


Fig. 1 MRI image with tumor and without tumor

2 Literature Survey

Nowadays computerization methods are used to detect medical imaging to detect the image through inner portions of the human body for diagnosis of medical. The image segmentation is one of the most important technology has a role in diagnosis, navigation, surgical planning, and different medical evaluations. Interest showing in existing segmentation is automatic and semi-automatic methods. Praveen and Agrawal et al. [1] proposed a hybrid method for brain tumor detection and classification through Magnetic Resonance images (MRI). There are four steps take place. The first one is image preprocessing includes filtering of noise, detection of the skull. The second one proceeds with using a grey level co-occurrence matrix with feature absorption of MR brain images. The third steps deal with the categorization of inputs into normal or abnormal using the Multilayer perceptron kernel and the Least Squares Support Vector Machine classifier. The final step utilizes a fast bounding box for the segmentation of tumor parts from the brain. 100 images are implemented for this experiment consists of 25 normal and 75 abnormal. The result of the accuracy is 96.63%.

Akram et al. [2], discussed new threshold-based segmentation methods act as channel sources for an automatic brain tumor diagnosis system. The conjunction of the beta mixture model and learning automata (LA) is used for the proposed segmentation method. The given MRI image approximates the histogram values through the fusion of Beta functions whose values are resolved through the LA tool. One pixel class values are shown for every beta function and the threshold values that are obtained through two closest beta functions which are used for segmentation. The segmentation is applied through binary classifiers which are based on support vector machine (SVM), Decision tree (DT) and K-nearest neighbor (KNN). From the TCIA dataset and the Harvard medical school, 79 MRI scanned images are inspected to evaluate the performance of the segmentation process. The outcomes prove that it provides great average values for image similarity indices and with an accuracy of 98% in tenfold cross-validation.

Present days, the most promising and demanding area is Medical image processing. The images are the most important thing for the development of image segmentation. The damaged region from magnetic resonance brain image is extracted, identified and segmented but still, it is a prolonged task for the medical experts. So it is important to use the computer-aided technique to beat this limitation. Lakshmi et al. [3], suggested the Adaptive threshold algorithm to overcome the problem of segmentation and also improve the veracity and performance in the medical segmentation process. The main important thing is deep learning CNN do image recognization, image classification, face recognization, and object detection to improve accuracy. After the execution proposed method outcome has checked and ratified based on some parameters. MATLAB software is used for the segmentation process.

To detect the brain tumor numerous approaches of image segmentations are approaching the modern medical field. There are many new types of equipment and machines are available in the modern medical field such as magnetic imaging scans (MRI), Computed tomography (CT) scan, X-ray is used for brain segmentation. Kavitha et al. [4], proposes a new method of segmentation. It merges the modified region growing and genetic algorithm for tumor detection. This method follows four-phase (i.e.) Preprocessing, Segmentation, Classification, and fitness calculation. To eliminate the noise present in the images the pre-processing phase is using. It uses the Gaussian filter to reduce the noise. After preprocessing, noise removed images are sent to the segmentation part. There Modified region growing method contains orientation constraints that are used to analyze the image. Then the tumor is classified as normal or abnormal using Back propagation neural network (BPNN). Finally, the fitness calculation is used to find the values from the MRI image of the segmented tumor.

Ali et al. [5], explaining that based on the doctor's knowledge and experience the brain tumor is done. For their help, they are using an automated tumor detection system to identify the brain tumor. This method consists of three stages (1) Preprocessing (2) Classification (3) segmenatation. In pre-processing the tumor classification is extreme learning machine local receptive fields (ELM-LRF). First, to remove the noise in the image the local and nonlocal smoothing method is used. Second, to identify whether the tumor is cancerous or non-cancerous the MR images are classified. Finally, the segmentation takes place. To reduce the physician's time using only the cranial MR images which have mass. The cranial MR image has the accuracy by using this technique is 97.18%. This method is more effective than earlier methods using computer-aided brain tumor detection.

Vinoth et al. [6] proposing the programmed division strategy based on Convolutional Neural Network (CNN). For differentiation of tumor, the kernels are used here. It is the deep neural network commonly applied to evaluate the performance of a visual imaginary. It is the multilayered perceptron of regularized versions. An MRI image is the imaging system to analyze the tumor but a large amount of information is produced in clinical practice. To analyze the certain characters of the image the extension of the work is processing. Detection of tumor and extraction of the tumor is done by using the MATLAB tool.

3 Proposed Methodology

In the proposed system, a novel brain segmentation algorithm is considered. The technique is the modified region growing algorithm is incorporated with the modified whale optimization algorithm which is utilized to segment the brain tumor segmentation as well as the tumor image is then classified by using the decision tree (DT) classifier technique. To investigate the least mean square (LMS) technique is utilized to preprocessing the MRI image. It also recognized the MRI image dataset that includes the brain tumor and without tumor images based on the image segmentation technique (Fig. 2).

BRATS dataset implemented as the MRI image. For, brain tumor segmentation, currently it is feasible to objectively contrast different glioma segmentation methods

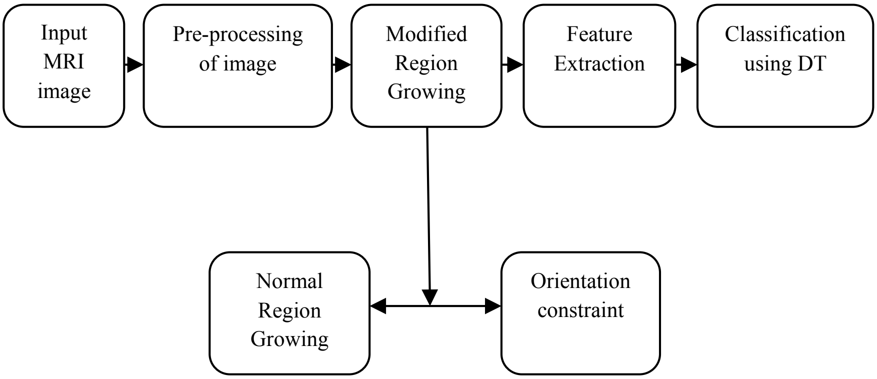
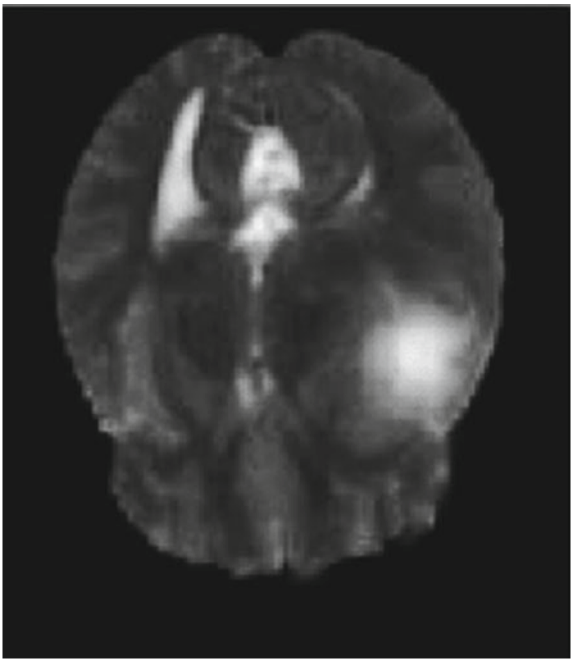


Fig. 2 Flow chart of brain tumor segmentation and classification

with this frequent dataset. This system to attain better outcomes for clasifying the brain tumor catergories using the BRATS dataset in which provides the testing and training process has been evaluated for nthe decision tree classifier. The 70% of the input data is used for the training purpose along with the 30% of the data is used for the testing purpose (Fig. 3).

The flow of the brain tumor system consists of BRATS dataset as Input MRI image, Pre-processing of an image using the LMS method, the Modified region growing for

Fig. 3 MRI image from the BRATS dataset



segmentation, Feature extraction, and Final classification using the DT technique is proposed. Initially, the image acquisition process consists of capturing scanned MRI images and MRI datasets. Read the MRI images from the input datasets. Subsequently, the LMS technique is initialized to using the pre-processing method which is carrying out the unwanted artifacts from the MRI image. Similarly, diverse techniques used for isolation and separation of the region of interest in the segmentation process and accessible the modified region growing for segmentation. Afterward, global features are extracted in the process of feature extraction and it can do the exact location of the brain tumor. Finally, to classify the MRI by using a DT classifier is utilized to identify the affected or non-affected MRI images and output image shows to the surgeon or specialist in diagnosing and making a final medical decision.

3.1 Pre-processing

The pre-processing method is one of the initial and complicated processes in the image processing. Least-mean-square (LMS) algorithm is an adaptive filter that is utilized to pre-processing. The pre-processing method is utilized to eliminate the unwanted artifacts from the MRI image depend on the Linear Mean Square (LMS) algorithm. LMS filter has training filter coefficients. Given below show the mathematical output of image.

The signal of power or energy is calculated by squaring as given below,

$$\mathrm{e} = \mathrm{X} + \mathrm{N} - \mathrm{y} \tag{1}$$

Both of the outcomes are getting expectations

$$\mathrm{E}\left(\mathrm{e}^2\right) = \mathrm{E}\left(\mathrm{X}^2\right) + \mathrm{E}(\mathrm{N} - \mathrm{y})^2 + 2\mathrm{EX}(\mathrm{N} - \mathrm{y}) \tag{2}$$

$$E\left(e^2\right) = E\left(X^2\right) + E(N - y)^2 \tag{3}$$

Adapting filter is used to diminish the error energy will not involve the signal energy. Hence, the minimum energy error is,

$$\mathrm{E}\left(\mathrm{e}^2\right)_{\min} = \mathrm{E}\left(\mathrm{X}^2\right) + \mathrm{E}(\mathrm{N} - \mathrm{y})^2_{\min} \tag{4}$$

$E(e - X)^2$ is also diminished since, (e − X) = (n − Y). So, the total amount of output energy is diminishing is equal to reducing the noise energy.

The LMS construct the least mean square of the error signal by varying the filter tap weight, the updating equation coefficient as expressed,

$$W_{k+1} = W_k + 2\mu e_k X_k \tag{5}$$

where μ is an appropriate step size that can be chosen as $0 < \mu < 0.2$. The Larger step size makes the coefficient to fluctuate widely.

3.2 Modified Region Growing

Modified region growing is a trouble-free image segmentation method based on the region. Initially, this system segmentation is verifying the adjacent pixels like "seed points" and the adjacent pixels are must added in the region based on assured situations. In a normal region growing technique, the neighbor pixels are observed only the intensity" constrain so that fixes the threshold level and those neighbor pixels are satisfied this threshold is chosen for region growing. The fault of the normal region growing is noise or variation of intensity and not differentiates the shading of the real images. To added constrain of "orientation" that is overcome the drawbacks in the normal region growing. It may consequence is holes or over-segmentation. In the modified region growing are utilized the two thresholds they are, intensity and the orientation. If both constraints are met with each other the region is grown after that divides the original image into 4, 18 and 24 grids. The result of the gridding is smaller grids that estimated easily carried out (Fig. 4).

In this proposed system, each of the grids is treated separately in which applied the region growing method. At the first step of the region growing for the grid created due to the selection of the seed point for the grid. The primary region initiates with the accurate place of the seed and also to locate the seed point of the grid histogram estimated carried out when the histogram originate for each pixel in the grid. For every grid, the histogram asset that comes most regularly is elected as the seed point pixel. A few of the seed pixel point is taken as the seed point for the grid. For the greyscale images, the asset of the images is from 0 to 255. After finding out the seed point, the region is grown from it. The neighboring pixels are evaluated with the seed

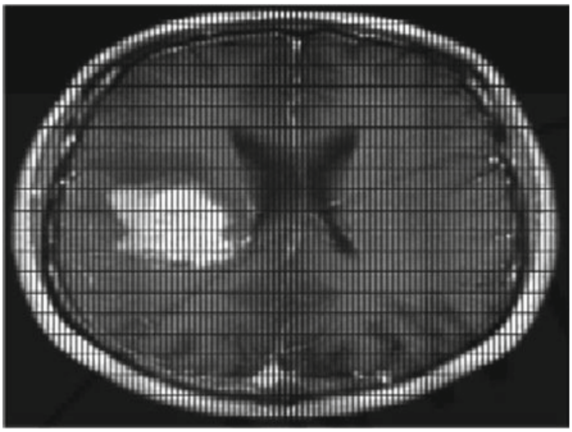


Fig. 4 MRI image with modified region growing segmentation technique

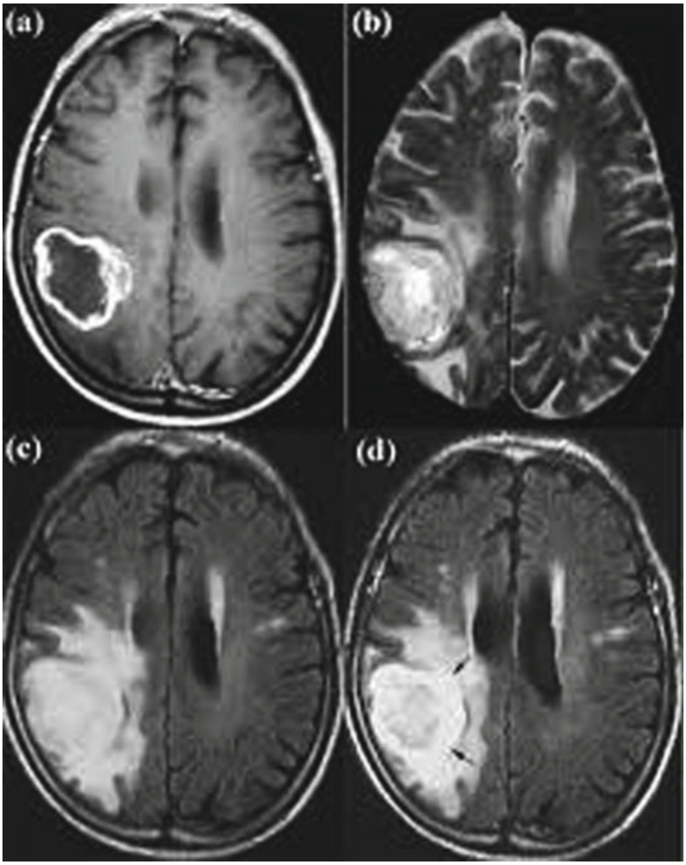


Fig. 5 Diferent types of MRI tumor image

point and if the neighbor pixel fulfills the constraints, afterward the region is grown moreover it is not grown to that pixel (Fig. 5).

3.3 Feature Extraction

Feature extraction is a unique form of dimensionality reduction. Feature extraction is utilized to converting the input file into a set of features vector. When the input data to an algorithm is taken much time to process and it is assumed to be disreputably redundant then the input data can be converted into a lessen illustration in a set of features vector. In this work, Feature Extraction is assisted to discover the exact location, shape, size, and composition of the brain tumor. It is obliging to make out the brain tumor where it is accurately located and assists in forecasting in the next stage.

3.4 Classification Using DT

In this method, the classification algorithm is based on a decision tree. A decision tree is a set of simple rules and also nonparametric because it does not require any assumptions about the distribution of the variables in each class. Each interior node includes a decision criterion based on one feature. Initially split into two parts, the feature is utilized among the highest relevance. This process is recursively repeated for each subset until no more splitting is possible since it pursued from a root to a leaf node the decision tree corresponds to a rule-based classifier. A benefit of decision tree classifiers is their simple structure, which allows for interpretation (most important features are near the root node) and visualization. A decision tree is assembled from a training set, which consists of objects, each of which is completely described by a set of attributes and a class label. The output of the tree is a leaf that is related to the class. When the tree images are misclassified the class label does not match in the class label.

4 Result and Discussion

In the proposed technique, the modified region growing algorithm with a modified whale optimization algorithm is used to segment the brain tumor segmentation and the tumor image is then classified by using the decision tree classifier technique.

The performance is calculated by utilizing accuracy, TPR, TNR and SSIM values with the comparison of the existing technique.

The accuracy, TPR and TNR values are computed based on the confusion matrix strategy. The accuracy is a measure of the degree of closeness of a measured or calculated value to its actual value. The classifier algorithm is estimated by utilizing True positive (TP), True negative (TN), False positive (FP), False negative (FN), accuracy, sensitivity, specificity, and f1 score. Accuracy defines several accurately predicated tuples. Sensitivity describes true positive that is number of tuples (positive) is positively predicted. Specificity defines true negative that is many tuples (negative) correctly rejected. F1 score is the average error of the precision and recall (Fig. 6).

$$\text{SSIM} = \frac{2(\text{True Positives})}{2(\text{TP}) + \text{False Positives} + \text{FN}} * 100\%$$

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negative} * 100\%$$

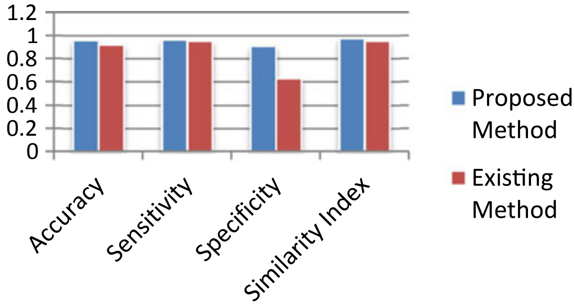


Fig. 6 Performance analysis of brain tumor image

$$Specificity = \frac{True\,Negatives}{True\,Negatives + False\,Positives} * 100\%$$

$$Accuracy = \frac{True\,Positives + True\,Negatives}{TP + TN + FP + FN} * 100\%$$

The performance analysis of the proposed technique has been simulated and analysed with the existing method. The purposed technique attained the better results when compared with the existing method. The existing method used traditional region growing techniue incorprate with traditional decision tree classifier without using any kind of optimization algorithm. The existing method has been simulated and evaluated by using the same dataset which is used to show the better outcomes. The performance of the system has been computed and verified.

5 Conclusion

The brain tumor segmentation and types of tumor is classified is the primary key of this paper. In this paper, the input image is collected from the BRATS dataset and then the input image is preprocessed by utilizing the median filter technique. The outer skeleton of the brain image is eliminated by utilizing the Sobel edge detection strategy. The extracted image is then segmented by the modified region growing algorithm and finally, the type of the tumor is classified by using the DT classifier. The performance analysis of the proposed system is measured by using accuracy, TPR, TNR, and SSIM. Better performance analysis is attained when compared with the traditional region growing algorithm. As a result, the proposed technique provides better outcomes in both segmentation and classification of the tumor.

References

1. Praveen GB, Agrawal A (2015) Hybrid approach for brain tumor detection and classification in magnetic resonance images. In: 2015 communication, control and ıntelligent systems (CCIS). IEEE, pp 162–166. Simpson H (2004) Dumb robots, 3rd edn. UOS Press, Springfield, pp 6–9
2. Akram E, Mosleh M (2019) Improving brain tumor diagnosis using MRI segmentation based on collaboration of beta mixture model and learning automata. Arab J Sci Eng 44(4):2945–2957
3. Lakshmi B, Kiruthika V, Feroz CA, Merlin JAJ (2018) Automated detection and segmentation of brain tumor using genetic algorithm. In: 2018 ınternational conference on smart systems and ınventive technology (ICSSIT). IEEE, pp 583–589
4. Kavitha AR, Chellamuthu C (2017) Brain tumour segmentation and detection using modified region growing and genetic algorithm in MRI images. Int J Med Eng Inform 9(3):269–283
5. Ali A, Hanbay D (2018) Deep learning based brain tumor classification and detection system. Turk J Electr Eng Comput Sci 26(5):2275–2286
6. Vinoth R, Venkatesh C (2018) Segmentation and detection of tumor in MRI images using CNN and SVM classification. In: 2018 conference on emerging devices and smart systems (ICEDSS). IEEE, pp 21–25

Music Detection Using Deep Learning with Tensorflow

Satish Chikkamath and S. R. Nirmala

Abstract Music is an expression through collection of harmonic frequencies whose medium is sound. Group of these frequencies will consist of various elements that create music or non music expression. The main objective of the work carried out is to detect the presence of music in a given audio file using the concept of transfer learning. The literature proves that music detection in an audio file can be done by extracting handcrafted audio features like (ZCR, entropy, AMR, LSTER) and train by using classifiers like SVM, Random forest. The advances in machine learning and deep learning architectures have opened the new path for music detection. End to end classification system performs feature extraction and classification jointly this process may lead to extract new unknown feature and contribute to improve the overall accuracy of the system, however to train the CNN networks from scratch we need huge dataset and its time consuming, hence the need of transfer learning ascends. We have used a tensor flow VGGish model released by google as feature extractor which is trained on Audioset data from YouTube videos and finally trained LSTM (Long short term memory) network, a special kind of RNN for classification.

Keywords Music signals · Transfer learning · VGGish · Tensor flow and LSTM

1 Introduction

Sound is the vibration of air particles, these air particles traverse through our ears making the sound. These vibrations of sound in air are called as sound waves. When a door is knocked or bench is bashed the object vibrates and generate sound waves which transmits through air. Similarly when guitar strings are plucked, the strings vibrate and generate sound waves what makes the difference between the two sound waves. To understand the difference let's look at the two waveforms generated by

S. Chikkamath (✉) · S. R. Nirmala
KLE Technological University, Hubballi, Karnataka, India
e-mail: chikkamath@kletech.ac.in

S. R. Nirmala
e-mail: nirmala.s@kletech.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_25

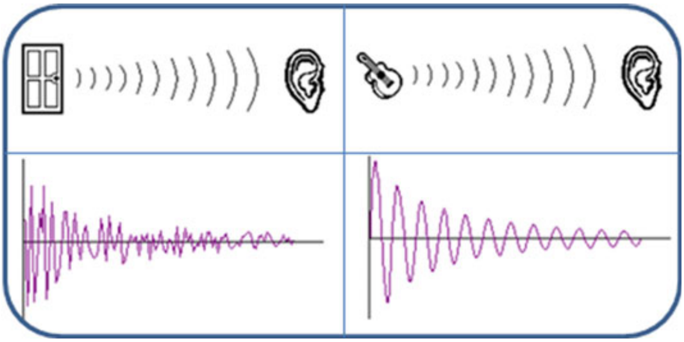


Fig. 1 Sound waves of two medium

two mediums bumpy and irregular which results in harsh sound and also we can notice the waves varies in amplitude from start to end.

On the other side of Fig. 1 we can observe the continuous series of waves (cycles) from guitar string, which are smooth and make a musical tone. The difference between a musical and non-musical sound can be justified by the regularity of the vibration. The vibrations of the musical sounds are regular. When we hear regular vibrations the ear detects the frequency of those vibrations and perceives it as pitch of a musical tone. Non-musical sounds produce irregular vibrations of changing frequencies, ears can still follow the pattern but there is no strong regularity from which we can make a musical tone. The two important properties of vibrations are amplitude and frequency. The work was carried out in makers space workstation [1], implementation results are discussed in later sections of the paper.

2 Literature Survey

Old methods of audio classification or music speech discrimination relies on extracting hand crafted features like Short-term and mid-term feature extraction, low short term energy ratio, amplitude modulation ratio and MFC coefficients. The main objective of feature extraction is to transform the original signal representation into feature vectors and also to reduce data dimensionality. To achieve the above mentioned goal of reducing data dimensionality we need to have deeper understanding of each feature extraction methodology so as to decide which feature helps us to have better classification accuracy. Application specific features are required for real time environments to have sufficient classification accuracies which consume time and labour.

The massive increase of multimedia data over internet has created the importance of online services. The most common online information in multimedia is music recently audio content analysis is important and widely used application [2]. The work proposed in [3] customs an audio indexing system which uses frequency tracking for various content levels of a sound track, this system does not need any prior knowledge. In [4] authors have designed a audio classifier based on automated audio analysis and fuzzy approach [5]. Proposes an optimized audio classification and segmentation algorithm which segments the given audio into four main types namely, environment sound, pure speech, Music and silence. Hybrid classification approach is used in this work i.e. bagged svm is used with artificial neural networks [6, 7]. Proposes speech music discrimination by building a hierarchical structure of decision trees by extracting a set of invariant features and refine the classification result by a strategy of context based state transform. The research work carried out in [8] proves that a multiple kernel learning approach can be carried out to select the optimal sub bands to discriminate the audio signals. The solutions offered by deep learning community have significantly benefitted the research on audio analysis. The work proposed in [9] have adopted convolutional neural network for the task of discriminating music with non-music signals. In comparisons with all the papers explained above, it exploits the capabilities of CNNs by fine tuning the network with relatively small spectrograms. The spectrograms generated are represented as images and these images will be used as input signals to classifier, the use of CNN methodology has proved the power over traditional hand crafted methodologies in speech-music discrimination task along with the efficiency of transfer learning.

In this work we have adopted concept of transfer learning for music detection. We have used tensor flow VGGish model as feature extractor. The VGGish model is trained on very large Audioset data of 632 audio event classes and a collection of 2,084,320 human-labelled 10-s sound clips drawn from YouTube videos. The model outputs 128D embeddings for every second of audio signal which are then, converted to Tfrecord files using Tf (Tensor flow) writer, once the pre-processing of the data is completed we can train the (LSTM) classifier. The Long short term memory (LSTM) network is used to classify the given audio signal into music or non-music signal. The detailed methodology carried out in the work and the results are discussed in the following sections.

3 Methodology

3.1 Concept Diagram

This section gives the information about the methodology carried out in the work, Fig. 2 illustrates the process flow carried out to classify the audio signal. Figure 3 illustrates the pre and post processing of input data. The detailed explanation of each block is discussed in the following section.

3.2 Audio Data Set

3.2.1 Training Data Set

Music class: GTZAN is a dataset created by Tzanetakis [10]. The dataset consists of 10 genres namely Reggae, Blues, Rock, Country, Metal, Disco, Hip Hop, Jazz, Popular, and Classical. Each class includes 100 tracks of 30 s long, so collectively 1000 tracks of pure music is available for training. The tracks are 22,050 Hz Mono 16-bit audio files in .wav format 80% of the class data is used for training and 20% is reserved for testing.

Non music class: As discussed in Sect. 1, non-music files are the one which don't have regular waveforms (frequency). This class has two parts, first part includes manual recordings of door slamming, bench banging, environmental sound recordings with noise. The second part includes other category of sounds like: car horn, siren, dog braking, engine idling, drilling, jack hammer, air conditioner. Overall this class has 1000 non music files in which 80% of the files are used for training and 20% used for testing.

3.3 Pre-processing of Data

The task of the program can be simplified and the results can be improved if input audio signals are appropriately manipulated and then fed into the network, in our work as can be seen in Fig. 2. The audio sample is resampled to 16 KHz and spectrograms are generated by applying short term Fourier transform with a window size of 25 ms and also generated spectrogram is mapped to 64 mel bins covering 125–7500 Hz.

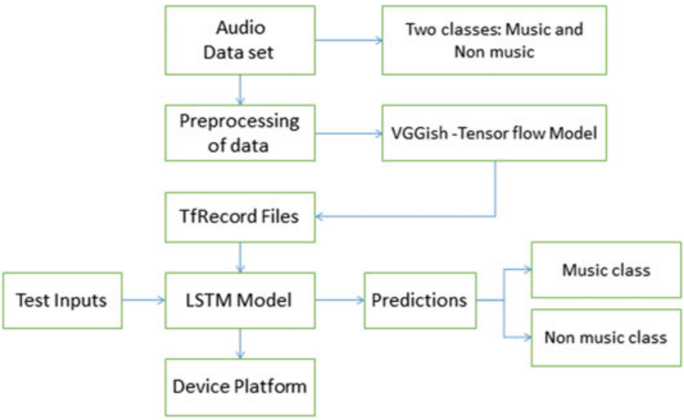


Fig. 2 Concept diagram for classification

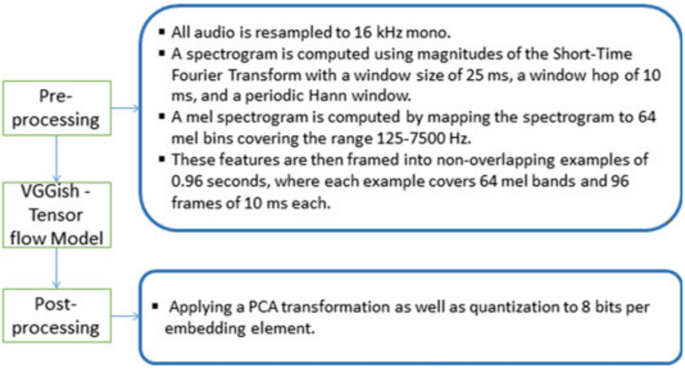


Fig. 3 Pre and post processing of input data

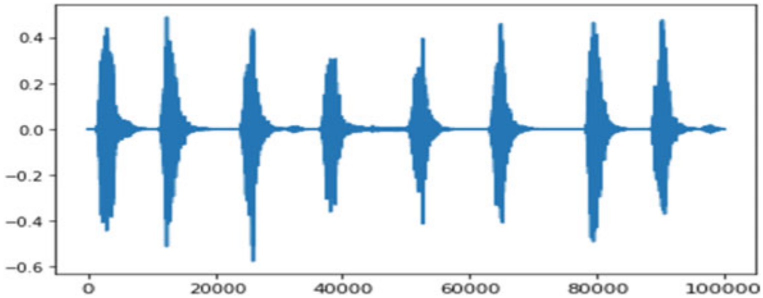


Fig. 4 Random sound for visualisation

These features are then framed into non overlapping examples of 0.96 s i.e. 96 frames of 10 ms each. Once these features are trained into VGGish model before releasing these features a PCA transformation is applied and quantised to 8 bits per embedding element. Finally these embedding elements are converted to Tensor flow record files by using TF writer. TF record files stores the data as a sequence of binary strings.

Sound preparation: As discussed above we have used VGGish tensor flow model as feature extractor, short explanation of the pre-processing step is discussed below

(i) A random example is used to visualise the work flow, the audio sample was resampled to 16 kHz mono as seen below (Fig. 4).
(ii) Using magnitudes of the STFT (Short Term Fourier Transform), spectrogram is computed with a periodic Hann window and window size of 25 ms (Fig. 5).
(iii) Map the generated spectrogram to 64 mel bins and compute mel spectrogram (Fig. 6).
(iv) By applying log with offset Compute stabilized log mel spectrogram (Fig. 7).

To frame non-overlapping examples of 0.96 s the extracted features are used and each example covers 96 frames of 10 ms each. To extract vector embedding, these examples are then passed into the VGGish model.

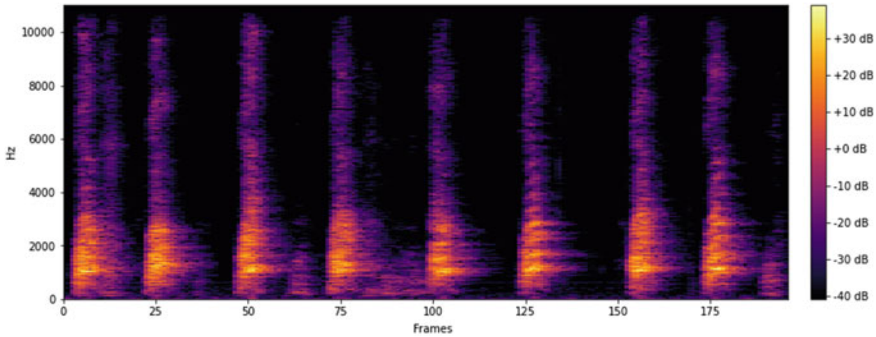


Fig. 5 Spectrogram

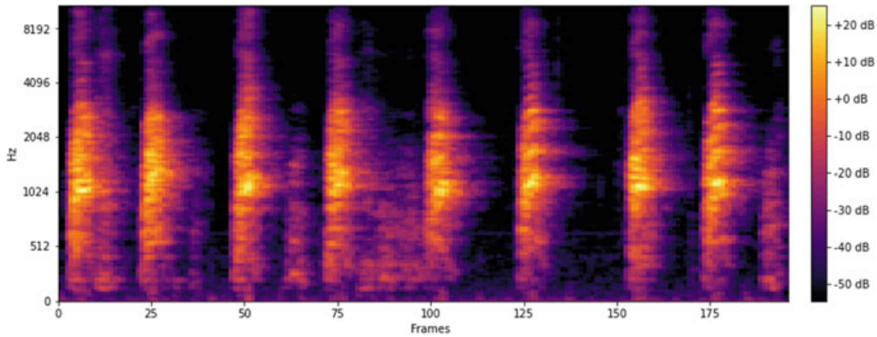


Fig. 6 Mel-spectrogram

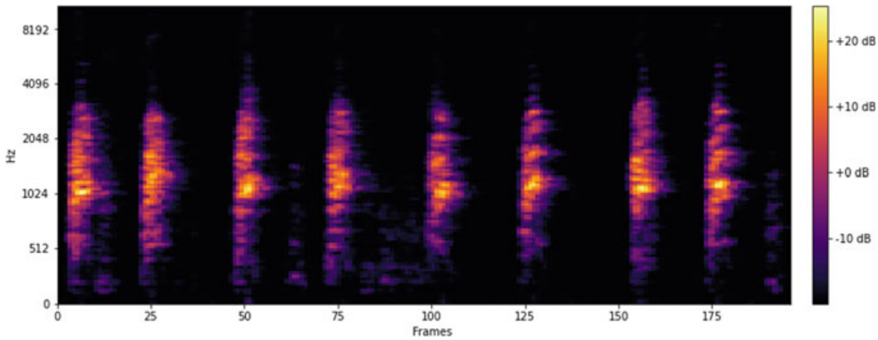


Fig. 7 Log mel spectrogram with offset

3.4 VGGish Tensor Flow Model and TF Record Files

There are many options and services for human speech recognition like google speech API, pyaudioanalysis and more, however these services are pretty good in classification of speech with music but not with classifying sound captured with microphone as music playing, non-music, animal sound.

We started the work using pyaudio analysis, a library to extract features from audio signal and perform classification using machine learning algorithms. The result accuracy was low as pyaudio was not flexible enough to handle variety of parameters, hence the use VGGish audio model as feature extractor was decided. This audio feature extractor was release by google in 2017 which represents sounds as sequence of vectors with 25 layers and parameter count of 72,141,184 and extracts 128 dimensional embeddings from 1 s audio signal.

TF record files: Tf record format is a simple format which stores a sequence of binary records, helps to read the data efficiently and serialize the data. This format is really useful if the data is streamed over a network. As binary data takes less space on disk there can be a significant impact on the pipeline which in turn reduces the training time of the network. The disadvantage of Tf record format is the training data has to be converted to TF format in the first place only.

Selecting the model: To train a classifier, Youtube-8m provides video level models and frame level models, in this work we have extracted features from audio file and converted those features into frames. The frame level LSTM model processes the features of each frame using multi-layered LSTM stack. The model was trained using a learning rate of 0.001 with 100 epochs and Batch size of 64.

4 Implementation and Results

Steps of Implementation in brief:

1. Prepared Dataset of 2 classes namely music and non-music
2. Extracted features from the audio files using TensorFlow VGGish model which are known as Tfrecord files
3. Training Tfrecord files using LSTM model which is trained on you Tube-8M dataset
4. Testing pre-recorded audio files and checking the prediction probability.

Confusion matrix is shown in Fig. 8 from which we can infer that music to music files (True Positives) are classified with 86% accuracy, true negative with accuracy of 76% and Fig. 9 shows the performance parameters of the classifier with F1 score of 0.82.

Fig. 8 Confusion matrix

		Actual class	
		Music	Non Music
Predicted class	Music	86%	24%
	Non Music	14%	76%

Fig. 9 Performance parameters

Test Acc.	Precision	Recall	F1score
0.80	0.86	0.77	0.82

5 Conclusion and Future Work

In this paper, we propose a simple methodology tensor flow of classifying audio files as music or non-music. Using the proposed method we would successfully classify the given files with accuracy of 80% and F1-score of 0.82. As we used the concept of transfer learning and VGGish feature extractor model, pre-processing time and training time of model was reduced drastically, however the trained classification model file is huge in size which is difficult to dump the model on any embedded platform. Model quantization and training VGGish network from scratch with reducing the number of layers may be taken as future work.

References

1. Chikkamath S, Shet R, Praveen P, Nalini CI, Kotturshettar BB (2020) Effective utilization of maker space for facilitating product realization course. J Eng Educ Transform 33(3):37–42
2. Purwins H, Li B, Virtanen T, Schlüter J, Chang S-Y, Sainath T (2019) Deep learning for audio signal processing. IEEE J Sel Top Sig Process 13(2):206–219
3. Le Coz M, Pinquier J, André-Obrecht R, Mauclair J (2013) Audio indexing including frequency tracking of simultaneous multiple sources in speech and music. In: 2013 11th international workshop on content-based multimedia indexing (CBMI). IEEE, pp 23–28
4. Kiranyaz S, Qureshi AF, Gabbouj M (2006) A generic audio classification and segmentation approach for multimedia indexing and retrieval. IEEE Trans Audio Speech Lang Process 14(3):1062–1081
5. Zahid S, Hussain F, Rashid M, Yousaf MH, Habib HA (2015) Optimized audio classification and segmentation algorithm by using ensemble methods. Math Probl Eng 2015, Article ID 209814:11. https://doi.org/10.1155/2015/209814
6. Khonglah BK, Sharma R, Prasanna SRM (2015) Speech vs music discrimination using empirical mode decomposition. In: 2015 twenty first national conference on communications (NCC). IEEE, pp 1–6
7. Wu Q, Yan Q, Deng H, Wang J (2010) A combination of data mining method with decision trees building for speech/music discrimination. Comput Speech Lang 24(2):257–272
8. Nilufar S, Ray N, Molla MKI, Hirose K (2012) Spectrogram based features selection using multiple kernel learning for speech/music discrimination. In: 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 501–504

9. George T (2002) Manipulation, analysis and retrieval systems for audio signals. Princeton University Princeton, NJ, USA
10. Dai J, Shan L, Wei X, Ni C, Liu W (2016) Long short-term memory recurrent neural network based segment features for music genre classification. In: 2016 10th international symposium on Chinese spoken language processing (ISCSLP). IEEE, pp 1–5

Weight Optimized Fuzzy Gravitational Classifier for Imbalanced Medical Datasets

P. O. Sinciya, V. Mary Amala Bai, and J. Jeya A Celin

Abstract Fuzzy set theory is an emerging tool that dealson uncertainties in classification problems. It has been applied to diverse fields of research because of its characterization by a set of parameters. Most of the instance based classification algorithms perform well for imbalanced datasets when combined with fuzzy approach. Considering this, fuzzy set theory is applied for an instance based gravitational classifier for both heart and diabetes disease prediction. An innovative approach to classify medical data based on fuzzy gravitational classifier is presented. Initially a set of n closest gravitational instances are selected from N samples. Then the binary class membership values of n instances are calculated. The fuzzy membership values of both positive and negative classes are calculated and are represented as a fuzzy set. The class which gives maximum membership value is assigned to the test instance. The proposed Fuzzy gravitational classifier (FGC) withboth feature selection and weight optimization performs well for both standard and imbalanced dataset. The experimental resultsshow improvement in performance than the existing complex methods. This method also reduces the overhead time and improves accuracy.

Keywords Fuzzy set · Comparison table · Gravitational search algoritm · Fuzzy gravitational classifier

1 Introduction

In recent years the imbalanced data classification have become an important research problem. In standard classification problems, the dataset contain equal distribution of instances for each class. Unlike this, imbalanced classification has an imbalanced class distribution [1, 2]. It may be binary class imbalance or multi class imbalance. Binary class imbalance means the number of samples in one class is more than the number of samples in another class. On the other hand, multi class imbalance

P. O. Sinciya (✉) · V. Mary Amala Bai
Rajadhani Institute of Engineering and Technology, Attingal, Kerala, India

J. Jeya A Celin
Hindustan College of Arts and Science, Chennai, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_26

means one or more classes might be outnumbered by the number of samples while the remaining classes contain smaller number of samples. Standard classification algorithms are not appropriate to handle imbalanced data problems because these models are over trained by majority class instances. Although standard classifier gives high classification performance across the entire dataset, the actual performance is very low because the classification rate of the smaller class is low.

In many real world imbalanced binary classification tasks, the dominant class is called positive class and subordinate class is named as negative class [3]. The basic measure of evaluating class imbalance is the Imbalance ratio (IMR). It is the ratio of the number of samples in positive and negative classes as in Eq. (1).

$$IMR = \frac{P}{N} \tag{1}$$

P is the number of samples in Positive class.
N is the number of samples in Negative class.

Another problem associated with the gravitational classifier is the overlapping of positive and negative examples in the class, in which most of the algorithms feel difficult to identify that small disjuncts [4]. In the conventional data gravitation classifier, all the gravitational instances get equal importance. It finds the relative distance between gravitational instances and the unknown sample. Cano [5] assigned featur weight for each class. These feature weights are optimized bycovariance matrix adaptation evolution strategy (CMA-ES) [6] algorithm. In this work, an extension of data gravitation classification algorithm in terms of fuzzy set theory is introduced. Zimmerman [7] introduced fuzzy set theory, to handle uncertainty. It allows partial membership of an object to different classes, and also considers the relative importance (closeness) of each neighbor instance tothat of the test instance. In the proposed Fuzzy Gravitational Classifier (FGC), this relative distance is changed to absolute distance. Thus the fuzziness influences the gravitational distance associated with each class label. The closeness factor associated with fuzzy sharing in FGC allows possibilistic classification capability and avoids small disjuncts due to class overlapping problem.

The content of the paper is organized as follows. Section 2 describes concepts on fuzzy set. Section 3 presents the proposed Fuzzy Gravitational Classification (FGC) algorithm. Section 4 gives experimental results and discussion. Finally Sect. 5 concludes the work.

2 Fuzzy Classification

2.1 *Fuzzy Set*

In a traditional binary classification, consider a universal set U, where class C is the subset of U. Input data x, x ∈ U, can either be an element or not an element of the known class C. This property of specifying whether an input pattern of universal set be a member or not the member of class C is defined by a fuzzy characteristic equation C: U → {0, 1} and is given in Eq. (2).

$$\delta_C(x) = \begin{Bmatrix} 1, \text{ if } x \in C \\ 0, \text{ if } x \notin C \end{Bmatrix} \tag{2}$$

In real world applications, the classes may overlap in boundary region [8]. Therefore it is uncertain to predict whether an input data belongs to class C or not. Fuzzy set is a theoretical model which handles this problem by replacing fuzzy characteristic equation by fuzzy membership function. It has the values in the range 0 to 1, C: U → [0, 1].

The learning of a binary fuzzy classifier for a set of input values x_i is an assignment of fuzzy membership μ_C on x_i, $x_i \in X, where\, X = \{x_1, x_2, \ldots x_n\}$, $\forall i = 1, 2, \ldots, n$. For a crisp membership, the input pattern X is divided in to n subgroups during training process. These are fuzzy membership values of X, $\mu_C(X)$ which can be arranged in n × N matrix format, Therefore $U = \mu_C(X)$.

Definition 1 *Fuzzy Soft set.* The family of all fuzzy sets of U is denoted by Ƒ (U). Let $A_i \subseteq E$. Then a pair (F_i, A_i) is called a fuzzy soft set over U, where F_i is a mapping given by F_i: A_i → Ƒ (U).

2.2 *Comparison Table*

Comparison table is a square table approach used to solve the problem of decision making in real time environment. For the set of parameters, if the decision has to be taken from the availability set U then it can be achieved through comparison table approach. The resultant comparison table for fuzzy set (H, Z) is shown in Table 1. Both row and columns represent the set of classes C_1, C_2, C_3, C_4, C_5. Each entry in the table h_{ij} = 1, 2, 3, 4, 5 represent the number of features for which the value of class C_i exceeds or equal to the value of class C_j. Then the sum of rows and columns are calculated and score of each class is determined as shown in Table 2. Based on the score value the unknown class label for a test attribute can be determined.

Here, the highest score is achieved by class C_5 which is 9. So test sample is assigned with class 5.

Table 1 Comparison table for (H, Z)

Classes	C1	C2	C3	C4	C5
C1	5	2	4	2	1
C2	3	5	4	2	3
C3	2	1	5	1	1
C4	5	4	4	5	2
C5	4	3	5	4	5

Table 2 Score matrix

Classes	Row sum (rs)	Column sum (cs)	Score (rs-cs)
C1	14	19	−5
C2	17	15	2
C3	10	22	−12
C4	20	14	6
C5	21	12	9

3 Fuzzy Gravitational Classification (FGC)

In some cases the highest gravitational instance may belong to more than one class. It happens when the classes overlap. Thus a better classifier is needed to handle the above mentioned problems. In this work, a hybrid fuzzy gravitational classifier is developed which overcomes the problem of imbalancing by combining fuzzy approach with weight optimized gravitational classifier (WOGC). The steps in fuzzy gravitational classification algorithm are given below:

3.1 *Find the Class Membership Function of Input Pattern*

The algorithm starts by applying a feature selection algorithm known as Joint feature Interaction MAximization JFIM [9] to select best set of features. Standard fuzzy classification algorithms finds membership only for majority instances. They do not give importance to minority instances. Hence the rate of misclassification of minority instances is very high as compared to majority instances. To avoid this problem we make the change that if the input pattern has the minority instances then set membership values for majority and minority classes as 0 and 1 respectively.

For a given standard and imbalanced medical dataset, $X_n = \{x_1, x_2, \ldots x_n\}$, the set of n patterns and its class label is denoted by $C_j = \{0, 1\}$. Its i gravitational instances from X_n are found using gravitational distance (g_d). Here the majority and minority class is expressed by C^- and C^+ and its corresponding values are 0 and 1. For each class membership function, X_i is defined as in Eqs. (3) and (4).

$$\mu_{C^+}(X_i) = \begin{Bmatrix} 0, & if\, \mathrm{C}_j = 1 \\ \left(\frac{N_{c+}}{n}\right) * \varepsilon, & if\, \mathrm{C}_j = 0 \end{Bmatrix} \tag{3}$$

$$\mu_{C^-}(X_i) = \begin{Bmatrix} 1, & if\, \mathrm{C}_j = 1 \\ \left(\frac{N_{c-}}{n}\right) * \varepsilon + (1 - \varepsilon), & if\, \mathrm{C}_j = 0 \end{Bmatrix} \tag{4}$$

Such that

$$\mu_c(X_i) = \mu_{C^+}(X_i) + \mu_{C^-}(X_i) = 1$$

The values of N_{c^-} and N_{c^+} are the number of majority and minority gravitational neighbors that belong to X_i. The minority samples X_i related to C^+ is ε when the number of negative samples are 1, thus ε should take the value lesser than 0.5. Here we set ε = 0.49.

3.2 Calculation of Membership Function

Handaga et al. [10] proposed a new method of numerical data classification based on Fuzzy soft approach. Roy and Maji [11] have used a comparison table approach for medical data classification. In this work, for handling both standard and imbalanced data, both positive membership $g_{c^+}(y_t)$ and negative membership function $g_{c^-}(y_t)$ are calculated for each class using Eqs. 5 and 6. The fuzzy membership values are calculated by finding gravitational distance (distance from the centroid) for both positive and negative classes. The membership function which gives maximum gravitation is evaluated using comparison table approach. The weight w_{jt} is optimized using weight optimized particle swarm algorithm (WOPS) [12].

3.3 Comparison Table for Positive and Negative Membership

In the proposed work, a comparison table is generated for both positive and negative membership function. Thus for $g_{c^+}(y_t)$ and $g_{c^-}(y_t)$, a 2 × 2 comparison table is generated. Then the score vector is computed for both positive and negative membership functions. Positive score vector is denoted as S_{c^+} and negative score vector as S_{c^-}. The difference in score vector gives the final score for each class. The class which gives maximum score value is assigned to the unknown instance.

3.4 Proposed Fuzzy Gravitational Classifier Algorithm

1. Apply Joint feature Interaction Maximization (JFIM) to select best set of features.
2. Find class membership of each instance $\mu_{C^+}(X_i)$ and $\mu_{C^-}(X_i)$ using Eqs. (3) and (4).
3. Compute the centroid vector C_i (for each class i), by calculating the average value of data present in the dataset D_i (set of instances of class i).
4. Represent the centroid vectors as a table of size $M \times N$ (M classes and N features) which can be considered as a fuzzy set (F, E). An entry in the table is g_{in}, $i = 1, 2, \ldots, I$ and $n = 1, 2, \ldots, N$.
5. Get a feature vector y_{jt} from unknown class.
6. Generate a fuzzy set using fuzzy membership functions $g_{c^+}\left(y_{jt}\right)$ and $g_{c^-}\left(y_{jt}\right)$ for each class j, {0, 1}, with its entry as y_{jt}, j = {0, 1} and t = {1, 2, …, T} (T number of features}, calculated using Eqs. (5) and (6).

$$g_{c^+}(y_t) = 1 - \frac{\mu_{C^+}\left(X_j\right)}{w_{jt} * \left|g_{in} - y_{jt}\right|^2} \tag{5}$$

$$g_{c^-}(y_t) = 1 - \frac{\mu_{C^-}\left(X_j\right)}{w_{jt} * \left|g_{in} - y_{jt}\right|^2} \tag{6}$$

   where w_{jt} is the weight of attribute t for class j. The optimized weight for w_{jt} is evaluated using WOPS.
7. Represent the score for each class in a matrix (S_{c^+} and S_{c^-}).
8. Determine the final score (S_i) for each class using comparison table approach.
9. Assign the unknown sample to the class if $s_i > S_{i+1}$ for $\forall i = \{1, 2\}$.

4 Results and Discussion

The experiments were conducted by considering two classes from Cleverand heart disease and Pima Indian diabetes datasets. These are benchmark datasets collected from KEELrepository [13]. The output obtained for Cleverand Heart diseaseand diabetes dataset with well known classifiers such as Fuzzy Nearest Neighbour (FNN) and Fuzzy Soft set Classifier (FSC) is given in Tables 3 and 4.

In classification scenario, evaluation measures are essential to assess the efficiency of classifier and give guidance to improve classifier modelling. Standard classifiers use accuracy as the common evaluation measure to assess the performance. But accuracy is not aright measure to evaluate class imbalance problem, since the rare class gets very little impact as compared to majority class.

Suppose for a given binary class classification problem, where class C_1 have 92% of instances and class C_2 have 8% of instances. In this case, if the classification

Table 3 Performance metrics for CHD imbalanced data set

Data set	CHD imbalanced data			
Classifier/metrics	FNN	FSC	WOGC	FGC
AUC (%)	84.18	85.22	89.46	91.61
G-mean (%)	86.12	86.28	87.72	91.33
F-measure (%)	83.31	83.27	86.41	87.48

Table 4 Performance metrics for diabetes imbalanced data set

Data set	Diabetes imbalanced data			
Classifier/metrics	FNN	FSC	WOGC	FGC
AUC (%)	80.38	81.81	87.32	87.92
G-mean (%)	82.47	83.45	86.39	88.24
F-measure (%)	79.16	79.59	84.72	82.45

algorithm is evaluated by an accuracy rate, it gives 92% of accuracy by assigning all unknown instances to majority class. This does not mean that the algorithm would be good at class predictions rather it tells that it is very easy to maximize classification accuracy rate when the class distribution is very imbalanced. Thus a more challenging measure is used to evaluate the performance.

For a given data set with unbalanced data distribution, the classification performance should be evaluated by different measures. They are,

4.1 Area Under Curve (AUC)

AUC is a widely used measure for imbalanced classification. It is derived from confusion matrix values a, b, c and d. It is given by Eq. (7).

$$AUC_i = \frac{1 + TP_r - FP_r}{2} = \frac{1 + \frac{a}{a+b} - \frac{c}{c+d}}{2} \tag{7}$$

4.2 F-measure

The two important measures used to calculate F-measure are True Positive rate (TPr) and Positive Predictive value (PPv).

In knowledge discovery, True Positive Rate is defined as the percentage of correctly retrieved instances, denoted as Recall (R_v) and is given in Eq. (8).

$$R_v = TP_r = \frac{a}{a+b} \tag{8}$$

Positive Predictive Value is defined as the percentage of relevant objects known for retrieval, denoted as precision P_v and is as in Eq. (9). The F-measure which uses R_v and P_v is given in Eq. (10).

$$P_v = Pp_v = \frac{TP}{TP + FP} \tag{9}$$

$$F\text{-}measure = \frac{2R_v P_v}{R_v + P_v} \tag{10}$$

4.3 G-mean

When the performance of both classes are considered, G-mean can be used. G-mean is given in Eq. (11).

$$G\text{-}mean = \sqrt{TP_r.TN_r} \tag{11}$$

The values for various performance metrics for imbalanced Cleveland data set is shown in Table 3.

The values for various performance metrics for imbalanced Diabetes data set is shown in Table 4.

The overall performance of MDC system for imbalanced data given in Tables 3 and 4 and are represented in Figs. 1 and 2.

‘

For imbalanced data, AUC is the most generally used metric to evaluate the performance of the classifier. This measure considers both true positive and true negative rates equally. To improve AUC rate for imbalanced data, class weight is

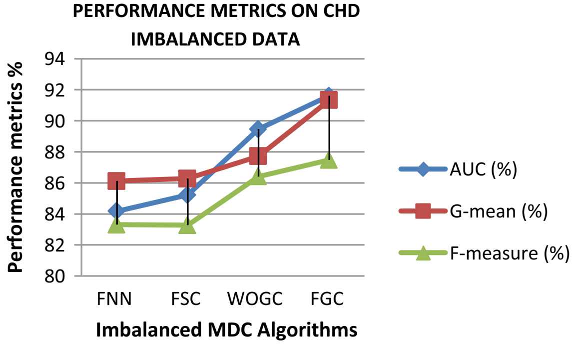
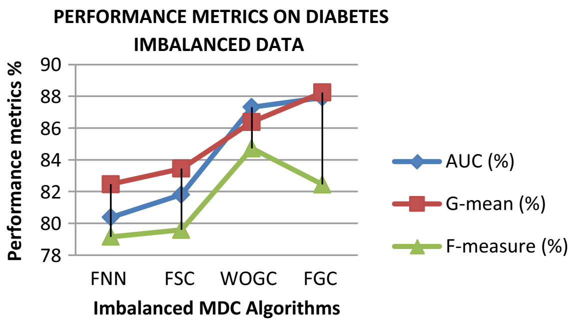


Fig. 1 Comparative performance analysis of cleveland heart disease imbalanced data set

Fig. 2 Comparative performance analysis of pima indian diabetes imbalanced data set



added for each feature and the best weight is calculated by WOPS. Both positive and negative classes get the same amount weight, which results in good AUC value. The proposed approach gives high AUC rate of 91.61 for imbalanced cleverand database and 87.92 for Pima Indian Diabetes database and G-mean values as 91.33 and 88.24 respectively. The results show the proposed method outperforms the other approaches.

5 Conclusion

Fuzzy set is an emerging technique which is applied for various applications in different fields. So far it is not successfully applied on real world imbalanced data. An efficient weight optimized fuzzy gravitational classifier is developed to handle the imbalanced data. It has been tested with imbalanced medical data including Cleverand heart disease and Pima Indian diabetes data from KEEL repository. Here the fuzzy membership functions are calculated using gravitational distance and it is evaluated by comparison table approach. It is observed that, for imbalanced data, the proposed weight optimized fuzzy gravitational classifier outperforms other well known classification algorithms. The performance has been compared with that of FSC classifiers and has been found to be interesting.

References

1. Seiffert C, Khoshgoftaar TM, Van Hulse J (2007) An empirical study of the classification performance of learners on imbalanced and noisy software quality data. In: Procedings of IEEE international conference on information reuse and integration, pp 651–658
2. Napierala K, Stefanowski J, Wilk S (2010) Learning from imbalanced data in presence of noisy and borderline examples. In: Proceedings of 7th international conference on rough sets and current trends in computing (RSCTC2010), pp 158–167
3. Peng L, Zhang H, Yang B, Chen Y, Qassrawi Y, Lu G (2012) Traffic identification using flexible neural trees. In: Proceeding of the 18th international workshop of QoS (IWQoS 2012), pp 1–5

4. García V, Mollineda RA, Sánchez JS (2008) On the k-NN performance in a challenging scenario of imbalance and overlapping. Pattern Anal Appl 11:269–280
5. Cano A, Zafra A, Ventura S (2013) Weighted data gravitation classification for standard and imbalanced data. IEEE Trans Cybern 1–16
6. Hansen N, Ostermeier A (2001) Completely derandomized selfadaptation in evolution strategies. Evol Comput 9:159–195
7. Zimmerman J (1996) Fuzzy set theory and its applications. Kluwer Academic, Boston, MA
8. Sarkar M (2002) Rough-fuzzy functions in classification. Fuzzy Sets Syst 132:353–369
9. Sinciya PO, Jeya A Celin J (2017) JFIM: a novel filter feature selection approach using joint feature interaction maximization. Int J Intell Eng Syst 10:203–213
10. Handaga B, Herawan T, Herawan T (2012) FSSC: an algorithm for classifying numerical data using fuzzy soft set theory. Int J Fuzzy Syst Appl 2:244–256
11. Maji PK, Biswas R, Roy AR (2003) Soft set theory. Comput Math Appl 45:555–562
12. Sinciya PO, Jeya A Celin J (2017) Weight optimized gravitational classifier for high dimensional numerical data classification. Int J Pure Appl Math 116:251–263
13. Alcalá-Fdez J, Fernandez A, Luengo J, Derrac J, García S, Sánchez L, Herrera F (2011) KEEL data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. J Multiple-Valued Logic Soft Comput 17:255–287

Method for Improvement of Product Sales Forecast for Long Horizon Using Hybrid Decomposition and Machine Learning on Multi-variate Time Series Data

Arvind Kumar Sharma, Sreekar Bathula, and Kaushik Saha

Abstract Forecasting product sales in a consumer product company is one of the most important and challenging tasks for business financial planning and inventory management. In this paper, we are proposing a novel forecasting scheme through unique data pre-processing pipeline and machine learning model using a combination of Ensemble Empirical Mode Decomposition (EEMD) and Extreme Learning Machine (ELM) for multivariate sales data comprising of macroeconomic indicators. EEMD decomposition is performed on sales time series and intrinsic mode functions are analyzed. EEMD is also employed to find important constituent parts of indicators as a data pre-processing step. Experimental results show that EEMD helps in improving forecasting performance. We are also using auto-regressive sale terms along with macroeconomic indicators and employing feature ranking to analyze the most impactful features for prediction performance. Final forecasting performance is further compared with state of art techniques used for multivariate sales prediction models namely Vector Auto-Regressive (VAR) and Least Absolute Shrinkage and Selection Operator (LASSO) models. The proposed forecasting scheme demonstrates better performance than the current state of the art techniques.

1 Introduction

Sales forecasting is very crucial for a consumer product company for inventory management, customer service, marketing, and overall financial planning. Sales forecasting directly impacts the company's strategy towards channel relationships, customer satisfaction, and monetary savings. For consumer product company, it is even more important and challenging at the same time due to highly dynamic market scenario,

A. K. Sharma (✉) · S. Bathula · K. Saha
Samsung R&D Institute, New Delhi, India
e-mail: arvind.s1@samsung.com

S. Bathula
e-mail: s.bathula@samsung.com

K. Saha
e-mail: kaushik.s14@samsung.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_27

rapid technological development and accelerating rate of product innovation, all of which leads to shrinkage of the product life cycle. Poor sales forecasting may lead to product backlog, inventory shortages, and unsatisfied customer demands. Therefore, it is very important to develop a forecasting model that is robust and effective. Present literature demonstrates humongous work towards generating a robust prediction model using statistical as well as modern machine learning-based algorithms [1]. Exponential smoothing and Autoregressive Integrated Moving Average (ARIMA) models are most popular among statistical methods in univariate time series forecasting setting [2] whereas Vector Auto-Regressive (VAR) model and Vector ARMA are among most popular multi-variate counter-parts. We have used the VAR method in our analysis for comparison with the machine learning approach. While there are a lot of machine learning based methods used for better sales forecasting but their accuracy improvement as compared to statistical methods is debatable at least for small to medium datasets [1]. There is also literature available that suggests ML-based methods may not perform well alone but they have the potential to improve the forecast accuracy in presence of better data pre-processing algorithms such as seasonal adjustment and EEMD [1, 3] but these methods are focused on univariate data. The reason for improved performance is attributed to the fact that time series data have certain inherent characteristics such as seasonality, trend, and various constituent signals which are not possible to detect by ML algorithms in presence of limited data size. While classical time series decomposition [4] has been around for a long time which is still an effective tool to visualize and deseasonalize the data, there are recent advancements in time series signal decomposition approaches and Empirical Model Decomposition (EMD) is one of the most popular among them because of its relevance for non-linear and non-stationary data. Ensemble EMD is an enhanced version of EMD which improves on the weaknesses of EMD [5]. EEMD is an adaptive data analysis method that is based on local characteristics of the data so it catches nonlinear, nonstationary oscillations more effectively. The majority of the work is focused on developing novel sales forecasting methods for univariate data because of the ease of modeling and data gathering. Whereas, most of the sales prediction work in a multivariate setting utilizes either macroeconomic indicators [6–9] or some indicators derived from sales data [10]. It has also been shown that economic indicators correlate well with a sales forecast. We have used 16 economic indicators along with product sales which constitute our multivariate dataset. Details of economic indicators and sources of data along with the source for motor vehicle sales data is mentioned in Table 1.

In our study, we have focused on improving forecasting performance on multivariate data by employing data pre-processing using EEMD decomposition and machine learning algorithm and demonstrated that machine learning based methods have the potential to out-perform statistical methods in presence of better data pre-processing techniques. There are several studies for augmenting forecasting models with macroeconomic indicators but most of them either use statistical forecasting methods [6, 7, 9] or classical machine learning approaches [8]. Some work exists on using either EEMD or ELM or combination of them [3, 10–15] but none of them are using EEMD and ELM along with feature ranking and autoregressive (AR) terms in

Table 1 Economic indicators used in sales forecasting

Economic indicators	Description	Frequency
Consumer price index (CPI)[a]	A measure that examines the weighted average of prices of a basket of consumer goods and services, such as transportation, food and medical care. It is calculated by taking price changes for each item in the predetermined basket of goods and averaging them	Monthly
Composite leading indicator (CLI)[a]	The composite leading indicator (CLI) is designed to provide early signals of turning points in business cycles showing fluctuation of the economic activity around its long term potential level. CLIs show short-term economic movements in qualitative rather than quantitative terms	Monthly
Business confidence index (BCI)[a]	This business confidence indicator provides information on future developments, based upon opinion surveys on developments in production, orders and stocks of finished goods in the industry sector. It can be used to monitor output growth and to anticipate turning points in economic activity. Numbers above 100 suggest an increased confidence in near future business performance, and numbers below 100 indicate pessimism towards future performance	Monthly
Narrow money (M1)[a]	M1 includes currency i.e. banknotes and coins, plus overnight deposits. M1 is expressed as a seasonally adjusted index based on 2015 = 100	Monthly
Broad money (M3)[a]	Broad money (M3) includes currency, deposits with an agreed maturity of up to two years, deposits redeemable at notice of up to three months and repurchase agreements, money market fund shares/units and debt securities up to two years. M3 is measured as a seasonally adjusted index based on 2015=100	Monthly
GDP2015s3[b]	Gross Domestic Product	Quarterly
PIM[b]	Population (In Millions)	Annual
CPI2015[b]	Consumer Price Index: All Items for India	Monthly
IUSFER[b]	India/U.S. Foreign Exchange Rate	Monthly
REII[b]	Ratio of Exports to Imports for India	Monthly
EPUII[b]	Economic Policy Uncertainty Index for India	Monthly
TREGI[b]	Total Reserves excluding Gold for India	Monthly
IVGI[b]	Imports: Value Goods for India	Monthly
PITMI[b]	Production in Total Manufacturing for India	Monthly
PMIGTI[b]	Production: Manufacturing: Investment goods: Total for India	Monthly
GDPC[c]	GDP (% change)	Quarterly
Sales data[d]	Motor vehicle sales data	Monthly

[a] https://data.oecd.org/india.htm
[b] https://fred.stlouisfed.org/tags/series
[c] https://data.oecd.org/india.htm
[d] https://www.ceicdata.com/en/indicator/india/motor-vehicles-sales

improving the forecasting performance. For the rest of the paragraphs in this section, we will discuss the most related literature for our work and how our approach is unique in comparison to them.

Macroeconomic indicators are used in sales forecasting in [6], AR terms of the sales components are used in the regression model and selection of the indicators is done based on importance. Characteristics of our data match quite well with data characteristics in this paper but we have approached it using better decomposition of data and advanced machine learning ELM algorithm. Sagaert et al. [7], Hülsmann et al. [8] and Sa-ngasoongsong et al. [9] also demonstrated improvement in sales forecasting using macroeconomic indicators where [7] explored only linear modeling technique (LASSO), Hülsmann et al. [8] compared more no of classical machine learning techniques but did not incorporate AR terms and feature ranking and [9] presented a comparative performance analysis for Vector Auto Regression (VAR) and Vector Error Correction (VEC) methods. Zhang et al. [16] is focused more on comparing univariate Singular Spectrum Analysis (SSA) with multi-variate VAR and demonstrate the superior performance of VAR over SSA. We have compared the performance analysis of our approach with LASSO and VAR to demonstrate forecasting improvement of our method.

EEMD and ELM combination has been proven effective in improving sales forecasting of computer product sales as compared with EMD and ELM combination, single Back Propagation Neural Network (BP NN), single Support Vector Regressor (SVR) and single ELM models [3] but they are using EEMD for single time-series sales data and impact of macroeconomic indicators on forecasting model is not analyzed.

EMD is used for improving prediction for financial time series data in [11, 13]. Chai et al. [11] compared Wavelet de-noising and EMD for data pre-processing before Least Square SVM (LSSVM) for single CSI300 index time series and found EMD and LSSVM combined performing better than the other whereas [13] utilized EMD along with SVR to improve the stock prediction for single highly non-linear, non-stationary S&P 500 Index time series data. Additionally, there are numerous other fields where EMD decomposition has shown an improvement in modeling accuracy [17–19]. However, we did not find any work that uses EEMD, ELM, AR terms along with feature ranking for improving sale forecasting accuracy using economic indicators to the best of our knowledge.

In the present study, we do seasonal adjustment of sales time series using classical decomposition before modeling. Product sale time series data is generally non-stationary and non-linear and we have established this fact by performing tests in Sect. 3. We have used EEMD decomposition to tackle the non-stationary characteristics of the sales time series. EEMD decomposition has been applied on sales to obtain IMFs to get better insights on underlying information of time series and its impact on ML model convergence. We have further used EEMD decomposition on all economic indicators to study if constituent IMFs of economic indicators provide better insights to explain changes in constituent IMFs of sales time series and its impact on ML model convergence. Extreme Learning Machine (ELM) algorithm is employed for developing final prediction models to address the non-linear characteristics of the sales time series. The proposed method has been compared with LASSO and VAR methods also to establish that the proposed method is superior to the state of the art counterparts.

2 Methodology

This section gives the details of the methods that are being used for the time series forecasting.

2.1 Classical Decomposition

The process of converting time series data into components like the trend, seasonal and residual is called the classical decomposition. Classical decomposition mostly consists of additive (1) and multiplicative (2) decomposition [4] approaches where the trend, seasonal and residual components are added and multiplied respectively to get the actual data value.

$$y = Trend + Seasonal + Residual \tag{1}$$

$$y = Trend * Seasonal * Residual \tag{2}$$

The trend is considered as a moving average of the data for a given period. After removing the trend component from data, the average of the monthly sales is taken for the given period. This will be the seasonal curve where the values are equal for each month for the considered period. Finally, the remaining component is the residual. The goal of this decomposition is to make it perfect to get the residual as minimum as possible. In our approach, the period was identified as 12. The inference from this is that the sales data shows a repetitive trend each year. This inference came out to be a valid upon experimentation, where the average residual was low when compared to other time-series data.

2.2 Ensemble Empirical Mode Decomposition

Ensemble Empirical Mode Decomposition (EEMD) is a combination of multiple EMDs. EMD is an adaptive method for data decomposition as its basis are derived from the data itself. The data is split into multiple Intrinsic Mode Functions (IMFs) depending on the data and a residual. Each IMF is derived from the data by extracting two envelopes defined by local maxima's and minima's respectively. After drawing these envelopes, a function h at any point of time is defined as the difference between the data and the mean of two envelopes. With h as a data, the steps (extracting envelopes and defining another h function) are repeated until the envelopes are symmetric. Each h function is defined as the IMF. To improve the accuracy of the EMD, the EEMD was introduced. This method of EEMD adds different white noises to the series multiple times and then decomposes the data using the EMD method. These multiple white noises are added in such a way that the mean of all the noises is zero.

2.3 Extreme Learning Machine

Extreme Learning Machine (ELM) is based on the neural networks concepts that have been broadly applied to problems such as classification, regression, and time-series forecasting. ELM can be a Single-Layer Feed-Forward Neural Network (SLFN) or a Multi-Layer Feed-Forward Neural Network (MLFN). The traditional neural networks are slower with the iterations and parameter tuning included in the backpropagation. In contrast to this, the data fitting mechanism in ELM solves the regularized least squares with the training data by making it faster than the neural networks. In ELM [20, 21], the input weights, and hidden biases are randomly generated and output weights are analytically determined by the least-squares method. If T is the target vector consisting of the outputs and, w is the weights vector that is being generated randomly and, H is the values of the hidden neurons. T can be calculated using the equation $T = Hw$. To correct these randomly generated values ELM uses the inverse of H to identify analytically the weights of the hidden neurons. Hence $w = H^{-1}T$, This makes the ELM faster than the neural network by giving more generalization power.

2.4 LASSO

Lasso (Least Absolute Shrinkage and Selection Operator) regression is a type of regression that uses shrinkage. The Ridge regression and lasso [22] are similar in minimization term which includes an additional regularization term along with the least squares. This regularization term in both the regressions puts constraints on the coefficients (w).

$$\sum_{i=1}^{n}\left(y_i - \sum_{j} x_{ij}\beta_j\right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{3}$$

The difference between these two regressions is the way it takes the coefficients. The Ridge regression puts a penalty on the squared coefficients whereas the lasso puts a penalty on the absolute coefficients given in Eq. (3), where λ is the regularization penalty and β terms are the coefficient terms.

Unlike ridge regression, the lasso can nullify the effects of the parameters by making it's coefficient equal to zero. This gives lasso an additional capability to minimize the number of parameters by assigning weights to only the important parameters. This capability of selecting only some parameters in lasso makes it more relevant to the time series prediction when the external parameters are included in the model.

2.5 Vector Autoregressive (VAR)

Vector Autoregressive method is a multivariate linear regression model. It is generally used for two major things one is the prediction and the other is the causality test. The prediction of value is simply based on a linear fit on the preceding values. The number of preceding values chosen to make the prediction is known as the lag. The autoregression model is a simpler form of vector auto regression [23] which is a linear regression model for multiple variables. The time series in auto regression is defined as

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_{lag} y_{t-lag} \tag{4}$$

where y_t the value of time series at time t. Lag number of preceding values included in the prediction, α_0 to α_{lag} are the regression coefficients. The time series in vector auto regression in a multi variable time series data is defined as

$$Y_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_{lag} Y_{t-lag} \tag{5}$$

where $Y_t = (y_{1,t}, y_{2,t} \ldots, y_{n,t})'$ is the matrix of size [n × 1] and n is the number of features.

$$A_0 = \begin{bmatrix} a_1^0 \\ \vdots \\ a_n^0 \end{bmatrix} \quad for\, t = 0 \quad and \quad A_t = \begin{bmatrix} a_{1,1}^1 & \ldots & a_{1,n}^t \\ \vdots & \ddots & \vdots \\ a_{n,1}^1 & \ldots & a_{n,n}^t \end{bmatrix} \quad for\, t \geq 1. \tag{6}$$

VAR estimates the values of coefficients that minimizes the error using Ordinary Least Squares (OLS) [24]. This makes the predicted value of one feature as a weighted average of preceding values of all the features. The number of preceding values of features to be included in the weighted averages is defined as lag. The lag can be estimated by Causality Test. This test elucidates the correlations between the features and examine whether the lagged values of one feature predicts the value of other features.

3 Experimental Results and Proposed Model

3.1 Dataset and Performance Criteria

All the experiments are done on monthly data of motor vehicle sales in India and economic indicators mentioned in Table 1. Data is gathered from January 2003 to December 2018. We have a total of 16 simultaneous time series including sales time series as shown in Fig. 1. Stationarity test results for each of these are shown in Table 2. Most of these indicators as mentioned in the table are nonstationary.

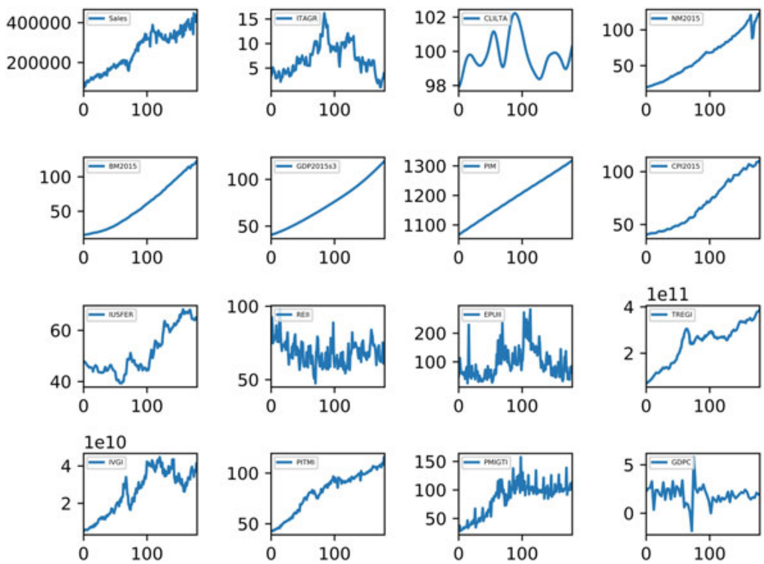


Fig. 1 Time series data (1 sale, 15 economic indicators)

Table 2 Augmented Dickey-Fuller test for stationarity [25]

Time series	Sales	ITAGR	CLILTA	NM2015	BM2015	GDP2015s3	PIM	CPI2015
No. of lags	13	12	7	13	11	12	14	12
P value	0.8675	0.5552	0.1265	0.9987	0.99	0.9984	0.3069	0.9958
Stationary	False	False	False	False	False	False	False	False
Time series	IUSER	REII	EPUII	TREGI	IVGI	PITMI	PMIGTI	GDPC
No. of lags	5	14	2	9	1	2	12	10
P value	0.9631	0.0097	0.1858	0.5017	0.5054	0.4377	0.3368	0.0056
Stationary	False	True	False	False	False	False	False	True

We are using 12 steps lag for economic indicators meaning lagged data till the current month for sales and economic indicators are used to forecast next year same month sales. The accuracy measures used in this paper are symmetric Mean Absolute Percentage Error (sMAPE). sMAPE is defined as follows:

$$sMape = \frac{2}{k}\sum_{t=1}^{k}\frac{|Y_t - \hat{Y}_t|}{|Y_t| + |\hat{Y}_t|} * 100\% \quad (7)$$

where k Y_t are actual observations, $\hat{Y}_t$ are the forecasts and k is the forecasting horizon.

3.2 Dataset Pre-processing

Product sales data has been converted to per-day sales by dividing monthly sales by no of days in the month. Further sales figures has been seasonally adjusted using classical multiplicative decomposition. Figure 2 shows multiplicative decomposition into seasonal, trend and residual components.

EEMD decomposition of seasonally adjusted sales is shown in Fig. 3.

Since different economic indicators have different value ranges, it is important to normalize them before using them in machine learning models. Z-score normalization has been done, the formula is mentioned below:

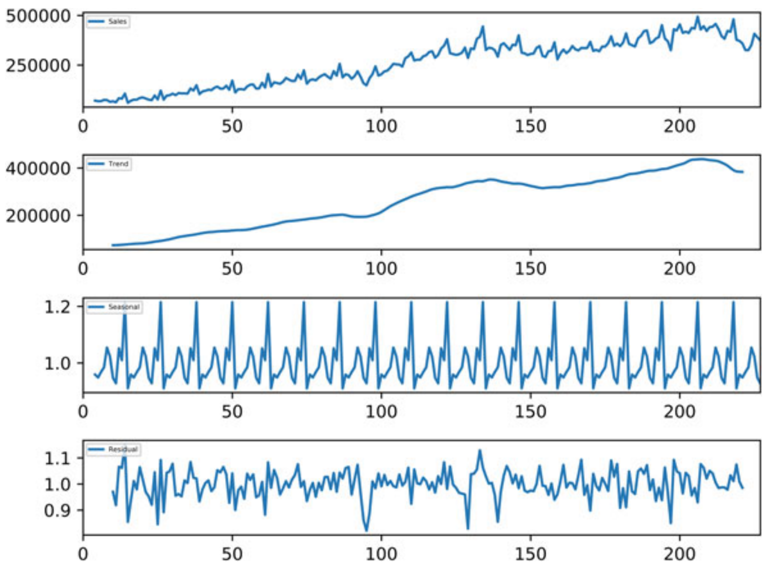


Fig. 2 Multiplicative decomposition

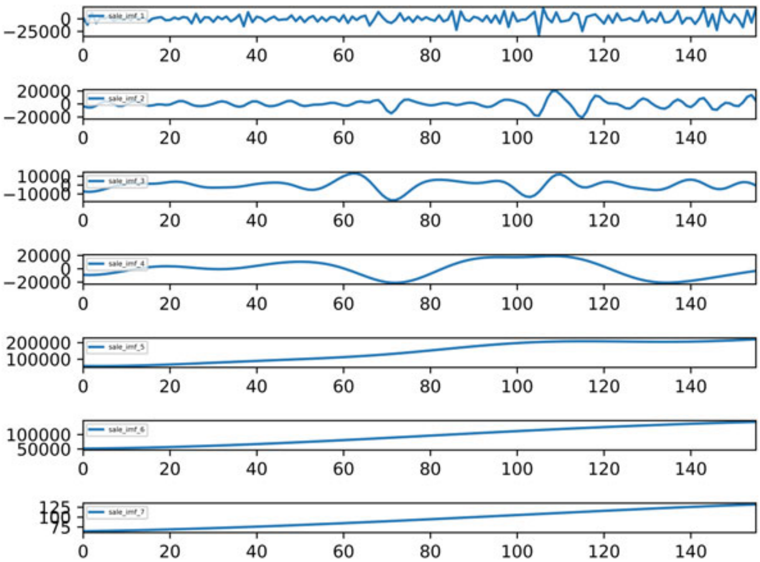


Fig. 3 EEMD decomposition

$$Z(x) = \frac{x - \mu}{\sigma} \tag{8}$$

where μ is mean and σ is the standard deviation. The selection of economic indicators is done using human judgment without any knowledge of impact of indicators on forecasting performance. So we analyzed feature importance of individual indicators on forecasting performance using the average importance measured by Recursive Feature Elimination (RFE) and Random Forest (RF) algorithms and selected top-ranked features only where it mattered in forecasting performance substantially. Further EEMD decomposition is done on economic indicators to extract hidden independent phenomena which can be useful in improving the forecasting accuracy further. It is important to mention that we performed seasonal adjustment and EEMD decomposition on training dataset only, so care has been taken, not to accidentally inject test data information into training data. All the results on the test dataset are thus unbiased and should be truly representative of model performance on unseen data.

3.3 Hyper-Parameter Tuning

Tunable hyper-parameters for ELM are mixing coefficient for distance, dot product input activations (alpha), the number of neurons in a hidden layer, and activation function at each neuron. Similarly for LASSO, since it is a linear regression problem, the only hyperparameter used is penalizing the number of features in a model is with the value α in order to achieve the regularization effect. Whereas in the case of VAR, the only hyperparameter it takes is the lag value which specifies the number of past values included in the linear equation to predict the succeeding value. The grid-Search algorithm is used to decide the best values for hyper-parameters for LASSO and ELM models. Walk-forward 7-fold cross-validation is employed during the training phase to have a better estimate of model performance on the validation set during the training. Model tuning curves for ELM are shown in Figs. 4 and 5.

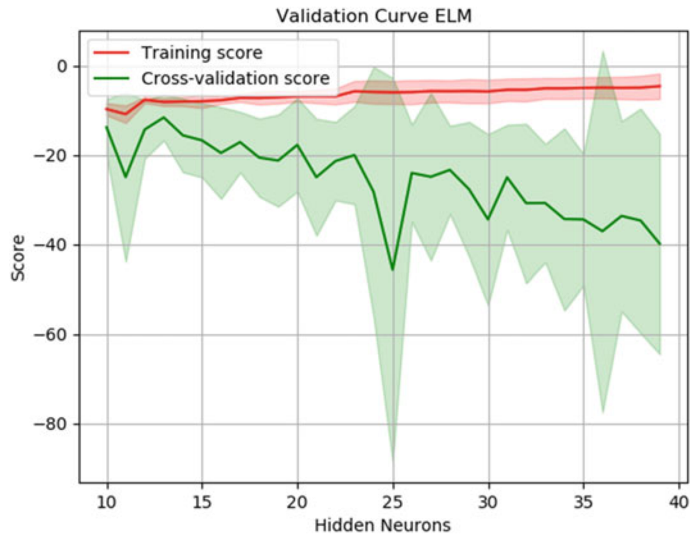
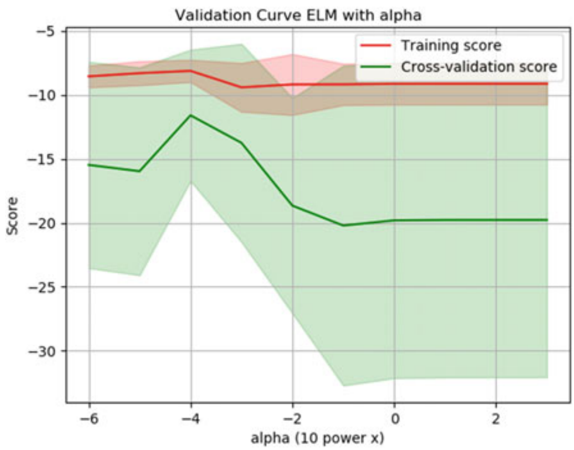


Fig. 4 ELM hidden neurons parameter tuning

Fig. 5 ELM alpha parameter tuning



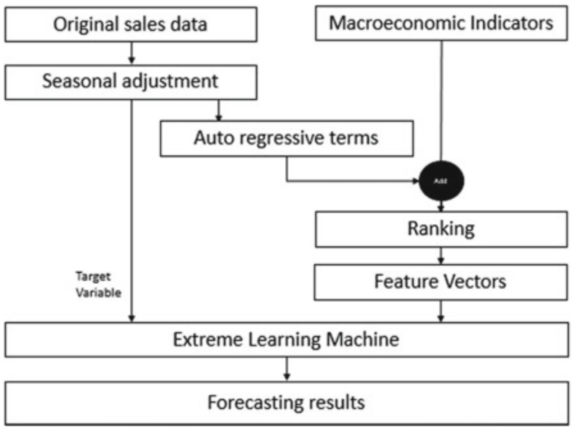


Fig. 6 Forecasting scheme for Model #1

3.4 Model Performance

We have used three schemes for forecasting model development depicted in Figs. 6, 7 and 8. Auto-regressive terms of seasonally adjusted sales along with macroeconomic indicators are used in ELM model in first forecasting scheme. In second forecasting scheme, sale component is broken into 7 IMFs using EEMD and separate ELM models are tuned for each of these IMFs. The sum of predictions of each of the ELM model is used as final sale forecast. In third forecasting scheme, EEMD decomposition of all economic indicators is also done, each economic indicator produced 5, 6 or 7 IMFs so all combined produced total of about 100 IMFs, these are further combined with auto-regressive sale terms which constitutes all the features used in this modelling scheme. Sale IMFs are modelled using separate ELM models using the features generated in previous step. Finally, sum of predictions of each of the ELM model is used as final sale forecast similar to second forecasting scheme.

Forecasting performances of all three schemes are depicted in Table 3. The impact of sale auto-regressive terms and feature ranking is established in each of the schemes. We can see cross-validation scores keep improving with the addition of feature ranking and sale auto-regressive terms in each of the schemes, the test scores also have improved in almost all cases. Since there are 7 ELM models in schemes 2 and 3, each for individual IMF, average of best CV scores for all models are used for arriving at the best CV score for the forecasting scheme. So CV score should be compared within forecasting schemes for the impact of feature ranking and sale auto-regressive terms only. It is observed that EEMD feature extraction of sale improves forecasting accuracy from 6.96 to 5.42 and subsequently, EEMD feature extraction of economic indicators improves forecasting accuracy from 5.42 to 5.02. The prediction graph corresponding to the EEMD with feature extraction of economic indicators (approach 3) is shown in Fig. 9.

As illustrated previously, there is significant improvement in sales forecasting if we extract EEMD features of economic indicators. This improvement can be

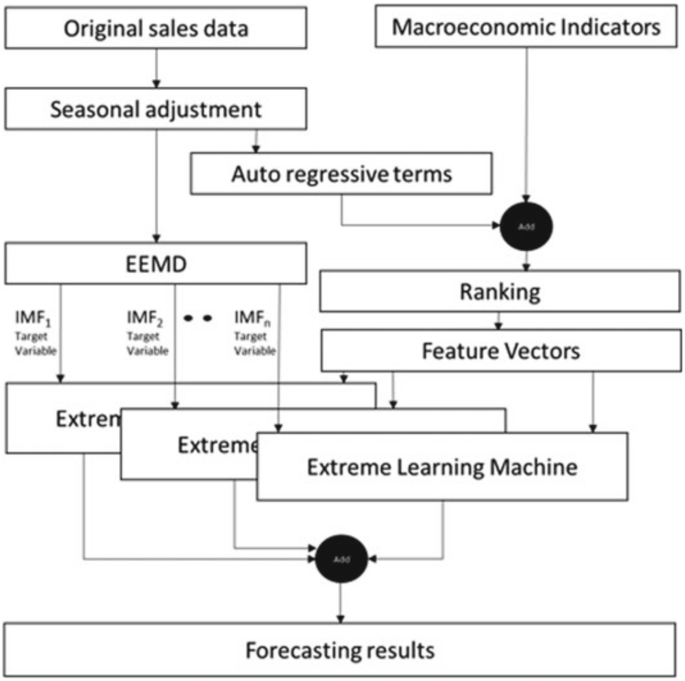


Fig. 7 Forecasting scheme for Model #2

attributed to the fact that EEMD features of a time series represent different independent causes which is hidden in the original time series so when we break all the indicators into EEMD components, we have a better chance of finding independent causes which can explain the variations in sales time series data. Table 4 compares performances of individual ELM models for sale IMFs, first thing to notice is that initial IMFs have more CV error than later IMFs. This is natural because initial IMFs represent high frequency oscillations in the time series which are relatively difficult to model than low frequency trend components represented by later IMFs. We also observe improvement in CV errors for ELM models of initial IMFs of scheme 3 (Table 5) over those of scheme 2 (Table 4), this substantiate the fact that high frequency oscillations are modelled more accurately with EEMD decomposition of economic indicators which explains the improvement of forecasting performance of scheme 3.

Literature survey shows LASSO and VAR as most successful forecasting methods for multivariate data setting so forecasting models are developed using these methods to compare the forecasting performance of proposed schemes. LASSO performs feature selection itself so feature ranking was skipped while sale auto-regressive terms were used to give it same playing ground as proposed approaches. Similarly feature ranking and sale auto-regressive terms were used in VAR model as well how-

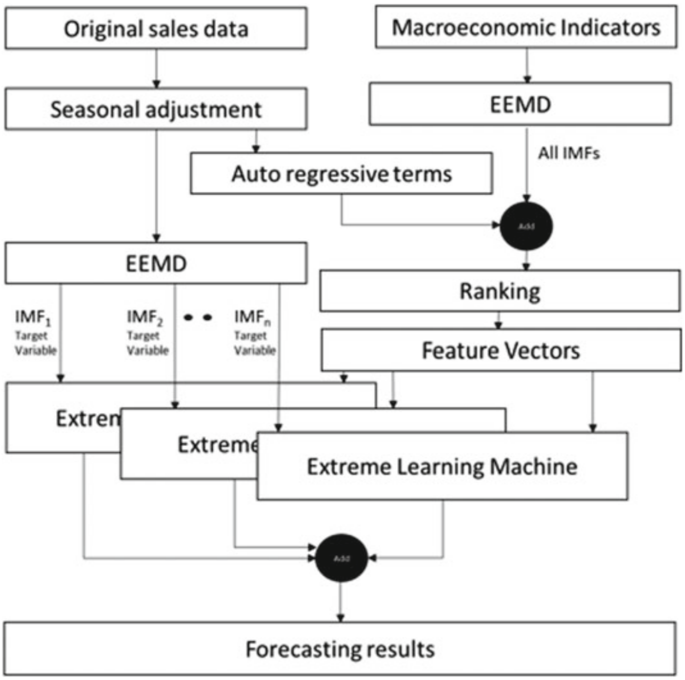


Fig. 8 Forecasting scheme for Model #3

Table 3 Performance comparison of three approaches

	Combination of ranking, sale AR terms	CV score	Test score
Approach 1	No Ranking, No AR, ELM	15.58	9.03
	Ranking (11), No AR, ELM	11.95	5.96
	Ranking (5), AR(6), ELM	11.60	6.96
Approach 2	No Ranking, No AR, Indicators with sale IMFS	20.50	4.83
	Ranking, No AR, Indicators with sale IMFS	14.43	7.06
	Ranking, AR, Indicators with sale IMFS	12.22	5.42
Approach 3	No Ranking, No AR, Indicators IMFs with sale IMFS	20.30	7.82
	Ranking, No AR, Indicators IMFs with sale IMFS	12.80	11.16
	Ranking, AR, Indicators IMFs with sale IMFS	11.20	5.02

Fig. 9 Prediction graph corresponding to the best model

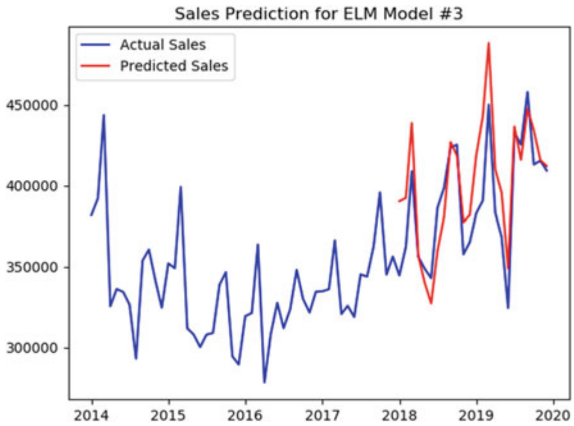


Table 4 Performance of forecasting schemes 2

IMF no.	# AR terms, # Top ranked features	ELM hyper parameters (α, # hidden neurons)	ELM CV
1	4.5	0.01, 10	13.7
2	7.3	0.0001, 12	21.76
3	5.11	0.001, 12	29.78
4	12.7	0.0001, 29	20.09
5	9.2	0.0001, 33	0.2
6	6.4	1, 19	0.005
7	7.0	0.1, 14	0.002

Table 5 Performance of forecasting schemes 3

IMF no.	# AR terms, # Top ranked features	ELM hyper parameters (α, # hidden neurons)	ELM CV
1	0.1	0.0001, 10	13.39
2	1.5	0.1, 10	20.78
3	1.2	0.0001, 13	27.77
4	6.45	0.0001, 36	16.38
5	5.10	0.0001, 39	0.11
6	5.11	0.0001, 31	0.007
7	2.7	0.0001, 34	0.005

Table 6 Performance comparison for all methods

Model	Test error
LASSO (α = 100,000, Lag = 2)	9.47
VAR (all features, Lag = 12)	6.49
ELM model #1	6.96
ELM model #2	5.42
ELM model #3	5.02

ever most optimal performance was observed when all features were used. Table 6 includes optimal working parameters for LASSO and VAR.

As observed in Table 6, all other methods provide a significant performance boost as compared to LASSO. This happens because nonlinear models are better able to describe nonlinear sales time series as opposed to linear models. VAR is a simple modeling scheme and works quite better in most of the scenarios but ELM supplemented with EEMD was able to outperform the VAR model. ELM model provides benefits of the versatility of neural networks in terms of modeling complex, nonlinear data with simpler and faster training methods, and EEMD was able to extract independent causes in nonlinear data which further supplement ELM to produce most accurate forecasts.

4 Conclusions and Future Work

Sales forecasting is a crucial element for marketing and inventory management and various business decisions in a company. This paper proposes a novel forecasting scheme for sales prediction by combining seasonal adjustment, EEMD decomposition, autoregressive sale terms, feature ranking, and ELM machine learning method on a multivariate dataset. Our experimental results show that the proposed forecasting scheme provides superior accuracy. ELM is a versatile machine learning method that is found to perform best for time series data of limited data points (few hundreds) in many studies majorly focused on univariate time series. Our study establishes that ELM can perform better than the state of the art approaches if supplemented with EEMD feature extraction, sales auto-regressive terms, and feature ranking. Sales auto-regressive terms are enhancing the forecast capacity as macroeconomic indicators may not be enough to capture all the underlying causes for changes in sales data. Furthermore, feature ranking is deployed to reduce the dimensionality of data as EEMD feature extraction, sale auto-regressive terms, and macroeconomic indicators together add a lot of features for the ELM model, some of which may not be very useful in explaining the changes in sales data and need to be discarded. Feature ranking also brings in interpretability to the forecasting model to the extent that we can know the impact of the individual feature on forecasting performance. Hence the inclusion of different economic indicators or any possible data can be decided based on its importance. The current study is focused on macroeconomic indicators only but there is no reason to limit features to macroeconomic indicators only, more potential features based on domain expertise can be decided and added to the model, some of these could be intrinsic to the company such as business confidential information, etc. and some could be extrinsic to the company such as user's reviews on the product, competitor sales, etc. Forecasting schemes can be further improved based on the addition of more machine learning models such as LSTM, GRU, etc. and a hybrid of these ML models for each sale IMF.

Acknowledgements Many thanks to Dr. Manish Sharma for his numerous valuable suggestions during this research and Mrs. Harshita Bhatia for her support in dataset collection. We would also like to express our gratitude to all the senior management of Samsung R&D Institute, Delhi especially Mr. Pankaj Mishra for supporting our work.

References

1. Makridakis S, Spiliotis E, Assimakopoulos V (2018) Statistical and machine learning forecasting methods: concerns and ways forward. PLOS ONE 13(3):1–26
2. Ramos P, Santos N, Rebelo R (2015) Performance of state space and ARIMA models for consumer retail sales forecasting. Robot Comp Integr Manuf 34:151–163
3. Lu C-J, Shao YE (2012) Forecasting computer products sales by integrating ensemble empirical mode decomposition and extreme learning machine. Math Prob Eng 2012:831201
4. Brockwell P, Davis R (2002) An introduction to time series and forecasting. Springer, pp 23–31. https://doi.org/10.1007/978-1-4757-2526-1
5. Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 01(01):1–41
6. Zhang C, Tian Y, Fan Z, Fan L (2020) Product sales forecasting using macroeconomic indicators and online reviews: a method combining prospect theory and sentiment analysis. Soft Comp 24(9):6213–6226
7. Sagaert Y, Aghezzaf E, Kourentzes N, Desmet B (2018) Tactical sales forecasting using a very large set of macroeconomic indicators. Eur J Oper Res 264(2):558–569
8. Hüsmann M, Borscheid D, Friedrich CM, Reith D (2012) General sales forecast model for automobile markets and their analysis. Trans Mach Learn Data Min 5:65–86
9. Sa-ngasoongsong A, Bukkapatnama STS, Kim J, Iyer PS, Suresh RP (2012) Multi-step sales forecasting in automotive industry based on structural relationship identification. Int J Prod Econ 140(2):875–887
10. Gao M, Xu W, Fu H, Wang M, Liang X (2014) A novel forecasting method for large-scale sales prediction using extreme learning machine. In: 2014 Seventh international joint conference on computational sciences and optimization, pp 602–606
11. Chai J, Du J, Lai KK, Lee YP (2015) A hybrid least square support vector machine model with parameters optimization for stock forecasting. Math Prob Eng 2015:231394
12. Lu C-J, Kao L-J (2016) A clustering-based sales forecasting scheme by using extreme learning machine and ensembling linkage methods with applications to computer server. Eng Appl Artif Intell 55:231–238
13. Nava N, Di Matteo T, Aste T (2018) Financial time series forecasting using empirical mode decomposition and support vector regression. Risks 6(1)
14. Lin S-L, Tung P-C, Huang NE (2009) Data analysis using a combination of independent component analysis and empirical mode decomposition. Phys Rev E 79:066705

15. Chen I-F, Lu C-J (2017) Sales forecasting by combining clustering and machine-learning techniques for computer retailing. Neural Comp Appl 28(9):2633–2647
16. Zhang Y, Zhong M, Geng N, Jiang Y (2017) Forecasting electric vehicles sales with univariate and multivariate time series models, the case of China. PLOS ONE 12(5):1–15
17. Majumder I, Behera MK, Nayak N (2017) Solar power forecasting using a hybrid EMD-ELM method. In: 2017 International conference on circuit, power and computing technologies (ICCPCT), pp 1–6
18. Wang W, Chau K, Qiu L, Chen Y (2015) Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. Environ Res 139:46–54
19. Yu L, Dai W, Tang L (2016) A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. Eng Appl Artif Intell 47:110–121
20. Huang G-B, Zhu Q-Y, Siew C-K (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. In: Proceedings of the IJCNN, vol 2. Budapest, Hungary, 25–29 July 2004, pp 985–990
21. Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501
22. Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Methodolog, 267–288. JSTOR 2346178
23. Montgomery DC, Jennings CL, Murat K (2015) Introduction to time series analysis and forecasting, 2nd edn, pp 343–350. 978-1-118-74511-3
24. https://en.wikipedia.org/wiki/Ordinary_least_squares
25. https://en.wikipedia.org/wiki/Augmented_Dickey-Fuller_test
26. Zhang X, Yin H, Wang C, Wang J (2015) Forecast the price of chemical products with multivariate data
27. Ahmad I, Basheri M, Iqbal MJ, Rahim A (2018) Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection
28. Cao J, Lin Z (2015) Extreme learning machines on high dimensional and large data applications: a survey
29. Zhu B, Shi X, Chevallier J, Wang P, Wei Y (2016) An adaptive multiscale ensemble learning paradigm for nonstationary and nonlinear energy price time series forecasting

Localization in Wireless Sensor Networks Using Link-Quality Values: A Real Test-Bed Implementation

Tuhin Majumder and Punyasha Chatterjee

Abstract Localization is the process of obtaining the location of a wireless node in indoor/outdoor setting or environment. In the recent age of Internet of Things (IoT), localization is important to provide communication, connection and efficient inter-networking between various devices. Global Positioning System is a popular approach to solve the localization problem. But unfortunately, it has considerable energy consumption, requires an unobstructed line of sight to satellites and costly if attached to every node in IoT. Localization using link-quality values can be an alternate solution. In this paper, we have implemented the localization technique using link-quality values for a Wireless Sensor Network on a real test-bed scenario.

Keywords Internet of things (IoT) · Wireless communication · Wireless sensor network (WSN) · Localization · Link-quality · Trilateration · SENSEnuts

1 Introduction

Internet of Things (IoT) is the extension of Internet connectivity from cyber world to physical devices and every-day objects. In the recent era of IoT, anything can be connected with anything else in the world and is important to business and other enterprises [1]. Embedded with electronics, communication facility, and other forms of hardware (such as sensors, actuators, processors), these physical devices can communicate and interact with each other, and they can be remotely monitored and controlled. There are so many applications of IoT which are necessary to build a safe, secure and smart world; some examples are security cruise control, involuntary car parking, driver less vehicles, distribution of road information, crime and emergency control etc.

T. Majumder
Electronics and Communication Engineering Department, Heritage Institute of Technology, Kolkata, India
e-mail: tuhinmaj1998@gmail.com

P. Chatterjee (✉)
School of Mobile Computing and Communication, Jadavpur University, Kolkata, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_28

The rapid growth in technology needs more accuracy in terms of localizing different objects in IoT. Global Positioning System (GPS) is a well-known approach to solve localization problems. Specifically, for vehicle localization, this is the most widely used technology. But unfortunately, it requires an unobstructed Line-of-Sight (LoS) connection to satellites [2] for which it cannot be functioned correctly at obstructed areas like tunnels or urban canyons or in indoors [3]. GPS is energy consuming and can be costly if implemented in a large network. In addition to this, GPS offers accuracy up to five meters at maximum [4]. It is highly desirable to design a localization technique [5] having low-cost, scalability, efficiency, and most importantly maximum accuracy.

In this paper, we have implemented a GPS-free localization technique, named *Trilateration* [6], on a real WSN test-bed scenario, using the link-quality values. The rest of the paper is organized as follows: Different types of GPS-free localization techniques are described in Sect. 2; Localization using Trilateration is discussed in Sect. 3; The whole experiment is illustrated in Sects. 4 and 5 includes conclusion and probable future opportunities.

2 Different Localization Techniques

Here, we give a brief study on some of the GPS-free localization techniques that have been proposed in different applications.

2.1 Angle of Arrival (AoA)

The angle between the propagation direction [7] of an incident wave and some reference direction or orientation is defined as *AoA* [8]. A current application of *AoA* is in the geodesic location finding or finding geolocation of cell phones. It doesn't need synchronization between the measuring units, and it works well in situations with Line of Sight. But the accuracy and precision decrease when there are signal reflections (e.g. *multi-path effects*) or the mobile target moves further from the measuring unit. So, it is not good at indoors. Moreover, it requires large and complex hardware.

2.2 Received Signal Strength Indicator (RSSI)

RSSI is measured based on the power present in the received signal [9]. A very well recognized formula is

$$RSSI(\text{dBm}) = -10n \log 10 \log 10(d) + S$$

where d is distance (meters), S is the received signal strength (dBm) at 1 m, and n is the *path-loss exponent*. The main reason of popularity of *RSSI* is that it doesn't require extra hardware and can be used on almost any devices having any type of wireless communication strategy [10, 11].

2.3 Time of Arrival (ToA)

ToA exploits time stamps, attached in transmitted packets to calculate the distance, the packet travels to reach destination [5]. The signal propagation time can be determined using synchronous clocks [12]. The mathematical model for *ToA* is given as

$$d = c * (t_{arrival} - t_{sent})$$

where t_{sent} is the exact time at which the signal is sent, $t_{arrival}$ is the time at which the signal arrives at the destination and c is constant i.e., the speed of the light [13].

2.4 Time Difference of Arrival (TDoA)

TDoA is believed to be the most accurate method [14]. The mathematical model for *TDoA* is given as

$$d = (S_{\text{radio}} - S_{\text{sound}}) * (t_{\text{sound}} - t_{\text{radio}} - t_{\text{delay}})$$

where S_{sound} and S_{radio} are the speed of sound and that of radio signals in air respectively. The fixed amount of time anchor nodes waits is t_{delay}, the time required for receiving a radio signal by an unlocalized node is t_{radio}. When unlocalised node detects chirps sent by anchor node, it saves the time t_{sound}. It works well with wideband signal and requires simple antenna, but the performance degrades as signal bandwidth decreases [15].

2.5 Link Quality Indicator (LQI)

In radio frequency transmissions, after each successful reception of data frame, *LQI* provides link quality information and signal strength to network and above layers [16]. It should be noted that higher *LQI* signifies better link quality degrading the data loss rate. *LQI* is measured in integers, ranging from 0x00 to 0xff (0–255). It indicates the signal quality, detectable by the receiver (between −100 and 0 dBm).

In this paper, we are focusing on solving the problem of localization using *LQI*. We are estimating the distance between two wireless nodes using their *LQI* values. To find the exact location of a node, we are implementing an easy yet very effective method i.e. *Trilateration.* In the next section, we will discuss the *Trilateration* technique in a brief.

3 Localization Using Trilateration

With the *Trilateration* technique, we can determine an absolute or relative location of an unknown point in 2*D* space provided the distances from that unknown point to at least three non-linear known (anchor) points are given.

Let, there are three anchor nodes *A*, *B* and *C* with known coordinates (i, j), (k, l) and (m, n). The coordinate of node *P* is unknown which we need to find out. The distances of node *P* from *A*, *B*, *C* are *a*, *b* and *c* respectively. The problem is illustrated in Fig. 1. We have drawn three circles centered at *A*, *B* and *C* with radii *a*, *b* and *c* respectively. Hence, *P* will always be at the convergence point of these three circles.

Let (x, y) be the coordinate of *P*. Therefore,

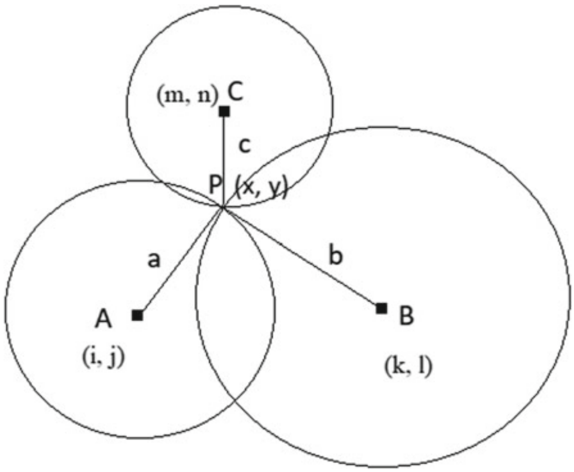
$$a^2 = (x - i)^2 + (y - j)^2 \tag{1}$$

$$b^2 = (x - k)^2 + (y - l)^2 \tag{2}$$

$$c^2 = (x - m)^2 + (y - n)^2 \tag{3}$$

From these three equations, we can find the expressions of *x* and *y* as given below:

Fig. 1 Finding co-ordinates of unknown node using three anchor nodes



$$x = \frac{\beta_1.\gamma_2 - \beta_2.\gamma_1}{\alpha_2.\beta_1 - \alpha_1.\beta_2} \quad (4)$$

$$y = \frac{\gamma_1.\alpha_2 - \gamma_2.\alpha_1}{\alpha_2.\beta_1 - \alpha_1.\beta_2} \quad (5)$$

where, $\alpha_1 = 2(i - k)$, $\beta_1 = 2(j - l)$, $\gamma_1 = (i^2 + j^2 - k^2 - l^2 + b^2 - a^2)$, $\alpha_2 = 2(k - m)$, $\beta_2 = 2(l - n)$, $\gamma_2 = (k^2 + l^2 - m^2 - n^2 + c^2 - b^2)$.

4 Experiment on Real Testbed

We have used the laboratory of School of Mobile Computing and Communication at the Jadavpur University for our experiment. We have used SENSEnuts motes [17] as shown in Fig. 2 for our experiment. The coding is done using C and Matlab in SENSEnuts GUI.

We have used four motes, among which one node acts as *PAN (Personal Area Network) coordinator* and the other three act as *coordinators*. *PAN coordinator* is nothing else but a sink node which processes the data, received by it and takes necessary decisions. It has two modes of operations: *PULL* mode and *PUSH* mode. It can send requests for data in *PULL* mode and it can receive data from all other neighbouring nodes/*coordinators* in *PUSH* Mode. In general, we assume that the signal transmission pattern of sensors is omnidirectional as shown in Fig. 3.

Objective: We assume that the positions of the three *coordinators* are known a-priori. So, they are behaving as *anchor* nodes. By applying *Trilateration* technique, and using the link quality values of the *coordinators*, the *PAN-coordinator* will automatically find its own position.

It is to be noted that, in Fig. 3, the *PAN-coordinator* is connected to the computer through an USB cable. So, the network is single hop. We have done it for the simplicity. But this experiment can easily be extended to find out the position of

Fig. 2 SENSEnuts motes

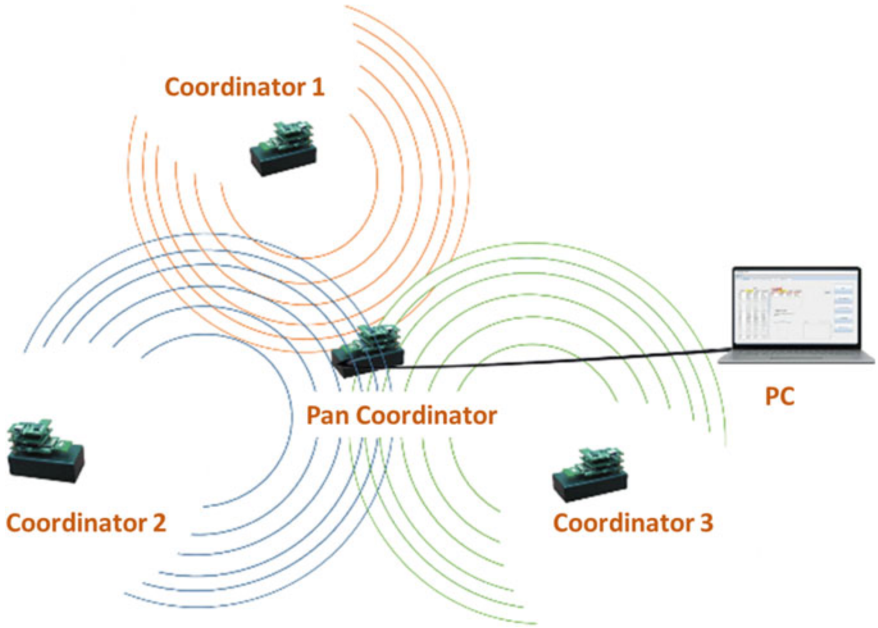


Fig. 3 *Coordinators* sending data to *PAN-coordinator*

any wireless mote in multi-hop network with a slight modification of the initial set-up.

The experiment is done in two steps:

1. Obtaining relationship between link-quality indicator and distance
2. Finding the location of unknown mote/*PAN coordinator* by applying *Trilateration.*

4.1 Link-Quality Indicator Versus Distance

A large piece of paper with fine measurement (in inches) blocks has been spread over the floor of the laboratory where the four motes are being placed. First, we switch-on one *coordinator* and the *PAN coordinator*. We keep the *PAN coordinator* in a fixed position and reposition the *coordinator* at different points on the paper and measure manually the respective distances from the *PAN-coordinator*. The *coordinator* starts sending its link quality value to the *PAN-coordinator* at every position. Finally, we get the Table 1 from the data received. We can see that, as the distance decreases, the *LQI* value increases. In Fig. 4, a graph is shown indicating the relationship between *LQI* and Distance. From the graph, by using https://www.mycurvefit.com/, we generate an equation, given below:

Table 1 Link quality indicator (LQI) versus distance

Distance (in.)	LQI
48	132
45	144
42	147
39	158
36	159
33	162
30	168
27	171
24	177
21	186
18	192
15	195
12	204
9	216
6	219
3	237

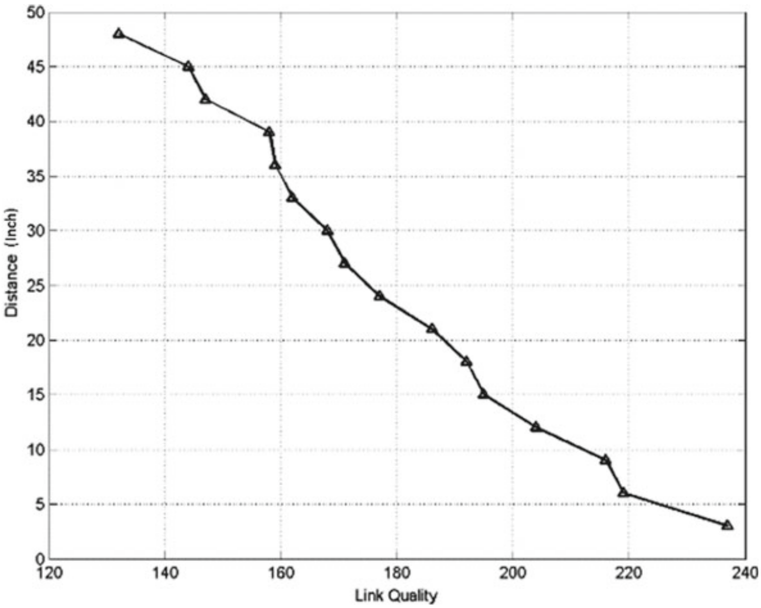


Fig. 4 Graph depicting *LQI* versus distance

$$\mathrm{d} = -2.523772 + \frac{(58.322792)}{\left(1 + \left(\frac{\mathrm{lq}}{174.281}\right)\right)^{7.144644}} \tag{6}$$

where *d* is the distance and *lq* is the *LQI* value. So, this is basically the relationship between the distance and *LQI*.

4.2 Finding the Location of Unknown Mote

Next, we fix the positions of all the four motes as shown in Fig. 5 and switch them on. All the three c*oordinators* parallelly start sending their link quality values along with their addresses to the *PAN-coordinator*. By applying Eq. 6, we convert the link quality values to distances and got the distances between the *coordinators* and the *PAN-coordinator,* as given in Table 2.

Finally, by applying *Trilateration* technique (using Eqs. 4 and 5), we generate the coordinate position of the unknown mote i.e. the *PAN-coordinator* as follows:

$$(\mathrm{x}, \mathrm{y}) = (25.9222\,\mathrm{in.}, 30.1869\,\mathrm{in.})$$

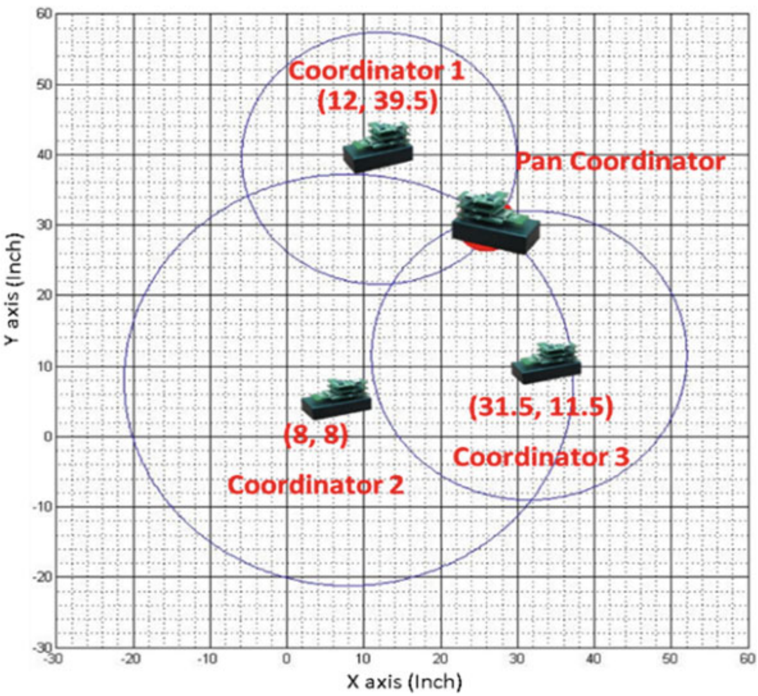


Fig. 5 Mote positions

Table 2 *LQI* and distances of *coordinators* w.r.t *PAN coordinator*

Mote	Address	LQI	Distance (in.)
Coordinator 1	D0 A6	190	17.9166
Coordinator 2	0B 70	170	29.2217
Coordinator 3	6B BD	185	20.5124

We observe that, the calculated point is located just little deflected from the original (27, 31) position of the *PAN coordinator*. This may be due to some interference or the position of the radio receiver of the mote etc. If we apply the experiment in a larger spectrum then this tiny deflection does not matter. So, it is a quite satisfactory result.

All the codes used in the experiment are available in GitHub [18].

5 Conclusion

In this experiment, we have observed an easy to implement, accurate and efficient method of localization technique, based on link-quality index and *Trilateration* technique, suitable for both indoor and outdoor environment. Yet there are some meagre problems such as the interference, battery percentage, sensing range limitations which can alter the link-qualities of the motes that would give an opportunity for further research. Also, we need to compare the accuracy of our results with respect to GPS technique.

References

1. Ramnath S, Javali A, Narang B, Mishra P, Routray SK (2017) IoT based localization and tracking. In: 2017 international conference on IoT and application (ICIOT), Nagapattinam, pp 1–4. http://doi.org/10.1109/ICIOTA.2017.8073629
2. Barani H, Fathy M (2010) An algorithm for localization in vehicular ad-hoc networks. J Comput Sci 6(2):168–172
3. Seidel SY, Rappaport TS (1992) 914 MHz path loss prediction models for indoor wireless communications in multifloored buildings. IEEE Trans Antennas Propag 40(2):207–217. https://doi.org/10.1109/8.127405
4. Thin LN, Ting LY, Husna NA, Husin MH (2016) GPS systems literature: inaccuracy factors and effective solutions. Int J Comput Netw Commun (IJCNC) 8(2). http://doi.org/10.5121/ijcnc.2016.8211
5. Zafari F, Gkelias A, Leung KK (2019) A survey of indoor localization systems and technologie. IEEE Commun Surv Tutorials 21(3):2568–2599. http://doi.org/10.1109/COMST.2019.2911558
6. Yagi Y, Yachida M (2002) Omnidirectional sensing for human interaction. In: Proceedings of the IEEE workshop on omnidirectional vision 2002. Held in conjunction with ECCV'02, Copenhagen, Denmark, pp 121–127. http://doi.org/10.1109/OMNVIS.2002.1044505

7. Peng R, Sichitiu ML (2006) Angle of arrival localization for wireless sensor networks. In: 2006 3rd annual IEEE communications society on sensor and ad hoc communications and networks, Reston, VA, pp 374–382. http://doi.org/10.1109/SAHCN.2006.288442
8. Tomic S, Beko M, Dinis R (2017) 3-D target localization in wireless sensor networks using RSS and AoA measurements. IEEE Trans Veh Technol 66(4):3197–3210. https://doi.org/10.1109/TVT.2016.2589923
9. Michael AO (2012) Standardization of attenuation formula for radio waves propagation through free space (LOS) communication links. Sci J Phys 2012(Article ID sjp-281):7. http://doi.org/10.7237/sjp/281
10. He W, Ho P-H, Tapolcai J (2017) Beacon deployment for unambiguous positioning. IEEE Internet Things J 4(5):1370–1379
11. Paul AS, Wan EA (2009) RSSI-based indoor localization and tracking using sigma-point Kalman smoothers. IEEE J Sel Top Sign Process 3(5):860–873
12. Ravindra S, Jagadeesha SN (2014) Time of arrival based localization in wireless sensor networks: a linear approach. ArXiv abs/1403.6697
13. O'Keefe B (2017) Finding location with time of arrival and time different of arrival techniques. ECE senior capstone project. https://www.sites.tufts.edu/eeseniordesignhandbook/files/2017/05/FireBrick_OKeefe_F1.pdf
14. Farid Z, Nordin R, Ismail M (2013) Recent advances in wireless indoor localization techniques and system. J Comput Netw Commun 2013(Art. no. 185138)
15. Alrajeh NA, Bashir M, Shams B (2013) Localization techniques in wireless sensor networks. Int J Distrib Sens Netw. http://doi.org/10.1155/2013/304628
16. Yang T, Yang Q, Cheng L (2015) Experimental study: a LQI-based ranging technique in ZigBee sensor networks. IJSNet 19:130–138
17. sensenuts.com (SENSEnuts IoT sensation). https://www.eigen.in/pdf/sensenutsuniversity.pdf
18. https://www.github.com/tuhinmajumder1998/find-mote-sensenuts

Twitter Spam Review Detection Using Hybrid Machine Learning Techniques

Hrishikesh Viswanath, Ravendra Singh, and Varun Gupta

Abstract This paper focuses on detection and classification of spam reviews in a data set of reviews scraped from Twitter. Detection of spam reviews is a major step in screening and blocking irrelevant information posing as a review and can further be used to blacklist users. It is not possible to conclusively classify anything as spam even by human beings. Expecting a machine to do so requires extensive training and exploration of different kinds of models that may offer a multi-dimensional look at the problem that is superior to the discerning ability of human beings. Models should be devised in such a way that they look beyond simple sentiment analysis. In this work, Random Forests and Support Vector machines are designed to work together as a hybrid model to classify reviews. With hybrid methods that utilize ensemble learning, it will be possible to exploit the advantages of multiple models and combine the results to classify a review as spam.

Keywords Spam detection · Cascaded models · Random forest · Support vector machines · Natural language processing · Artificial intelligence

1 Introduction

Product reviews play an important role in imparting bias in favor of or against the said product in the mind of the user. Not only does it impact the sales of the product, but it also reflects on the reliability and credibility of the review site. The opinions of users available on various review sites are important for potential shoppers. On the business end, opinions help vendors gain an insight into customer interests and furthermore, help in changing marketing strategies. The problem of determining whether a review is spam is not objective. A review that is flagged as spam may not be so, in a similar manner, spam reviews may get flagged as being authentic. This discernment varies from person to person [1]. The ability to judge if a review is spam or not is influenced by various features such as usage of capital letters, excessive use of superlative adjectives and punctuation symbols. Certain words like

H. Viswanath (✉) · R. Singh · V. Gupta
Department of Computer Science and Engineering, PES University, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_29

'Offer' are indicators of what might be considered as spam [2]. Many customers are influenced by online reviews, which act as a catalyst in imparting bias in the minds of the customers. It is of primary concern to prevent customers from being wrongly influenced by these reviews. Furthermore, spam reviews lead to discrediting the review website. Any user who repeatedly encounters spam content or malicious data in the review site will eventually avoid the site. Lastly, vendors use review sites as a feedback mechanism to improve their customer relationships and better understand the needs of the buyers. It is therefore important that they be given right data lest they misinterpret the market.

2 Literature Analysis

Various Techniques have been proposed to identify opinion spam. This is usually considered a problem of binary classification where every review is either a spam review or a ham review [3].

The situation regarding spam reviews is such that it is quite confusing for the human mind to judge between spam and ham. Expecting a computer algorithm to do what humans fail to is one of the challenges. The methods proposed in Li et al. [4] used supervised and semi-supervised models to classify reviews. However, they failed to give F1 scores higher than 0.583 and that became the primary metric to measure the effectiveness of the models. In these situations, one cannot expect simple sentiment analysis or rudimentary binary classification algorithms to classify reviews [5]. Relying on accuracy to be the sole benchmark for evaluating any model can be quite inefficient. Accuracy as a measure of quality fails when the data set is flooded with biased reviews. One cannot expect that there be a uniform dataset available for testing in real time and it is futile to disregard imbalance in the data set. A more appropriate measure of efficiency would be F1 score, which is a statistical measure of how often positive values are classified positive. It measures the efficiency of the model assuming that it contains a single class, which is the primary goal.

A paper by Shojaee et al. [6] states that opinion mining and its effectiveness relies on the credibility and availability of sentiments during analysis. The need to filter out defective opinions from a spammer, and performing various experiments to identify spammer behavior. The absence of an annotated data-set also poses a problem to test the validity of the techniques in spam detection. In their paper, they show a technique to annotate data-sets (of reviews) for detecting spam by providing meta information about both the reviewers and the reviews to the annotators for effectively annotating spam.

A paper by Barbado et al. [7] describes a feature framework for identifying spam reviews that have been processed in the consumer electronics domain. They show the construction of a data-set for classifying and identifying the spam reviews in the said domain in 4 unique cities based on techniques of scraping, using a feature framework for reviewing spam detection, the development of a classification method for spam reviews based on the mentioned proposed framework and the analysis and evaluation of the results for all the cities as given in the paper.

A paper by Sun et al. [8] presents a review of the NLP (Natural Language Processing) techniques and methods based for opinion mining. The authors introduce various standard Natural Language Processing techniques which are mandatory to preprocess textual data. They investigate the approaches of opinion mining in review data-sets for various kinds of levels and different situations. They then introduce comparative opinion mining and deep learning based techniques for opinion mining. Opinion summarization and other advanced topics are introduced later in the comprehensive study. Finally, the authors have a discussion on some of the challenges and open problems that are faced during and related to opinion mining.

A journal by Dhingra et al. [9] gives insight about the various techniques to classify spam such as Support Vector Machine, Naive Bayes, URL analysis and so on. The paper discusses classification and spam identification for techniques used mainly on the social media platform Twitter, as it is considered to be one of the most sensitive Social Media platform and vulnerable to public opinion.

A paper by Baraithiya and Pateriya [10] proposes an ensemble classifier based technique to identify spam reviews. The ensemble classifier uses a Naive Bayes based classifier, SVMs (Support Vector Machines), and the KNN Mutual (k-Nearest Neighbor) classifiers to classify the reviews. This technique as given in their paper is primarily processed on the “Yelp” datasets. The results of the classifier are shown to have a higher accuracy than that of traditional techniques on both the classified datasets.

A paper by Lui and Baldwin [11] shows that transductive learning (or cross domain learning) is an important factor to consider when building a great general-purpose system for language identification. The authors develop a feature selection method that generalizes this across various domains. The results of their paper demonstrate that their method provides good improvements for language identification using the transductive transfer learning system. The authors implement the method as described and show that their system is in fact much faster than the currently available popular and standalone language identification systems, while maintaining enough competitive accuracy.

A paper by Li et al. [4] exploits various machine learning algorithms to detect spam in consumer reviews. The authors build a collection of spam reviews by analyzing reviews extracted using web crawling. The first step was to analyze the effect that various features had on the detection of spam reviews and to simultaneously observe that the reviewer that was classified as a spammer consistently published spam reviews. This method is key and provides another perspective to detect spam in reviews: they help identify if the writer of a particular review is in fact a spam reviewer. Based on such observations, the authors are able to suggest a two-view, semi-supervised technique called “co-training,” to exploit the huge number of currently unlabeled review data-sets. The result of this experiment performed by the authors conclude that their proposed solution is pretty effective and that the developed learning methods achieve significant improvements when compared to the industry standards in heuristics of spam review classification.

An article by Kale et al. [12] says that the goal of the authors is to identify reviews which are likely to be spam reviews by taking into account some common give-aways such as the discontinuous textual flow of the review, inadequate information pertaining to the product, inappropriate language not related to the specific context of the product/service which represents honest feedback for the customers that review a given product on the website. The authors mention that they have noticed that all the previous research in this domain is primarily focused on the extraction, the classification and the summarization of user opinions and differentiating between spam reviews and ham ones. But, their aforementioned system aims to identify discontinuous and/or irregular text flow, vulgar or inappropriate language or not related to the context in any specific way and to find any similarities between the reviews posted by review spammers and to create a standard template out of them.

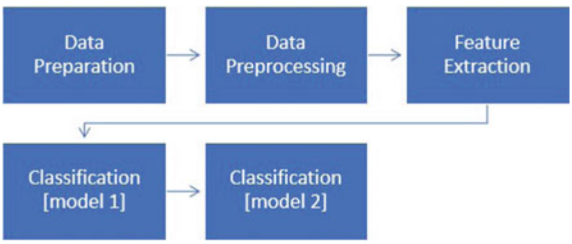
A paper by Dixit and Agrawal [5] pre-re-views and re-views the substantial research that has happened on Spam Review Detection techniques. The paper provides a state of the art analyzing and deducing some of initial attempts made to study spam detection in reviews, and to improve upon it by proposing better techniques.

3 Dataset Preparation

The flow of the implementation is as depicted below (Fig. 1).

The dataset that forms the corpus for training the models is built by scraping twitter reviews. Any tweet is considered to be a review if it contains the hashtag "#review" or a partial match for the term review within the tweet. The data set used for training needs to be balanced. Care is taken to ensure that it contains a reasonable number of tweets of both the categories—spam and ham. A part of the data set comprised reviews that are pre-processed and classified. This is done to ensure that the models don't over fit to exclusively Twitter reviews. The data scraped off Twitter is dubbed "Live dataset".

Fig. 1 Classification process



3.1 Pre-processing and Feature Extraction

The data obtained needs to be preprocessed to remove unwanted features such as stop words and proper nouns. Preprocessing also includes stemming [13, 14], the process of reducing words to their roots and lemmatization [14], which refers to removal of inflectional endings such as -ing and -es.

4 Models

The classification is a 2-step process. The preprocessed data is fed to the first classifier initially. The data points that are wrongly classified are fed to the second model for classification. This process is analogous to proof reading. The two models are chosen in such a way that they do not fail in the same situations. Support Vector machine and Random Forest are used as the components of the hybrid model. Random Forest, which uses ensemble learning [15] technique based on bagging [15, 16], is chosen as the first model due to its tendency to not over fit when presented with large datasets that are inherently noisy. SVM neither performs well on large datasets which may be noisy, nor is it immune to overfitting. However, the number of trees created by the RF algorithm increases with the increase in the size of the data. Therefore, both space and time are conserved by passing the wrongly classified data points to the random forest model rather than the entire dataset.

4.1 Support Vector Machine

A support vector machine is a supervised machine learning algorithm that, in this scenario, is used for binary classification [17].

The support vector machine, when used for binary classification, aims to find a way to separate the two classes of the data in such a way that it constructs a hyperplane [18] that is farthest from either of the classes (Fig. 2).

This is achieved by first finding the data points of the two opposite classes that are closest to each other, i.e. support vectors. The hyper plane is constructed such that it is at the farthest distance from the two support vectors belonging to the opposite classes.

The maximum distance of the hyper plane between the two support vectors is first calculated by differentiating the Lagrangian [19], which is a function of bias (b) and weights (w). In order to account for outliers and data points that lie on the wrong side of the margin, a slack variable (ξ) is taken into consideration. ξ is zero if the variable lies on the right side of the margin.

The equation is denoted by

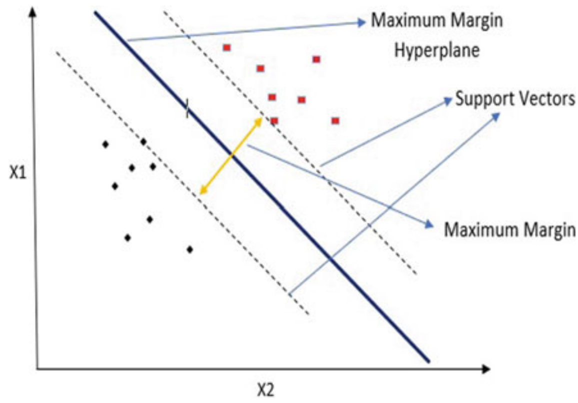


Fig. 2 Support vector machine

$$y_i(wx + b) \geq 1 - \xi \tag{1}$$

The penalty imposed on data points that lie on the wrong side of the margin is denoted by

$$\min\left(\frac{1}{2}w^2 + c\sum_i \xi_i\right) \tag{2}$$

The Lagrangian function is given by

$$L(w, b, \xi, \alpha, \beta) = \frac{1}{2}w^2 + c\sum_i \xi_i - \sum_i \alpha_i[y_i(wx_i + b) - 1 + \xi_i] - \sum_i \beta\xi_i \tag{3}$$

Differentiating the Lagrangian function with respect to w, b and ξ_i will provide the relationship between the Lagrangian parameters. In the above equation, a non-zero value of α refers to the support vectors.

$$\nabla_w L = 0 \rightarrow w = \sum \alpha_i y_i x_i \tag{4}$$

$$\nabla_b L = 0 \rightarrow \sum \alpha_i y_i = 0 \tag{5}$$

$$\nabla_{\xi_i} L = 0 \rightarrow c - \alpha_i - \beta_i = 0 \tag{6}$$

The dual of the Lagrangian is determined by maximizing the Lagrangian with respect to α after substituting the values derived in the above equations.

$$(dual)\max_{\alpha}\sum\alpha_i - \sum\sum\alpha_i\alpha_j y_i y_j x_i x_j \tag{7}$$

Radial Basis function [20] is used to generate the hyperplane by raising the data points to a higher dimension because text contains an arbitrary number of words. All the words are taken into account when constructing the hyperplane. The contribution of each word is considered variable. As the number of variables increases, it becomes harder at lower dimensions to determine relationships between them in order to split them into classes. Radial basis function is denoted by the equation

$$\phi(X_1, X_2) = Ce^{\langle X_1.X_2\rangle} \tag{8}$$

C controls regularization [21]. As C increases, the model tends to over fit. The exponent term, raised to dot product of the two vectors is responsible for projecting the data points to a higher dimension. By representing the exponent term as a series, the model is able to raise the dot product of the vector to any arbitrary dimension.

$$\phi(X_1, X_2) = c\sum_{n=0}^{\infty}\frac{\langle X_1.X_2\rangle^n}{n!} \tag{9}$$

The issue of dimensionality arises owing to the large number of words in the data set. A simple way to convert text to numerical values is to create a table of Term Frequency-Inverse Document Frequency (TF-IDF) [17] values corresponding to each word in the document. Support vector machine requires the data points to be presented as vectors and the set of TF-IDF scores for the complete document forms the vector that is processed by the model.

TF-IDF is a measure of how important a word is in the document. IDF is an inverse measure of how often the word occurs in the document, and is primarily used to weed out stop words such as "the". Words that occur frequently are given less importance, hence the term inverse frequency. TF IDF scores are calculated as follows

$$IDF(word) = \log_e\left(\frac{n+1}{n_c+1}\right) \tag{10}$$

Where, n corresponds to the number of documents and n_c refers to the number of documents in which the word occurs.

$$TF(word) = \frac{count(word)}{\sum words} \tag{11}$$

$$TF - IDF(word) = TF(word) * IDF(word) \tag{12}$$

The TF-IDF scores are L2 normalized as denoted below

$$TF - IDF(word) = \frac{TF - IDF(word)}{\sqrt{\sum_{word} TF - IDF(word)^2}} \tag{13}$$

4.2 Random Forest

Random Forests Algorithm is a supervised learning algorithm that builds a forest out of decision trees. The principle behind this action is that when a large number of decision trees classify the data together in ensemble fashion, the model outperforms a single tree attempting to do the same [15, 22]. Each decision tree takes as input, a subset of the larger dataset, whose elements are picked randomly with replacement. This method is called bagging or bootstrap aggregating [16]. Care is taken to ensure that the data fed to each decision tree is of the same dimensions. Since each tree operates on a different subset of the data, its outcome varies with the outcome of the other trees (Fig. 3).

A single decision tree classifies the result through voting (in this case, either spam or ham). However, with random forests, a large number of decision trees are used to collectively classify the result through voting.

The technique of using Random Forests to classify spam reviews was inspired by a similar implementation in a paper by McCord and Chuah [23].

A forest of 250 trees is generated and trained with a hybrid data-set comprising both pre-classified hotel reviews and manually scraped product reviews off the internet.

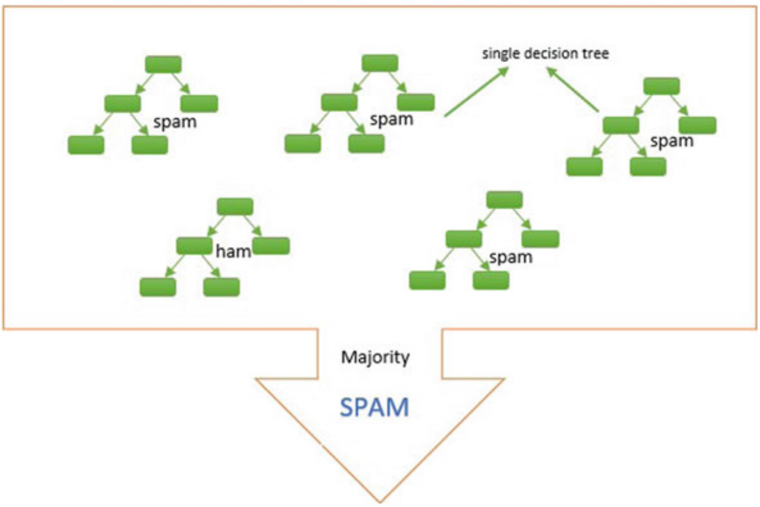


Fig. 3 Random forest classifier

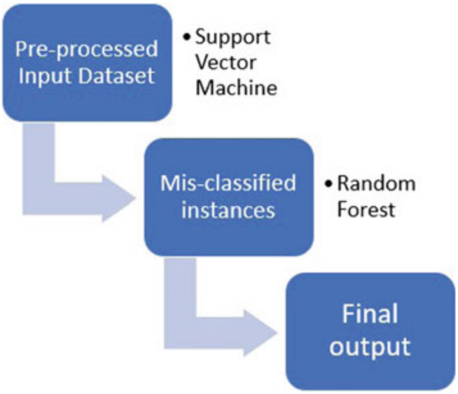


Fig. 4 Hybrid model—cascading SVM and random forest classifier

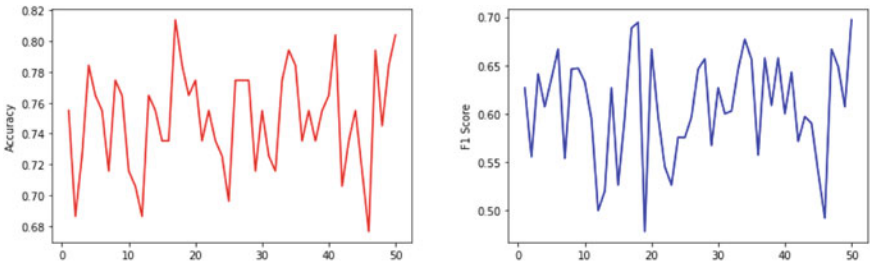
4.3 Cascading the Classifiers

The hybrid model is designed to improve the F1 scores obtained by the previous models. The idea behind setting up a hybrid model architecture is to mitigate the inefficiency of one model by using another model as a fallback mechanism (Fig. 4).

Chances of the second model failing in the exact same instances where the first model had failed is quite slim and the overall hybrid model outperforms the two individual models as proven by the F1 scores obtained.

5 Performance and Results

The models were tested repeatedly with disjoint subsets of the test dataset. This is done to simulate the variation and diversity in live data. The subsets may be balanced, biased, noisy and may contain irrelevant data. The variation in performance (Accuracy and F1 scores of the models) is depicted graphically (Fig. 5).



(a) Plot of Accuracy of SVM Model against instances of Twitter data

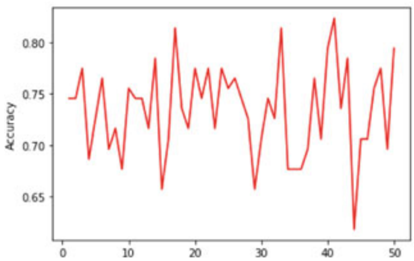
(b) Plot of F1 Scores of SVM Model against instances of Twitter data

Fig. 5 Performance of support vector machine model

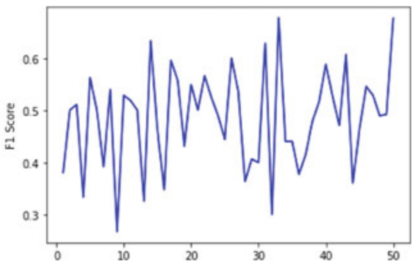
It can be inferred from the graphs that the variations in accuracy is consistent in Random forest and Support vector machine. In either case, the accuracy varies between 65 and 82%. With this information, it is not possible to conclude whether one of the models is better than the other.

The graph of F1 scores provides a much better insight into their performance. Support Vector machine has a higher range of F1 Scores, ranging between 0.5 and 0.7 while Random Forest has a lower range between 0.27 and 0.7. The variance in F1 scores is much higher in random forest than in support vector machine. This indicates that the model performs as well as support vector machine when tested with unbiased data but performs worse than the former when tested with skewed data.

However, When the Random Forest classifier is cascaded after the SVM classifier, exceptional improvement in performance is observed. The hybrid model has lower variance in accuracy and F1 scores than the individual models and the lower bound for accuracy, and F1 scores are significantly higher than either of them. This implies that the hybrid model performs both consistently and more accurately than the individual models (Figs. 6 and 7).

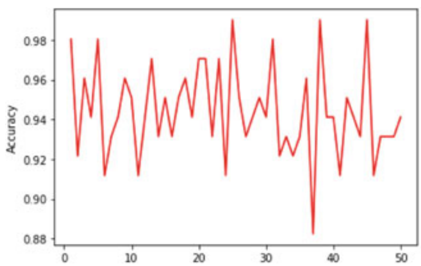


(a) Plot of Accuracy of Random Forest Model against 50 instances of Twitter data

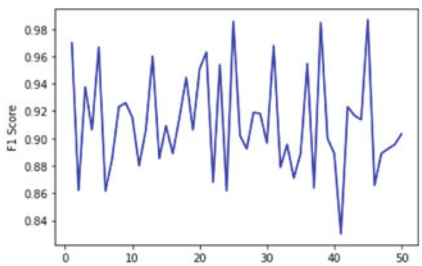


(b) Plot of F1 scores of Random Forest Model against 50 instances of Twitter data

Fig. 6 Performance of random forest model



(a) Plot of Accuracy of the cascaded model against 50 instances of Twitter data



(b) Plot of F1 Scores of the cascaded model against 50 instances of Twitter data

Fig. 7 Performance of cascaded model

Table 1 Performance of the models

Model	Accuracy (%)	F1 score (%)
Support vector machine	72.54	58.82
Random forest model	75.71	60.12
Hybrid model	96.07	93.51

The overall performance of the model, which is the aggregate of the performance on 50 instances of test data is denoted in Table 1.

It is pertinent to conclude that the hybrid model is able to better classify spam review data set than each of them can individually. This is possible because the premise that two models would not fail in the same scenario proved to be true. The second model compensated for the failure of the first model and they are able to better classify the data set together.

6 Future Work

The models explored in the paper fail to consider emotional aspects of the reviews such as sarcasm and anger. Approaches in integrating these models with tools that detect emotions would improve the accuracy of the models.

When the hybrid model is used to classify live data, it is impossible to know which data points have been wrongly classified by the first model. Heuristic approaches are a possible fix to this issue and are to be explored in determining where the first model fails.

Acknowledgements The members of the team would like to express gratitude to professor Ashwini M Joshi, who provided idea and feedback. This work was done while the authors were with PES University.

References

1. Wu T, Wen S, Xiang Y, Zhou W (2018) Twitter spam detection: survey of new approaches and comparative study. Comput Secur 76:265–284
2. Coomer G, Johnston N (2013) Systems and methods for creating text signatures for identifying spam messages. uS Patent 8,353,035
3. Mukherjee A, Venkataraman V, Liu B, Glance N et al (2013) Fake review detection: classification and analysis of real and pseudo reviews. Technical report UIC-CS2013-03. University of Illinois at Chicago
4. Li FH, Huang M, Yang Y, Zhu X (2011) Learning to identify review spam. In: Twenty-second international joint conference on artificial intelligence
5. Dixit S, Agrawal A (2013) Survey on review spam detection. Int J Comput Commun Technol 4:0975–7449

6. Shojaee S, Azman A, Murad M, Sharef N, Sulaiman N (2015) A framework for fake review annotation. In: Proceedings of the 2015 17th UKSIM-AMSS international conference on modelling and simulation. IEEE Computer Society, pp 153–158
7. Barbado R, Araque O, Iglesias CA (2019) A framework for fake review detection in online consumer electronics retailers. Inf Process Manage 56(4):1234–1244
8. Sun S, Luo C, Chen J (2017) A review of natural language processing techniques for opinion mining systems. Inf Fusion 36:10–25
9. Dhingra A, Mittal S, Kaur H, Singh A (2015) A review on techniques of spam classification in twitter. Int J 1(5):57–59
10. Baraithiya H, Pateriya RK (2019) Classifiers ensemble for fake review detection. Int J Innov Technol Explor Eng 8(4):730–736
11. Lui M, Baldwin T (2011) Cross-domain feature selection for language identification. In: Proceedings of 5th international joint conference on natural language processing, pp 553–561
12. Chaitanya K, Dadasaheb J, Tushar P (2016) Spam review detection using natural language processing techniques. Int J Innov Eng Res Technol 3(1)
13. Frakes WB (1992) Stemming algorithms, pp 131–160
14. Korenius T, Laurikkala J, Jarvelin K, Juhola M (2004) Stemming and lemmatization in the clustering of Finnish text documents. In: Proceedings of the thirteenth ACM international conference on Information and knowledge management, pp 625–633
15. Dietterich TG et al (2002) Ensemble learning. In: The handbook of brain theory and neural networks, vol 2, pp 110–125
16. Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140
17. Tong S, Koller D (2001) Support vector machine active learning with applications to text classification. J Mach Learn Res 2:45–66
18. Noble WS (2006) What is a support vector machine? Nat Biotechnol 24(12):1565–1567
19. Suthaharan S (2016) Support vector machine. In: Machine learning models and algorithms for big data classification. Springer, Berlin, pp 207–235
20. Scholkopf B, Sung K-K, Burges CJ, Girosi F, Niyogi P, Poggio T, Vapnik V (1997) Comparing support vector machines with Gaussian kernels to radial basis function classifiers. IEEE Trans Signal Process 45(11):2758–2765
21. Schölkopf B, Smola AJ, Bach F et al (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge
22. Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, RigolSanchez JP (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogramm Remote Sens 67:93–104
23. Mccord M, Chuah M (2011) Spam detection on twitter using traditional classifiers. In: International conference on autonomic and trusted computing. Springer, Berlin, pp 175–186

Predicting Expected Time of Arrival of Shipments Through Multiple Linear Regression

Prasad C. Mahajan, Arvind W. Kiwelekar, Laxman D. Netak, and Akshay B. Ghodake

Abstract Handling uncertainty in information is one of the biggest challenges faced by logistics management and freight forwarding agencies. The uncertainty may be in user needs, availability of raw materials, Estimated Time of Arrivals (ETA) of shipments, or occurrence of any natural calamities. These agencies are increasingly adopting intelligent ways of handling uncertainties in the information. In this paper, we describe a machine learning-based approach to deal with the uncertainty in the Estimated Time of Arrivals (ETA) of shipments. The approach uses multiple regression techniques to predict the ETA of a shipment at the destination port. The approach uses data from two different sources. First is the in-house internal data about shipments maintained by a company. Second is the marine traffic data collected through the Automatic Identification System (AIS). The approach combines these two data to predict the ETA of a shipment. The approach predicts the ETA of the shipment with 89% accuracy. The prediction model described in the paper is being in use and the predicted ETA has been found helpful to effectively manage resources in the downstream supply chain such as human sources at port, transporter and customers inquiries.

Keywords Automatic identification system · Marine transportation · Predictive analytics · Machine learning · Logistics

P. C. Mahajan (✉) · A. W. Kiwelekar · L. D. Netak
Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, Maharashtra, India
e-mail: awk@dbatu.ac.in

L. D. Netak
e-mail: ldnetak@dbatu.ac.in

A. B. Ghodake
ATA Freight Line India Pvt. Ltd., Pune, Maharashtra, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_30

1 Introduction

Knowing information about the Expected Time of Arrival (ETA) of a shipment is critical for logistics management and freight forwarding agencies [2]. Estimates about ETA are commonly used to answer queries of customers, to issue contracts to liners, to allocate human resources at ports, to manage warehouses, and overall to run the business smoothly. Accurate estimates of ETA can also improve resource utilization and reduce operating costs.

Estimating ETA is one of the extensively researched problems in transportation and supply chain industry [4–6]. However, for logistics management and freight forwarding agencies, the techniques and models developed to estimate ETA in a general context are often inadequate. Such models fail to utilize organization-specific information and knowledge gathered over the years, dealing with specific liners or carriers. Hence, the estimates of ETA that merely rely either on historic or streamed data of traffic fail to capture the operating behaviour of carriers or liners that these agencies deal with regularly.

Though the task of estimating ETA is crucial for logistics management and freight forwarding agencies, it poses numerous challenges to arrive at an acceptable value of ETA [7]. First, the majority of the logistic business rely on verbal communication. Second, logistics and freight forwarding business is enacted through various contractors and subcontractors. Third, the information required to predict ETA is often scattered and present in unstructured documents.

In this paper, we present a solution for estimating ETA of shipments specific for the agencies involved in the business of freight forwarding that prefers marine transportation to transfer goods. The approach uses two different data sources. First is the data about maritime traffic acquired through the Automatic Identification System (AIS) [8]. Second is in-house company data that capture the operating behaviour of liners or carriers. We use the multiple regression analysis techniques to predict the value of ETA of a shipment at a port.

2 Approach

The conceptual diagram in Fig. 1 captures the main elements of our approach. The purpose of this diagram is to make explicit the tacit knowledge assumed in various business processes carried out by logistics and freight forwarding agencies. It states that:

1. Vessels carry Shipments.
2. The position of a vessel is defined by Automatic Identification System (AIS). The AIS is a self-reporting navigational system mandatory to be installed onboard every ship. More information about AIS can be found in [1].
3. An AIS data provider can be a commercial or a non-profit entity which supplies AIS data. Logistic management or a freight forwarding agency need to subscribe to

an AIS data provider to have historical or up-to-date AIS data. The *MarineTraffic* (marintraffic.com) is one of the commercial AIS data providers.

4. A vessel is identified by AIS data attributes. Each ship has a unique vessel name. The latitude and longitude of a location define the current position of a boat in the ocean. When a vessel transmits AIS messages, the message is timestamped with $FileTimeStamp$. The $createdTimeStamp$ is the time when an AIS data provider records the received AIS message in the system.
5. Logistic management and freight forwarding agencies internally maintain shipment data. A shipment is defined by attributes such as a unique shipment number, the port-of-loading (POL), the port-of-delivery (POD), a carrier, a vessel name, the actual time of departure and arrival. The flag $isTransshipped$ denotes whether the shipment is a multi-hop shipment. A sample of shipment data is shown in Table 1.
6. The fused data is the merged data consisting of external of AIS data and internal shipment data agreeing on a common vessel name.

The fused data is then processed to make it ready to input format for building a multiple regression model.

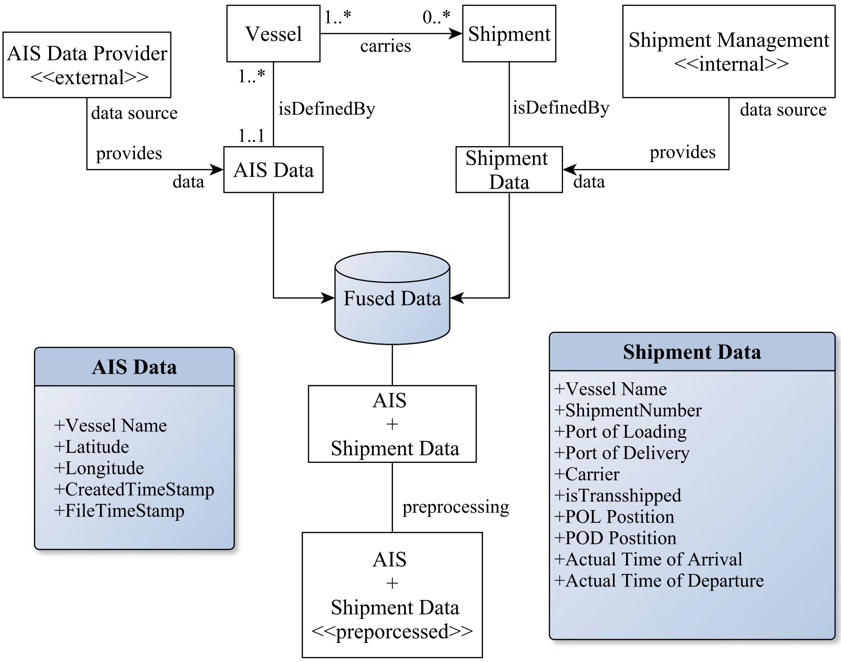


Fig. 1 Main elements in our approach

Table 1 A sample of shipment data

Vessel name	Ship No.	POL	POD	Carrier	IsTS	ATA	ATD	POL Cord	POD Cord.
ABC	X002	HAMB	IZMIR	Hapag-Lloyd	Yes	10-05-2016	16-04-2016	53.55, 9.99	38.41, 27.03
PQR	X790	MUND	HAMB	Hapag-Lloyd	No	05-05-2016	14-04-2016	22.75, 69.70	53.54, 9.99
XYZ	X211	MUND	BARC	UASC	No	17-05-2016	28-04-2016	22.75, 69.70	41.38, 2.17
LML	X376	MUND	AMBA	Safma	Yes	25-04-2017	09-04-2017	22.75, 69.70	40.96, 28.67
SQR	X017	AMBA	JEDD	UASC	No	22-05-2017	27-04-2017	40.96, 28.67	21.57, 39.15

3 Model Development

This section describes the steps adopted to build the prediction model using multiple regression technique [3].

3.1 Data Pre-processing

The fused data consist of data from two sources. The first one is external AIS data which reflects the vessel's current position in the ocean. The AIS periodically transmits vessel information. As shown in Table 2, an AIS message contains the longitude and latitude of the vessel's current position and not the distance travelled by vessel. The first objective of the prepossessing step is to calculate the distance travelled by a vessel from the port-of-loading (POL) when an AIS provider receives the AIS message from the vessel. We used two latest AIS messages received from the vessel of interest for model building. We refer the corresponding distances travelled by ship as $L1$ and $L2$. The $L1$ is the most recent distance, and $L2$ is the second

Table 2 A sample of AIS data

Name	Latitude	Longitude	CreatedTS	FileTS
ABC	21.95	68.00	17-04-2016	16-04-2016
PQR	22.63	69.17	15-04-2016	14-04-2016
XYZ	19.71	70.36	28-04-2016	28-04-2016
LML	14.20	52.11	14-04-2017	14-04-2017
SQR	21.80	68.93	27-04-2017	27-04-2017

Table 3 Calculating distance covered

```
calculate_distance<-function(long1, lat1, long2, lat2){
rad <- pi/180

#convert longitude and latitude to radian

lat1_rad <- lat1*rad
long1_rad <- long1*rad
lat2_rad <- lat2*rad
long2_rad <- long2*rad

#calculate distance between latitudes and longitudes
dlon <- long2_rad - long1_rad
dlat <- lat2_rad - lat1_rad

#Haversine formula

a <- (sin(dlat/2))^2 + cos(a1)*cos(b1)*(sin(dlon/2))^2
c <- 2*atan2(sqrt(a), sqrt(1 - a))
R <- 6378137 #Radius of the earth
d <- R*c #final distance
return(d)
}

for (everyVessel in dataset){
    for (everylocation for everyVessel){
        L=calculate_distance(long1, lat1, long2, lat2)

    }
}
```

most latest distance covered by the vessel of interest. The method implemented in R programming language to calculate the distance travelled by a vessel form the given longitude and latitude is given in Table 3. The second objective of the prepossessing step is to associate shipment data and vessel data that agree on the common value of vessel name and project out the attribute such as POL, POD, carrier and ATA from the shipment date. The sample of processed data is shown in Table 4.

3.2 Data Analysis

We started the data exploration and analysis activity with the intuitive idea that expected time of arrival of a shipment depends on the distance travelled by a shipment so far, port-of-landing, port-of-delivery, and carrier who owns the vessel. To validate our intuition, we checked the strength of correlation between input variables

Table 4 A sample of Data after pre-processing

POL	POD	Carrier	L1	L2	ATA	Predicted ETA
1	2	1	28.39	34.07	20-05-2016	20-05-2016
2	1	2	7.46	8.95	30-06-2016	01-07-2016
3	2	1	8.10	9.72	01-06-2017	30-05-2017
1	3	2	17.04	20.45	12-02-2018	13-02-2018
2	1	2	15.62	18.74	03-03-2017	03-03-2017

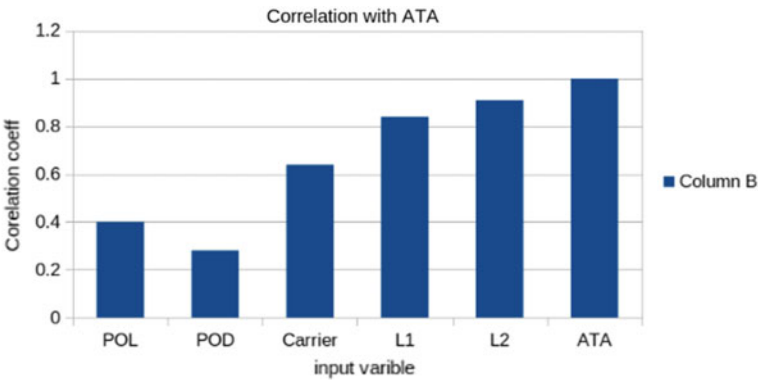


Fig. 2 Correlation of input variables with ATA

and the output variable, i.e. ATA. In our case, the input variables are POL, POD, Carrier, $L1$ and $L2$. As defined in the previous section, $L1$ and $L2$ are the distances travelled by vessel as reported through two latest messages. Figure 2 shows the correlation between input variables and ATA. The highest correlation is observed between L2 (0.91)and ATA, while the lowest correlation is observed between POD(0.28) and ATA. The other input variables, i.e. L1 (0.84), carrier (0.64), and POL(0.4) also demonstrate a positive correlation with ATA. Hence, we selected all five input variables to build the multiple linear regression model..

3.3 Multiple Regression Model

The multiple regression analysis is predication technique that identifies a relationship among input and output variables. The general format for the multiple regression technique is:

$$y = \beta_0 + \beta_1 x X_1 + \beta_2 x X_2 + \cdots + \beta_n x X_n$$

where y: is output variable, in our case it is ATA

Table 5 Regression coefficients

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.1823593	0.030148	31.103	<2e−16	***
POL	0.0035776	0.0006051	5.913	3.39e−09	***
POD	0.0012442	0.0002230	5.584	2.36e−08	***
Carrier	0.0076308	0.0013385	5.701	1.20e−08	***
L1	0.0676596	0.0044488	15.209	<2e−16	***
L2	1.1038425	0.0040762	270.802	<2e−16	***
Signif. codes	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				
Residual standard error	3.69 on 53,012 degrees of freedom				
Multiple R squared	0.9324	Adjusted R squared		0.9323	
F-statistic	1.461e+05 on 5 and 53,012 DF		*p*-value	<2.2e−16	

$X_1, X_1, \ldots, X_n$ are input variable, in our case POL, POD, carrier, L1 and L2
$\beta_0, \beta_1, \ldots, \beta_n$ are regression coefficients.

The purpose of model building is to find the values of $\beta_0, \beta_1, \ldots, \beta_n$ from the training data sets.

We have used the method ‘*lm*’ from the R programming language to build the multiple regression model. The model parameters obtained when the model is trained on 53,018 data items are shown in Table 5.

3.4 Model Evaluation

We used the summary information shown in Table 5 and obtained through $summary()$ command in R. This information tells us how the model fits our data and how dependent and independent variables are related to each other. The most important information is in the last two columns of the Table, i.e. *p-value* and *significance level*. A lower p-value indicates that independent variables are statistically significant. Generally, the threshold used for p-value is 5%. In our case, the p-value is below the threshold for all the independent variables, so they are statistically significant. The Multiple R-squared value is a measure of how well our model as a whole performs or in other words, how t strongly the dependent variables and independent variables are related. The closer the value to 1.0, the better is the model. Since the R-Squared value is 0.9324, we know that model explains nearly 0.93% of the variations in the dependent variables.

4 Conclusion

Accurately predicting ETA of a shipment is a challenging and significant activity for the overall functioning of logistics and supply chain management organizations. This paper describes a novel approach of combining data external to an organization and internal organization-specific data to predict the expected arrival time of a shipment. The paper describes the approach based on Multiple Linear Regression and observes that port of loading, the port of delivery, intermediate distance travelled by vessel, and vessel operator are statistically significant parameters to predict ETA of a shipment. The model reported in this paper predicts the ETA with an acceptable degree of error. It will be interesting to explore the performance of other machine learning techniques such as Support Vector Machines and Neural Network for the task of predicting ETA.

Acknowledgements The authors thank Murat Timucin and Swapnil Bhate for valuable discussion on Machine Learning, Predictive Analysis and Marine Transportation in general. ATA Freight Line Pvt. Ltd. supports this research. The authors would like to thank the ATA Freight Pvt. Ltd. Pune for providing access to AIS data. The article reveals the authors' views solely, and neither the ATA Freight Line Private Limited nor Dr. Babasaheb Ambedkar Technological University Lonere is responsible for the expressed views.

References

1. Bole AG, Wall AD, Norris A (2013) Radar and ARPA manual: radar, AIS and target tracking for marine radar users. Butterworth-Heinemann
2. Cavinato JL (2004) Supply chain logistics risks: from the back room to the board room. Int J Phys Distrib Logist Manage
3. Edwards AL (1985) Multiple regression and the analysis of variance and covariance. WH Freeman/Times Books/Henry Holt & Co
4. Karbassi A, Barth M (2003) Vehicle route prediction and time of arrival estimation techniques for improved transportation system management. In: IEEE IV2003 intelligent vehicles symposium. proceedings (Cat. No. 03TH8683). IEEE, pp 511–516
5. Sun D, Luo H, Fu L, Liu W, Liao X, Zhao M (2007) Predicting bus arrival time on the basis of global positioning system data. Transp Res Rec 2034(1):62–72
6. Wall Z, Dailey D (1999) An algorithm for predicting the arrival time of mass transit vehicles using automatic vehicle location data. In: 78th annual meeting of the transportation research board. Citeseer, pp 1–11
7. Wattx (2020) Predicting ETA's for maritime logistics (Online). https://wattx.io/portfolio/container-shipping.html. Accessed 15 Feb 2020
8. Yang D, Wu L, Wang S, Jia H, Li KX (2019) How big data enriches maritime research—a critical review of automatic identification system (AIS) data applications. Transp Rev 39(6):755–773

Learning via Long Short-Term Memory (LSTM) Network for Predicting Strains in Railway Bridge Members Under Train Induced Vibration

Amartya Dutta and Kamaljyoti Nath

Abstract Bridge health monitoring using machine learning tools has become an efficient and cost-effective approach in recent times. In the present study, strains in railway bridge member, available from a previous study conducted by IIT Guwahati has been utilized. These strain data were collected from an existing bridge while trains were passing over the bridge. LSTM is used to train the network and to predict strains in different members of the railway bridge. Actual field data has been used for the purpose of predicting strain in different members using strain data from a single member, yet it has been observed that they are quite agreeable to those of ground truth values. This is in-spite of the fact that a lot of noise existed in the data, thus showing the efficacy of LSTM in training and predicting even from noisy field data. This may easily open up the possibility of collecting data from the bridge with a much lesser number of sensors and predicting the strain data in other members through LSTM network.

Keywords LSTM · Railway bridge · Strain · Prediction · Health monitoring

1 Introduction

Strains in critical members of railway bridges under the action of moving train load are regularly monitored to understand the health of the structure. The strain in a bridge member can be a direct indicative whether the bridge is undergoing any possible degradation. The strains in bridge members are also monitored to appreciate the effect of any increase in axle load of the moving vehicle. However, any standard railway bridge comprises of a large number of members and a good number of these members are instrumented to acquire strain data while a train passes over the bridge. It is thus easily appreciated that the entire process of fixing strain gauges

A. Dutta (✉)
Indian Institute of Information Technology Guwahati, Guwahati, India

K. Nath
Indian Institute of Technology Guwahati, Guwahati, India
e-mail: kamaljyoti@iitg.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_31

and collection of data is expensive as well as time consuming. Thus, an alternative strategy is considered to address this issue of strain data collection, which is very intricately associated with structural health monitoring.

Monitoring of bridge using machine learning has become very popular among researchers. Shu et al. [1] implemented a damage detection algorithm for railway bridge based on back propagation Artificial Neural Network (ANN) using the statistical properties of the dynamic responses of the structure as input for the ANN. Chalouhi et al. [2] presented a method that uses machine learning to detect and localize damage in railway bridges. They applied the strategy to a historical bridge and validated the proposed algorithm. The proposed method can be used to detect inconsistent responses that can be labelled as possible damage. Neves et al. [3] presented a model-free damage detection approach based on machine learning techniques. Artificial neural networks are trained with an unsupervised learning approach with input data in the form of accelerations gathered from the bridge. Malekjafarian et al. [4] proposed a new two-stage machine learning approach for bridge damage detection using the responses measured from sensors placed on a passing vehicle. Dutta [5] applied a density-based clustering technique on railway bridge strain data acquired from field to identify the data density and noisy data elements, which can be subsequently used in decision making for structural health monitoring. In recent years, LSTM has demonstrated noteworthy performance on a good number of real-world applications such as machine translation [6], speech recognition [7] and video classification [8]. There has also been an increasing attention towards using LSTMs for time series prediction such as air pollution forecasting [9], traffic flow prediction etc. [10].

In the present study, learning via LSTM is proposed for predicting strains in railway bridge members utilizing already available strain time history of the same bridge for a good number of earlier conducted tests. The study presented here uses real data from the bridge and hence the study is more challenging due to the presence of noise and other associated uncertainty in data collection. Highly satisfactory agreement is observed between the predicted strain time history with ground truth for a good number of cases and hence provides an ideal opportunity to explore the application of such machine tools.

2 Description of Test Bridge and Data

The considered bridge is a twenty span truss bridge between Jalpaiguri Road and New Domohoni station under Alipurduar Division in West Bengal, India. The spans of the bridge are all equal and is 45.72 m carrying a single broad gauge track. The steel superstructures rest on concrete piers and abutments at two ends. One of the spans is taken for study and is shown in Fig. 1. A train with known axle loads (named here as Test train) is allowed to travel at different speeds. Strains of some critical members were measured. Some of the typical sensors attached to the bridge are shown also in Fig. 2. Further, data were also collected for passenger trains, whose

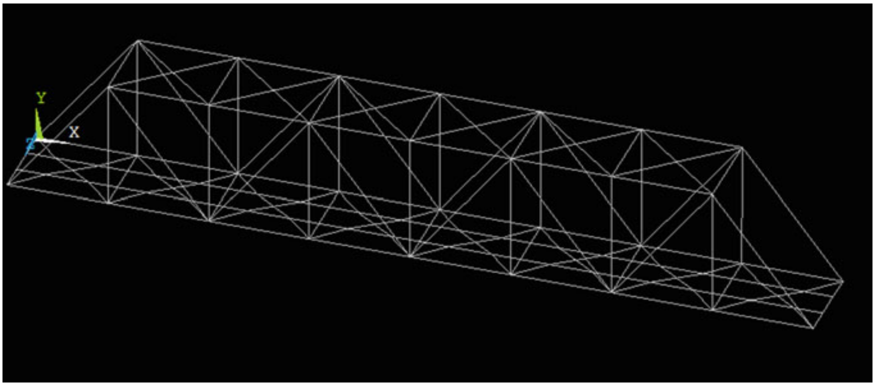


Fig. 1 One typical span of the railway truss bridge

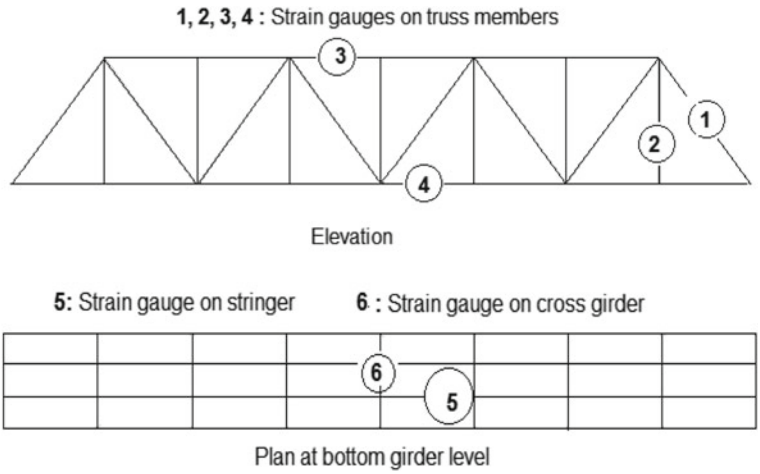
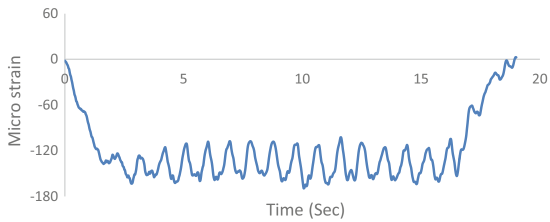


Fig. 2 A typical span of the bridge showing the sensor locations

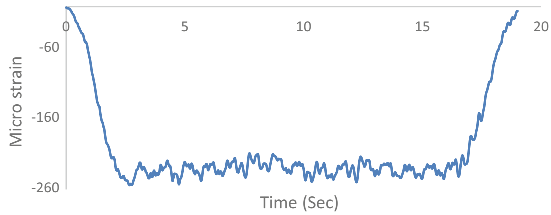
axle loads are not exactly known. The data collection continued over four years and hence significant amount of measured data are available.

Typical plots of measured strains are shown in Fig. 3. These are for different speed of test train as well as for service train. The data were collected at the sampling rate of 0.025 s for train moving at different speeds. Thus, it may be observed that the amount of time series data in Fig. 3 are different as the train was made to move at different speeds. In the case of service train, the axle load for engine is much higher than that of the rest, which can also be observed from the initial higher values of strain. It may be appreciated that the strain in a particular member is correlated to strain in other members of the bridge on the same span. However, the measurement of all these strains at different locations are quite involved and expensive. Thus, the

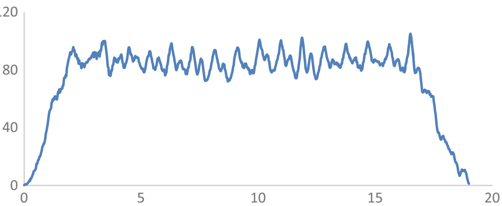
Fig. 3 Typical plots of strain time histories corresponding to different sensor locations, speed and train



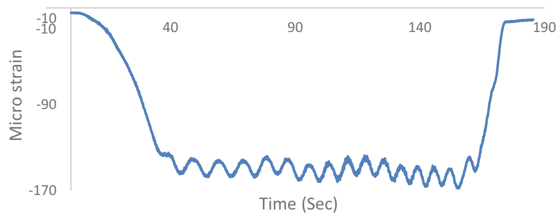
(a) Location 1 (Test train at 50 kmph)



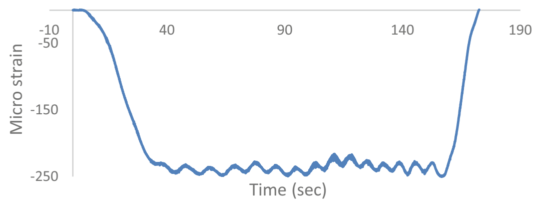
(b) Location 3 (Test train 50 kmph)



(c) Location 4 (Test train at 50 kmph)

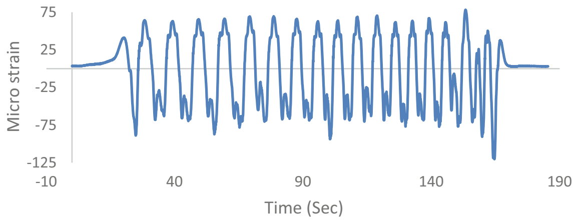


(d) Location 1 (Test train at 5kmph)

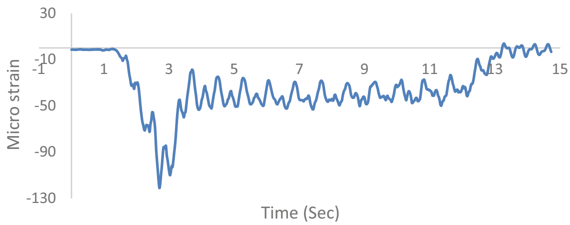


(e) Location 3 (Test train at 5kmph)

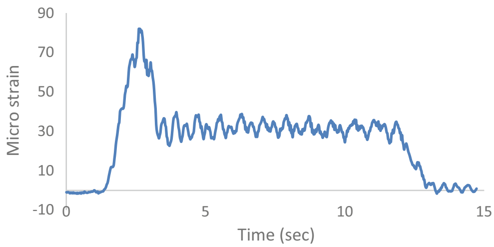
Fig. 3 (continued)



(f) Location 5 (Test train at 5kmph)



(g) Location 1 (Passenger train at 5kmph)



(h) Location 4 (Passenger train at 5kmph)

objective of the present study is to predict the strain in some of the critical bridge members using the strain data of any one of the members of the bridge.

3 LSTM Network

RNNs are a generalization of feedforward neural networks, one that allows the neural network also to learn by remembering the past information. However, in case of long-term dependencies, RNNs face vanishing and exploding gradient problems. That is why a gated variant of the RNN, termed as LSTM has been used for the work in this paper. The gates in these units are used to control the flow of information that passes through the state of the cell. The chosen type of gated cell (LSTM) was introduced two decades ago [11] and has now gained popularity in the context of language modeling. However, the work done in the present paper attempts to exploit the advantage of LSTMs when it comes to time series forecasting. As a result, an attempt has been made to predict the various forms of strain the bridge members would experience using the time series data.

The formulation of LSTM cells is as defined by [11, 12]. Assuming i_t, f_t, o_t, c_t and h_t to indicate the values of the input gate, forget gate, output gate, memory cell and hidden state at time t in the sequence respectively and x_t be the input of the system at time t, the architecture of the LSTM cell can be defined as follows

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{1}$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \tag{2}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{3}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{4}$$

$$h_t = o_t \odot \tanh(c_t) \tag{5}$$

4 Time Series Forecasting Using LSTM

The strain data collected from different bridge members contains multiple features that vary over time. The work done in this paper uses LSTM to learn from the time series data. Since the aim is, given a time varying feature as the input to the model, the strains in other members are to be predicted, the time series data is converted such that it resembles supervised learning. Therefore, the data is transformed into a sliding window format such that every sample of the time series data is of the form

Input: $[x_i; x_{i+1}; \ldots x_{i+T-1}]$
Output: y_{i+T-1}.

where T is the sequence length, x_i is the input feature and y_i is the feature to be predicted.

Following this, the time series model is used for the desired forecasting task. The strain corresponding to a location (taken as 1 in the present study) acts as our input, while the targets are strains at different locations considered one at a time that vary according to the cases as explained in the next section.

5 Prediction of Strain Using LSTM

A number of cases of prediction of strains are studied to understand the LSTM and its efficiency for each of the cases. The training is time consuming as different associated parameters are to be tuned to achieve the best possible prediction using LSTM. Parameters adopted for different cases are mentioned below the plots of each of the cases. In order to evaluate the accuracy of prediction, Root Mean Square Error (RMSE) is calculated between the predicted and target strain time history. While a smaller value of RMSE is indicative of a better performance in prediction, the magnitude of RMSE alone does not very clearly convey the extent of perfection achieved in such prediction. In view of this, a ratio is calculated for L_2 norm of error vector between Target and Predicted strain time history to L_2 norm of Target strain time history. If this ratio is deducted from unity and expressed in percentage, provides a measure of accuracy, which is physically interpretable. In the present study, both RMSE and % accuracy are evaluated for all the cases as detailed below.

5.1 Prediction Case-1

The first problem is the prediction of strain time history corresponding to location 3 by training a LSTM model with strain time history corresponding to location 1. The training and predictions are carried out with strain measured from the bridge due to the passage of test train at 60 kmph. The predicted strain time history and the target strain time history measured in field are shown in Fig. 4. Very good agreement is observed between the predicted strain time history and the ground truth, which is the available measured strain time history corresponding to the same location 3 of the existing bridge.

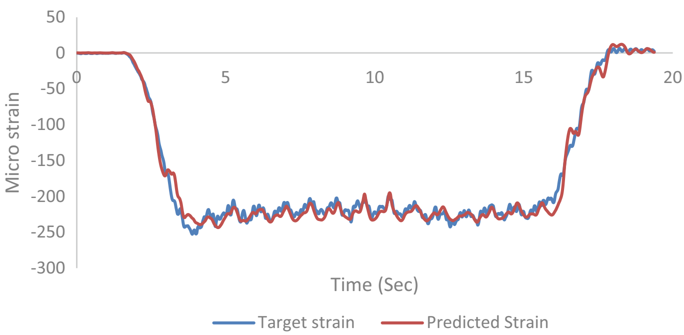


Fig. 4 Target and predicted strain time history at location 3 for test train at 50 kmph (20 LSTM neurons, 1 hidden layer with 30 neurons, window size = 50)

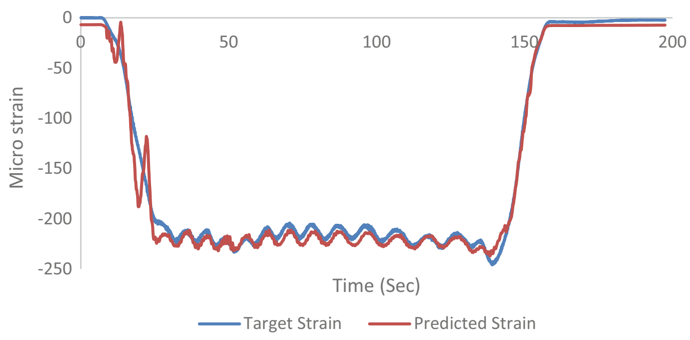


Fig. 5 Target and predicted strain time history at location 3 for test train at 5 kmph (10 LSTM neurons, 1 hidden layer with 30 neurons, window size = 50)

5.2 *Prediction Case-2*

The second problem is similar to the first case, where the training and prediction corresponds to the test train moving at a velocity of 5 kmph. The prediction of strain time history corresponding to location 3 is done by training a LSTM model with strain time history of location 1. Since the train is moving at a much lesser speed, the data size is quite different in this case as may be seen from Fig. 2d, e. The predicted strain time history and the target strain time history (ground truth) are shown in Fig. 5. Very good agreements are observed in this case as well.

5.3 *Prediction Case-3*

As may be seen from Fig. 3c, f, the patterns of strain time history corresponding to locations 4 and 5 are quite different than what have been observed corresponding to locations 1 and 3. In order to appreciate the applicability of LSTM, it is tried to predict strain time history corresponding to location 4 and 5 using the strain time history data corresponding to location 1. These are shown in Figs. 6 and 7 respectively. In this cases as well, very good agreements are observed between predicted strain time history and ground truth for the respective cases.

5.4 *Prediction Case-4*

In all the above-mentioned three cases involving test trains, wheel loads of engine and wagons are not much different as may be evident from Fig. 3g, h. Next, it is tried to utilize strain time history data from bridge members due to passenger train induced vibration, where engine wheel loads are much higher than those of wagons.

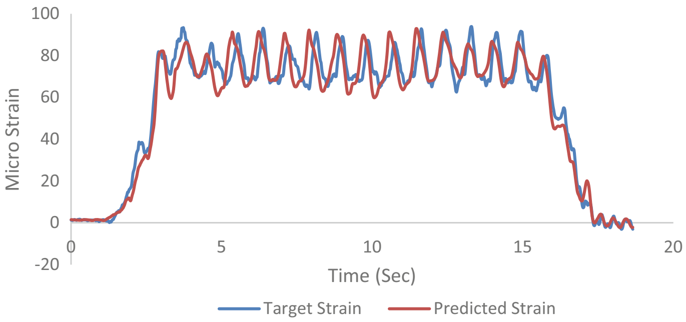


Fig. 6 Target and predicted strain time history at location 4 for test train at 50 kmph (20 LSTM neurons, 1 hidden layer with 50 neurons, window size = 50)

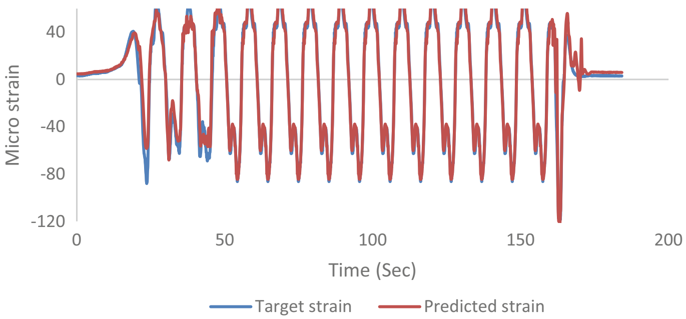


Fig. 7 Target and predicted strain time history at location 5 for test train at 5 kmph (Stacked LSTM with 80 neurons followed by another 60 neurons, window size = 50)

The pattern of strain time history corresponding to locations 1 and 4 (Fig. 3g, h) are thus quite different. The strain time history corresponding to location 4 is predicted using the strain time history corresponding to location 1 and very good agreement is observed between predicted strain time history and ground truth as shown in Fig. 8.

Thus, in all the cases, strains in member "1" have only been used for training and strain in many other members are predicted. Table 1 shows the values of RMSE and % Accuracy achieved in prediction using LSTM. It may be noted that all the cases studied here are done using field measured data, which is always mixed with lots of noise and has built in uncertainty. However, the level of accuracy attained using LSTM in prediction based on field measured data is noteworthy. This has huge practical implications. If a network can be trained with already existing data representing strain time histories in important bridge members, the strain time history obtained subsequently from one identified bridge member can be cost-effectively utilized to extract the state of strain in other important bridge members. These strain data can be effectively used to monitor the health of the bridge in future using measured data from only one strain gauge fitted in member "1" for example.

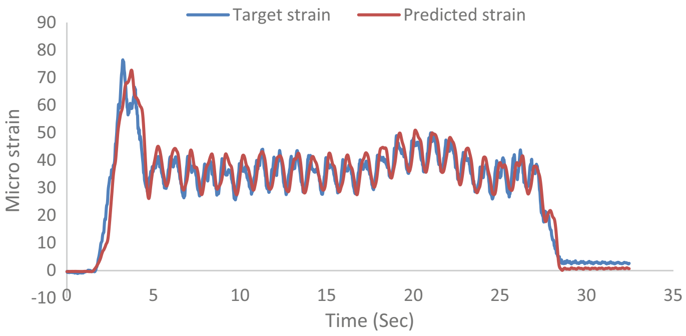


Fig. 8 Target and predicted strain time history at location 4 for passenger train (Stacked LSTM with 80 neurons followed by 60 neurons, window size = 60)

Table 1 Details of accuracy in prediction using LSTM

Case	RMSE	% Accuracy
Case 1 (Fig. 4)	8.929	95.19
Case 2 (Fig. 5)	9.361	94.66
Case 3 (Fig. 6)	7.326	88.71
Case 3 (Fig. 7)	7.027	84.66
Case 4 (Fig. 8)	4.451	86.96

6 Conclusion

A successful attempt is made for predicting strains in bridge members using LSTM, utilizing actual data from the bridge. The actual data were collected while trains of different axle loads were passing over the bridge at different speeds. While the collection of such data are very important to monitor the health of the bridge, the process is expensive and much more time consuming than training a LSTM model. Thus, introduction of LSTM and successful prediction of strain will certainly open up opportunities for engineers to carry out similar exercise for structural health monitoring in a more efficient and cost-effective manner.

Acknowledgements The authors gratefully acknowledges the support of Department of Civil Engineering, IIT Guwahati for providing access to the test data of Bridge no. 40 under Alipurduar Division over river Tista. The bridge data are corresponding to project no. CE/C/AD/225 entitled "Instrumentation of bridges for running of CC+6+2 Tonne loaded BOXN Wagons train on NF Railway" in Civil Engg. Deptt., IIT Guwahati.

References

1. Shu J, Zhang Z, Gonzalez I, Karoumi R (2013) The application of a damage detection method using artificial neural network and train-induced vibrations on a simplified railway bridge model. Eng Struct 52:408–421
2. Chalouhi EK, Gonzalez I, Gentile C, Karoumi R (2017) Damage detection in railway bridges using machine learning: application to a historic structure. Procedia Eng 199:1931–1936
3. Neves AC, González I, Leander J, Karoumi R (2017) Structural health monitoring of bridges: a model-free ANN-based approach to damage detection. J Civil Struct Health Monit 7:689–702
4. Malekjafarian A, Golpayegani F, Moloney C, Clarke S (2019) A machine learning approach to bridge-damage detection using responses measured on a passing vehicle. Sensors 19:1–19. http://doi.org/10.3390/s19184035
5. Dutta A (2015) Bridge strain data analysis using density based clustering technique. In: Proceedings of the 2015 IEEE international conference on research in computational intelligence and communication networks (ICRCICN), Kolkata, India
6. Sutskever I, Vinyals O, Le QV (2014) Sequence to sequence learning with neural networks. In: Advances in neural information processing systems, pp 3104–3112
7. Graves A, Mohamed A-R, Hinton G (2013) Speech recognition with deep recurrent neural networks. In: 2013 IEEE international conference on acoustics, speech and signal processing, pp 6645–6649. http://doi.org/10.1109/ICASSP.2013.6638947
8. Ng JY-H, Hausknecht M, Vijayanarasimhan S, Vinyals O, Monga R, Toderici G (2015) Beyond short snippets: deep networks for video classification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4694–4702
9. Freeman BS, Taylor G, Gharabaghi B, Thé J (2018) Forecasting air quality time series using deep learning. J Air Waste Manag Assoc 68:866–886. https://doi.org/10.1080/10962247.2018.1459956
10. Tian Y, Pan L (2015) Predicting short-term traffic flow by long short term memory recurrent neural network. In: 2015 IEEE international conference on smart city/socialcom/sustaincom (SmartCity), pp 153–158. http://doi.org/10.1109/SmartCity.2015.63
11. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780
12. Géron A (2017) Hands-on machine learning with scikit-learn and tensorflow: concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc., USA

Smart Transportation Using Fog Computing

Rizwana Shaikh and Masooda Modak

Abstract In the era of all smart applications with connected things, there is a need to tackle the issues of data preprocessing, increased latency in data and the overhead in traffic. Intelligent transportation systems (ITS) is one such application. Fog computing can be used as a support or replacement of cloud computing to solve the issue of data latency and traffic overhead. The paper proposes Fog computing based smart transportation application prototype that can be achieved by providing inter-fog communication. The communication over a distributed network is presented to cover a large geographical area. The inter-fog communication in the smart transportation could ensure the edge nodes managing the traffic lights and the street light control to be locally handled by the respective fog nodes with real time traffic reports. It can also provide suggestions for alternate routes at the time of congestion.

Keywords Smart transportation · Fog computing · Cloud computing · Edge nodes

1 Introduction

Fog computing is a decentralized computing structure emerged from the base of cloud computing. Cloud Computing is the distributed system architecture that provides data, applications and other services from distant location. Various delivery models of cloud computing can be purchased from a cloud service provider on demand. Cloud service provider also provides the maintenance and upgrades of the services and the infrastructure needed by the cloud consumer. Cloud computing process data and provides services based on service request. The services and demand is increasing exponentially. The computing preformed at the cloud server is increasing, and hence network traffic.

The main goal of fog computing is to bring basic computing and analytical services to the network level. The processing and decision can be taken at the level of the node network instead of moving to distant cloud. It improves the performance by placing the computing and communication resources closer to network. Network data and

R. Shaikh (✉) · M. Modak
SIES Graduate School of Technolgy, Navi Mumbai, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_32

the resources are more needed at this level where they capture the data and needs processing. It reduces the distance that data needs to be transferred to the cloud thus improving the network bandwidth that required for data communication to the cloud. This will improve the overall efficiency and performance of the network traffic.

Fog computing has emerged as a promising technology that can bring the cloud applications closer to the physical IoT devices at the network edge. Fog computing overcome the cloud computing limitations and bottlenecks by providing a decentralized architecture. It is the extension service provided to support the cloud computing. It works on behalf of cloud computing by collaborating with one or more edge nodes. Edge nodes are the devices providing the substantial amount of computing infrastructure that is controlled locally.

Smart transportation is developed on the base of smart infrastructure that includes automated traffic signals, tolls and fare collection. Smart services offer different benefits, from smart parking and vehicle locating systems, to route diversion alerts. A central command controller can tie together the smart transportation ecosystem, with real-time and updated data, handling passenger information, traffic signals and incident management.

Fog computing has a lots of benefits. As the computing resources that are needed are closer to the data that is being created, the amount of data transfer and communication bandwidth requirements are eliminated. Only small amount of data is transferred to cloud. It improves the network latency and hence the traffic. Fog computing provides the intermediate channel between the cloud and the IOT devices. Maximum processing is done at the edge node and very little data is pushed to the cloud.

Here we are attempting to use fog computing architecture to provide solution for smart transportation.

2 Literature Survey

Smart transportation is the area of research in vehicular network. Various solutions are proposed by the researcher in this field. An intelligent hybrid system is proposed by the author in [1]. It uses various technologies to provide real time traffic control using fog computing. The TIMON project initiative [2] presents information in real time and the services based on cloud through an application on mobile or web-based platform and a mobile application.

Industry strengthens the users by allowing them to exploit the data analytics capacity of cloud computing that can be used for the purpose of forecasting, predictions, analysis, etc. They also can help in reducing the downtime during routine maintenance, storage on centralized servers and management on remote servers. But nowadays the cloud computing technology is facing issues in cloud support based on subscription or large amount of data to be pushed on cloud. This is due to the high increase in IoT based applications and missing proper network bandwidth usage and allocation. This also leads to high latency because of network connection speed and

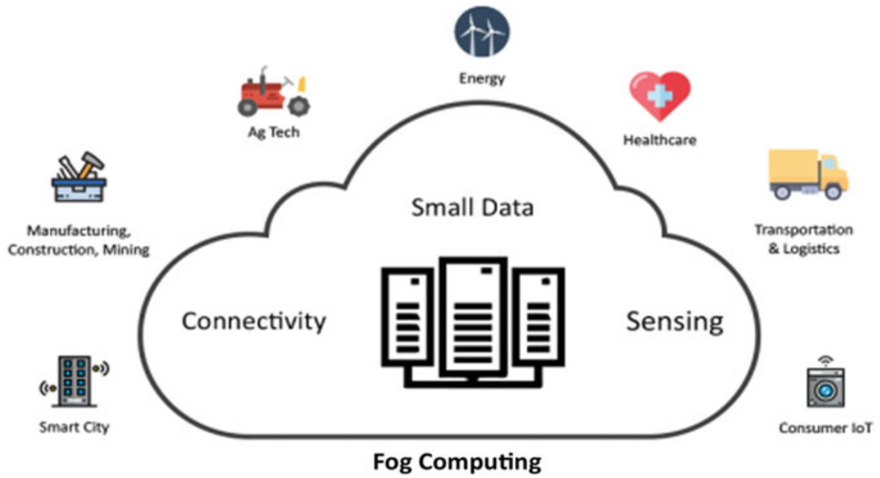


Fig. 1 Fog computing

network bandwidth overutilization. These shortcoming of cloud leads the foundation for fog computing [1–8].

2.1 Fog Computing

Fog computing is also known as fogging or fog networking. It is a computing infrastructure in decentralized environment in which the actual data or the computing power or the storage capacity or sometimes even the applications are spread in a locations between the cloud and the source of data. Fog computing is an extension of cloud computing which presents the benefits of cloud computing near the location where the data is created.

The aim of fog computing is to decrease the amount of transportation of data to be done on cloud for analysis, storage and processing. The takeaway is improved efficiency, even though it may be used for security reasons as shown in Fig. 1.

2.2 IOT-Internet of Things

Connecting any device to the internet and to other connected devices is the main idea of IoT. The IoT is a huge network of people and things that are connected to each other. It means that these all connected things share and collect the data about how and where they are used, the environment in which they are used. The various devices which may have a sensor can connect to the IoT network. The IoT network

Fig. 2 IOT in smart transportation

then has the capability of integrating the data from all connected devices. Not only that, various analytics techniques can be applied on these collected and integrated data to serve the various applications (Fig. 2).

There are lots of applications of IoT in transportation eg Traffic analysis, any incident reporting, using fog networking to send the traffic data for broadcasting etc.

2.3 Fog Computing in IOT

Despite the broad utilization of cloud computing, some applications and services still cannot benefit from this popular computing paradigm due to inherent problems of cloud computing such as unacceptable latency, lack of mobility support and location-awareness. IoT has established with invent of wearable computing and smart devices like smart city with vehicles connected using huge wireless sensor network. Fog computing delivers computing service directly at the network edge. They can deploy new services and applications in the near future. Improving the efficiency and decreasing the data transported to the cloud for forecasting, predictions, analysis and storage is the ultimate aim of fog computing to gain better efficiency. Fog computing applications include smart city, smart buildings, vehicle networks and software-defined networks. So fog computing is a layer added in between to improve the efficiency and latency problem (Fig. 3).

3 Proposed System

As the advancements in the IOT field is increasing, various IOT-enabled devices are getting the popularity for new applications. Smart transportation is the emerging area where the IOT based devices are used for communicating and collaborating the traffic information. As the cloud computing has the network and latency related issues, so a solution based on fog computing for IOT enabled traffic control is proposed. Here we

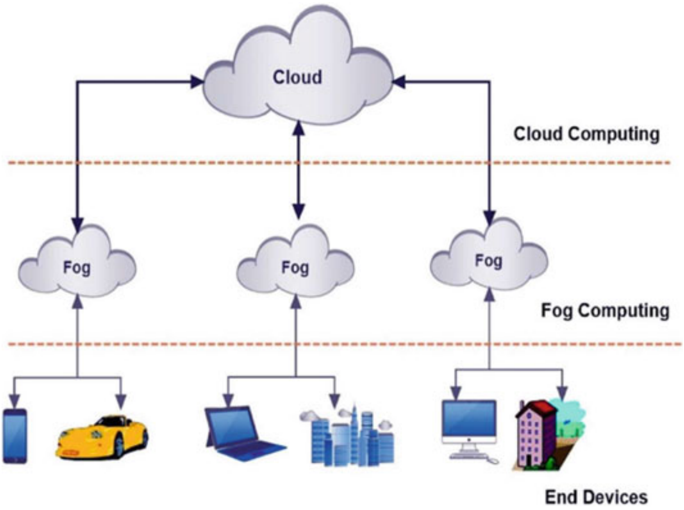


Fig. 3 IOT in fog and cloud computing

are providing a simple solution for real time traffic monitoring using fog computing and IOT.

Fog computing has been used to assist smart transportation applications. It is achieved by performing the computation and the analysis process at the edge node. Multiples edge nodes located at distributed geographic locations are processing the data received through various IOT devices. These devices collect the traffic data and process them locally to give the data about the traffic signal. Other accidental and whether data also are processed locally to control the traffic. This locally processed data are given to multiple nodes and finally to the cloud for the storage. The local communication reduces the overhead done at the cloud side and hence improves the efficiency and latency.

4 Implementation

Sensors are placed in various devices at multiple locations. These sensors will be communicating with the regional unit with the required data. These data will be processed by the microprocessor or microcontroller. Raspberry pi is used as a computing node to give the data the fog node or edge node. The unit can be acting as a processing element for the data. Likewise many such units can be connected to one device making as mini-cloud and the processing element called router or smart router where the processing itself can take place instead of transmitting to cloud. These local processing decisions can be communicated to other router via actual cloud. Figure below shows the infrastructure required for fog computing and its working (Fig. 4).

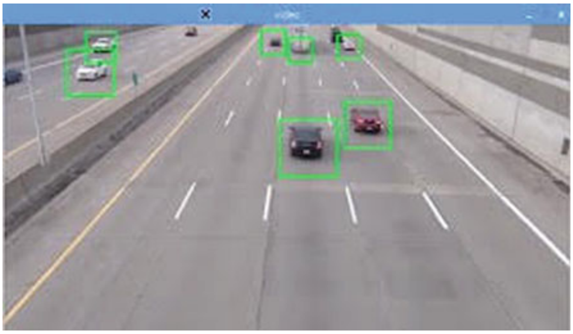


Fig. 4 Counting of vehicles

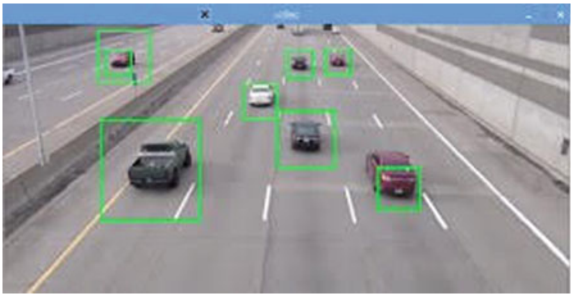


Fig. 5 Analyzing traffic

The implementation starts with the road traffic count based on computer vision and with deep learning algorithms. We have used only Python and OpenCV with the idea of motion detection with help of background subtraction algorithm. Object detection during motion is captured using the video (Figs. 5 and 6).

Once the count of vehicles is evaluated, the multiple lanes connected to one fog node will be getting the count value. The fog node accordingly gets the traffic at each lane and determine the signal time of that lane.

5 Conclusion

The use of Fog Computing is not the replacement but supports the cloud computing to make the reduced latency and response time. The processing node at the edge of fog computing or edge nodes perform the operation of processing that is critical in traffic monitoring and accidental Analysis. Here the processing information is passed to edge nodes and decision is taking to control the traffic and helps in smart transportation. In future the idea can be communicated to the government authorities to perform the full implementation of this prototype.

Fig. 6 Snapshot of the code

```
#import libraries of python opencv
import cv2
import numpy as np

#create VideoCapture object and read from video file
cap = cv2.VideoCapture('cars.mp4')
#use trained cars XML classifiers
car_cascade = cv2.CascadeClassifier('cars.xml')

#read until video is completed
while True:
    #capture frame by frame
    ret, frame = cap.read()
    #convert video into gray scale of each frames
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)

    #detect cars in the video
    cars = car_cascade.detectMultiScale(gray, 1.1, 3)

    #to draw arectangle in each cars
    for (x,y,w,h) in cars:
        cv2.rectangle(frame,(x,y),(x+w,y+h),(0,255,0),2)

    #display the resulting frame
    cv2.imshow('video', frame)
    #press Q on keyboard to exit
    if cv2.waitKey(25) & 0xFF == ord('q'):
        break
#release the videocapture object
cap.release()
#close all the frames
cv2.destroyAllWindows()
```

References

1. Elhoseny M et al (2018) A hybrid model of internet of things and cloud computing to manage big data in health services applications. Future Gener Comput Syst 86:1383–1394
2. Hu P et al (2017) Survey on fog computing: architecture, key technologies, applications and open issues. J Netw Comput Appl 98:27–42
3. Bonomi F et al (2012) Fog computing and its role in the internet of things. In: Proceedings of the first edition of the MCC workshop on mobile cloud computing. ACM
4. Sanjeev S, Thusu S (2017) A survey of fog computing and its applications. Int J Adv Res Ideas Innov Technol 3(2)
5. Bachmann K (2017) Design and implementation of a fog computing framework. Diss. master's thesis, Vienna University of Technology (TU Wien), Vienna, Austria
6. Elmroth E et al (2017) Connecting fog and cloud computing. IEEE Cloud Comput 4(2):22–25
7. Basu S et al (2018) An intelligent/cognitive model of task scheduling for IoT applications in cloud computing environment. Future Gener Comput Syst 88:254–261
8. Tao F et al (2014) IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing. IEEE Trans Ind Inf 10(2):1547–1557

Detection of Learning Disability: A Survey

Masooda Modak, Prachi Gharpure, and Sasikumar

Abstract Learning Disabilities (LD) are generally disabilities in which individuals with average or above average intelligence are affected. The ability to learn is affected and this may be true for one's lifetime. Some children may have single learning disability or some may have many overlapping learning disabilities. Learning disability may include disabilities in various areas related to reading, language and mathematics. With the right help at right time, right assessment and remediation, children with LD can and do learn successfully and become winners in the society later. It is a great challenge to identify and diagnose and assist children with learning disability. Based on these facts, it is suggested that the early diagnosis of learning disability in children is essentially important to identify and suggest remedial solutions to the parents. Many times the student himself is not aware that a particular symptom of LD is applicable to him. Sometimes the students are reluctant or hesitant to accept that they possess one or more symptoms of LD. Many researchers have contributed to the detection of these learning Disabilities. This paper provides a survey of the various research work done with this focus. Once the disability is diagnosed the student will be able to learn according to his/her learning requirements/preferences that may lead to positive performance of the LD learner.

Keywords Learning disability · Detection of learning disability · Neural network · Machine learning

M. Modak (✉)
Research Scholar-S.P.I.T, Andheri, Assistant professor-SIES Graduate School of Technolgy, Navi Mumbai, India

P. Gharpure
SVKM's NMIMS University, Indore Campus, Indore, India

Sasikumar
CDAC, Santacruz, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_33

1 Introduction

Learning Disabilities (LD) are generally disabilities in which individuals with average or above average intelligence are affected. The ability to learn is affected and this may be true for one's lifetime. Some children may have single learning disability or some may have many overlapping learning disabilities. The IQ of the learning disabled is usually high. Learning disability may include disabilities in various areas related to reading, language and mathematics. The problems associated with the LD have always been a matter of worry and concern for parents and the school authorities. With the right help at right time, right assessment and remediation, children with LD can and do learn successfully and become winners in the society later It is a great challenge to identify and diagnose and assist children with learning disability. Based on these facts, it is suggested that the early diagnosis of learning disability in children is essentially important to identify and suggest remedial solutions to the parents.

With proper recognition and support, the learning disabled can successfully complete the assigned task. Learning difficulties do not disappear with age or training. Inspite of the manifested learning difficulties, the LD students could develop compensatory strategies or preferences in learning to provide assistance while they are studying. However, despite their efforts, when compared to their peers, affected students still show significant difficulties in reading tasks. Surprisingly, not all students whose performance is affected by learning disability are diagnosed and/or assisted before starting their studies. Therefore, there are many students with learning difficulties who have not been diagnosed with LD by means of an official psycho assessment procedure. Consequently, a considerable number of students may be using e-learning platform without having expected learning skills. There are many e-learning sources for learning disabled. But they do not consider and respond to the individual learning characteristics. Every learning disabled may be suffering from a different learning disability. For example, one dyslexic student may be having problem of letter and number reversals while another dyslexic student struggles to decode written text. Many researchers have contributed to the detection of these learning Disabilities. This paper provides a survey of the various research work done with this focus.

2 Literature Review

Learning Disability (LD) according to the WHO's International Classification of Diseases the LD are known as "Specific developmental disorders of scholastic skills" [1–4]. They could be classified as Dyslexia, Dysgraphia, Dyscalculia, Dyspraxia.

A complete psychological assessment can be used to diagnose the presence of LD. A variety of standardized tests are given to thousands of people. The psychologists will analyze at how people think, remember, solve problems, understand and express information [5].

2.1 Detection of Learning Disability

There are various formal and informal tests available for the detection of LD. The formal detection method involves the child undergoing various tests such as intelligence tests, vocabulary test, achievement tests, etc. The findings of such tests are recognized by the schools and colleges to provide the special education to the LD child. But this requires many sessions with the recognized LD examiner. It is a time consuming process because of which parents are reluctant to go for such testing. The informal tests on the other hand are specifically for the early identification of LD, so that timely remediation is provided to the LD children.

2.1.1 Formal Method of Diagnosing LD

Several types of tests are needed to diagnose the presence of learning disability. The Individuals with Disabilities Education Act (IDEA) mandates that single test do not conclude the presence of LD.

1. Intelligence tests like Wechsler Preschool and Primary Scale of Intelligence (WIPPSI), Wechsler Intelligence Scale for Children (WISC), and the Wechsler Adult Intelligence Scale (WAIS).
2. Achievement tests like Woodcock-Johnson Tests of Achievement (WJ), the Wechsler Individual Achievement Test (WIAT), the Wide Range Achievement Test (WRAT), and the Kaufman Test of Educational Achievement (KTEA).
3. Visual-motor integration—The Bender Visual Motor Gestalt Test and the Developmental Test of Visual Motor Integration
4. Language testing—The Clinical Evaluation of Language Fundamentals (CELF), Goldman Fristoe Test of Articulation, the Test of Language Development.

2.1.2 Various Computational Techniques for Detection of Different Types of LD (Informal Testing)

They can be classified into approaches Eye movements, questionnaires, Local characteristics checklists (LCC),Test performance, EEG, games, Table 1 summarizes the various research carried on with respect to the detection of LD. Each research work is focusing on some symptoms or attributes of learning disability.

Thus, it can be summarized that methods to detect the learning Disability can be classified into five categories,

Table 1 Various research work carried on informal detection of learning disability

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
1	**Eye movement**	**Dyslexia detection from eye movements using artificial neural networks**	Eye movements of 76 school children were measured using VOG technique	Feed forward Back propogation algorithm **The classifier gave 90% correct identification for presence of dyslexia and 90.5% correct identification for absence of dyslexia**	The author implemented a neural network based classifier to detect dyslexia through eye movements using **videooculographic (VOG). Further feature selection was applied to extract significant features** They found that the feature subset related to verbal tasks gave the best results supporting the expectation that dyslexia can be detected from the reading task The author has used static features in the presence of fixed light. The results would be misleading if the tests are conducted under moving light
2	**Questionnaires**	**Effects of feature selection on the identification of students with learning disabilities using ANN**	Data set containing 159 samples, correctly identifies 18 out of the 76 students with LD without any false positive case Also five-fold validation using the four feature sets that achieve Average CIR 70%	Four feature selection algorithms [Brute-force (BF), Greedy with forward selection (GF), Greedy with backward elimination (GB) and Genetic algorithm (GA)]	The same authors in the [6] improved the classification accuracy and time by applying feature selection on the 16 features in the previous work But the downside remained the same as in the previous work, i.e, the reluctance of the parents/teachers to accept the prediction due to the black box prediction employed

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
3	**Game**	**Early identification of dyslexia and the use of computer game-based practice to support reading acquisition**	**Two studies in the age group of 6 and 7 years consisting of 12 and 41 children**		A grapho game was developed by the authors to identify the children who could develop a risk of dyslexia The game intends to reveal the connection between graphemes and phonemes, i.e, letters and sounds Two studies were conducted with 41 and 12 children. the game gave inspiring results. Although these game were intended as preventive training they gave improvement in the child's grapheme-phoneme concepts and reading words
4	**EEG**	**Detecting complexity abnormalities in dyslexia measuring approximate entropy of electroencephalographic signals**	57 participants were included which detected 38 as dyslexic and 19 control between 2 and 13 years	An SVM classifier was employed based on Approximate entropy (ApEn) to distinguish between dyslexic and controlled groups The work achieved sensitivity of 89.47% and specificity of 57.89%	The authors used entropy of EEG signals, to detect abnormalities in dyslexia. They found that Wavelet Entropy could be used as quantified measure for the detection of dyslexic pattern of brain

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
5	**Checklist**	**Machine learning approach for prediction of learning disabilities in school-age children**	For evaluating the LD, the author have used a checklist comprising of some 16 frequent symptoms of LD which could be difficulty while reading, writing, doing calculations, difficulty in paying attention, getting distracted, etc	16 attributes with PCA reducing them to 7	The research work employs rough set approximation which does its work as a four step process on 513 samples with 16 attributes each. The first step involves decision table development, second step is the decision space approximation, i.e., rough classification, attribute reduction is the third step and finally the last step is the production of minimal set of attributes giving the same classification as before. Six rules that predict the learning disability are generated by the algorithm

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
6	**Test performance**	**Developing prognosis tools to identify learning difficulties in children using machine learning technologies**	134 children in 7–12 years old 94.64% classification accuracy is achieved	PCS, SOM, Bayesian classifier	The mental attributes profiling system was used to get a multimodal evaluation of learning potential of the students The large no, i.e, 226 variables obtained through a series of battery tests were replaced by the first 66 principal components using PCA, an then SOM algorithm is applied to identify clusters in which the population belongs

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
7	**Test performance**	**A neuro-fuzzy approach to diagnose and classify learning disability**	Dataset consists of 240 cases, out of which 160 are of normal children and 80 of LD children	Single-layer perceptron and learning vector quantization (a) PLEDDOR: Accuracy = 84%, Sensitivity = 80%, Specificity = 80% (b) DICQOLD: Accuracy = 91.8%, Sensitivity = 84.5%, Specificity = 84.5%	The authors have used the two models classify a child as learning disabled or non-learning disabled. Once diagnosed with learning disability, fuzzy based approach is used further to classify them into types of learning disability that is dyslexia, dysgraphia, and dyscalculia. The author has used two methods namely PLEDDOR using single layer perceptron and DICQOLD using learning vector quantization The author have generated a rule base for classifying the diagnosed LD But the dataset used is static and is purely based in the child's test performance in the subjects English comprehension, reading and mathematics

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
8	**Game**	**A serious game for predicting the risk of developmental dyslexia in pre-readers children**	Dataset comprises of 24 children from last year of kindergarten		The authors have focused on development of functions in a child related to the eye-hand coordination, visual search ability, rapid identification of visual and acoustic inputs and the capability to associate visual and auditory stimuli

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
9	**EEG**	**An SVM based algorithm for analysis and discrimination of dyslexic readers from regular readers using ERPs**	50 participants in the age group of 24–35 diagnosed 20 dyslexic and 30 as regular readers TP for dyslexic was **78.7% and for regular readers was 82.84%**	A single SVM classifier proved to be inaccurate and therefore an ensemble of SVM classifiers were employed and a majority vote procedure was used to make a decision	A classification task as regular or dyslexic readers was achieved using recorded channels of EEG with method of Event Related Potentials (ERP) The authors did not indicate whether they considered the inclusion and exclusion criteria of participants (age, gender, etc.)
10	**Questionnaires**	**A computational approach for screening dyslexia**	The 52 participants were selected from specialized centers 29% (15 individuals) diagnosed as dyslexic. shown accuracy rate of 80%	A backpropagation neural network was implemented using the 52 sets of records, divided as follows: 70% of the cases were for training stage and 30% for testing purposes. After several performances (5000 epoch) we observed the organization in two groups, dyslexics and non-dyslexics	The classifying attributes were collected through personal data, parents, personal history, language, education, disease, trauma, disorders, family history, sociability and behavioral The parameters selected as the classifying attributes were very much static and biased for the reasons that the human (parent/child) may not be willing to reveal the difficulty they are facing in reading. Also the difficulty experienced may not be known to the individual himself

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
11	**Eye movements**	**Hidden Markov models for analysis of eye movements of dyslexic children**	Eye movement of 76 subjects are recorded The probability of remaining in the fixation state for stimulus A had the best sensitivity of 80% and specificity of 62% and best saccades for stimulus C sensitivity of 70% and specificity of 58%	Hidden Markov model	Eye movement are used for dyslexia detection by the application of Hidden Markov model. The children are given the non verbal task like 5 stimulus activities like looking through the dots, reading text quietly, counting the smiling faces, remembering the things in a given picture and looking through the numbers 1–15 in sequence. Two features fixation and saccades were captures for every stimulus
12	**EEG**	**Classification of dyslexic and normal children during resting condition using KDE and MLP**	3 dyslexic and 3 non dyslexic in age 4–7 years of age The study gave a 90% accuracy rate since the dataset used is very small it is not very encouraging	90% accuracy with MLP	The authors used MLP to classify the EEG signals. KDE (Kernel Density Estimation) was used to extract the features from EEG signals. The EEG signals were captured to measure the brain activities during the subjects eyes closed and opened The sample size as participants was very less and this work too lacked the specification of the inclusion or exclusion criteria for participants

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
13	**Questionnaires**	**Learning disability prediction tool using ANN and ANFIS**	16 attributes of 1020 students are preprocessed cy applying attribute reduction using PCA and they are reduced to 7. The intention is not to delete the other attributes but to arrange them in their order of importance which helps in better detection of LD	Principal component analysis, backpropagation network	The aim is to design an useful and controlling tool, using the methods of artificial neural network and adaptive neuro-fuzzy inference system. These systems are used to measure the percentage of LD with which the students are affected. Thus, this method is not only detecting LD but also used to measure it by giving a class such as minor or major The author presents a detection tool where the children under testing have to click on the learning difficulties they suffer with. The children themselves may not be aware of the learning difficulty they may be suffering with, or the child may himself feel that he faces such difficulty but it may not be so. Thus, there may be false positives or false negatives in such cases

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
14	**Eye movements**	**Detection of dyslexia from eye movements using anfis and Bbwpe feature extraction methods**	Detection rate of proposed ANFIS with BBWPE is high when compared to existing ANN and ANFIS with MAR	A feed-forward neural network using back propagation algorithm was used for a supervised learning	Reading and non reading tasks were given to 76 school children for capturing the eye movements. using videooculographic (VOG) technique. Time and frequency domain features were extracted and various feature selection methods were performed to select subsets of significant features The work relies on the eye capture movement using Videooculographic, the method which mainly has good performance if the subject is not wearing glasses, in a well lighted environment. In the absence of these constraints the detection could be wrong

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
15	**Eye movements**	**Detecting readers with dyslexia using machine learning with eye tracking measures**	97 subjects with normal or corrected-to-normal vision; 48 of then with diagnosed dyslexia	Support vector machine binary classifier Achieving 80.18% accuracy	The samples were given a text to read and while reading the eye tracking was done along with the user preference regarding the font using questionnaires. The questions were in the form of comprehension at the end of the text

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
16	**Questionnaires**	**Feature selection: a novel approach for the prediction of learning disabilities in school-aged children**	500 sample data 16 signs or symptoms of LD **No. of features 16 reduced to with same classification accuracy (%) 98.6**	Proportional Rough Set (PRS) relevance method for feature reduction and Sequential Minimal Optimization (SMO) for classification	The authors employed feature selection. But before that, the various features are checked independently along with class to determine its importance in the data. They are then arranged in order of their significance and the less significant attributes are deleted as long as the classification accuracy is not changed or improved The author has created a data set by manually querying the children from various sources including 3 special schools. The signs and symptoms were marked as true/false. The genuineness of the data set depends on the response given by the children. The response could be incorrect due to various reasons, including the child being hesitant to accept the fact of possessing the asked symptoms

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
17	**Eye movement**	**Screening for dyslexia using eye tracking during reading**	**97 high risk samples and 88 low risk subjects forming the control group**	Predictive modeling and statistical resampling techniques	An ophthalmological examination was done to capture the eye movements to check if there any differences in the two groups related to basic and oculomotor functions The authors developed classification models from eye tracking records. They used predictive modeling and statistical resampling techniques. Even though the duration of capture was less than one minute, the authors proved that the models could correctly identify and distinguish between the high risk from the low risk subjects
18	**Game**	**Diagnosing risk of dyslexia with a game**	The experiment was performed on 243 samples including adults and children attaining a n accuracy of 83%	Binary SVM	The aim is to detect early the presence of dyslexia through machine learning by developing a game called DYTECTIVE. The game analyses the language principles and attainment, also specific language skills related problems

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
19	**Questionnaires**	**Hybrid classification approach HDLMM for learning disability prediction in school going children using data mining technique**	15 attributes of 1020 samples FS, FR, Deep learner model, Markov-TP rate is 99.09	GA employed for feature reduction and Markov model classification along with deep neural network	The authors used adaptive genetic algorithm for feature selection and feature reduction and hybrid model of deep learner neural network and Markov chain model process is used for classification The author has used a synthetic dataset instead of real world dataset. Thus the accuracy of prediction of LD on real world sample could be compromised if the synthetic data isn't nearly matching to a real-world data set

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
20	**Test performance**	**Lexa: A tool for detecting dyslexia through auditory processing**	56 children between age 6 and 15	The author built a decision tree with all the relevant features obtained through the four tasks and got a classification accuracy of 89.2% which dropped to 53.8% when only the attributes related to phonological processing were retained and others removed	The phonological processing tasks of the children were analysed The work involved four task of the child going through a battery of tests, followed by test of phonological awareness, followed by prosodic sensitivity and lastly auditory processing

(continued)

Table 1 (continued)

S. No.	Approach	Name of paper/publication	Number of attributes (symptoms/accuracy)	Algorithm	Findings/detection
21	**App/Game**	**An Android application for helping in the identification of children with reading difficulties**	46 second class children of a primary school	The author have used the threshold for the parameters efficiency and effectiveness set considering the scores under the first percentile	21 sentences are given as part of the test. The child constructs the sentences by the given words that appear as buttons. Many distracters are given along with the correct words, so as to confuse the reader because they differ for the correct word only in some syllable or a letter As the test progresses, level by level in four level, the length and the complexity of the orthographic structure go on increasing
22	**Game**	**Screening dyslexia for English using HCI measures and machine learning**	267 participants For training machine learning model achieving the detection accuracy of 84.62%	LIBSVM, a binary classifier was employed in the Gaussian Support Vector Machine (SVM) setup	The authors developed a computer based game Considering the language errors dyslexics make and cognitive skills like working memory, perceptive functions, executive functions, etc The 226 performance measures based on the above mentioned skills were captured through the game attempts by 267 samples, a statistical model was trained to predict readers with or without dyslexia

1. **Static Questionnaire/Local Characteristics Checklist** [7–17]

The results from the work proposed from [12, 13, 15] were suffering from the same drawback that the teachers and parents were reluctant to accept the prediction result because of the black box prediction logic and that the rules or logic of prediction was not very clear.

The accuracy, however obtained by Costa et al. [15] in which 15 out of 55 were detected as having LD was 80% by applying NN. Also a good result was obtained by David and Balakrishnan [16] who applied PCA for reducing the attributes from 16 to 7 with 99% accuracy. Also the work implemented by Sabu [8] applied feature reduction from 16 attributes to 4 with the same classification accuracy as 98.6%.

2. **Test Performance** [18–23]

The major disadvantage in this approach is that the test performance solely is the judging criteria for a child, which is very much biased.

However the authors in [20] performed a test on 240 samples detecting 80 LD samples. And applied LVQ and achieving accuracy of 95.8%.Th authors in [22] used two algorithms LVQ and SLP and achieved an accuracy of 84% and 91.8% respectively.

3. **Electroencephalogram (EEG)** [24–26]

Electroencephalogram, commonly known as EEG, is a technique that can be used to monitor and detect brain functions. EEG is a record of the oscillations of brain electric potential recorded from electrodes on the human scalp The electrical activity of the brain for various stimuli can be identified via the electrodes placed on the scalp.

The author Frid in [25] applied SVM classifiers on EEG signals Obtained from 50 participants in the age group of 24–35 diagnosed 20 Dyslexic and 30 as regular readers. TP for Dyslexic was **78.7% and for regular readers was 82.84%. The author in** [26] **attained 90%accuracy with MLP but the no. of samples were very less**.

4. **Eye Movements** [27–31]

The author in [30] Martin has recorded Eye movement of 76 subjects mainly the fixations and saccades on attained a good sensitivity of The probability of remaining in the fixation state for stimulus A as 80% and best saccades for stimulus C sensitivity of 70% and specificity of 58%.

In another attempt [27], using eye movement approach, the author captured 76 school childrens' eye movements. using videooculographic (VOG) and employed a Fuzzy membership function with Adaptive Neuro-Fuzzy Inference System-PSO which gave a classification accuracy of 94.6%. Another good results were obtained by Rello and Ballesteros [28] in which the The samples were given a text to read and while reading the eye tracking was done along with the user preference regarding the font using questionnaires. 97 subjects out of which 48 were detected to have dyslexia achieving accuracy of 80.18%.

5. **Game/App** [32–36]

The game Dytective [34], was first tested on 243 people and attained 83% accuracy in a held-out test set with 100 participants using Support Vector Machines. The authors improved the performance by extending the scope of the game by including the HCI measures covered by 226 performance measures. the model trained by SVM classifier and achieving an accuracy of 84.62% [36].

3 Conclusion

The existence of Learning Disability may affect the individuals throughout their life. LD usually exists for a lifetime. The research work carried on till now focuses on the static attributes of the child. Some other work also deals with questionnaire where the students are answering them. Through the literature survey, it can be seen that the identification of Learning Disability is done using a static questionnaire, or with the involvement of parents in the screening process. There exists a dependency on the response given to such questionnaires, which may not be answered genuinely. The reason for this is that the child may himself be not aware that he may be facing such a difficulty. He may also be reluctant accepting this fact. Also the EEG techniques that are used to detect LD may not be very comfortable for children to undergo. Its again a test that a child undergoes for LD detection. The Eye tracking measures gives good accuracy but depending solely on these measures could not be very helpful because, all eyes cannot be tracked. Also the eye movement data is only tracked which further needs to be related to the context and it is not very simple and easy to interpret the results.

Thus, we need a system that will not depend on the response to such questionnaires and the involvement of the parents. We need a system that understand the learning behavior of a student in the environment in which its happening so as to detect the learning disability LD using static characteristics of the child or using the static questionnaires. Once the disability is diagnosed the student will be able to learn according to his/her learning requirements/preferences that may lead to positive performance of the LD learner.

References

1. Gartland D, Strosnider R (2007) Learning disabilities and young children: identification and intervention. Learn Disabil Q 30(1):63
2. Shah HR, Trivedi SC (2017) Specific learning disability in Maharashtra: current scenario and road ahead. Ann Indian Psychiatry 1(1):11
3. Common signs of learning disabilities. Retrieved from http://www.ldonline.org/ldbasics/signs
4. Types of learning disabilities. Retrieved from https://ldaamerica.org/types-of-learning-disabilities/

5. Logsdon A. Types of tests used to diagnose learning disabilities. Retrieved from https://www.verywellfamily.com/learning-disability-tests-in-public-schools-2161894
6. Guo B et al (2015) Predicting student's performance in educational data mining. In: 2015 International symposium on educational technology (ISET). IEEE
7. Kohli M, Prasad TV (2010) Identifying dyslexic students by using artificial neural networks. In: Proceedings of the world congress on engineering, vol 1
8. Sabu MK (2015) Feature selection: a novel approach for the prediction of learning disabilities in school-aged children, pp 127–137
9. Margaret Mary T, Hanumanthappa M (2017) Hybrid classification approach HDLMM for learning disability prediction in school going children using data mining technique. J Theor Appl Inf Technol 95(13):2989–2998
10. Akhil R, Soori A, Shankar D, Mrinal Krishnan M, Poorna BR (2017) Detecting specific learning disabilities (decision tree). In: 2017 international conference on networks & advances in computational technologies (NetACT). 978-1-5090-6590-5/17©2017 IEEE
11. Wu T-K, Huang S-C, Meng Y-R (2006) Identifying and diagnosing students with learning disabilities using ANN and SVM. In: The 2006 IEEE international joint conference on neural network proceedings. IEEE
12. Wu T-K, Huang S-C, Meng Y-R (2006) Effects of feature selection on the identification of students with learning disabilities using ANN. In: International conference on natural computation. Springer, Berlin
13. Wu T-K, Meng Y-R, Huang S-C (2006) Application of artificial neural network to the identification of students with learning disabilities. IC-AI
14. David JM, Balakrishnan K (2010) Machine learning approach for prediction of learning disabilities in school-age children. Int J Comput Appl 9(11):7–14
15. Costa M et al (2013) A computational approach for screening dyslexia. In: Proceedings of the 26th IEEE international symposium on computer-based medical systems. IEEE
16. David JM, Balakrishnan K (2014) Learning disability prediction tool using ANN and ANFIS. Soft Comput 18(6):1093–1112
17. Ambilili A (2016) A framework for learning disability prediction in school children using naïve Bayes-neural network fusion technique. J Inf Knowl Res Comput Eng 04(01). ISSN: 0975-6760
18. Jain K (2009) Computational diagnosis of learning disability. Int J Recent Trends Eng 2(3):64
19. Loizou A, Laouris Y (2011) Developing prognosis tools to identify learning difficulties in children using machine learning technologies. Cogn Comput 3(3):490–500
20. Manghirmalani P, Panthaky Z, Jain K (2011) Learning disability diagnosis and classification—a soft computing approach. In: 2011 world congress on information and communication technologies. IEEE
21. Manghirmalani P, More D, Jain K (2012) A fuzzy approach to classify learning disability. Int J Adv Res Artif Intell 1(2):1–7
22. Jain K, Mishra PM, Kulkarni S (2014) A neuro-fuzzy approach to diagnose and classify learning disability. In: Proceedings of the second international conference on soft computing for problem solving (SocProS 2012), Dec 28–30, 2012. Springer, New Delhi
23. Poole A, Zulkernine F, Aylward C (2017) Lexa: a tool for detecting dyslexia through auditory processing. In: 2017 IEEE symposium series on computational intelligence (SSCI). IEEE
24. Andreadis II et al (2009) Detecting complexity abnormalities in dyslexia measuring approximate entropy of electroencephalographic signals. In: 2009 annual international conference of the IEEE engineering in medicine and biology society. IEEE
25. Frid A, Breznitz Z (2012) An SVM based algorithm for analysis and discrimination of dyslexic readers from regular readers using ERPs. In: 2012 IEEE 27th convention of electrical and electronics engineers in Israel. IEEE
26. Karim I, Abdul W, Kamaruddin N (2013) Classification of dyslexic and normal children during resting condition using KDE and MLP. In: 2013 5th international conference on information and communication technology for the Muslim World (ICT4M). IEEE

27. Gomathi PM, Nasira GM (2015) Detection of dyslexia from eye movements using anfis & Bbwpe feature extraction methods. Res Inventy Int J Eng Sci 5(01):20–29
28. Rello L, Ballesteros M (2015) Detecting readers with dyslexia using machine learning with eye tracking measures. In: Proceedings of the 12th web for all conference. ACM
29. Macaš M et al (2005) Dyslexia detection from eye movements using artificial neural networks. In: IFMBE proceedings, vol 11, no 1
30. Macaš M, Lhotská L, Novák D (2013) Hidden Markov models for analysis of eye movements of dyslexic children. In: 2013 18th international conference on digital signal processing (DSP). IEEE
31. Benfatto MN et al (2016) Screening for dyslexia using eye tracking during reading. PLoS one 11(12)
32. Lyytinen H et al (2007) Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. Nordic Psychol 59(2):109–126
33. Gaggi O et al (2012) A serious game for predicting the risk of developmental dyslexia in pre-readers children. In: 2012 21st international conference on computer communications and networks (ICCCN). IEEE
34. Rello L et al (2016) Dytective: diagnosing risk of dyslexia with a game. In: Proceedings of pervasive health, p 16
35. Francese R, Monaco C, Nicoletti C (2018) An android application for helping in the identification of children with reading difficulties. In: Proceedings of the 4th EAI international conference on smart objects and technologies for social good
36. Rello L et al (2018) Screening dyslexia for English using HCI measures and machine learning. In: Proceedings of the 2018 international conference on digital health

EEG Based Emotion Investigation from Various Brain Region Using Deep Learning Algorithm

Vaishali M. Joshi and Rajesh B. Ghongade

Abstract Emotion detection using Electroencephalogram (EEG) is a reliable approach for the formulation of human brain signals as this does not depend on humans' posturing or pretentiousness. Analyzing every brain region generates noise, redundant information. This increase more complexities in calculations. Ultimately, handling unwanted and excessive data degrades the detection accuracy as it calls for use of additional and complex hardware. Therefore to overcome this challenge, this paper proposes a detailed study and understanding of different brain regions which contribute most to emotion generation and which are very much vital for emotion analysis. This work provides guidelines in neuroscience and establishes that a particular brain region is vital which provides sufficient information and EEG data for emotion recognition. In this paper features of differential entropy, Hjorth parameters, and power spectral density (PSD) are used to train bidirectional long short term memory (BiLSTM) network. The performance of the proposed technique is evaluated on the DEAP benchmark database for different brain regions namely the prefrontal region, frontal region, temporal region, parietal region, and occipital region. The results show that the frontal brain region has the highest emotion recognition rate compared to other regions and is vital for emotion recognition.

Keywords Emotions · Differential entropy · Hjorth parameters · Power spectral density · Bi-LSTM · Brain region

1 Introduction

Emotions are a result of various factors such as surroundings, lifestyles, prevailing economical conditions, social status, likings, preferences etc. It is a psycho-physiological phenomenon that creates positive or negative impacts from humans on routine communication. For humans, emotions are very important as they have the ability to express through facial expressions against actual emotions [1]. Emotions

V. M. Joshi (✉) · R. B. Ghongade
Electronics Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, Maharashtra 411043, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_34

play a very crucial role in humans interpersonal interactions, cognition, perception, and reputation. Considering emotion's enormous impact on every angle of life including health, it is of paramount importance to monitor the emotion and take corrective actions. For example, for a deaf and dumb, emotion is one of the most powerful tools to understand the physiological state. Emotion recognition has a variety of applications in various areas such as intelligent systems, brain-computer interaction, e-learning, advertisement, and the entertainment industry. For the past decade, neuroscience researchers across the globe are working continuously on this subject and this is a popular study amongst the research community.

2 Related Work

While aiming to achieve emotion recognition, the methodology comprises two major broad domains namely feature extraction from the EEG signal and constructing an accurate predictive model for emotion recognition.

Xing et al. [2] presented a linear EEG mixing model based on stack auto to improve accuracy performance. They achieved an accuracy of 81.10% for valence and 74.38% for arousal with this approach on the 'Database for Emotion Analysis using Physiological Signals' (DEAP) database. Wang et al. [3] proposed data augmentation-based emotion recognition using deep learning. With this approach, they achieved 75% accuracy on the 'SJTU Emotion EEG Dataset' (SEED). Jirayucharoensak et al. [4] used power spectral density (PSD), principal component analysis (PCA), and deep neural network and achieved validation accuracy of 52.05% and 53.42% for arousal and valence on DEAP database. Pandey et al. [5] used VMD features and deep neural network classifier using DEAP database and achieved recognition accuracy for valence 62.50%, arousal 61.25%. Joshi et al. [6] used differential entropy as a feature and BiLSTM classifier and achieved 74% accuracy using DEAP database. Joshi et al. [7] in another work on SEED database achieved accuracy of 74.91% using differential entropy and BiLSTM classifier. Li et al. [8] have worked on two databases namely DEAP and SEED. Using DEAP database they achieved recognition accuracy of 59.06% and on SEED database they achieved accuracy is in order of 83.33%. They have used support vector machine (SVM) classifier. Cimtay et al. [9] used convolutional neural network to train model and median filter to remove false detections. They worked on SEED and DEAP database. On SEED database they achieved an accuracy of 86.56% and 78.34% for two and three emotion classes respectively. On DEAP database they achieved an accuracy of 72.81%. Deep learning is high end recent technique used in machine learning. It is found that models with has deep architectures achieves improved result compare to old methods [10–12].

3 Proposed Method

This paper aims to find the brain region for effective and accurate emotion analysis. Figure 1 explains that implemented operational workflow. Workflow has mainly four phases as database collection, band selection, feature extraction, and emotion classification.

3.1 Collection of the EEG Signal Database

DEAP [13] database for emotion recognition is used in the proposed work. Thirty-two participants have watched 40 videos of one minute each. This experiment is conducted on Biosemi Active Two system which has 40 electrodes. Thirty-two electrodes were used for recording EEG signals and balance electrodes were used for peripheral physiological signals. Data was collected using a10-20 international electrode placement system at a sampling rate of 512 Hz. This data was downsampled to 128 Hz and bandpass filtered in the range of 4–45 Hz.

3.2 Feature Extraction Method

In this work, three different features namely differential entropy (DE), Hjorth parameters, power spectral density (PSD) have been used to examine the performance of the EEG-based emotion recognition human–computer interaction (HCI) system. EEG signal is decomposed into four frequencies. For each band, features are extracted. Theta, alpha, beta, and gamma frequency bands are-theta: (4 = 8 Hz), alpha: (8 = 13 Hz), beta: (13 = 30 Hz), and gamma: (30 = 44 Hz).

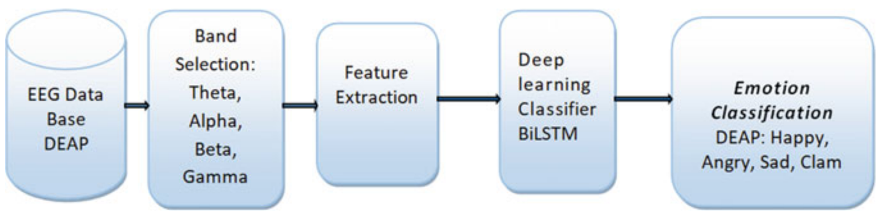


Fig. 1 Block diagram of proposed work

Table 1 Hjorth parameters

Activity	Mobility	Complexity
$Ax = variance(v(t))$ $v(t)$ is a signal	$Mx = \sqrt{\frac{variance(\dot{v(t)})}{variance(v(t))}}$ $\dot{v(t)} = \frac{dv}{dt}$	$Cx = \frac{Mx(\dot{v(t)})}{Mx(v(t))}$

3.2.1 Differential Entropy

Differential entropy (DE) [14] is used to measure the continuous random variable's entropy. The DE with normal distribution is illustrated as,

$$\mathrm{DE} = \frac{1}{2}\log_e 2\pi e s^2$$

where s^2 variance, and π and e are constants.

3.2.2 Hjorth Parameters

Hjorth parameters [15] are used for time-domain analysis in signal processing. These feature parameters are activity (Ax), mobility (MX), and complexity (Cx), which are often used in EEG signal analysis signals as given in Table 1.

3.2.3 Power Spectral Density

The power spectral density (PSD) [16] is defined as the average energy distribution per unit time over frequency bands.

3.3 Classification Method

As explained above deep learning algorithm-BiLSTM network is implemented for performance assessment of the proposed feature extraction method.

Long-Short Term Memory Network (LSTM) [17] is a special class of recurrent neural network (RNN). In order to overcome the problem of long-term dependency in RNN, LSTM is introduced. Long sequences are difficult to learn from standard RNN and cause the problem of vanishing/exploding gradient. LSTM has the ability to extracts the temporal and spatial features from EEG signals. Figure 2 shows BiLSTM architecture. BiLSTM network formed from the input layer, BiLSTM layer, fully connected layer, and classification layer. BiLSTM network has a forward and a backward LSTM layer. BiLSTM network has the ability to remember more information as compared with unidirectional networks. The LSTM cell is composed of three

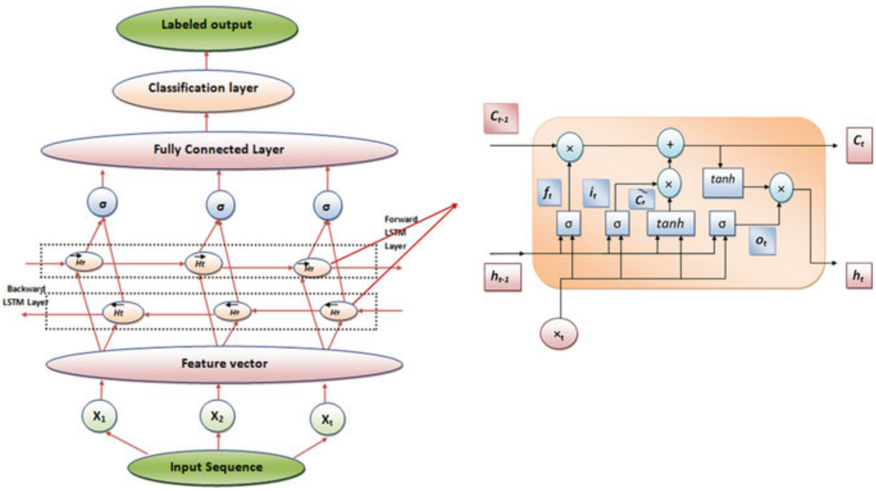


Fig. 2 BiLSTM architecture

Table 2 Hyper parameters

Hyper parameter	Value
Initial learn rate	1.0000e−03
Hidden units	200
Gradient decay factor	0.9

control units, namely input gate (i_t), output gate (o_t), and forget gate (f_t) respectively as shown in Fig. 2.

4 Result and Discussion

4.1 Hyper Parameter

The hyper parameters used for BiLSTM are shown in Table 2.

4.2 Experimental Result

Work is related to the subject-independent approach. Test data and training data are different i.e. same subject's data is not used for training and testing purposes. In this work features of DE, Hjorth parameters, and PSD of 25 subjects are used to train the BiLSTM classifier and the other 5 subject data are used to test the classifier.

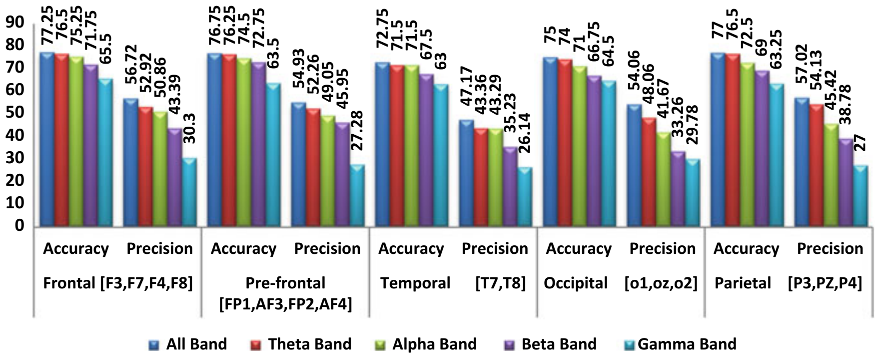


Fig. 3 Performance comparison for different brain region using differential entropy

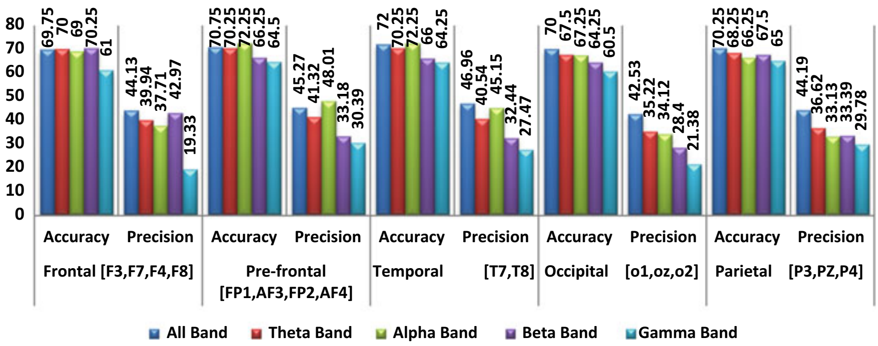


Fig. 4 Performance comparison for different brain region using Hjorth parameters

Figures 3, 4, and 5 demonstrate experimental outcomes using features of differential entropy, Hjorth parameters, and power spectral density with a BiLSTM network classifier.

From Figs. 3, 4, and 5 it can be seen that using differential entropy the highest accuracy obtained is 77.25% with all bands (concatenation of four bands). From the obtained results it can be clear that the frontal brain region with electrodes F3, F7, F4, F8 gives higher accuracy. The result obtained by the proposed method is also compared with that of the other research in this field and is summarized in Table 3.

5 Conclusion

Traditionally bulky multichannel machines were used to cover all the brain regions. The data generated through each electrode is not useful to its full extent and therefore it is prudent to avoid the collection of redundant data. From the results, it is found that

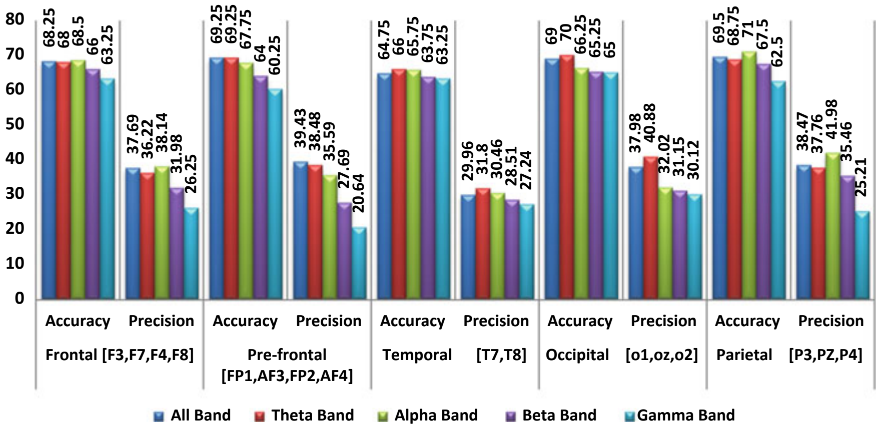


Fig. 5 Performance comparison for different brain region using PSD

Table 3 Performance comparison with related method

S. No.	Name of author	Method	Name of class	Average accuracy (%)
1	Cimtay and Ekmekcioglu [9]	Convolutional neural network (CNN)	Positive, negative	72.81
2	Pandey and Seeja [18]	Convolutional neural network	Valence, arousal	61.50, 58.50
3	Rayatdoost and Soleymani [19]	Spectral topography, convolutional neural network	Valence, arousal	59.22, 55.70
4	**Proposed method**	**Differential entropy, BiLSTM**	**Happy, calm, sad, angry**	**80.5, 75.5, 71, 82**

the frontal brain region generates sufficient information for emotion recognition. If we analyse frontal brain region data using differential entropy feature and BiLSTM classifier, we can achieve higher accuracy compared to established methods. We can also conclude that by using lesser electrodes we can reduce signal noise and avoid analysis of redundant information. This helps to achieve higher emotion recognition accuracies. This approach helps to avoid use of bulky machines. We can use a portable and handy emotion detection machine that can have a lesser number of electrodes and achieve higher accuracies. In the future using other deep learning architecture like CNN, DNN will be explored to analyze the performance of emotion recognition system.

References

1. Zhang Z, Yan J, Lu G, Li H, Sun N, Ge Q (2017) Multiscale adaptive local directional texture pattern for facial expression recognition. KSII Trans Internet Inf Syst 11
2. Xing X, Li Z, Xu T, Shu L, Hu B, Xu X (2019) SAE+ LSTM: a new framework for emotion recognition from multi-channel EEG. Front Neurorobot 13:37
3. Wang F, Zhong S-H, Peng J, Jiang J, Liu Y (2018) Data augmentation for EEG-based emotion recognition with deep convolutional neural networks. In: International conference on multimedia modeling, pp 82–93
4. Jirayucharoensak S, Pan-Ngum S, Israsena P (2014) EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. Sci World J 2014
5. Pandey P, Seeja K (2019) Subject independent emotion recognition from EEG using VMD and deep learning. J King Saud Uni-Comput Inform Sci (2019)
6. Joshi VM, Ghongade RB (2020) Optimal number of electrode selection for EEG based emotion recognition using linear formulation of differential entropy. Biomed Pharmacol J 13
7. Joshi VM, Ghongade RB (2020) Emotion detection with single channel EEG signal using deep learning algorithm. Int J Recent Technol Eng 8
8. Li X, Song D, Zhang P, Zhang Y, Hou Y, Hu B (2018) Exploring EEG features in cross-subject emotion recognition. Front Neurosci 12:162
9. Cimtay Y, Ekmekcioglu E (2020) Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition. Sensors 20:2034
10. Joshi VM, Ghongade RB (2020) IDEA: intellect database for emotion analysis using EEG signal. J King Saud Univer-Comput Inform Sci
11. Joshi VM, Ghongade RB (2021) EEG based emotion detection using fourth order spectral moment and deep learning. Biomedical Signal Process Control 68:102755
12. Joshi VM, Ghongade RB (2021) Subject noncontingent EEG-based emotion detection using deep learning algorithm. In: Data Engineering and Intelligent Computing pp 67–75. Springer, Singapore
13. Koelstra S, Muhl C, Soleymani M, Lee J-S, Yazdani A, Ebrahimi T et al (2011) Deap: a database for emotion analysis; using physiological signals. IEEE Trans Affect Comput 3:18–31
14. Duan R-N, Zhu J-Y, Lu B-L (2013) Differential entropy feature for EEG-based emotion classification. In: 2013 6th international IEEE/EMBS conference on neural engineering (NER), pp 81–84
15. Hjorth B (1970) EEG analysis based on time domain properties. Electroencephalogr Clin Neurophysiol 29:306–310
16. Song T, Zheng W, Lu C, Zong Y, Zhang X, Cui Z (2019) MPED: a multi-modal physiological emotion database for discrete emotion recognition. IEEE Access 7:12177–12191
17. Hochreiter S, Schmidhuber J (1997) LSTM can solve hard long time lag problems. In: Advances in neural information processing systems, pp 473–479
18. Pandey P, Seeja K (2020) Subject independent emotion recognition system for people with facial deformity: an EEG based approach. J Ambient Intel Humanized Comput 1–10
19. Rayatdoost S, Soleymani M (2018) Cross-corpus EEG-based emotion recognition. In: 2018 IEEE 28th international workshop on machine learning for signal processing (MLSP), pp 1–6

Electric Motor Drive Using Single Input Fuzzy Logic Controller for Husk Extraction in Rice Mill Industry

A. Jagadeesh, K. Deepa, and K. Sireesha

Abstract This paper deals with the Single Input Fuzzy Logic Controller (SFLC) deployed for controlling the speed of a DC motor used in Rice mills for separating the husk layer surrounding the paddy to get the rice grain with perfect shape and size. System utilizes the power from the Photo voltaic (PV) array with a boost converter to meet desired specifications of DC motor. Different operating procedures like Proportional Integral Derivative (PID) logic controller procedures can be deployed for the speed control of DC motor. This paper illustrates the fuzzy operation technique with single input variable. PID logic controller requires knotty mathematical models. While Fuzzy Logic Controller (FLC) function is based on rule base knowledge. The DC motor's speed can be tuned to a reasonable extend to function in a settled manner. In summary, this paper demonstrated the speed regulation of both controllers (SFLC and Two input FLC) designed for the DC motor and results are discussed with the help of SIMULINK program in MATLAB.

Keywords Single variable fuzzy logic controller (SFLC) · Fuzzy inference system (FIS) · Proportional integral derivative (PID)

1 Introduction

The necessity for the speed control of DC motors that have been widely used in rice mills [1], lathe machines, robotic industries, weaving machines with the emerging technologies in the power electronics is the burning desire for the usage of Fuzzy Logic Controller (FLC) technique due to its precise and continuous control characteristics. Basically, a perfect running electric motor drive must have speed regulating

A. Jagadeesh (✉) · K. Deepa · K. Sireesha
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, Bengaluru, India

K. Deepa
e-mail: k_deepa@blr.amrita.edu

K. Sireesha
e-mail: k_sireesha@blr.amrita.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_35

characteristics with dynamic varying load. Usually DC drives, due to their coherence in nature are compatibly low in cost [2] and reliable. Therefore, has been a dorsum of industrial utilizations. DC motors have a numerous evolving purposes which are well known for its use in controllable speed machines. Armature control technique was the popular and commonly applied technique for low rated DC motors. Nevertheless**,** increase in efficiency and new technologies in power electronic converters made the controllability easier and cheaper, and also the higher current carrying capabilities of static power converters brought a vital coinage in the contribution of electrical drives. Many papers have proposed disparate types of control paradigms depending on their motto and applications of DC motor and its purpose. Few systems use the conventional control paradigms like Proportional-Integral-Derivative (PID) [3, 4], Proportional-Integral (PI) controller. This discourse discusses on armature voltage control method, using the concept of single input fuzzy technique [5, 6].

This proposed system can be implemented in rice husk extruder industries. This extruder machine contains DC motor running at constant speed that extracts the layer of husk, controller in this system maintains constant speed during variations in load safeguarding the rice grain. The Armature voltage and the Field Flux speed control are the popular methods for speed control of DC motor. In this work, armature voltage control is applied to separately excited DC motor. Speed control can be achieved using various techniques or algorithms such as PID Controller, Fuzzy Logic Controller (FLC), Fuzzy-Genetic Algorithm (FGA), Artificial Neural Networks (ANN) [7], Fuzzy-Ants Colony (FAC). The main motive of this paper is achieving the desired voltage from the PV cell [8] and the boost [9] circuit deploying Single Variable Fuzzy Logic Controller (SFLC) to extract the maximum output power from the motor during different load and environment conditions. The perfectly tuned PI controllers cannot respond to the dynamic changes but perfectly defined knowledge base fuzzy logic controller [10, 11] could mitigate the dynamic changes happening in the environment and stabilizes the required output. The novelty of parallel tuning of input and output scaling factors of the FLC [12, 13] for rated speed has been achieved with the help of gain blocks and its simulation results are being discussed. The paper has been categorized into five sections. Section 2, describes the specifications of the proposed system. Section 3, discuss about design of controller's in proposed system. In Sect. 4 the simulation results for the given system is dealt about. And through Sect. 5 paper gets concluded.

2 System Specifications

2.1 DC Motor

This system uses a separately excited DC motor which was widely used in Rice mills. The specifications of separately excited DC motor are given in Table 1. DC motor

Table 1 DC-wound motor specifications

Motor specifications	Value	Motor specifications	Value
Motor rating	5 HP	Armature inductance (La)	0.016 H
Rated voltage	240 V	Rated speed (N)	1750 rpm (183.25 rps)
Rated current	15 A	Inertia (J)	0.05 (kg m^2)
Armature resistance (Ra)	0.78 Ω	Viscous friction (Bm)	0.01 (N m s)

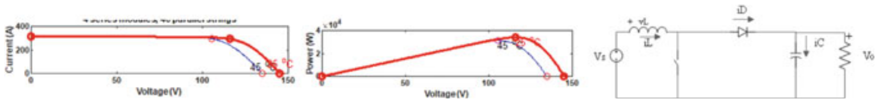


Fig. 1 Current and power versus voltage plot of photo voltaic (PV) cell and boost converter

used in this discourse is powered using PV [14] array which is discussed in the next section.

2.2 *Photo Voltaic (PV) ARRAY*

The Photo Voltaic (PV) array is a module which converts light energy into electrical energy [14]. Modeling of a PV, requires weather data i.e. (irradiance and temperature) as input variables. The obtained output is in terms of power, current or voltage. However, the characteristics I and P versus V are plotted in Fig. 1 and the maximum power [14] is 28.3 kW. Any change in the input parameters, reflects major changes in outputs. A precise input parameter is essential for modeling the PV module. In the following paper irradiance and temperature parameters of the PV module is peculated from a constant block. The single diode with array structure containing both series and parallel resistors is chosen for precise accuracy. Obtained voltage is fed to DC motor using Boost converter.

2.3 *Boost Converter Design*

This system requires boost converter [9] (in Fig. 1), to have smooth operating voltage from the Photo Voltaic (PV) cell.

When the switch is closed. the diode is reverse biased [9]. Applying Kirchhoff's Voltage Law in the closed path of the source, inductor, and switch, the voltage across the inductor is given by Eq. (1). The inductor current is increasing linearly hence the derivative is constant and can be attained by Eq. 2.

$$v_L = V_s = L\frac{di_L}{dt} \rightarrow \frac{di_L}{dt} = \frac{V_s}{L} \tag{1}$$

$$\frac{\Delta i_L}{\Delta t} = \frac{\Delta i_L}{DT} = \frac{V_s}{L} \rightarrow (\Delta i_L)_{closed} = \frac{V_s DT}{L} \tag{2}$$

The diode is forward biased by the energy stored in the inductor so as to provide a path for the flow of inductor current during the open condition of the switch [9]. The inductor voltage is given by Eq. 3 at the steady state (Output voltage constant) when the switch is open, current follows the linear path (rate of change of inductor current is constant). The derivative of inductor current is inferred from Eq. (4)

$$v_L = V_s - V_o = L\frac{di_L}{d_t} \rightarrow \frac{di_L}{d_t} = \frac{V_s - V_o}{L} \tag{3}$$

$$\frac{\Delta i_L}{\Delta t} = \frac{\Delta i_L}{(1-D)T} = \frac{V_s - V_o}{L} \rightarrow (\Delta i_L)_{open} = \frac{(V_s - V_o)(1-D)T}{L} \tag{4}$$

from $(\Delta i_L)_{close} + (\Delta i_L)_{open} = 0$, by substituting the values of $(\Delta i_L)_{close}$ and $(\Delta i_L)_{open}$ from Eqs. 2 and 4 output voltage is obtained as

$$Vo = \frac{vs}{1-D} \tag{5}$$

Design parameter of boost converters for the specifications input voltage (120 V), desired output voltage (240 V), duty ratio = D = 0.5 (using Eq. 5), switch frequency (20 kHz), load as DC motor (refer Table 1), voltage ripple (5%), current ripple (2%) are calculated (Table 2).

$$C = \frac{Vo * D}{\Delta Vo * R * Fs} = \frac{Io * D}{\Delta Vo * Fs} = \frac{Vo * D * Ts}{\Delta Vo * R} = 0.3\,\text{mF},$$

$$L = \frac{Vg * D}{\Delta IL * Fs} = \frac{Vo * (1-D) * D}{\Delta IL * Fs} = 0.243\,\text{mH}$$

Table 2 Values of parameters in PV and boost converter

Photo voltaic (PV)		Boost converter	
Temperature	25 °C	Capacitor	0.3 mF
Irradiance	1000 W/m^2	Inductor	0.243 mH

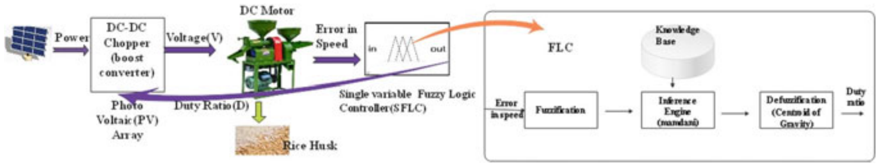


Fig. 2 Block diagram

3 Paradigms of Control Schemes

3.1 *Modelling of the Fuzzy Controller System*

Fuzzy logic mimics the human decision-making methodology as shown in Fig. 2. It trades in with indistinct and imprecise data. In other words, fuzzy is a logic that is used to describe fuzziness. Features of membership function and their membership values are Core: $\mu_A(y) = 1$, Support: $\mu_A(y) > 0$, Boundary: $1 > \mu_A(y) > 0$.

3.2 *Defuzzification*

Process of transferring Fuzzy sets to crisp sets. Defuzzification methods to calculate crisp outputs are (1) *Maxima Methods* are Height Method, First of Maxima (FoM), Last of Maxima (LoM), Mean of Maxima (MoM) (2) *Centroid Methods* are Center of Gravity Method (CoG), Center of Sum Method (CoS) (3) Weighted average method. In this work, Centroid of Gravity (CoG) method is implemented to get the crisp i.e. duty ratio values. According to center of gravity to provide a crisp value for the fuzzy set Centroid of Area (CoA) Method is used.

$$x* = \frac{\int x * \mu c(x)dx}{\int \mu c(x)dx} \tag{6}$$

To find the defuzzified outputs, the total area of membership function is split into certain sub-areas whose centroid is evaluated using the Eq. 6. Among other methods, COG gives out precise crisp values.

3.3 *Rule Base of Fuzzy Systems*

The fuzzy rule-based system makes use of linguistic variables as their antecedents and consequents making it widely used in modeling complex systems. Some Graphical Techniques of Inference Engine are Mamdani Systems, Sugeno Model, Tsukomoto model. Max–min methodology followed by mamdani model gives out the input as precise scalar values. Mamdani inference system is intuitive and has widespread acceptance and is well suited to human input. Hence, this proposed system is designed using mamdani model.

3.4 Configuration of Single Variable Based FLC

Fuzzy logic controller modeling involves various steps and design blocks, for example rule base, fuzzification and defuzzification. The FLC in this design has one input, (error in speed) and one output (duty ratio). There are 8 fuzzy rules which are being particularly used in modeling for the fuzzy logic controller. The COA is applied for the defuzzification. This controller uses mamdani for fuzzy inference system (FIS), employing fuzzy sets in sequential part. The block diagram in Fig. 2 represents the overall system and single input FLC used in the speed regulation, designed using the fuzzy tool box. In Fig. 3 it can be observed that single input i.e. error variable is only considered as input for Fuzzy Inference System (FIS). Using the rule viewer shown in Fig. 4, Variation of output with respect to input can be observed. This helps in predicting the decisions taken by fuzzy with variation of input.

Figure 4 shows the surface for error and duty ratio scaling factors which are input scaling factors for the main SFLC. The surface in Fig. 4 represents a curve (Single Input and Single output). Input for the system is error. Triangular Membership functions (trimf) i.e. five trimf are chosen in Fig. 5 for the designing of Fuzzy Logic Controller (FLC). Error is scaled using the gain block and range is selected as [−1, 1] different parameters (prams) are set for each membership function. Output of fuzzy system depends on this params and gives out the duty ratio, the membership functions

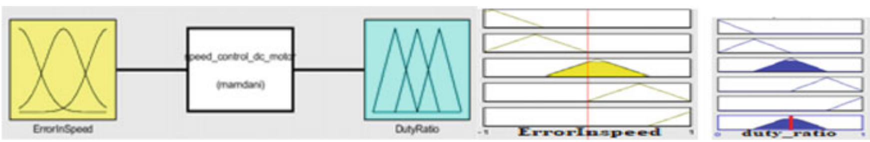


Fig. 3 Fuzzy logic designer tool box and rule viewer of fuzzy inference system (FIS)

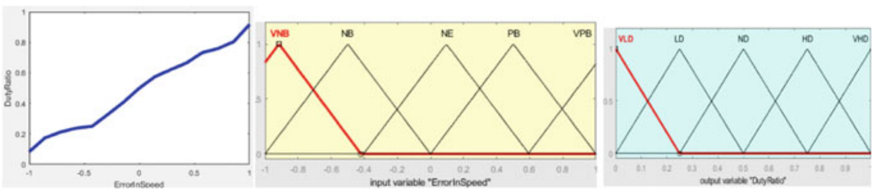


Fig. 4 Surface for output scaling factor, membership function for error and output variable

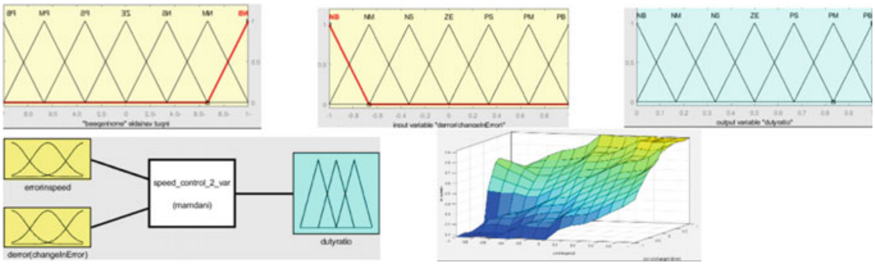


Fig. 5 Two input fuzzy logic tool box, surface for input and output scaling factor and membership function for the error in speed, derror (change_in_error), duty ratio variable

can be observed in Fig. 4 which are trimf and params are taking accordingly to get perfect output. The universe of discourse for error in speed has been divided into five fuzzy sets namely VNB, NB, NE, VPB, PB respectively every linguistic data is subset of these members with some degree of membership changing in the range of [−1, 1] for error and [0, 1] for output duty ratio. Where VNB = Very negative big, NB = Negatively big, NE = No error, VPB = Very positively big, PB = Positively big.

3.5 Configuration of Two Variable Based FLC

The conventional controllers like PI, PID are used to generate gate pulses. To achieve desired output for the plant these controllers need to be perfectly tuned and if tuned, the output is not dynamically responsive. This is the major drawback for the conventional controllers. The controllers like FLC, ANN based controllers, machine learning technique-based controllers etc. are used to overcome the drawback. But for any FLC to produce the desired response for any dynamic changes in the environment, one should properly define the rules or tune the scaling factors at both input and output side. Figure 5 is the block with two input FLC and visualizes the surfaces for error and derror (change in error) scaling factors which are input scaling factors for the main FLC. The fuzzy variables are described as Error: represents the deviation of the speed output of the DC Motor. Change in Error: represents the deviation of present error to the previous error of the speed output of DC motor. Speed: represents the speed of the machine at the every instant. Figure 5 show the membership functions of the input variables of the FLC error ΔN, derror (change in error) variables respectively. Figure 5 respectively depicts the output variable membership function. This are tuned with help of trimf and setting the params. The universe of discourse for error and change in error has been divided into seven fuzzy sets (Table 3) namely NM, NS, ZE, PS, PM, PB respectively where NB and PB are trapezoidal remaining being trimf.

Table 3 Rule base for two input FLC

Change in error↓	Error⟶						
	PB	PM	PS	ZE	NS	NM	NB
PB	PB	PB	PB	PB	PM	PS	ZE
PM	PB	PB	PB	PM	PS	ZE	NS
PS	PB	PB	PM	PS	ZE	NS	NM
ZE	PB	PM	PS	ZE	NS	NM	NB
NS	PM	PS	ZE	NS	NM	NB	NB
NM	PS	ZE	NS	NM	NB	NB	NB
NB	ZE	NS	NM	NB	NB	NB	NB

4 Results

The proposed technique both SFLC and two input FLC Fig. 6 is simulated using MATLAB-Simulink environment and results are discussed in this section. Fuzzy controller is designed as mentioned above and feedback is given to the gating signal of mosfet in terms of duty ratio.

4.1 *Simulation Results of SFLC*

The system has been simulated for varying loads ie from half load to full load. Motor being operated at its half load torque 10 N-m results in the speed variation, the proposed controller maintains the speed, nearer to the rated speed 183.25 rps (ref Table 1) using the rule base knowledge. Corresponding branch currents (Fig. 7a, b) are also displayed Current and Speed plots have a dynamic change at initial points at t

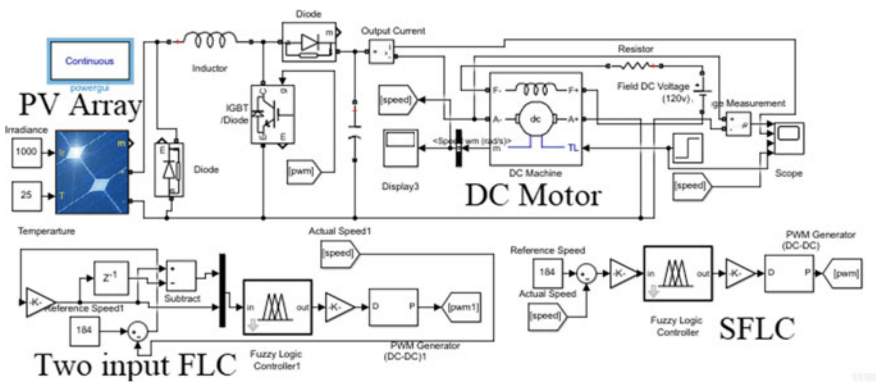


Fig. 6 Simulated block of single variable FLC

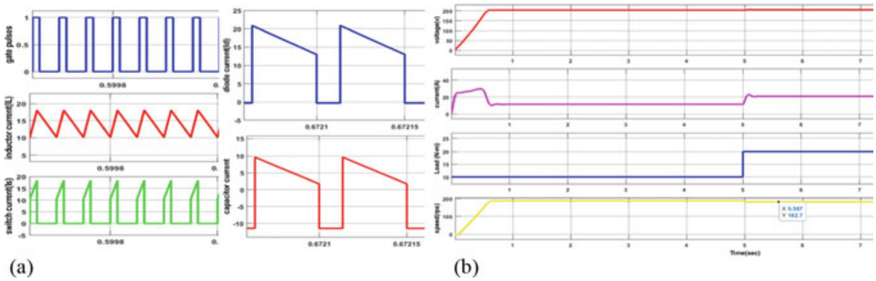


Fig. 7 **a** Inductor and switch currents with respect to gating signal, **b** capacitor and diode currents **c** speed characteristics using SFLC

= 0.1 and t = 0.2 s while the speed and load current reaches to the stable values from t = 0.6 s (Fig. 7c). Initially, motor is allowed to operate at rated load, rice industries operates their machine at normal load, but due to the external disturbances motor is forced to operate at different load which results in drastic changes in speed that damages rice grain. External disturbances are mimicked by changing load from Half load (Torque = 10 N-m) to Full load and the simulated results are taken where the speed is maintained constant using the SLFC and current values reaches the stable values from t = 5.1 s. Figure 7c summarizes, the speed of the system with varying load torque = 10–20 N-m is shifted at t = 5 s, Stability of speed at t = 5+ and t = 5− is attained using the controller.

4.2 *Simulation Results of Two Input FLC*

In this paper, two input FLC is also simulated (Fig. 7), in order to compare the results of SFLC. As mentioned above it is designed by considering error and change in error variables scaled down to range [−1, 1] using gain blocks. Same Load parameters have been deployed to this model and the respective speed plots are discussed. The speed and load current of the DC motor with the two input FLC operated at the half load refer Eq. 7 speed is maintained nearer to the rated speed and is observed that the speed reaches the stability at time t = 0.7 s. Even the Inductor, mosfet, capacitor, diode currents (Fig. 8a, b) dynamically raise to high values during the initial starting of the motor which could be neglected for the rice mill industries. Similar to the case in SFLC to mimic the external disturbances load is being varied and speed plots are taken. Dynamic changes in the load and speed regulation for two input FLC is been tested at t = 5 s (Fig. 8c) and Settling time for the load currents and speed at t = 5.12 s (Fig. 8c) instance, the plots reaches the stable positions attaining the rated values and respective comparison table of SFLC and two input FLC speed parameters is explained in Table 4. From Table 4, change in load due to some external factors, the speed is maintained nearer to the rated speed in case of half load (1.4% drop) and full load (2.11% drop) which makes the system more sophisticated and effective.

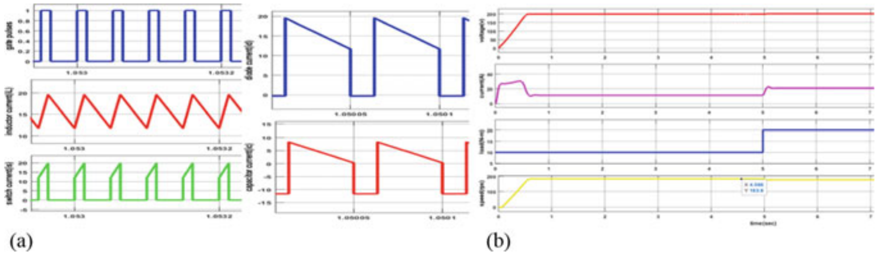


Fig. 8 **a** Inductor, switch currents with respect to the gating signal, **b** capacitor and diode currents, **c** speed plot with respect to variation in load using two input FLC

Table 4 Comparison of speed regulation between SFLC and two input FLC

Model type	Speed		Speed regulation		Peak overshoot (A)	Transient settling time (s)	Load settling time (s)
	Load		Load				
	50%	100%	50%	100%			
SFLC	181.3	1.4%	182.6	0.76%	30.31	0.6296	5.151
FLC	180.1	2.11%	180.7	1.79%	30.36	0.6364	5.177

Table 4 also clearly illustrates that the use of SFLC will not compromise regarding the speed regulation and also infers that the both methodologies in speed control of DC motor has similar peak overshoot, settling time and performance, stating that the two input FLC can override with SFLC, results in reducing the memory and run time, settling time with same output and efficiency.

5 Conclusion

Advantages of the proposed system, designed with only one variable as the input parameter where the vagueness and time delay for two inputs system is eliminated. Having the precise control during varying load parameters which help in proper and safe extraction of husk in rice mill industry. SFLC also provides the same regulation to that of two input FLC. The sole purpose of the system is safer extraction of husk in rice mill industry by having a precise Speed control using Single Variable Fuzzy Logic Controller is Illustrated, whose input and output scaling factors has been tuned consecutively and the proposed system uses PV Cell to promote Renewable energy resources. Power Electronic Converter (DC Choppers) which helps to Boost up the PV voltage and bridges between the Fuzzy Controller and Separately Excited DC Motor. Also it has very High regulation than the convectional PI Controller method which requires Mathematical Models and Gain adjustments. The better results have been obtained for the proposed system. And have been compared with the results obtained by the two input fuzzy controller and discussed the same. It can be inferred that proposed system has the best results and the whole control system is cost effective as the number of parameter (inputs) has been decreased to one.

References

1. Tir Z, Malik O et al (2017) Implementation of a fuzzy logic speed controller for a PMDC motor using a low-cost Arduino platform. In: 5th international conference on electrical engineering 2017, pp 1–4
2. Muruganandam M et al (2009) Modeling and simulation of modified fuzzy logic controller for DC motor drives. In: International conference on control, automation, communication and energy conservation 2009, pp 1–6
3. Elsrogy WM, Fkirin MA, Hassan MAM (2013) Speed control of DC motor using PID controller based on AI techniques. In: International conference on control, decision and IT (CoDIT), Hammamet, pp 196–201
4. Gubara W, Elnaim M, Babiker SF (2016) Comparative study on the speed of DC motor using PID and FLC. In: Conference of basic sciences and engineering studies, Khartoum, pp 24–29
5. Anuvinda V, Deepa K (2015) Fuzzy based flexible converter for satellite electronics. In: International conference on circuits, power and computing technologies [ICCPCT-2015], Nagercoil, pp 1–5
6. Muruganandam M, Madheswaran M (2009) Performance analysis of FLC based DC-DC converter fed DC series motor. In: Chinese control and decision conference, pp 1635–1640
7. Deepa K, Jeyanthi R, Mohan S, Vijaya Kumar M (2014) Fuzzy based flyback converter. In: International conference on advances in electrical engineering (ICAEE), pp 1–4
8. Karthika P et al (2016) PV based speed control of DC motor using interleaved boost with SiC MOSFET and fuzzy logic controller. In: International conference on communication and signal processing, pp 1826–1830
9. Hart DW (2011) Power electronics. Tata McGraw-Hill Education, New York
10. Cheung YM et al (1996) Motor speed control by using a fuzzy logic model reference adaptive controller. In: Sixth international conference on PE and variable speed drives, pp 430–435
11. Thepsatorn P et al (2006) DC motor speed control using fuzzy logic based on LabVIEW. In: SICE-ICASE international joint conference, Busan, pp 3617–3620
12. Sreelakshmi S, Deepa K (2015) Fuzzy based bidirectional converter. In: 2015 international conference on circuits, power and computing technologies [ICCPCT-2015], Nagercoil, pp 1–6
13. Ross TJ (2010) Fuzzy logic with engineering applications, 3rd edn. Wiley, Hoboken
14. Montecucco A, Knox AR (2015) MPPT converter based on the open-circuit voltage method for thermoelectric generators. IEEE Trans Power Electron 30(2)

States Categorization in India Based on Health Security Capacity with Machine Learning Technique

Ashmita Roy Medha, Malaya Dutta Borah, and Zakir Hussain

Abstract The rising pandemic, COVID-19, has horrendously shaken the entire world, especially India by indicating an enormous number of deaths, affected cases, and a moderate degree of relieved cases. To fight against the infectious disease, WHO (World Health Organization) has proposed for the intermediate direction of essential health parameters. It is important to comprehend the status of health affairs of each state individually with the goal that obligatory advances can be taken to prevent the deadly ailment in India. In this work, we have introduced a few points of health-care parameters for each state in India and furthermore demonstrated the degrees of each state on the basis of health measures and arranged them by using the Machine Learning Technique, K-Means Clustering Algorithm. The states are grouped into six categories and the process of measuring the number of clustered groups is done by Elbow Method. These states are presented by scatter-plot graphs based on various health boundaries. The states can be compared by the clustered points, knowing the position at which cluster a particular state belongs to and what is the status of that state according to the health parameters. Altogether, the study will be a controlling uplifting record for the health movement activists everywhere throughout the nation to limit the harm for the current pandemic and for working up a mainstream development to acknowledge health as a primary right for the citizens of India.

Keywords Machine learning · K-means clustering algorithm · Elbow method · COVID-19 · Categorized states · Health parameters

---

A. R. Medha (✉) · M. D. Borah · Z. Hussain
Department of Computer Science and Engineering, National Institute of Technology Silchar, Silchar, Assam 788010, India
e-mail: ashmita_ug@cse.nits.ac.in

M. D. Borah
e-mail: malayaduttaborah@cse.nits.ac.in

Z. Hussain
e-mail: zakir_rs@cse.nits.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_36

1 Introduction

Coronavirus outbreak is not easy to restraint, as the infection has not been recently recognized in people and the immunization for relieving the disease has not been found at this point. In the time of the major pandemic, it is difficult to give the health privileges to every people of each state in India, as various states have various measurements. One thing that should be possible is to gather the current health-security parameters of each state and make arrangements for actions to give the individuals essential needs and health safety so they don't affect the coronavirus. Thus, the states need to be divided into certain groups as per the health facilities, total cases, and deaths in various diseases, transportation, and other general requirements. For the process which is clarified above, we have gathered information for various health measures and set them up to groups state-wise. We have utilized the K-Means Clustering method for grouping the states and chose the all-out group numbers by the Elbow Method. From our work, the estimation for finding a way to guarantee the health measures for the individuals of each state, which states are developed and underdeveloped, can be predicted. The K-Means Clustering strategy unequivocal that a gathering of comparative information will be shaped into groups dependent on their basic structure. It is commonly the most known and utilized grouping technique. The strategy is for discovering cluster patterns from a data collection that is described by the best comparability inside a similar bunch and the best difference between various groups. The present reality uses of this algorithm are customer profiling, astronomy, market segmentation, search engines, computer vision [1]. The dataset we have prepared, is unlabeled and so we have used unsupervised K-Means Clustering technique to predict necessary measures.

The health parameters for any country has isolated into certain elements. For instance, Crude demise rate, Life expectancy, Infant death rate are gathered in health pointers; Healthcare delivery related, Health strategy markers, Health improvement exercises are the parts of the health framework; Prevalence, Incidence signify Morbidity marker [2]. Alongside the most importantly, India has arranged in some health records, for example, Prevent (Antimicrobial Resistance, Biosecurity, Zoonotic disease), Detect (Laboratory systems, Epidemiology workforce, Real-time surveillance and reporting), Respond (Exercising response plans, Emergency response operation, Risk communication), Health (Healthcare access, Health capacity in clinics, hospitals and community care centers, Communications with healthcare workers during a public health emergency), Norms (Financing, Joint External Evaluation and Performance of Veterinary Services) and so on [3]. As grouping might be possible from various perspectives and can be executed from different quantities of information, we have gathered information on the basis of described health parameters.

The ongoing pandemic has indicated that the country with the world's strongest economy is likewise not completely equipped for mitigating and controlling the pestilence diseases like COVID-19. This obviously shows the greater part of the nations is not completely prepared to handle the health crisis. Most of the countries are missing required health security parameter esteems. On the off chance that we think about the instance of India specifically, it is seen that a portion of the states are progressively equipped for containing the continuous pandemic and some are most certainly not. In this way, our research will concentrate on the order of the states of India dependent on the health security limits thinking about certain parameters. We have applied machine learning technique to categorize the states.

By this work, the concerned will get the thought regarding the limits of preventing, detecting, responding and also the operational readiness for diseases like COVID-19. The targets of this research are:

- To mark out reasonable parameters to guarantee the health security of the countries and so for the states.
- To make a dataset considering existing health security capacities of various states of India.
- To apply machine learning technique to group the states of India based on health security capacities.

1.1 Motivation

The reason for this study is to predict the future prerequisites required for the people of a particular influenced district. As there is an absence of equity developments for each state, an expected diagram can be made for giving the health credits to the people both infected and non-infected of the irresistible malady. The K-Means Clustering methodology has been used in many fields like document clustering, image compression, etc. and also a future prediction can be described. So, the idea has come from the related works of K-Means Clustering method. Categorizing the states with a large dataset and making a prediction, can be done properly by using the proposed method. By this, the exponential development of the demise rate for coronavirus can be diminished progressively. The topic is presently significant everywhere throughout the world and mostly for developing countries like India. To solve the recent issue about COVID-19, this work will be helpful.

2 Literature Survey

COVID-19 is a contagious disease brought about by a newfound coronavirus. The infection and sickness were begun in Wuhan, China, in December 2019. On 30 January 2020, the primary instance of coronavirus disease was found in India. To prevent or battle the deadly infection, India has made numerous strides. As the vaccination isn't designed at this point, a few precautionary measures and fundamentals steps need to take and make the health security framework more grounded and powerful to guarantee that the infection can't spread. Different sorts of researches have been done on this objective and a tremendous measure of information has been gathered from numerous nations to get the prediction and the potential ways not to spread the virus. Kandel [4] presented an overview among 182 nations, a few rates of preventing limits in different levels, showing whether the nations are prepared for 5 indices (prevent, detect, respond, enabling function, and operational readiness) or not. The local assessments are relied upon to comprehend the national status limits comparing to COVID-19. Also, there is a recommendation that COVID-19's nearby like bat coronavirus.

It is fundamental to have a precise expectation of new cases due to COVID-19 with the goal that the vital readiness by emergency clinics and essential activities by the organization can be taken ahead of time.

The quantity of patients has conceded till now, the quantity of cases has been detected and the number of staff and specialists must be calculated to give the individuals equivalent health care safety [5].

To battle with the ongoing pandemic, new innovation and information must be applied. Thus, a few strategies have been taken everywhere throughout the world among them Machine learning technique is conventional. Artificial Intelligence (AI) can be applied with the scales molecular, clinical and societal perspective to estimate the structure of SARS-CoV-2-related proteins, identify existing drugs, potential vaccine targets, improve diagnosis, medical imaging, track for the virus in alternate ways and find the similarities and differences of the growth of the pandemic between various regions [6].

Kapur [7] proposed a perfect way to deal with battle COVID-19 is to deal with the prevention procedure. Detection and isolation are significant parts remembered for the prevention procedure rather than immunization. Also, the new and innovative testing is Pooled Testing which can be driven both in rural and urban regions in India. Likewise, to improve the health facilities, a comparison of the death rate, active rate and recovered rate between India's active states and some other developed and developing countries may be productive. The sound technique for India to go digital is to find how its population consumes, deliver and measures the impact of the health services. The blueprint of NHDB (The National Digital Health Blueprint) looks at managing the privacy and security of the digitized health data [8]. The Global Health Security (GHS) Index is the principal far-reaching appraisal and benchmarking of health security and related capacities over the 195 countries. Quick Rapid Response, Detection and Reporting are the greatest facts in GHS indexes and the lowest are Risk Environment and Prevention. The GHS Index looks to light up readiness and limits holes to increment both political will and financing to fill them at the national and universal levels [9].

A portion of the proposed approaches anticipates the total number of COVID-19 tainted cases, the total of every day new cases, the complete number of deaths and the total of day by day new deaths. The total of cured people is also anticipated. Among them, the prediction of values utilizing the support vector regression model with Radial Basis Function is valuable and can be executed by utilizing Python. From the strategy the activities like Data Pre-processing, Support Vector Regression, Visualization, Model Performance Evaluation, and Prediction are important. In view of the ongoing patterns, future patterns have been anticipated utilizing a vigorous AI model, the support vector regression [10].

While there is an exponential growth of rate of death for coronavirus it is important to evaluate the adequacy of preventive measures, predicting future outbreaks and potential control methodologies, thus, a Bats-Hosts-Reservoir-People transmission fractional-order COVID-19 model can be introduced in which the Picard successive approximation technique and Banach's fixed point theory can be applied for confirmation of the presence and strength standards of the model. Numerical calculations of the information gathered from cases in March 2020 can be used by the iterative Laplace transform method. Moreover, the information is utilized to give a profound understanding of the different boundaries of the model utilizing the Caputo-Fabrizio fractional derivative operator [11].

Jo [12] targeted separating the novel coronavirus disease (COVID-19) spread in South Korea using the defenseless contaminated recouped by the Susceptible-Infected-Recovered (SIR) model. Specialists have gathered the information from Korea Centers for Disease Control and Prevention (KCDC) and acknowledged that each boundary in the SIR model is a function of time with the objective that they can figure significant parameters, for instance, basic reproduction number, all the more gently. Using neural networks, a method has been proposed to locate the best time-varying parameters and the solution for the model simultaneously.

The political economy of social insurance benefits in India has various dimensions. Numerous methods, different sorts of proprietorship patterns and various types of conveyance structures make up an intricate majority that makes the improvement of a composed framework troublesome. Allopathy, Ayurveda, homeopathy, Unani, Siddha, among others, are various frameworks of medication accessible in the nation [13]. The readiness for a medical issue is taken from health spending plan 2019–20, which varied to different states. In contrast with other countries, the state measures of complete lockdown and approving just fundamental exercises are a portion of the opportune estimates that have helped India in gaining an edge against the spread of the pandemic. The national population density middle factor of 395 people for every sq. km (determined with an anticipated population in 2019) like Tamil Nadu, Uttar Pradesh and Maharashtra have recorded more than 1000 cases regardless of a committed expenditure which is higher than average national health expenditure across states. In this manner, the experience of return migration in the work area and occupied industrial spaces in these states can be conceivable and estimated [14].

There is a worldwide proposition, produced for the WHO to all the more likely comprehend health care worker view of infection prevention and control techniques to prevent Covid-19 transmission. A few choices can be accommodated sampling and for information assortment. The developers of this protocol suggest that users embrace the protocol in accordance with what is doable in their context. All administrative and authoritative endorsements should be acquired by the user, and information ought to be prepared in accordance with national information security guidelines. The information assortment apparatus may likewise be refined to guarantee logical propriety. Healthcare parameters ought to be concentrated appropriately and basically look at that are for all intents and purposes valuable and flexible for helping staff accomplish compelling patient consideration by upgrading persistent fulfilment [15, 16].

Lakshminarayanan [17] signified that the enormous harm of both socially and economically for serious medical problems like malaria, tuberculosis, leprosy, high maternal and child mortality and lately, HIV virus, mental illness, road accidents, and so on is a typical image of India. Alongside this, vector-borne, food-borne, water-borne diseases, air quality is the public health risks related to environmental change. Hence, the government of India has given a plan which incorporates the epidemiological evolving condition, demographic transition and environmental changes for Public Health in India and furthermore played a role in establishing effective integration and convergence of health services and influencing architectural revision in the health care delivery system to strengthen the health system by propelling The National Rural Health Mission (NRHM).

Numerous meetings, workshops, training and conferences have been held to fabricate the limit of the public health workforce to prevent, identify and respond to the country over by conducting globally. Not just that, the quantity of creating projects which include training and helping the health experts, health laborers and nurses and furthermore helping in analysis in labs. Additionally, it is indicated that the permeable idea of worldwide land borders expands the possible hazard for the spread of public health threats, for example, the transmission of infectious disease across administrative and political limits [18].

India is worried to decrease the awful effect of COVID-19 and so, the government has made a few strides that will alert the people about the pandemic and control of spreading the infection everywhere throughout the nation. The basic approach is the lockdown. The nation has experienced lockdown 1.0 (25 March to 14 April) lockdown 2.0 (15 March to 3 May), lockdown 3.0 (4 May to 17 May), and is in lockdown 4.0 from 18 May. Up until now, India has seen a ~165× increment in cases notwithstanding the all-inclusive times of lockdown. The adequacy of the lockdown in controlling the spread of the infection and the accomplishment to which the lockdown has picked up the health targets is solid which has smoothed the curve of the exponential growth of the impact of COVID-19 [19]. An exponential growth rate model is expressed dependent on affirmed cases, recovered cases and deaths and furthermore an overview has appeared for the 21 days lockdown in view of COVID-19. The initial curve for affirmed and demise cases developed exponentially, however, for lockdown the curve turned out to be level. The mathematical model was utilized to compute the average reproduction number and herd immunity [20].

3 Proposed Work

3.1 Methodology

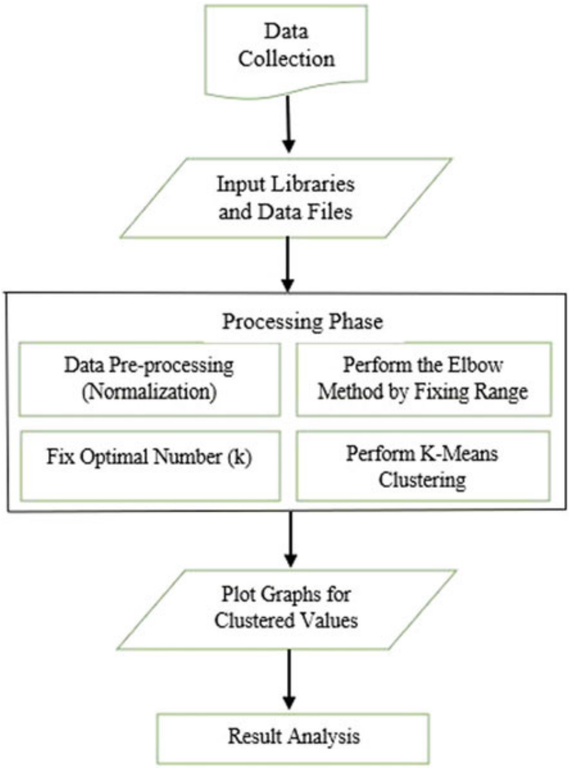
The dataset consists of both qualitative and quantitative attributes. The states are qualitative attribute and the other parameters include in quantitative attribute. Some data are independent and some are dependent. In the pre-processing step, normalization has done to scale the data in same measures and use the values for clustering the states properly.

The clustering analysis strategy is one of the primary expository strategies in data mining. In this work, we have performed the K-Means clustering algorithm [21] on a particular dataset and clustered them according to the result of performing the Elbow Method. A flowchart for the work is given in Fig. 1.

Unsupervised K-Means Clustering

The interest in sorting out the sharp expanding information and taking in important data from information, which makes clustering methods are generally applied

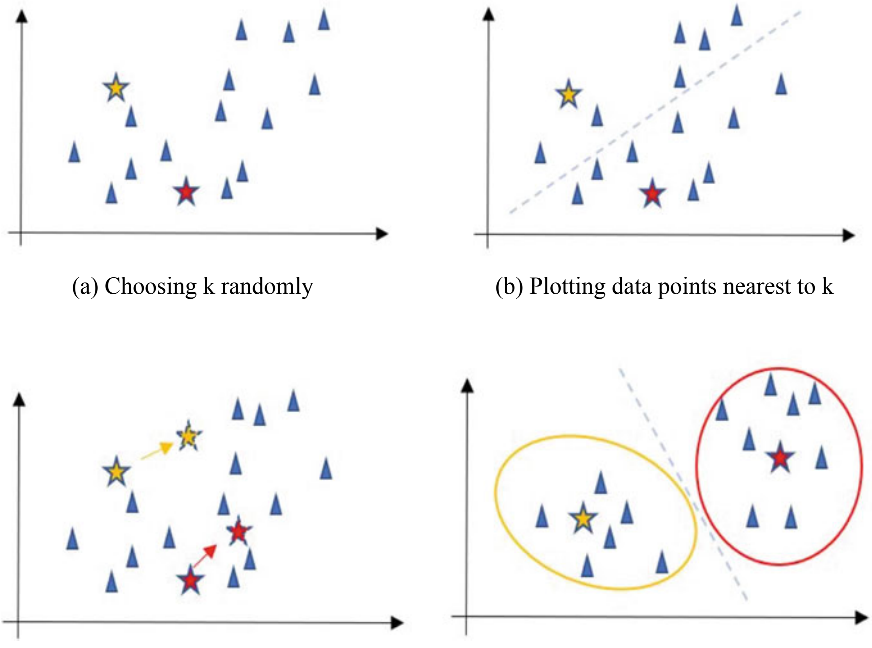
Fig. 1 Flowchart of the proposed work



in numerous applications. K-means is a basic, numerical, unsupervised, non-deterministic, iterative technique. Unsupervised calculations make inductions from datasets utilizing just input vectors without alluding to known or named results [22]. The algorithm can be done in some steps:

1. The process starts with k centroids by putting them at random places. For example, we have taken $k = 2$ centroids randomly (Fig. 2a).
2. The distance of every point from centroid is calculated and clustered accordingly (Fig. 2b).
3. The centroids are then adjusted so that they become the center of gravity for the given cluster (Fig. 2c).
4. Steps 2 and 3 are repeated until none of the cluster assignments change (Fig. 2d).

Now, fixing the optimal value for k is the main issue in k-means clustering. Randomly chosen the number of k can create an issue in clustering. The data points might not be clustered in groups perfectly. That's why, **we can apply the Elbow method to get the dataset clustered properly**.



(a) Choosing k randomly

(b) Plotting data points nearest to k

(c) Putting the centroids by calculating distance from every points

(d) Clustered data points

Fig. 2 K-means clustering process

Elbow Method

The Elbow method is used to find the optimal number of clusters for our dataset. To do this, first, we have imported k-means class from Scikit learn. Then, we have plotted the Elbow method graph after calculating the WCSS (Within Cluster Sum of Squares) for 36 different numbers of clusters, as we have dataset about 36 states. The equation for WCSS is:

$$\sum_{C_k}^{C_n}\left(\sum_{d_i\, in\, C_i}^{d_m} distance(d_i, C_k)^2\right) \tag{1}$$

where, C is the cluster centroids and d is the data points on each cluster.

We set the range for loop from 1 to 36 because of finding the optimal number for clustering. The number of iterations have happened from 1 to 36 and an optimal number has been found where the Elbow point has started. That point is taken as the number in which the states are clustered. The parameters in KMeans class are n_clusters which is the number of clusters we have decided, init which is the random initialization method and we have set here k-means++, max_iter is the maximum number of iterations can be defined into the final clusters when the algorithm is running and we set 500 as we have a big dataset, n_init is the number of times the algorithm will run with different initial centroids and we have kept its default value

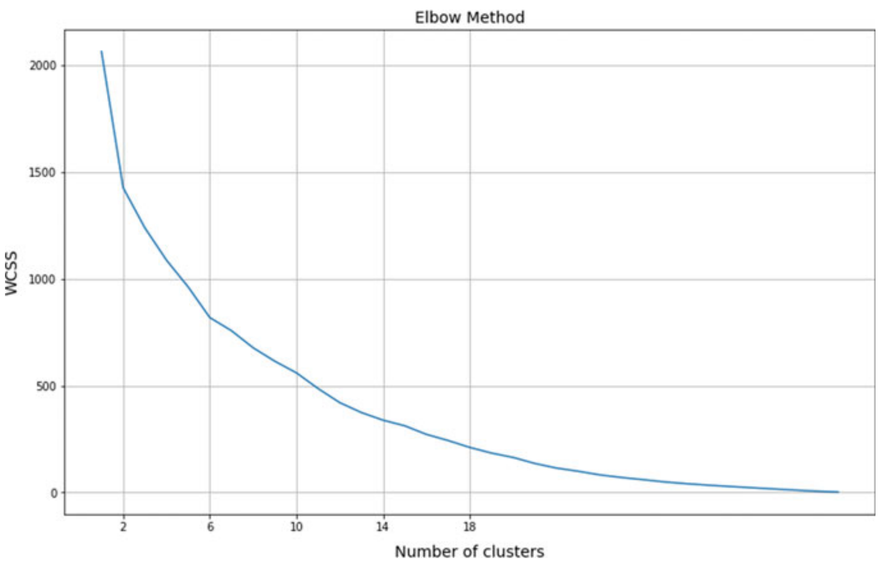


Fig. 3 Finding the optimal number of clusters by the Elbow method

10, random_state fixes all the random factors of the kmeans process. After doing the Elbow Method, we get the optimal number of clusters is 6 as shown in Fig. 3. Then, we find the states following the clusters and plot the dataset according to our choice.

3.2 Prepared Dataset

There are 58 observations of 36 states (28 states and 8 union territories). The observations are numerical data type which include number of TB cases (registered, microbiologically confirmed, treatment success rate) for two different years 2017 and 2018, number of immunization for the years of 2018–19 and 2019–20, number of health works, doctors and nurses, beds available in public facilities, expenditure in medical sector, rural and urban development expenditure, etc. The rows are arranged on basis of different states as shown in Figs. 4 and 5.

	Prevent								Detect		Enabling Capacity				
	Antimicrobal Resistance										Health Infrustructure				
									Laboratory systems		2019				
State	Malaria								Total num	Total number of Lab Tests Conducted from April-2019 to Febru					
				2018 (January - December)											
	Cases	Microbiologically confirmed	Treatment success rate among microbiological confirmed cases	Total patients registered for treatment	Microbiologically confirmed $	Treatment success rate among microbiological confirmed cases #	No. of General Clients detected sero-positive for HIV	Percentage of sero-positive for HIV among General Clients			Total number of Primary Health Centres	Total number of Community Health Centres	Total number of Subdivisional Hospitals	Total number of Public Health Facilties	Total number of Beds available in public facilities
Andaman & Nicobar Islands	505	130	0.84	558	286	0.71	9	0.15	1571256	4,12,857	27	4	NA	34	1,246
Andhra Pradesh (State)	16972	41106	0.89	91124	48070	0.89	16041	1.4	36323960	73,550	1,417	198	31	1,666	60,799
Arunachal Pradesh (State)	1546	1332	0.65	3419	1780	0.57	29	0.17	272174	19,670	122	62	NA	199	2,320
Assam (State)	5281	17334	0.78	42896	20566	0.7	1119	0.67	14177622	45,433	1,007	166	14	1,220	19,115
Bihar (State)	4020	25797	0.72	104930	44161	0.6	8425	1.79	7179949	6,897	2,007	63	33	2,146	17,796
Chandigarh (State)	114	1160	0.87	5696	2570	0.62	590	0.77	7218818	6,83,956	40	2	1	47	3,756
Chhattisgarh (State)	140727	15375	0.89	43031	17097	0.83	2128	0.97	7908434	30,959	813	166	12	1,023	14,354
Dadra & Nagar Haveli (State)	290	226	0.9	849	459	0.84	65	0.33	999592	2,90,825	9	2	1	13	568
Daman & Diu (State)	38	126	0.93	497	222	0.77	29	0.63	256783	1,05,565	4	2	NA	8	298
Uttar Pradesh (State)	32345	115252	0.64	420434	198457	0.57	11482	1.08	1554562	778	3,277	671	NA	4,122	58,310
Uttarakhand (State)	508	6277	0.78	22362	10260	0.71	612	0.73	3040536	30,145	275	69	19	383	6,660
West Bengal (State)	31265	50320	0.86	104145	66585	0.79	4879	0.4	37941436	41,568	1,374	406	70	1,905	51,163

Fig. 4 A small portion of prepared dataset

												Clusters
Delhi(State)	-0.365926	-0.388674	0.404782	0.142432	0.439898	-0.321097	-6.364092e-01	0.148579	-0.316370	-0.311312	-0.390892	0
Goa (State)	-0.364711	-0.741699	-0.755497	-0.762245	0.439898	-0.596365	-5.962909e-01	-0.566946	-0.316370	-0.311312	-0.746872	2
Gujarat (State)	0.241916	0.633023	1.358144	1.559793	0.805351	1.679679	1.410143e-01	4.283741	-0.134763	-0.184410	0.622792	5
Haryana (State)	-0.284067	-0.178935	-0.117813	0.005516	-0.291009	-0.232645	-1.071760e-01	-0.386099	-0.316370	-0.300386	-0.192029	0
Himachal Pradesh (State)	-0.373618	-0.619179	-0.478169	-0.555402	0.927169	-0.575266	-2.569510e-01	0.080097	-0.316370	-0.311312	-0.621332	2
Jammu & Kashmir (State)	-0.371539	-0.475892	-0.601528	-0.621251	0.439898	-0.556718	-5.381967e-01	-0.566946	-0.290294	-0.311312	-0.495877	0

Fig. 5 Clustered dataset

3.3 System Requirements

Software: Operating System: Windows 10.

Python: Jupyter, Spyder.
Tools: Pandas, Matplotlib, Seaborn, Scikit-Learn.

Hardware: Ram: 4 GB.

Processor: 64bit.

4 Experimental Results

From the Elbow method, we have found the optimal number is 6 which can be helpful for clustering the dataset. By this, we have categorized the states into clusters from 0 to 5 numbering as Cluster 1, Cluster 2, Cluster 3, and so on. The graphs shown in Figs. 6 and 7 can be helpful to a comparison between different states health parameters from the dataset.

From the diagrams, a little contrast can be seen among the confirmed cases and the treatment rate for the two specific years (2017 and 2018) which implies that there are not very changes in health parameters in these two years. Likewise, the clustered states can be anticipated clearly by the color contrasts and which states at a high position and which is in the lower position can likewise be assessed. Cluster 2 is the cluster number for state Uttar Pradesh, which indicates that the confirmed cases rate is higher than the treatment rate in both graphs. But it is noticed that the treatment rate has slightly been developed according to confirmed cases in 2018. Thus, from

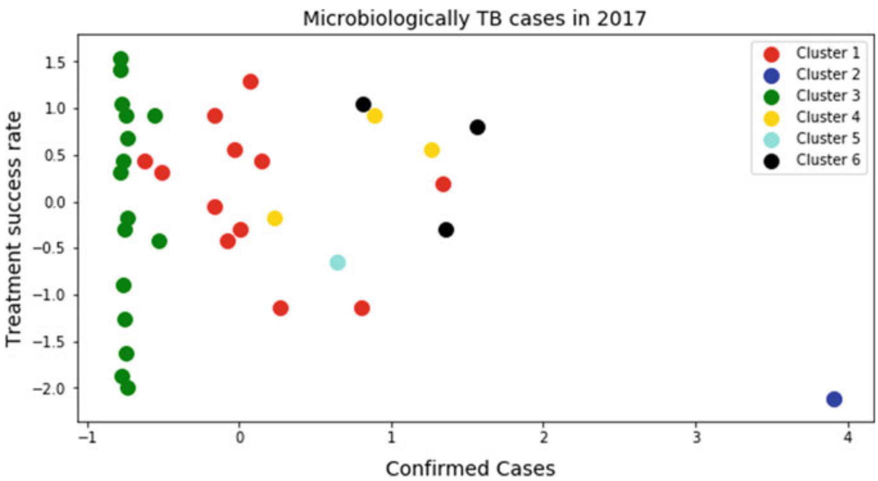


Fig. 6 Microbiologically confirmed cases versus treatment rate for the year of 2017

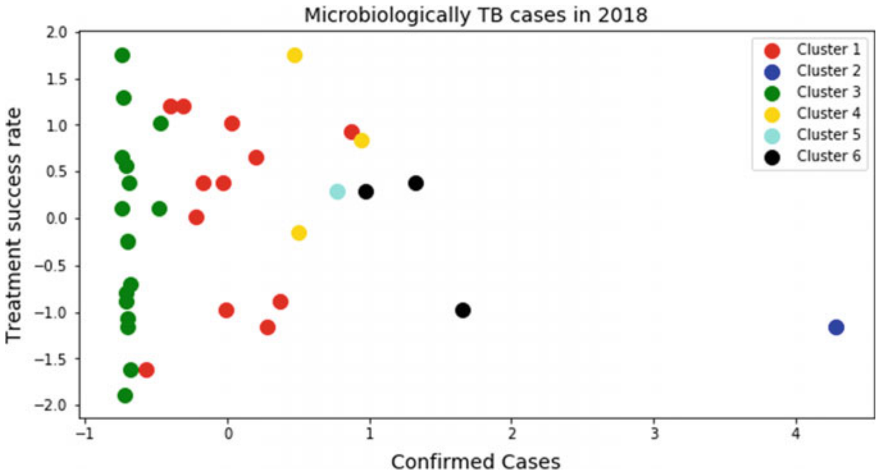


Fig. 7 Microbiologically confirmed cases versus treatment rate for the year of 2018

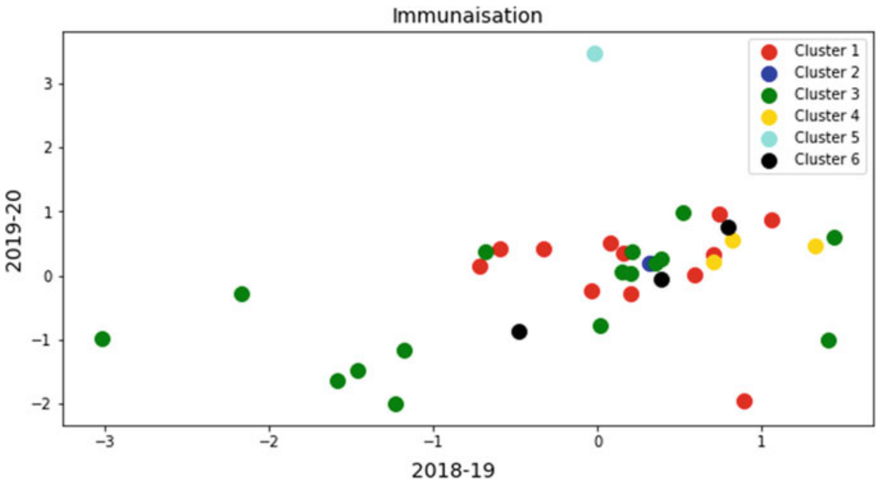


Fig. 8 Immunization in each state for 2018–19 and 2019–20

the Fig. 8, it is clear that fundamental advances need to be taken care to improve the health parameters in the downfallen states, especially in Uttar Pradesh.

In the above graph, the growth of immunization within 3 years, has been shown properly. The only state which has improved at a higher rate is Tamil Nadu as it is grouped by Cluster 5. On the other hand, there are some states which are in Cluster 3, lagging behind on the field of immunization. The backward states need to be taken care of in first stage so that the people of that states may not be affected easily by coronavirus.

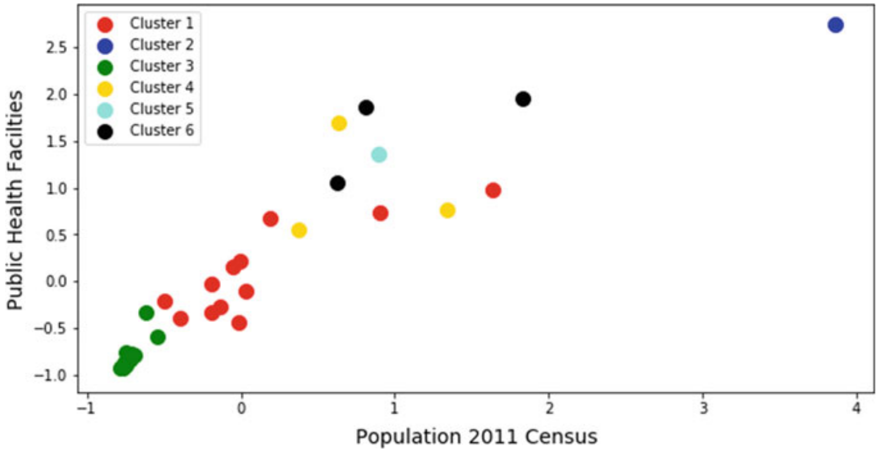


Fig. 9 Population census in 2011 versus public health facilities

Another graph as in Fig. 9 can be shown by doing clustering for the public facilities per population in each state of India. Although the public facilities is not at satisfying rate in conformity with population, the progression can be done to maximize the facilities at least for the poor people. Most of the states are in average level for getting public health facilities, but the states which are included in cluster 1 and cluster 3, need to develop more public facilities in health sector. The overall estimation can be taken to arrange necessary movements for providing basic rights to all people and also to get help from the uplifted states.

5 Conclusion and Future Work

In this work, the motive in utilizing the k-means clustering on a major dataset is to examine the condition of the health parameters state-wise with the goal that the government can take successful developments for each state. The health system management is fragile in comparison to different countries. Then again, India is a developing country thus the monetary development isn't so high. The clustering can assist with anticipating the solidness of each state in healthcare areas. If the people of the feeble condition states get equitable facilities keeping safe from coronavirus, they won't move to other states for clinical issue. The dataset incorporates the year 2020 as well as past years (1990–2019) with the goal that an assumption can be made to improve the current foundation of the health sector just as different sectors.

In future work, other Machine Learning techniques can be applied to get progressively precise plots to anticipate perfectly. To scale the number of affected people and communications with others, how the infection is spreading, a few AI strategies must be applied. Screening test by face checks, utilizing chatbots for screening and conclusion, anticipating fundamental and basic medications all these should be possible by Machine learning procedure. Utilizing cloud security, it is helpful to protect datasets and prevent to give unethical data to researchers and specialists. More information must be gathered on the health factors as well as the economic and social components. The more valuable information, the soon recuperation expectation can be estimated.

References

1. Jain AK, Dubes RC (1988) Algorithms for clustering data. Prentice-Hall. ISBN: 013022278X
2. NITI Aayog, Government of India. Accessed on: May 28, 2020. [Online]. Available: https://niti.gov.in/content/life-expectancy
3. Socio-Economic Statistical Information About India-Health, Indiastat. Accessed on: May 27, 2020. [Online]. Available: https://www.indiastat.com/health-data/16/stats.aspx
4. Kandel N, Chungong S, Omaar A, Xing J (2020) Health security capacities in the context of COVID-19 outbreak: an analysis of international health regulations annual report data from 182 countries. Lancet 395:1047–1053. https://doi.org/10.1016/S0140-6736(20)30553-5
5. Bhatnagar MR, Member S (2020) COVID-19: mathematical modeling and predictions, pp 20–22. http://doi.org/10.13140/RG.2.2.29541.96488
6. Bullock J, Luccioni A, Pham KH, Lam CSN, Luengo-Oroz M (2020) Mapping the landscape of artificial intelligence applications against COVID-19, pp 1–32
7. Kapur K (2020) # Covid19 testing: a new normal, pp 1–6
8. Sheetal Ranganathan, "Towards a Holistic Digital Health Ecosystem in India," *ORF Issue Brief No. 351*, April 2020, Observer Research Foundation.
9. Chang C, Mcaleer M (2020) Alternative global health security indexes for risk analysis of COVID-19, pp 1–18. http://doi.org/10.3390/ijerph17093161
10. Parbat D, Chakraborty M (2020) A python based support vector regression model for prediction of COVID19 cases in India. Chaos Solitons Fractals. 138:109942. https://doi.org/10.1016/j.chaos.2020.109942. Epub 2020 May 31. PMID: 32834576; PMCID: PMC7261465
11. Shaikh AA, Shaikh IN, Nisar KS (2020) A mathematical model of COVID-19 using fractional derivative: outbreak in India with dynamics of transmission and control, pp 1–21. http://doi.org/10.20944/preprints202004.0140.v1
12. Jo H, Son H, Jung SY, Hwang HJ (2020) Analysis of COVID-19 spread in South Korea using the SIR model with time-dependent parameters and deep learning. medRxiv. 2020.04.13.20063412. http://doi.org/10.1101/2020.04.13.20063412
13. Priya R (2005) Public health services in India: a historical perspective. In: Review of Healthcare In India, 1 st edn, Publisher: CEHAT, Editor: Leena Gangolli, Ravi Duggal, Abhay Shukla, ISBN: 81-89042-40-8
14. Shashank B (2020) Health budget vs Covid-19 impact: How are state governments faring ? pp 1–9
15. Hofer U (2013) Novel coronavirus spreads. Nat Rev Microbiol 11:222–222. https://doi.org/10.1038/nrmicro2998
16. Azam M, Rahman Z, Talib F, Singh KJ (2012) A critical study of quality parameters in health care establishment: developing an integrated quality model. Int J Health Care Qual Assur 25:387–402. https://doi.org/10.1108/09526861211235892
17. Lakshminarayanan S (2011) Role of government in public health: current scenario in India and future scope. J Fam Community Med 18:26. https://doi.org/10.4103/1319-1683.78635
18. Newsletter Q (2017) Global health security agenda (GHSA) in India, p 5
19. Krishnan S, Deo S, Manurkar S (2020) 50 days of lockdown: measuring India's success in arresting COVID-19, pp 1–12

20. Dwivedi LK, Rai B, Shukla A, Dey T, Ram U (2020) Assessing the impact of complete lockdown on COVID-19 infections in India and its burden on public health facilities. A situational analysis paper for policy makers. International Institute for Population Sciences, Mumbai, pp 1–20. http://doi.org/10.13140/RG.2.2.35840.87047
21. Maklin C (2020) K-means clustering python example, towards data science, Dec 29, 2018. Accessed on: June 25, 2020. [Online]. Available: https://towardsdatascience.com/machine-learning-algorithmspart-9-k-means-example-in-python-f2ad05ed5203
22. Shi N, Liu X, Guan Y (2010) Research on k-means clustering algorithm: an improved k-means clustering algorithm. In: 3rd international symposium on intelligent information technology and security informatics, IITSI 2010, pp 63–67. http://doi.org/10.1109/IITSI.2010.74

CovidSORT: Detection of Novel COVID-19 in Chest X-ray Images by Leveraging Deep Transfer Learning Models

Srikanth Tammina

Abstract Coronaviruses are a cluster of viruses belonging to the family of Coronaviridae, which infect animals and humans. Coronaviruses related to humans can cause mild disease very similar to common flu, which others cause more severe acute diseases. The new COVID-19, which was first detected in the city of Wuhan in Hubei province, China in December 2019. This new coronavirus that previously has not been identified emerged in China when a cluster of Pneumonia cases was reported. Signs and symptoms include respiratory symptoms namely, chronic mucus, fever, lingering chest pain, stubborn cough, breathing noisily. In more severe cases, the virus can lead to cause pneumonia, acute respiratory distress syndrome (ARDS), and sometimes death. The Health care system and the global economy has been severely disrupted since the Covid-19 pandemic began. An early diagnosis to identify the infection is very crucial to mitigate the stress on the health care system and health care providers. A chest X-ray is performed on patients to detect any inflammation in the lungs of a human. The objective of this paper is to leverage artificial intelligence models coupled with image augmentation techniques to accurately classify the chest X-ray images into two classes namely, Pneumonia and Normal. In this research, a new framework, CovidSORT, is proposed for detecting pneumonia infected lungs using chest X-ray images. The proposed framework is developed using deep transfer learning models namely, Inception-V3, VGG16, VGG19, ResNet-50, DenseNet-121 and MobileNetV2 which were pre-trained on ImageNet which led to quicker model training. Additionally, these models are fine-tuned with image augmentation techniques for better accuracy. The research concludes that the ensemble model built on the majority voting approach from these models in identifying pneumonia has achieved a classification accuracy of 96.83%. The above framework can be used by radiologists to corroborate and identify COVID patients quickly.

Keywords Deep learning · Image augmentation · Machine learning · Transfer learning · Convolutional neural network · Inception-v3 · Image classification · Ensemble · Chest X-ray images

S. Tammina (✉)
Indian Institute of Technology Hyderabad, Hyderabad, Telangana 502285, India
e-mail: ee12b1036@iith.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_37

1 Introduction

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Coronavirus disease can occur as a common cold and can lead to severe pneumonia disease such as MERS, SARS, etc. [1]. World Health Organization (WHO) originally named this contagious disease as Novel Coronavirus-Infected Pneumonia (NCIP) and very recently recommended to name this infectious disease as COVID-19. As of August 01, 2020, more than 18 million confirmed cases have emerged, and more than 700,000 people have succumbed to this infectious disease. The coronavirus COVID-19 is currently affecting 213 countries and territories around the world and 2 international conveyances. In the month of January, the world Health Organization has officially declared a global health emergency [2]. Phylogenetic analysis has proven that the recent virus was most closely associated (89.1% nucleotide similarity) to a group of SARS-like coronaviruses [3].

According to the Centers for Disease Control and Prevention (CDC), Pneumonia is a leading cause of hospitalization among adults and children in the United States with a high medical cost. Pneumonia kills more children than any other contagious disease. In the United States in early twentieth century, it is observed that pneumonia killed 470 of every 10,000 children under 5 years [4]. World Health Organization (WHO) reported that pneumonia killed 808,694 children who are under the age of 5 in 2017. In third world countries like India and Africa, pneumonia still persists in widescale and accountable for major child deaths. In first world countries, 25% of children have a history of pneumonia each year throughout 5 years of their lives [5].

The practice of deep learning has been expanding in the medical field, namely, medical image analysis, cancer treatments and brain tumors etc. [6]. Deep learning is well suited in MIA because it can be used to extract pertinent knowledge from it. Deep learning allows the establishment of start-to-end classification models to achieve state-of-art outcomes using input data, without any manual interaction from humans to obtain characteristics from the data [7]. CNN has been very successful in many image classification problems. Many researchers and scientists applied Convolutional Neural Networks model in various problems such as brain tumor or lesions classification [8], pneumonia detection from chest X-ray images [9], breast cancer detection [10], skin cancer detection [11], and lung image patches with lung interstitial disease [12].

2 Materials and Methodology

Computer vision is a field of artificial intelligence that uses deep learning models to train the computers to accurately identify and classify the object. There are various sets of problems dealt with under computer vision such as image classification, object detection, image colorization, image localization, image reconstruction, and image resolution. Among the defined problems, image classification is touted as the central problem and forms the root for other computer vision problems. The primary challenge in using machine learning to classify images is the manual extraction of features from the input images. This is a tedious and cumbersome process if the input image dataset is huge. In recent years, technology has revolutionized the approach towards image classification problems. Deep neural networks can extract relevant features from the base features through training the weights in the network. While developing any classification or prediction model, one seeks to find the optimum weights $W \in R^M$ and optimal bias of the model $b \in R$ to make accurate predictions.

$$y = W^T \varnothing(x) + b$$

With deep neural networks, one need to know the base feature vector $X \in R^D$ and leave the complicated feature representation $\varnothing$ to the network itself.

$$\varnothing : R^D \rightarrow R^M$$

The hidden layers in the deep neural networks determine more complicated feature transformations. One of the leading deep neural network architectures that focus on image classification problems is Convolutional Neural Network. CNN's are leveraged in many pattern recognition problems, for example, motion recognition [13], facial recognition [14], target classification [15]. A typical CNN structure has 3 consecutive layers, namely, the convolutional layer that extract salient features from the input images with the mathematical operation named convolution, a pooling layer is to reduce the number of parameters in order to reduce the computational complexity yet preserving spatial dependencies and improve the performance, and a fully connected layer which is placed as a classifier based on the feature maps extracted from the pooling layers. Instead of training a CNN from scratch which takes a lot of time to train the whole network, a more practical approach is to construct a model using already pre-trained models on a huge dataset. In this paper, we proposed a methodology that uses pre-trained models, namely, InceptionV3, ResNet-18, VGG-16, DenseNet-12, VGG-19, ResNet-50, and MobileNetV2 coupled with image augmentation techniques to classify COVID-19 chest X-ray images in contrast to normal chest X-ray images.

2.1 Material

In this research, images of chest X-ray both anterior and posterior are carefully chosen from retroactive group of pediatric patients between 1 and 4 years old from Guangzhou Women and Children's Medical Center. All the chest X-ray images are taken from patients as a part of their routine visit to nearest clinical labs. The data folder from this dataset is categorized into 3 folders, namely, train, test and validation. Each folder again classified into two subfolders, namely, Pneumonia and Normal. Collectively, there are 5910 chest X-ray image and 2 classes (Pneumonia/Normal).

Fig. 1 Normal X-ray image depicts without any areas of abnormal opacification

One of the classes i.e., Pneumonia in the dataset is categorized into two more classes (Bacterial/Viral Pneumonia) (Figs. 1, 2 and 3).

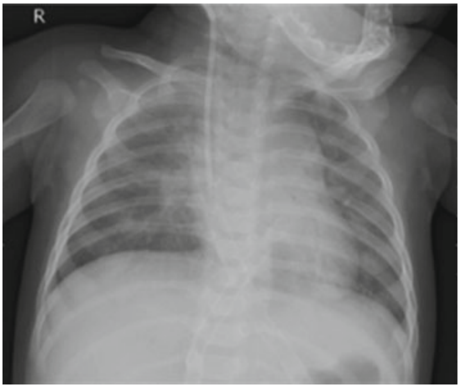


Fig. 2 Depicts viral pneumonia with more diffuse “interstitial” (DIL) pattern

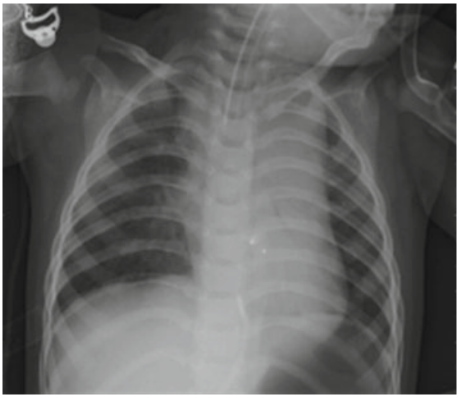


Fig. 3 Bacterial pneumonia exhibits a focal lobar consolidation in the right lung

Table 1 Dataset partition for training, validation and testing

	Category (# of images)			
Classification	Training	Validation	Testing	Total
Normal	1076	300	200	1576
Pneumonia	3172	762	400	4334
Total	4248	1062	600	5910

There are 4334 Pneumonia and 1576 Normal Chest X-ray images in the repository. Out of 4334 X-ray images, 4248 images are assigned for training the models, 1062 images are given for validating the model and 600 images are assigned for testing the model. Subsequently, for the further analysis of chest X-ray images, each chest radiograph has been screened for quality control by eliminating all poor grade quality or indecipherable scans. Consecutively, the images are then sent to two expert physicians for grading before being fed into deep learning model (Table 1).

2.2 Methods

See Fig. 4.

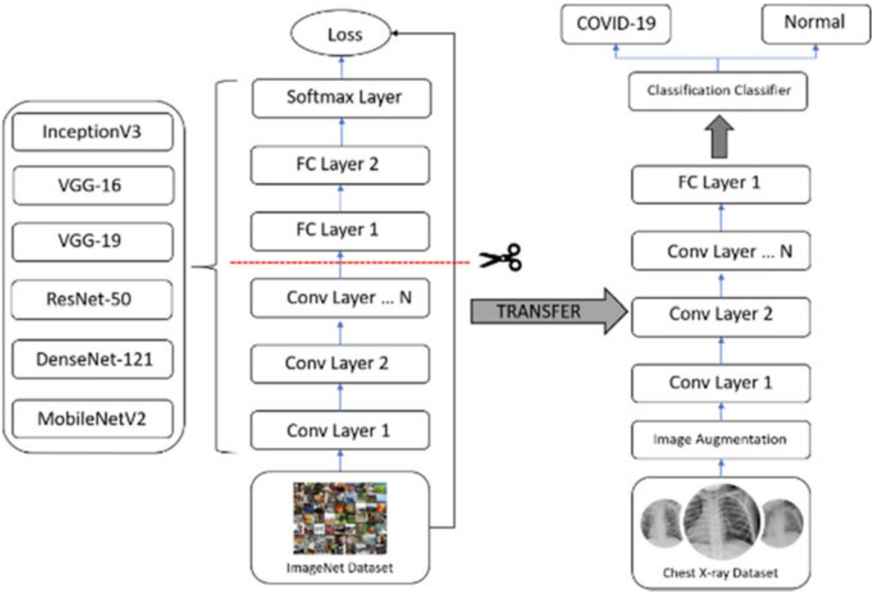


Fig. 4 Proposed CovidSORT framework leveraging transfer learning models to classify chest X-ray images

2.3 Image Pre-processing and Data Augmentation Techniques

ImageNet pre-trained models need an input size of 224 × 224 shape. The testing and validation dataset have been resized into 224 × 224 × 3 and center crop to that shape. Once the images are resized and prepared for our network then data augmentation techniques are performed on the input dataset. Data augmentation, very similar to imagination, that facilitates us to achieve a comprehensive understanding of the world [16]. Similarly, Image augmentation generates new set of images from fewer set of input images. Hence, we can take a inadequate dataset of chest X-ray images and augment them into objects of varied sizes by zooming in and out, shearing, flipping them vertically, or horizontally to develop rich, a diverse dataset with variations. First, we flipped the images horizontally and vertically, then sheared the image by 0–20°. Horizontal axis flipping has proven effective on datasets such as CIFAR-10 and ImageNet. Finally, we added noise which is sampled from standard Gaussian distributions element wise to images. Introducing noise into the images consists of inputting a matrix of random values sampled from the distribution [17]. Adding noise to images has proved useful and assist CNNs to learn robust and relevant features (Fig. 5).

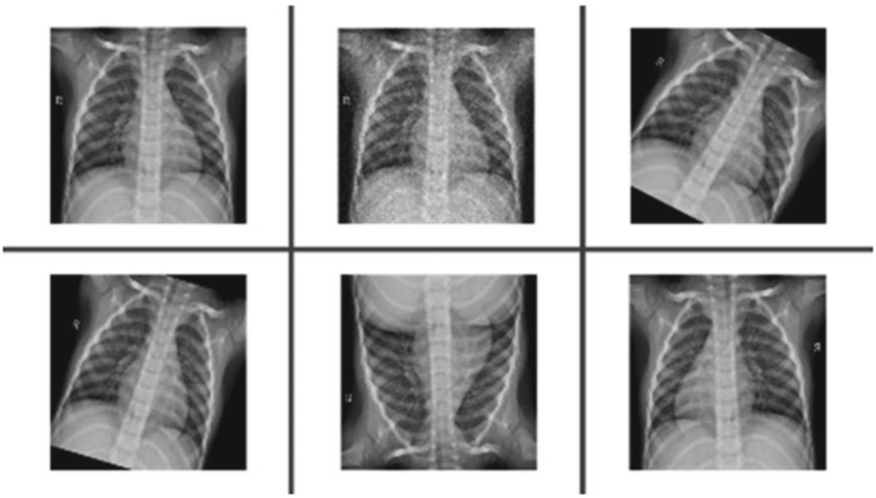


Fig. 5 (From Top clockwise) normal image, Gaussian noise image, rotated image, sheared image, vertically flipped, and horizontally flipped along the axis

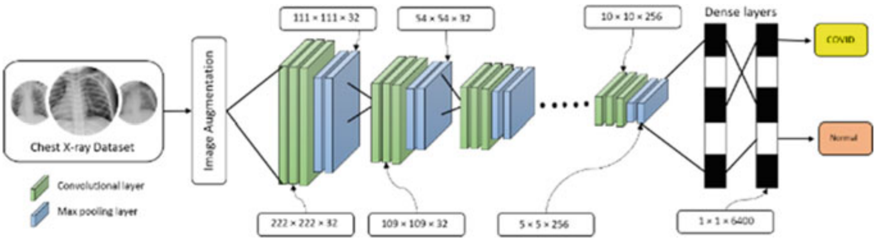


Fig. 6 Plain vanilla convolutional neural network with trainable layers

2.4 Plain Vanilla Convolutional Neural Network

The most prominent network architecture in deep learning models is Convolutional Neural Network. CNN architecture comprised of two entities namely, feature extraction layer (convolutional layer and max pooling layer) and classification layer (fully connected layer). Feature extraction layers outputs pertinent features from the input image and reduce the dimensions of the image yet preserve the most essential features in the image. Classification layer categorizes into two or more discrete classes according to business problem. The top layer consists of set of weight filters (feature extractors) which extracts features of varied dimensions using convolution operation. The max pooling layer is used to reduce the complexity in the network and spatial size of the feature maps in order to reduce the number of weight parameters and complexity in the network. Following that, fully connected layer act as a binary/multi classifier that bins the output into several categories. We use various nonlinear activation functions like Softmax, Relu, leaky Relu and sigmoid functions. The main objective of using nonlinear activation functions is to handle nonlinearity in the network (Fig. 6).

We have developed a Convolutional Neural Network with 7 convolutional layers which has 805,467 trainable parameters. We have used ReLU piece-wise linear activation function after every convolutional layer. As it is a two class classification problem, Sigmoid function is used after a fully connected layer to classify the image into either COVID-19 or normal image. For the loss metric, we employed binary cross entropy and training is performed using Adam optimizer. To tackle overfitting, we initialized a dropout of 0.25.

2.5 InceptionV3

InceptionV3 is a convolutional neural network architecture from an inception family. There are three module layers inside InceptionV3 architecture. Modules are Inception A, Inception B, and Inception C. These inception modules are designed effectively and efficiently to diminish the count of weights parameters in the network yet without decreasing the Inception network efficiency. In each inception module, several small

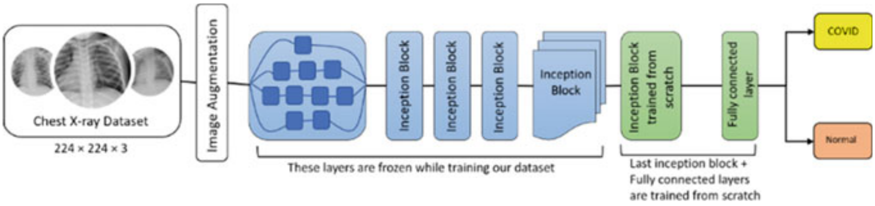


Fig. 7 InceptionV3 pre-trained neural network architecture

convolutional layers with size 3 × 3, 1 × 1, 1 × 3, 3 × 1 are used to factorize the convolutional layers and reduce the number of weight parameters [18]. By factorizing the convolutions, we can efficiently reduce the grid size as compare to AlexNet or VGGNet. In Inception module A, the kernel with 5 × 5 size is replaced with two 3 × 3 kernels. One 5 × 5 kernel produce 25 parameters, whereas, two concatenated 3 × 3 kernels produce 18 parameters which result in a reduction of parameters by 28%. In Inception module B, n × n is further factorized into the combination of 1 × n and n × 1 convolutions [19]. InceptionV3 implements n = 7 for module B, so it is replaced with two 1 × 7 and 7 × 1 convolutions. In Inception module C, 3 × 3 kernel filter is factorized into parallel 1 × 3 and 3 × 1 filters. In this paper, we used a pre-trained InceptionV3 model, froze layer till model_6, and trained the layers from model_7 in the Inception block and trained the fully connected layer from scratch. We first flatten the output layer and used a fully connected layer with 1024 hidden units and the ReLU activation function. Finally, added a sigmoid layer for classification. InceptionV3 required an image shape of 299 × 299, but the image dataset is scaled to 224 × 224 yet the results are satisfactory. RMSprop optimizer is used to train the Inception network with a learning rate = 0.0001 (Fig. 7).

2.6 VGG-16

At ICLR conference in 2015, K. Simonyan et al. from the Oxford University has published a paper "Very DCN Networks for Large-Scale Image Recognition" in which VGG-16, a CNN model, is proposed. This model is trained on 14M images belonging to thousand classes. VGG-16 architecture is an improved version over AlexNet by replacing kernels with multiple 3 × 3 kernel size one after the another. VGG-16 is a pretrained model of 16 weight layers trained on four NVIDIA Titan Black GPUs and it took 2–3 weeks for training. The input to the VGG-16 is the ImageNet dataset with 1000 classes and 14 million images with shape 224 × 224 RGB [20]. This image is sent through pile of convolutional and max pooling layers. The receptive field of these filters is 3 × 3 and the padding is 1 pixel to restore the spatial resolution of the image. Max pooling is applied with a 2 × 2 filter size window, with stride 2. This stack of convolutional layers is passed to three fully-connected layers (FC). This fully connected layer is trained from scratch based on the problem

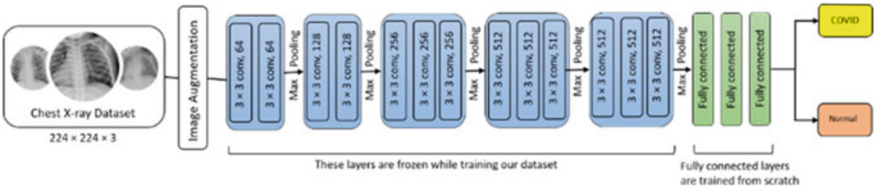


Fig. 8 VGG-16 pre-trained neural network architecture

statement. The first 2 layers have 256 channels each and 3 layers perform 2-way binary classification and thus contain 2 channels (one for each class i.e. COVID-19 and Normal). All hidden layers are stocked with a softmax activation function to induce non-linearity in the network. The dropout is set to 0.5. For this problem, we freeze the top layers and extracted the weights of the network trained on the imagenet dataset and only train the fully connected layer from scratch (Fig. 8).

2.7 VGG-19

VGG-19 is another variant of the VGG model with more depth in the convolutional layers in contrast with VGG-16. VGG-16 and VGG-19 are useful due to the simplicity of the convolutional layers i.e. 3 × 3 and increased depth level in the VGG-19. A total of 19 layers with 16 convolutional layers and 3 fully connected layers have been trained on the imagenet dataset. To reduce the volume size, max pooling layers are used. All layers are 2 × 2 pooling layers with stride 2. During the training phase, convolutional and pooling layers coupled are trained together to extract pertinent features from the input images. For classification in our problem, we freeze the top layers and trained the fully connected layer. The first two layers have 256 channels each and third layers perform 2-way classification and thus contain 2 channels (one for each class i.e. COVID-19 and Normal). All hidden layers are equipped with a softmax activation function to induce non-linearity in the network. The dropout is set to 0.5 to avoid overfitting. For training, we have used the RMSprop optimizer for VGG-19 and Adam optimizer for VGG-16. Weight parameters are transferred from 16 convolutional layers and weights in fully connected layers are trained from scratch (Fig. 9).

Fig. 9 VGG-19 pre-trained neural network architecture

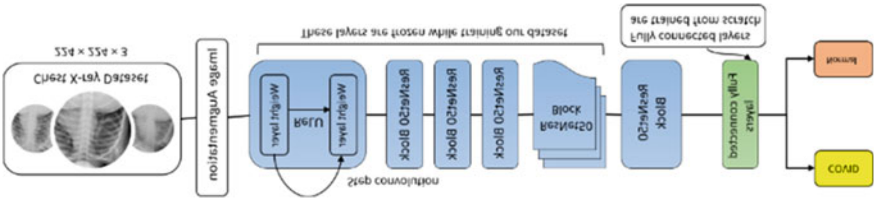


Fig. 10 ResNet-50 pre-trained neural network architecture

2.8 ResNet-50

In the work of Pan and Yang [21], a source domain $D_{Source} = \{X_S, P(X_S)\}$ with task $T_S = \{Y_S, f_S(\cdot)\}$ and a target domain $D_{Target} = \{X_T, P(X_T)\}$ with task $T_T = \{Y_T, f_T(\cdot)\}$ are given with $D_{Source} \neq D_{Target}$ and/or $T_S \neq T_T$. In transfer-learning, the weights learned in D_{Source} and T_{Target} is leverage to learn a prediction function $f_T(\cdot)$ in D_{Target} [22]. ResNet employs a concept of skip connection in the network. In a normal stack of a convolutional layers, we also add the original input to the output of the convolutional block. The advantage of using a skip connection in the network is to mitigate the problem of vanishing gradients by allowing an alternate short cut path for the gradient to pass through the network. The ResNet-50 architecture replaces each 2 layer block in the 34-layer net with the 2 layer bottleneck block which results in 50 layers ResNet [23]. In this research paper, we used ResNet-50 architecture and we employed the off-the-shelf approach. Extracting the feature weights from the ResNet-50 architecture which is trained on the imagenet dataset and we fine-tuned the weights of the last classifier which is a fully dense layer. For training the model, we used the stochastic gradient optimizer with a learning rate of 0.001. Since the problem is a binary classification problem, we used binary cross entropy function to compute the quantity of loss that the model should seek to minimize during training (Fig. 10).

2.9 DenseNet-121

The bottleneck challenge in using Convolutional Neural Network is when the depth of the neural networks increases then the path from the input to output becomes so big, that the information across the network gets vanished before reaching the other side. Thus, Dense Convolutional Neural Network (DenseNet) is proposed by Gao Hung et al. to ensure highest information gush without much vanishing gradients inside the network. In conventional convolutional networks with N layers have N connection between layers. Whereas, in DenseNet neural network have N(N + 1)/2 connections. For each layer in the DenseNet, the feature maps of all the prior layers are utilized as input, and the output of those feature maps are send as inputs into all successive layers [24]. The advantages of using DenseNet are to mitigate vanishing gradient

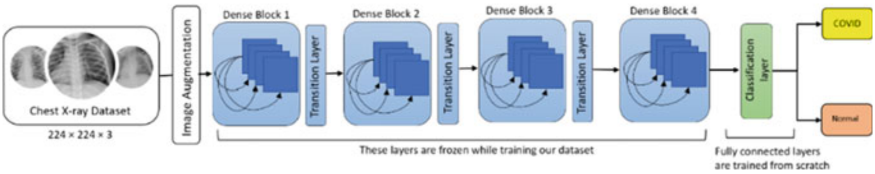


Fig. 11 DenseNet-121 pre-trained neural network architecture

problem, and significantly reduce the number of parameters in the network which in turn reduces the computational complexity. DenseNets are separated into dense blocks and transition layers. In dense blocks, feature maps dimensions will always remain constant whereas, the number of filters will modify between feature maps. In the transition layer, batch normalization is performed for downsampling with 1 × 1 convolution shadowed by 2 × 2 pooling layer. In between two adjoining dense blocks this transition layer is positioned. We employed DenseNet-121 by freezing the top layer and training the fully connected layer from scratch. For training, we used an RMS optimizer with a learning rate of 0.001. Since the problem is a binary classification problem, we used binary cross entropy function to compute the quantity of loss that the model should seek to minimize during training (Fig. 11).

2.10 MobileNetV2

In 2017, a group of Google AI researchers published a paper on new mobile architecture, MobileNetV2. MobileNetV2 uses depth-wise separable convolutions instead of standard 2D convolutions. This depth-wise convolution was introduced in MobileNetV2 to lessen the computational complexity in reducing the quantity of parameters. The fundamental idea is to replace the conventional mathematical convolutional operation with depth-wise convolution. In MobileNetV2, convolution is split into two layers, the first layer, depth wise convolution, we apply a two dimensional depth filter at each depth of input image and second layer, pointwise convolution, is to create new features by using 1 × 1 kernel iterating through every single point [25]. MobileNetV2 has inverted residual blocks in contrast with standard residual blocks used in ResNet-18,50,101. Residual blocks employ a concept of skip connection which connects the beginning and the end of the convolutional block. Whereas MobileNetV2 uses an inverted residual layer which performs three separate convolutions. First, a point wise 1 × 1 convolution is used to map low dimensional feature map to higher dimensional space and ReLU is applied. Next, to achieve spatial filtering, a depth wise convolution is applied using 3 × 3 filters. Finally, a point wise 1 × 1 convolution is used to project back to lower dimensional space. For training the model, we used an RMSprop optimizer with a learning rate of 0.0001. Since the problem is a binary classification problem, we used binary cross entropy function to

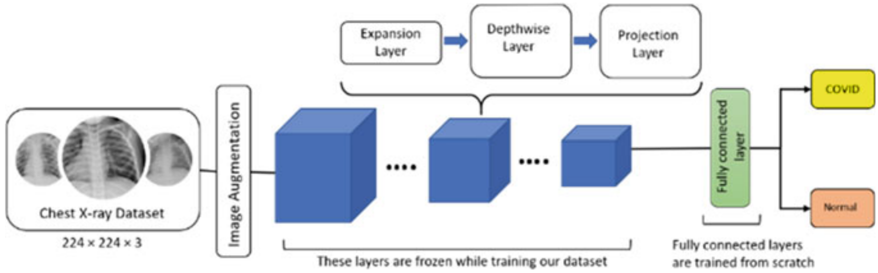


Fig. 12 MobileNetV2 pre-trained neural network architecture

compute the quantity of loss that the model should seek to minimize during training (Fig. 12).

3 Results

We performed experiments on input chest X-ray images to classify COVID-19 and Normal images using six different pre-trained transfer learning models. To combine the effect of all six learning models and predict even more accurately, we employed a voting ensemble classification approach which uses majority voting to conclude the final prediction of test images. We also added plain vanilla convolutional network output into ensemble classification to improve accuracy.

Training and validation of the models are performed on Intel core i7 9th generation processor 8 GB RAM with NVIDIA GEFORCE MX250 GDDR5 graphics. The average training and validation time of each transfer learning model is around 38 min. For training, we used the RMSProp optimizer for VGG-19 and Adam optimizer for remaining models. The binary cross entropy loss function is employed on all the models. Drop out is initialized at 0.5 and learning rate of 0.001 and 0.0001 is given for training. All the models are trained for 100 epochs/iterations. VGG16 has achieved the highest training accuracy of 99.32% and validation accuracy of 98.76%. Among 7 models, ResNet-50 has performed least with training accuracy of 92.8% and validation accuracy of 91.23%. Below shows the plot of training accuracy versus epochs and training loss versus epochs (Figs. 13 and 14).

All the 7 models are validated on test data which consist of 600 images. Below are the comparative table of testing metrics followed by confusion matrix for all 6 pre-trained models (Table 2).

The confusion matrices on the test dataset for six pre-trained models are as follows. It can be seen that among those, VGG-16 model has performed best. VGG-16 has achieved an accuracy of 96% with precision of 95% and Recall of 92.1%. Precision of ResNet-50 model (95.5%) is better than VGG-16 precision. Whereas, Recall of DenseNet121 model (95.69%) is higher compared to other models (Figs. 15, 16 and 17; Table 3).

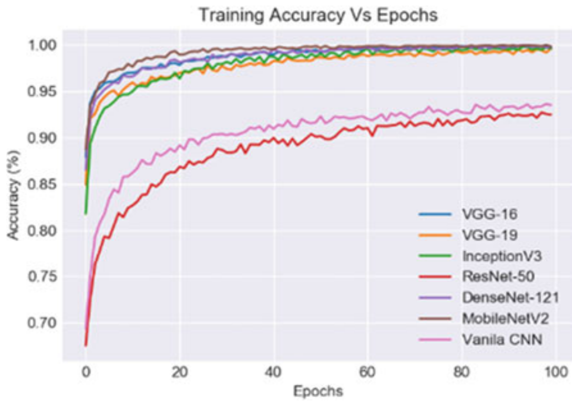


Fig. 13 Training accuracy against epochs (100)

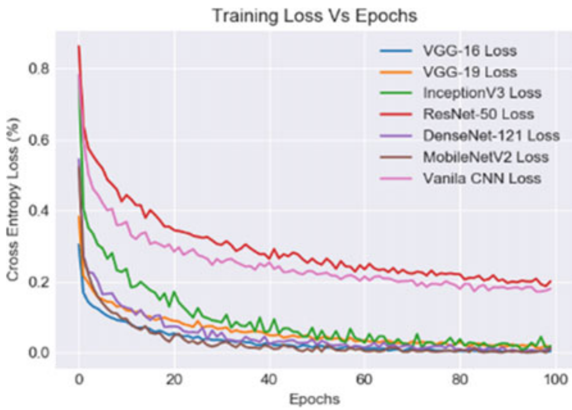


Fig. 14 Training loss against epochs (100)

Table 2 Comparative analysis of metrics of six pre-trained models on test dataset

	Confusion matrix metrics (%)					
Models	VGG-16	VGG-19	InceptionV3	ResNet-50	DenseNet-121	MobileNetV2
Accuracy	96	95.6	92.8	91.6	95	94.66
Precision	95	92	91	95.5	89	89
Recall	93.1	94.8	87.9	81.27	95.69	94.68
F1-score	94	93.4	89.4	87.81	92.22	91.75
Epochs	100	100	100	100	100	100

Deep neural networks used extensively and provide accurate results yet their decision formulating method is uncertain and hard to interpret by analysts. To validate the findings qualitatively, we generated importance maps using RISE (Randomized input sampling for explanation of black box models). RICE generates an importance map which shows which pixel or patch of the image has contributed how much to

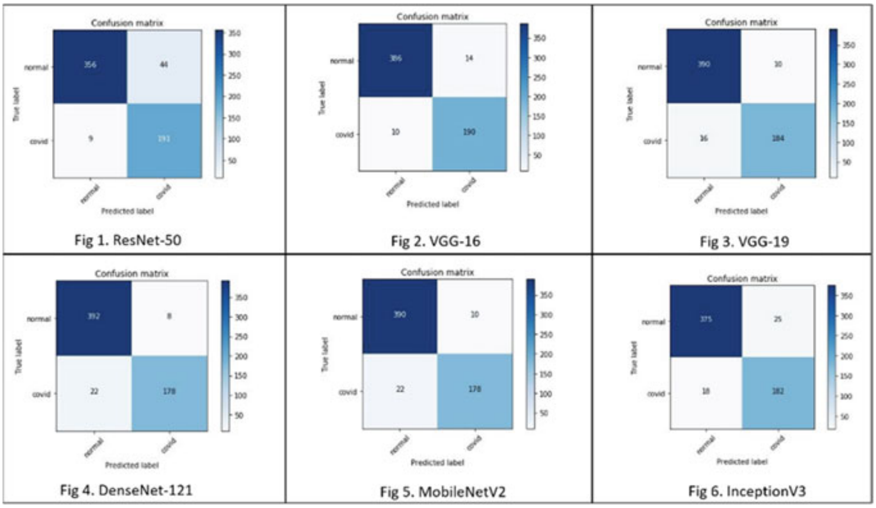


Fig. 15 Confusion matrices of test data images (600) on all six transfer learning models

Fig. 16 Sensitivity and specificity plot for all 6 pre-trained transfer learning models

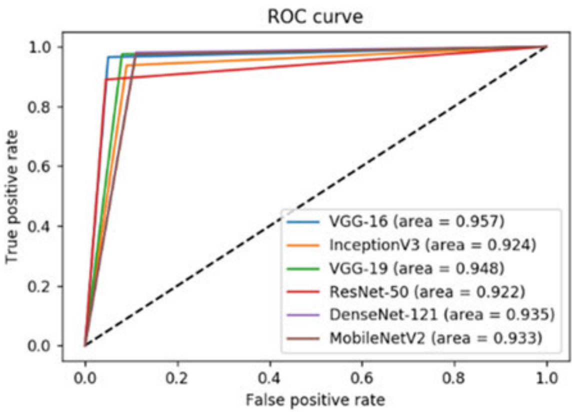


Fig. 17 Confusion matrix of CovidSORT model on test dataset (Accuracy 96.83%)

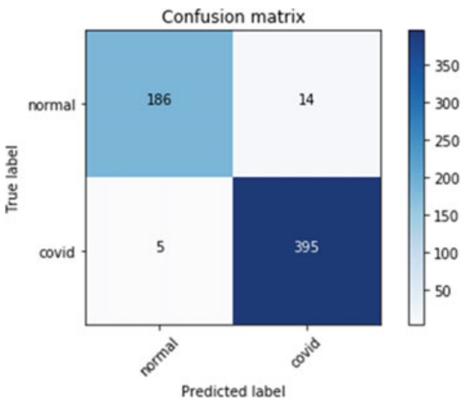


Table 3 Comparative analysis of metrics of ensemble model on test dataset

	Metrics (%)				
Model	Accuracy	Precision	Recall	F1-score	Epochs
Ensemble	96.83	98.75	96.57	97.65	100

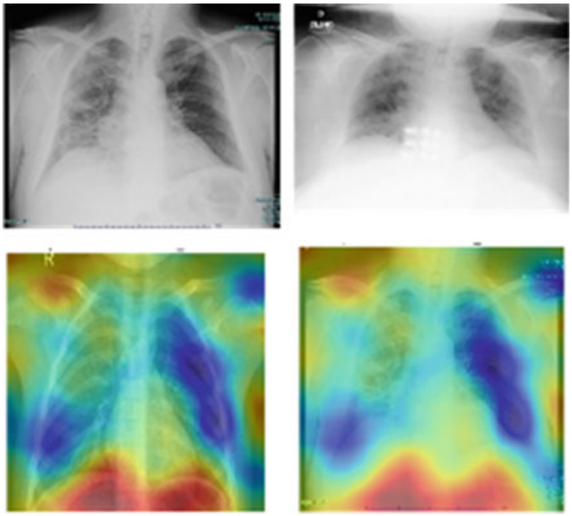


Fig. 18 RISE-generated saliency maps for two representative COVID-19 images

the prediction. RICE calculates the weight of the pixel contribution by randomly masking various versions of the input image and calculate the equivalent outputs [26]. In this approach, measurement of the important region of an image is done by obscuring it by adjusting pixel intensities to zero, distorting the area on the image or including Gaussian noise, and observe how much this affects the neural network. The fundamental idea behind the RISE method is that masks that store semantically salient parts of the image will lead to more accurate classification and hence a higher weight in the final mask for the respective class [27] (Fig. 18).

4 Conclusion

The proposed architecture CovidSORT has attained the highest accuracy of 96.83% by using an ensemble voting approach leveraging 6 pre-trained models. CovidSORT, a deep learning centered model, able to efficiently and effectively detect and categorize COVID and Normal images from a group of chest X-ray scannings. This model can be employed at hospitals by radiologists to support them in the identification of COVID-19 using X-ray chest radiographs. Emerging viral diseases such as SARS, COVID present abundant bottleneck challenges to the health systems, first line workers, government organizations, and citizens of the society. Initially, the gold standard to detect coronavirus in patients is by RT-PCR testing but later there are many complaints that RT-PCR testing has produced less optimal results. Chest X-rays are painless, non-invasive tests and it is most widely favored for the reason that they are quickly available for Pneumonia diagnosis. By using the model proposed in this research, hundreds of images can be classified into COVID-19 and normal in a matter of seconds. Chest X-ray is a rapid and painless imaging test compare to computer tomography (CT). CT releases some amount of radiation during diagnosis which is harmful to patients. Hence, with the combination of chest X-ray images and CovidSORT, it is more efficient in identifying the status of the patient during this global pandemic with monumental rise in the number of diagnoses every day. The CovidSORT model has achieved the highest accuracy of 96.83%, Precision of 98.75%, Recall of 96.57%, and F1-score of 97.65%. This model can be converted into software and installed in the lab computers or can be used in remote places where there is a huge deficit of radiologists to diagnose chest X-ray images. These models can be employed to detect other chest-related diseases namely, Asthma, Bronchiectasis, Emphysema, bacterial pneumonia, and tuberculosis. The greater number of chest images are feed into the model, the better the accuracy will improve. As the usage of this CovidSORT model is increased day by day, the model will be more robust and accurate by feeding further such images from lab reports.

References

1. Ahmad S, Hafeez A, Siddqui SA, Ahmad M, Mishra S (2020) A review of COVID-19 (coronavirus disease-2019) diagnosis, treatments and prevention. EJMO 4(2):116–125
2. Yuen K, Ye Z, Fung S et al (2020) SARS-CoV-2 and COVID-19: the most important research questions. Cell Biosci 10:40
3. Wu F, Zhao S, Yu B et al (2020) A new coronavirus associated with human respiratory disease in China. Nature 579(7798):265–269
4. Preston SR, Haines MR (1991) Fatal years—child mortality in late 19th century America. Princeton University Press, Princeton, pp 4–5
5. Rudan I, Tomaskovic L, Boschi-Pinto C, Campbell H (2004) Global estimate of the incidence of clinical pneumonia among children under five years of age. Bull World Health Organ 82:895–903
6. Suzuki K (2017) Overview of deep learning in medical imaging. Radiol Phys Technol 10:257–273
7. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
8. Talo M, Yildirim O, Baloglu UB, Aydin G, Acharya UR (2019) Convolutional neural networks for multi-class brain disease detection using MRI images. Comput Med Image Graph 78:101673
9. Rajpurkar P, Irvin J et al (2017) Chexnet: radiologist-level pneumonia detection on chest X-rays with deep learning
10. Celik Y, Talo M, Yildirim O, Karabatak M, Acharya UR (2020) Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. Pattern Recog Lett 133:232–239. ISSN 0167-8655
11. Esteva A, Kuprel B, Novoa RA et al. Dermatologist-level classification of skin cancer with deep neural networks

12. Gaál G, Maga B, Lukács A. Attention U-net based adversarial architectures for chest X-ray lung segmentation
13. Bobić V, Tadić P, Kvaščev G (2016) Hand gesture recognition using neural network based techniques. In: 2016 13th symposium on neural networks and applications (NEUREL). IEEE, pp 1–4
14. Lawrence S, Giles CL, Tsoi AC, Back AD (1997) Face recognition: a convolutional neural-network approach. IEEE Trans Neural Netw 8(1):98–113
15. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D et al (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
16. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6:60
17. Francisco JM-B, Fiammetta S, Jose MJ, Daniel U, Leonardo F (2018) Forward noise adjustment scheme for data augmentation. arXiv preprints
18. Szegedy, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2015) Rethinking the inception architecture for computer vision. 1512.00567, arXiv
19. Ganesh Samarth CA, Bhowmik N, Breckon TP (2019) Experimental exploration of compact convolutional neural network architectures for non-temporal real-time fire detection. 1911.09010
20. Simonyan K, Zisserman A (2015) Very deep neural networks for large-scale image recognition. In: International Conference on Learning Representations
21. Pan SJ, Yang Q (2010) A survey on transfer learning. IEEE Trans Knowl Data Eng 22:1345–1359. https://doi.org/10.1109/TKDE.2009.191
22. Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A. Comparison of deep learning approaches for multi-label chest X-ray classification. In: Computer vision and pattern recognition. arXiv:1803.02315
23. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778
24. Huang G, Liu Z, van der Maaten L, Weinberger KQ (2016) Densely connected convolutional networks. 608.06993, arXiv
25. Howard A, Zhmoginov A, Chen L-C, Sandler M, Zhu M (2018) Inverted residuals and linear bottlenecks: mobile networks for classification, detection and segmentation
26. Petsiuk V, Das A, Saenko K (2018) RISE: randomized input sampling for explanation of black-box models. 1806.07421, arXiv
27. Mangal A, Kalia S, Rajgopal H, Rangarajan K, Namboodiri V, Banerjee S, Arora C (2020) CovidAID: COVID-19 detection using chest X-ray. 2004.09803

A Real Time Radio Spectrum Measurement Campaign for Machine Learning Based Spectrum Inference in Cognitive Radio Network

Mudassar Naikwadi and Kishor Patil

Abstract Cognitive Radio (CR) has matured as a technologically viable solution for dynamic spectrum access. Spectrum sensing is the first vital function of CR. Spectrum measurement and modeling coupled with prediction technique improves spectrum sensing efficiency. Machine learning techniques have evolved for predicting future spectrum occupancy information. In this work we have done a real time spectrum measurement campaign targeting frequency range (700 MHz to 2.7466 GHz) at Pune for two week days. We have quantitatively assessed the spectrum utilization in this band for different services like TV broadcasting, cellular, Wi-fi, ISM etc. The overall spectrum utilization in this entire band is very low 16.1227%. This data was analyzed for time and frequency diversity assessment. After correlation analysis of measured data it will be used as training and testing data for a multi-dimensional spectrum inference technique based on machine learning for Cognitive Radio Network.

Keywords Cognitive radio · Spectral occupancy measurements · Spectrum inference · Machine learning

1 Introduction

Cognitive Radio (CR) is a technologically viable solution for conciliating the contrasting situation of increased spectrum demand and spectrum underutilization [1]. Spectrum measurement is the first and vital step for studying the dynamic spectrum usage. In order to have a standardized and uniform platform for acquiring measured spectrum data a comprehensive evaluation methodology is essential [2]. A real time measurement of radio spectrum occupancy percentage gives a valuable insight about the usage of static licensed spectrum being utilized by various popular communication services like TV broadcasting, mobile cellular communication, Wi-Fi, Broadband Wireless Access and ISM etc. [3]. A number of such real time spectrum occupancy measurement campaigns have been conducted worldwide in the past in

M. Naikwadi (✉) · K. Patil
Department of E&TC Engineering, Sinhgad Academy of Engineering, Pune, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_38

countries like Europe, UK, Germany, Spain, Singapore, New Zealand, China, Qatar [4–13] etc. presenting real time understanding of dynamic spectrum usage.

Fundamentally every measurement campaign needs to clearly define the frequency dimension (usually frequency span and frequency points to be measured), time dimension (usual pointers are sampling rate and measurement duration), space (site location), antenna parameters (type, range etc.). The empirical real world spectrum data measured gives a clear picture of spectrum activity but it is difficult to address the time and frequency diversity aspect through it without proper analysis. It is also not practically feasible and economically viable to capture spectrum data from all geographical locations to address space diversity issue. The statistical characteristics of spectrum usage can be determined from long term measurements but such campaigns are not feasible. Machine learning (ML) is a very powerful often heuristic tool that has received increasing attention recently. ML doesn't have any prerequisites or assumptions on data thus in many cases provides higher accuracy than conventional probabilistic and statistical tools. In order to accommodate different dimensions of measured data a multi-dimensional spectrum inference/prediction technique based on machine learning can help solve this problem.

In this work we have used an optimally designed measurement setup with due consideration to proper methodological aspects for occupancy analysis of the measured data. We intend to use this valuable real world spectrum occupancy data analytics for future prediction of spectrum availability. The remaining part of this paper is organized as follows. In Sect. 2 we have explained spectrum occupancy measurement set up and methodology. We have summarized results in Sect. 3. In Sect. 4 conclusions and future work has been outlined.

2 Measurement Setup and Methodology

Our spectrum occupancy measurement campaign setup consists of a discone antenna, spectrum analyzer and data processing and analysis machine (laptop/PC) as shown in Fig. 1. In this work handheld spectrum analyzer Rohde and Schwarz R&S FSH3 [14] has been used which is capable of measuring from 100 kHz to 3 GHz (Table 1).

Fig. 1 Block diagram

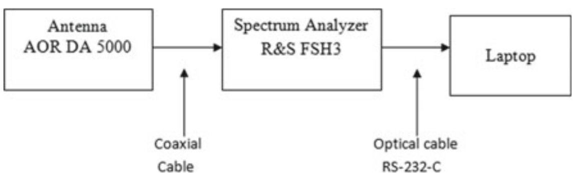


Table 1 Spectrum analyzer configuration used for data collection

Resolution bandwidth (RBW)	200 kHz
Trace detector	RMS
Sweep time	1.5 min
Frequency span	60 MHz
Range	10 dB/V
Reference level	−20 dB
Threshold	PFA method
Preamplifier	ON (gain = 20 dB)
Measurement duration	2 week days
Center frequency	Constantly changing

2.1 Measurement Setup and Campaign Location

The measurement setup used in this campaign consists of a broadband discone antenna AOR DA5000.

In order to sense the signals with proper strengths the antenna was installed on the rooftop Fig. 2 of a commercial cum residential building Konark Puram (latitude: N 18° 28′ 14.88″, longitude E 73° 53′ 23.82″) in Kondhwa, Pune, India. This location has an educational site along with some small scale industries and commercials. This is a strategic location as shown in Fig. 2 having transmitter base stations of various broadcast and cellular services in the vicinity of tens to hundreds of kilometers. The spectral behaviour in both static bands namely TV broadcasting etc. as well as comparatively more dynamic like GSM, ISM etc. has been studied in this work.

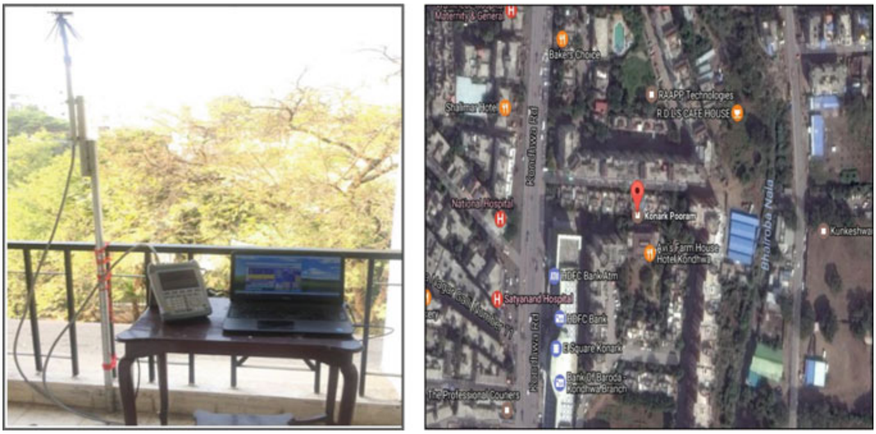


Fig. 2 Measurement set-up and location

2.2 Data Calibration Methodology

The band from 700 MHz to 2.7 GHz spectrum is selected for scanning. For data collection, program is written in the Remote Control software of FSH (R&S FSH-Remote) using the commands of FSH-Remote software. For analyzing the whole band from 700 MHz to 2.7 GHz span of 60 MHz was chosen. This span is further divided into 34 sub bands with a frequency resolution of 200 kHz which gives a total of 301 measurement points in each sub band. The center frequency for each sub band is dynamically selected by program giving a total 10,234 measured values in one complete scan. This complete scan takes about 1.5 min. The program was executed for the number of times the loop is given for complete 48 h (2 days). The files are collected in .csv format. This spectrum scanning for whole spectrum is carried out at the same location for 48 h (2 days) continuously. The noise measurements were taken for 2 h in the same location by replacing the antenna with a matching load of 50 Ω. The data is then processed and viewed for further analysis and comparison. Selecting the right decision threshold to differentiate signal from noise is the most important step towards getting the occupancy data. There are three different methods used to decide the decision threshold and widely employed in spectrum measurement campaigns viz MaxNoise, m-dB and PFA method [2]. After experimentally assessing the impact of the decision threshold using these different techniques, PFA 1% criteria was selected as a optimal decision threshold for computing the occupancy statistics in our campaign.

2.3 Mathematical Analysis of Spectrum Occupancy

Let $\Phi_{f(n),t(k)}$ denote the spectrum occupancy at time sample $t(k)$ and frequency point $f(n)$:

then

$$\Phi_{f(n),t(k)} = \begin{cases} 0, \ldots \rho_r < \tau \\ 1, \ldots \rho_r > \tau \end{cases} \tag{1}$$

where ρ_r is the received PSD measured for frequency point $f(n)$ and at time index $t(k)$ and τ is decision threshold determined by PFA 1% criteria. The average spectrum occupancy Φ of a band is then given by:

$$\Phi = \frac{1}{KN} \sum_{k=1}^{K} \sum_{n=1}^{N} \Phi_{f(n),t(k)} \tag{2}$$

where, N is the number of the frequency points in a particular band while K is the corresponding total number of time samples for a frequency point [8].

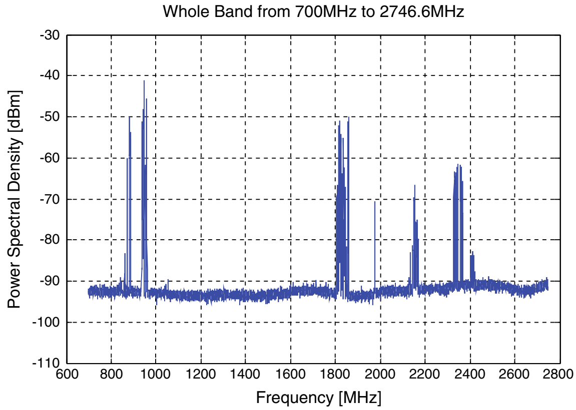


Fig. 3 Power spectral density versus frequency plot for entire band (700 MHz to 2.7466 GHz)

3 Results, Discussion and Analysis

3.1 Frequency Versus PSD Plot

PSD (Power Spectral Density) plot indicates the peak power of the received signal at each discrete frequency. Figure 3 shows the PSD plot for two days reading, this shows that utilization of spectrum is almost similar for two days. As seen downlink frequencies are utilized all the time, because control channels are allotted to downlink frequencies. This location has an educational site along with some small scale industries and commercials. The bands utilized are CDMA, GSM 900, GSM 1800, LTE 2100 ad LTE 2300 cellular systems and ISM. Majority of the spectrum is occupied by the Mobile communication signals. This gives an indication of crowding of spectrum in these bands and under utilization in other services like broadcasting etc.

3.2 CCDF (Complementary Cumulative Distribution Function) Plot

This plot gives graphical representation of the range of possible values of primary user signal and their frequency of occurrence. A CCDF curve can also indirectly give an estimate of utilized spectrum occupancy. Figure 4 indicates that signal lie between −100 and −90 dBm for most of the time. CCDF can give an estimate of signal occupancy in the desired spectrum band.

CCDF plot has been used as measure to differentiate these major telecom services. It gives a graphical visualization of extent of service utilization.

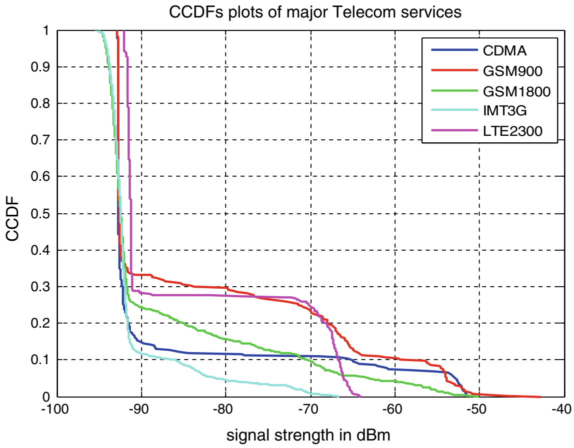


Fig. 4 CCDF plots for major telecom services

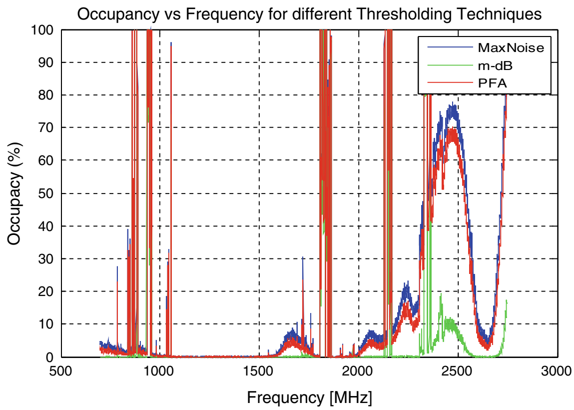


Fig. 5 Occupancy versus frequency for whole band for different thresholding techniques

3.3 Occupancy Versus Frequency with Different Threshold Values Using Different Techniques

It can be inferred from Fig. 5, that spectrum utilization is varied over threshold values and frequency bands. So GSM900, CDMA, GSM1800, BWA and ISM bands heavily utilized for long period of time which can be observed with the help of corresponding bands occupancy.

3.4 Spectrum Occupancy Band Wise Statistics

Different wireless services exhibit a different level of spectral activity as can be readily inferred from observed spectral measurement statistical analysis. Here the

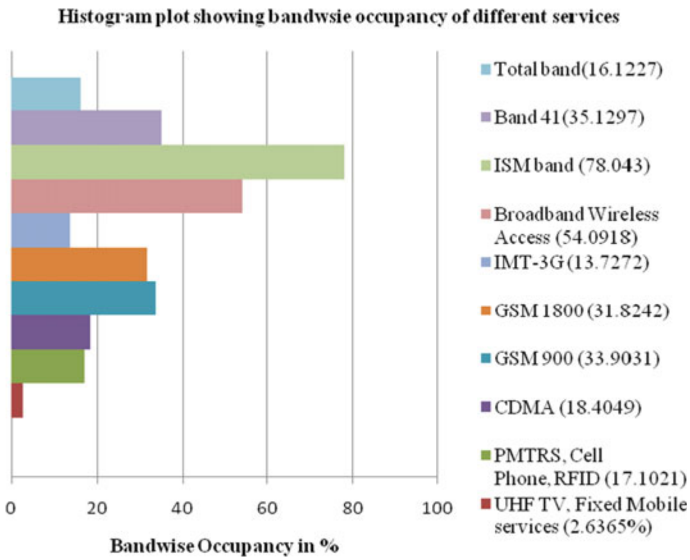


Fig. 6 Histogram plot for bandwise occupancy in %

occupancy statistics band wise by setting threshold level −91.343 dB is illustrated. The histogram plot of Fig. 6 shows visualization in percentage usage of different bands. It can be inferred from the observations and further analysis that spectrum utilization exhibits non uniformity in time and frequency dimensions. Some fraction of observed spectrum exhibits high utilization while other part has under or very low utilization. It can be seen that point to multipoint cellular communication services exhibit a considerable amount of spectrum utilization. It can be easily inferred from observed statistics that there is considerable very low utilization of spectrum band (700–2746.6 MHz) which account for only 16.1227% which means almost 83.8773% spectrum is unutilized or underutilized. Very high spectrum occupancy rate of about average 75% is being observed in Industrial Scientific and Medical (ISM), Broadband Wireless Access (BWA), LTE 2100, 2300 and 2500 bands. Moderate and considerable occupancy is being observed in popular bands like GSM with average of about 30%. The traffic load pattern of PUs in the cellular band is significantly dynamic and non-stationary, this is basically due to enormous number of mobile phone users accessing this GSM band.

4 Conclusions and Future Work

A real world data measurement campaign targeting spectrum band from 700 MHz to 2746.6 MHz was carried out for duration of two week days (48 h). The captured data was processed for occupancy statistics. The occupancy for the entire band was found to be varying with respect to thresholding technique employed. The spectrum occupancy was compared by using three threshold techniques MaxNoise, PFA and m-dB. The PFA technique gives optimal occupancy of 16.1227%. The results show that the spectrum is available for application of cognitive radio as most of bands are vacant and can be used for unlicensed users. This measurement gives useful insights into the spectrum utilization of popular communication services. As a part of future investigation the detailed occupancy analysis band wise in the entire band from 700 to 2746.6 MHz will be carried out taking time and frequency diversity into consideration. This will give an insight into trends in band by band spectrum utilization with respect to frequency and time dimensions. A machine learning based approach for considering time and frequency diversity will be applied for predicting the future spectrum occupancy states. This real world spectrum occupancy data will form the basis input for training and testing of a multi-dimensional spectrum inference algorithm employing machine learning.

References

1. Mitola J, Maguire GQ (1999) Cognitive radio: making software radios more personal. IEEE Pers Commun 6(4):13–18. https://doi.org/10.1109/98.788210
2. López-Benítez M, Casadevall F (2009) Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio. In: 2009 European wireless conference, Aalborg, pp 199–204. http://doi.org/10.1109/EW.2009.5357973
3. Patil K, Prasad R, Skouby K (2011) A survey of worldwide spectrum occupancy measurement campaigns for cognitive radio. In: 2011 international conference on devices and communications (ICDeCom), Mesra, pp 1–5. http://doi.org/10.1109/ICDECOM.2011.5738472
4. Lopez-Benitez M, Umbert A, Casadevall F (2009) Evaluation of spectrum occupancy in Spain for cognitive radio applications. In: VTC Spring 2009—IEEE 69th vehicular technology conference, Barcelona, pp 1–5. http://doi.org/10.1109/VETECS.2009.5073544
5. Wellens M, Wu J, Mahonen P (2007) Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio. In: 2007 2nd international conference on cognitive radio oriented wireless networks and communications, Orlando, FL, USA, pp 420–427. http://doi.org/10.1109/CROWNCOM.2007.4549835
6. Mehdawi M (2016) Spectrum occupancy measurements and lessons learned in the context of cognitive radio. http://doi.org/10.13140/RG.2.1.5117.7367
7. Patil K, Barge S, Skouby K, Prasad R (2013) Evaluation of spectrum usage for GSM band in indoor and outdoor scenario for dynamic spectrum access. In: 2013 international conference on advances in computing, communications and informatics (ICACCI), Mysore, pp 655–660. http://doi.org/10.1109/ICACCI.2013.6637250
8. Agarwal A, Sengar AS, Gangopadhyay R, Debnath S (2016) A real time measurement based spectrum occupancy investigation in north-western India for cognitive radio applications. In: 2016 international conference on wireless communications, signal processing and networking (WiSPNET), Chennai, pp 2035–2039. http://doi.org/10.1109/WiSPNET.2016.7566499
9. Shirgan S, Bombale U (2018) Analysis of wideband microstrip antenna based spectrum occupancy measurement campaign for cognitive radio application. Radioelectron Commun Syst 61:55–63. https://doi.org/10.3103/S0735272718020024
10. Şeflek İ, Yaldız E (2017) Spectrum occupancy measurements at university campus in Turkey. Int J Electron Electr Eng 5:1–6. http://doi.org/10.18178/ijeee.5.1.1-6

11. Islam MH et al (2008) Spectrum survey in Singapore: occupancy measurements and analyses. In: 2008 3rd international conference on cognitive radio oriented wireless networks and communications (CrownCom 2008), Singapore, pp 1–7. http://doi.org/10.1109/CROWNCOM.2008.4562457
12. Xue J, Feng Z, Zhang P (2013) Spectrum occupancy measurements and analysis in Beijing. IERI Procedia 4:295–302. https://doi.org/10.1016/j.ieri.2013.11.042
13. Qaraqe KA, Celebi H, Gorcin A, El-Saigh A, Arslan H, Alouini M (2009) Empirical results for wideband multidimensional spectrum usage. In: 2009 IEEE 20th international symposium on personal, indoor and mobile radio communications, Tokyo, pp 1262–1266. http://doi.org/10.1109/PIMRC.2009.5450192
14. https://scdn.rohdeschwarz.com/ur/pws/dl_downloads/dl_common_library/dl_brochures_and_datasheets/pdf_1/FSH_Technical_Information_en_v11.pdf

Application of Disruptive Technologies in Intelligent Transportation System

P. Santhiya, Immanuel John Raja Jebadurai, and Getzi Jeba Leelipushpam Paulraj

Abstract Intelligent Transportation system (ITS) forms the backbone of smart cities in developed and developing countries. They improve the economy, saves time and provides a safer journey. ITS exploits various technologies viz., Internet of Things, Cloud computing, Fog computing, Artificial Intelligence, Machine learning, deep learning and mathematical models to offer traffic management, route planning, object and lane detection, accident prevention and many more. This paper review various literature that applies these technologies in ITS. İt also discuss various issues and challenges in ITS.

Keywords Intelligent transportation system · Deep learning · Machine learning · Internet of Things · Cloud computing

1 Introduction

Intelligent Transportation system (ITS) offers a faster, efficient and safer mode of transport by smarter use of connected transport network. The services offered include lane and object detection, adaptive cruise control, monitoring driver data and intelligent vehicle navigation [1]. These services are efficiently offered by disruptive technologies viz., Internet of Things, Fog computing, Cloud computing, Artificial Intelligence, Machine learning, deep learning. Figure 1 shows the Gartner chart displaying various emerging technologies of 2019. From the figure, it is understood that most of the technologies are in the innovation trigger or in the peak of inflated expectations. Such technologies go in hand with an Intelligent transportation system and offer sophisticated customer experience in moving from one place to another. Few contributions have been discussed in this paper which will enable researchers to understand the application of these disruptive technologies in ITS.

P. Santhiya (✉) · I. J. R. Jebadurai · G. J. L. Paulraj
Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India

G. J. L. Paulraj
e-mail: getzi@karunya.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_39

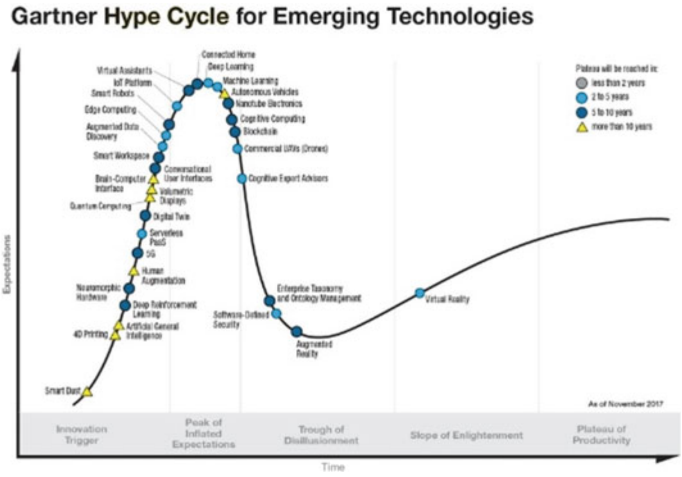


Fig. 1 Gartner cycle for emerging technologies [2]

The rest of the paper is organized as follows. Section 2, reviews various literature that implements technologies viz., Internet of Things [3, 4], Cloud computing [5], Fog computing [6], Artificial Intelligence [7, 8] and Machine learning [9]. Section 3, insights into the issues and challenges faced by ITS. Section 4, concludes the paper.

2 Review on Application of Distruptive Technolgies

2.1 Internet of Things (IoT)

IoT based roadside unit has been proposed in [10]. The proposed architecture is equipped with a speed control unit, pollution based traffic control unit and weather information unit. Various sensor report viz., speed of the vehicle and weather information has been collected using the onboard sensor module and communicated to the central server through the roadside unit. The central server analyzes the sensor data and takes a decision based on the collected value in terms of pollution and speed based traffic control. A prototype has been developed and tested for real-world road traffic. However, the proposed architecture could be improved by proposing a methodology to automatically detect the speed of the vehicle using sensor and image data. A lightweight smart parking technique has been proposed in [11]. The technique aims at allotting the best parking lot by optimizing the distance between the parking lot and the destination. The online broker agent interacts with the other traffic information sources and offers the best route plan. It also reduces the interaction of the driver with the smartphone thereby enforcing driver's safety. An Analytical Hierarchy based approach has been proposed to offer the best parking lot for the user.

In [12], the authors have proposed an IoT based vehicle tracking system using Long Range (LoRa) services). A software tool has been developed which collects data from the microcontrollers using LoRa technology. The gateways relay messages between the end devices and store the messages in the central server. Experiments have been conducted in real-time and the LoRa technology achieves a vehicle tracking system with high accuracy.

2.2 Cloud Computing

A Cloud-based driver shockwave analysis has been proposed in [13]. From the gathered big data of the vehicle, the high threat backward shockwave is predicted and analyzed using predicted backward shockwave analysis approach. The results are processed and analyzed using 3 tier cloud infrastructure. The analyzed result is reported to the vehicles in the high threat areas. Performance analysis has been carried out in error prediction, accuracy, vehicle velocity and some goodput vehicles. In [14], the authors give an overview of data infrastructures for storing and processing the data collected from the transport systems. The authors have proposed Hadoop infrastructure for compute and storage capabilities required for the connected transportation system. The architecture consists of a lower layer that collects data from data science or embedded data pipelines. The machine learning layer models the data using data modeling tools Spark and iterative solvers. İt includes six layers viz., Low-level infrastructures, Data Collection and Ingest, Data processing engines, machine learning and analytics, Data science and connected transport applications.

An accident prevention system using secure cloud and IoT has been proposed in [15]. The emergency section button enables a secure cloud-based virtual eye that captures the videos of the event to the nearest patrons using an ad-hoc network and web server. In addition to this, the audio and video footage are also shared with authorized relatives in terms of emergencies. In addition, the footages are also stored in the database for further investigation.

2.3 Fog Computing

Fog computing offers decentralized low latency services suitable for Gaming, Augmented reality and smart city applications. Vehicular Fog Computing (VFC) deals with intelligent collaboration between vehicles to offer smart services [16]. In [16], moving and parking vehicles are used as communication and computation infrastructures. The mobility models are framed and computation capacity was calculated. The proposed technique showed high accuracy compared to vehicular cloud computing.

A Fog based real-time Intelligent transportation system Big Data analytics has been proposed in [17]. The technique considers real-time analytics and computation of the internet of vehicles. It has three dimensions viz., internet of vehicles, real-time analytics and intelligent computation. The intelligent computation dimension enabled with fog devices offers a computational facility for real-time analysis of big data collected from the internet of vehicles. The real-time analytics dimension has three layers viz., batch layer, speed layer and serve layer. The batch layer collects data from vehicles and exploits cloud technologies for computation and storage. The speed layer is responsible for analyzing the data in real-time. The serve layer answers all the queries of the vehicles in real-time with the help of fog devices. The internet of vehicle dimension has the perception layer that collects data from the vehicles using sensors onboard. The infrastructure network layer has the roadside unit, access point and base station for collectively collecting and analyzing data from the perception layer. The communication layer communicates the data to the cloud for batch processing. The artificial intelligence layer is responsible for taking quick and fast decisions using fog devices. The application and business layer provides a user interface and visualization for the user.

2.4 Artificial Intelligence

An optimized smart traffic governance system has been designed in [18]. The system counts the number of vehicles on every lane on the crossroads. The heavy lane is detected using image processing based tools and the traffic signal is controlled to reduce traffic congestion on heavy lanes. The data is also stored in the database which can be viewed from a well-organized dashboard that displays the vehicle count, user per day and user per month. The website is used by the users to well understand the road that they are traveling and also used by the government to understand and implement road safety policies.

2.5 Machine Learning

Few challenges in autonomous driving have been solved using machine learning models. Long term location prediction for collision avoidance using Gaussian process regression (GPR) has been proposed in [19]. The proposed architecture consists of two modules offline module and a real-time module. The Gaussian process regression model in the offline module is used to preprocess the trajectory data collected from the roadside static sensor. The real-time module uses the output of the GPR to predict the future trajectory of the remote vehicle. The proposed method is experimented with a real-world dataset and shows improved accuracy. Experimenting with the proposed technique using different clustering technique has been suggested as future work.

Autonomous driving conditions are highly uncertain. The driver's intent of the driver is modeled using a partially observable Markov decision process (POMDP) in [20]. The other vehicles serve as hidden variables. The solution for POMDP is an optimal acceleration path of the vehicle considering the intension of the other vehicles. In the first stage, the algorithm is converged using additional search heuristics. In the second stage, more uncertainties are included to form more planning. In the third stage, simulations show that the proposed technique formulate an optimal conservative path even in intersections.

3 Issues in Next Generation ITS

Intelligent trasnportation system has wide applications that enable safer and smarter transportation. However, it can be improvised by addressing few issues and challenges discussed below:

- A huge amount of data is being generated by ITS. The data needs to be analyzed for prediction, classification and detection applications. This requires the data to move from vehicles to roadside units. Compute intensive data are migrated to the cloud for storage and computation. This enables data leakage, man in the middle attack, spoofing and denial of service attacks. Various approaches have been proposed by literature. However, with the emerging trends of attacks, foolproof techniques have to be developed.
- Most of the AI, machine learning and deep learning algorithms that are implemented in applications viz., automated cars and self-driving cars require high computational hardware architectures such as GPUs which is expensive. This increases the necessity of lightweight learning algorithms that could run in an embedded microcontroller. This improves the size, cost and energy efficiency of the applications.
- Connectivity plays a major role in ITS. The fast-moving vehicles require uninterrupted connectivity to the vehicles and roadside units. Routing algorithms that consider mobility, handoff issues have to be developed for quality of service in ITS applications.

4 Conclusion

This paper focuses on discussing the application of disruptive technologies viz., Internet of Things, Cloud computing, Fog computing, Artificial intelligence, machine learning and deep learning in the Intelligent transportation system. Few works of literature have been discussed in each area that includes applications viz., smart traffic management, smart parking, route planning, event detection, object and lane detection. Issues and challenges have also been discussed that include security, design of embedded algorithms and design of quality of service routing algorithms.

References

1. T. U. S. National and H. Traffic, Progress and challenges in intelligent vehicle area networks
2. Koon J (2019) The top 7 trends that will drive embedded technologies in 2019. Fierce Electronics. [Online]. Available: https://www.fierceelectronics.com/embedded/top-7-trends-will-drive-embedded-technologies-2019
3. Chand V (2019) Survey on the role of IoT in intelligent transportation system, no September 2018, pp 936–941
4. Sookhak M, Gani A, Buyya R (2013) A survey on vehicular cloud computing. J Netw Comput Appl, 1–20
5. Darwish TSJ, Bakar A (2018) Fog based intelligent transportation big data analytics in the internet of vehicles environment: motivations, architecture, challenges, and critical issues. IEEE Access **6**:15679–15701
6. Javadzadeh G (2019) Fog computing applications in smart cities: a systematic survey. Wirel Networks 0123456789
7. Li J, Cheng H, Guo H, Qiu S (2018) Survey on artificial intelligence for vehicles. Automot Innov 1(1):2–14
8. Yuan T et al (2019) Harnessing machine learning for next-generation intelligent transportation systems : a survey
9. Veres M (2019) Deep learning for intelligent transportation systems : a survey of emerging trends, pp 1–17
10. Al-Dweik A, Muresan R, Mayhew M, Lieberman M (2017) IoT-based multifunctional Scalable real-time enhanced road side unit for intelligent transportation systems. Can Conf Electr Comput Eng, 1–6
11. Rizvi SR, Zehra S, Olariu S (2019) ASPIRE: an agent-oriented smart parking recommendation system for smart cities. IEEE Intell Transp Syst Mag 11(4):48–61
12. Salazar-Cabrera R, Pachón De La Cruz A, Madrid Molina JM (2019) Proof of concept of an IoT-based public vehicle tracking system, using LoRa (long range) and intelligent transportation system (ITS) services. J Comput Networks Commun 2019
13. Chang BJ, Chiou JM (2020) Cloud computing-based analyses to predict vehicle driving shockwave for active safe driving in intelligent transportation system. IEEE Trans Intell Transp Syst 21(2):852–866
14. Luckow A, Kennedy K (2017) Data infrastructure for intelligent transportation systems. Elsevier Inc.
15. Mishra KN (2020) A novel integration of smart vehicles and secure clouds for supervising vehicle accidents on roads/highways. Sādhanā 0123456789
16. Hou X, Li Y, Chen M, Wu D, Jin D, Chen S (2016) Vehicular fog computing: a viewpoint of vehicles as the infrastructures. IEEE Trans Veh Technol 65(6):3860–3873
17. Darwısh TSJ, Bakar KA (2018) Fog based intelligent transportation big data analytics in the internet of vehicles environment: motivations, architecture, challenges, and critical issues. IEEE Access 6
18. Sukhadia A, Upadhyay K, Gundeti M, Shah S, Shah M (2020) Optimization of smart traffic governance system using artificial intelligence. Augment Hum Res **5**(1) (2020)
19. Goli SA, Far BH, Fapojuwo AO (2018) Vehicle trajectory prediction with gaussian process regression in connected vehicle environment. In: IEEE intelligent vehicles symposium, June 26–30, pp 550–555
20. Hubmann C, Schulz J, Becker M, Althoff A, Stiller C (2018) Automated driving in uncertain environments: planning with interaction and uncertain maneuver prediction. IEEE Trans Intell Vehicle 3

Performance Analysis of Software Enabled Accelerator Library for Intel Architecture

Gaurav Mohindru, Koushik Mondal, and Haider Banka

Abstract With the advancement of the General Purpose Graphics Processing Unit (GPGPU) for parallel processing, the Central Processing Unit (CPU) based architecture industry focused its development on software accelerator based performance maximizer so as to provide the parallel environment in a software way. This set of accelerator libraries support all stages of data science or data analysis pipelines viz. preprocessing, transformation, analysis, modelling, validation, and decision-making. The recent trends of large dataset based computation models are to use all the available features to create rule-bases for validation rather than feature summarization or reduction based on subset selection statistical techniques. Deep learning allows these computational models that are composed of multiple processing layers to learn and train from the large dataset with multiple levels of abstraction without reduction in features. Intel's Data Analytics Acceleration Library (Intel DAAL) enables applications to process large datasets to make faster and better predictions with available Intel processors in a lab environment. In this paper, we will discuss the enabling factors of Intel's DAAL and how it optimizes the performance in a Python computational environment with and without Anaconda frameworks.

Keywords GPGPU · Intel DAAL · CPU · Deep learning · Data analysis · Anaconda · Outlier detection · Performance evaluation

1 Introduction

CPU-based industry, in the last decade, extensively worked on optimization of performance analyzers for producing better throughput with the available architectures. The different available options viz. compiler switches, call graph, sampling, proofing,

G. Mohindru · H. Banka
Department of Computer Science and Engineering, IIT (ISM), Dhanbad, India
e-mail: haider@iitism.ac.in

K. Mondal (✉)
Department of Computer Centre, IIT (ISM), Dhanbad, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_40

block ordering, handling data cache, etc. are used to boost up the parallel performance of the processors. Thread Building Blocks (TBB) library is used for shared memory parallel programming and heterogeneous computing and the MPI library is used for distributed memory programming to perform better on cluster computing environments. Subsequently, for handling large datasets for in-memory computation, the canonical addressing scheme was put in place for 64 bit architecture along with linear and physical addressing [1]. These all techniques [2] are essentially used to handle big data induced Machine Learning (ML) and Internet-of-Things (IoT) based applications. The Integrated Performance Primitives (IPP) library is used for SSE2, SSE3 instruction sets. This library includes more than 500 computer vision, 300 cryptographic, 1300 signal processing, and 2500 image processing primitives that are optimized for CPU performance [3]. Intel's Math Kernel Library (MKL) is another performance optimizer that features vectorized and threaded functions to maximize performance for available CPU architectures. Even after all efforts, the number of cores per CPU is still 64 with 128 threads with maximum throughput close to 100 GFlops whereas a single GPU has 2500 processing cores with 1000 TFlops of speed. To enable powerful performance through the software library, Intel, in its latest February 2020 release, introduced Data Analytics Acceleration Library (DAAL) that helps end-users to create limitless applications in ML and IoT domains that deliver impressive performance in CPU architecture, especially in Intel hardware [4]. To provide a robust analytical framework, Anaconda supports Intel distribution of Python and Intel Performance Libraries in Conda package manager and Anaconda Cloud [5]. The DAAL library includes highly optimized analytical and machine learning functions that simultaneously ingest data and compute the output for maximizing the throughput and using the same Application Programming Interface (API) for application development on different operating systems. These APIs also allow to communicate easily with node to cluster to edge to support end-to-end analytics in a better way. Intel-based Movidius Neural Compute Stick (NCS) is a prominent edge device widely used in image classification [6]. The cross-device communication technology of Intel's DAAL API helped the end-users to build an end-to-end framework effectively. DAAL library also consists of three major components viz. data management; algorithm and services to handle machine learning-based applications efficiently. In this present work, we will discuss how to integrate Intel DAAL and Intel Python in the Anaconda framework and will compare its throughput with bare metal python in the machine.

2 Literature Survey

In post-2006, the profounding break down of scaling effects of Moore's Law, Dennard Scaling, and Koomey's Law regarding doubling the number of transistors used in a chip, constant power density of a chip, and rate of doubling of transistors in a chip is about 1.57 years respectively resulting inability to increase clock frequencies [7], forced to engage the CPU makers to develop multi-core processors. Due to that, Instruction Level Parallelism (ILP) is slowly replaced with Data Level Parallelism (DLP), Task Level Parallelism (TLP), and Request Level Parallelism (RLP). To exploit the DLP, the CPU-based industry focused on software accelerator-enabled applications with different programming frameworks. DLP is usually supported by Single Instruction Multiple Data (SIMD) architecture and it is an energy-efficient option as a single instruction can initiate multiple operations at one go. The advancement of the Intel DAAL library basically supports this SIMD architecture at large with different implementation variations. A semi-stochastic parallel mathematical model, which is neither purely stochastic nor purely batch, works fine with SIMD architecture-based programming model for handling data-level parallelism [8]. The recently launched Comet Lake series (i9-10900K) also supports this working principle for speed up through hardware-supported software libraries. Intel distribution of Python is one such package under the Anaconda framework for speed up of core computational packages to achieve faster python application performance by accelerate Numpy, SciPy, scikit-learn, mpi4py etc. After integration with MKL and DAAL. Machine learning developers used to take advantages of faster scikit-learn packages for different classification and clustering algorithms. High-Performance Analytics Toolkit (HPAT) enables the researchers to speed up pandas [9] and NumPy packages with a compiler-based framework. The DAAL software accelerator library not only speed up the performance of the developed model but also helps in preprocessing of the data viz. handling missing and noisy value, identify and manipulate dates, normalize and scale data, fix inconsistencies and typos, read files with encoding problems, etc. through data management component. This data management segment helps to quantize a model from fp32/bf16 to int8 and to analyze the performance speedup among different data types. The algorithm segment consists of modules that implement algorithms for machine learning, data mining, and data modeling in three computational modes viz. online processing, batch processing, and distributed processing. Different distributed algorithms [10, 11] in the Intel DAAL library vary from matrix decomposition, classification, and clustering, regression to association rules that are work seamlessly in cross-device platforms either with Hadoop/Spark cluster or with MPI based cluster environment by maintaining all low-level data-exchange protocols. The services component of the Intel DAAL library includes utilities and classes that are used across Data Management and Algorithms components. These classes help in memory allocation, error handling, and shared pointers management. Intel, in its latest System Studio software suite, allows end-users to integrate Intel DAAL for developing quick prototyping of big data, machine learning, and IoT-based applications under this all-in-one development tool suite. The 9th and 10th generation processors are also integrated with Intel Performance Maximizer which enables to boost up the speed up process by enabling overclocking hyper-intelligent tool and supported by the Performance Tuning Protection Plan (PTPP). Anaconda framework, along with all these features of Intel CPU architecture and accelerator libraries helped us to build and speed up the development of different applications in big data, IoT [12], and machine learning domains [13].

Speed up in the performance of parallel computation is largely dependent upon how much of the job must be performed sequentially. Amdahl's Law characterizes the speedup S that can be achieved by n processors collaborating on an application where p is the fraction of the job that can be executed parallelly, $S(n)=1/((1-p)+(p/n))$. The main challenge to speed up in processing large scientific datasets is mainly dependent upon the data movement between disk and main memory and parallel and non-parallel blocks in a program [14]. To reduce the data movements between main memory and disk, creating and managing standard datasets for machine learning environments is crucial. Intel's DAAL software framework, by binding all the three components viz. Data Management, Algorithms, and Services reduced the data movement to a minimum so as enable the speed-up process in execution. There are structured, unstructured, stream, and multi-modal datasets publicly available (like UNICEF dataset, World Bank dataset, etc.) which helped us to assess our model's effectiveness, in terms of training and validation, for minimizing the error which may be less than the pre-defined threshold value. Due to the advancement of the machine and deep learning-based data modeling, another important aspect that is crucial during model selection or design is the tuning of hyperparameters. Hyperparameters are the parameters, like learning rate, number of layers, number of clusters, etc., which govern/control the whole training process. It is either model-specific hyperparameters or optimizer hyperparameters. The learning rate, number of epochs, batch size, etc. fall under optimizer hyperparameters, and the number of hidden units, first hidden layer, and number of layers comes under model-specific hyperparameters [15]. Grid search, Random search, and Bayesian optimization are the popular techniques available for hyperparameter optimization. The Intel distribution of Python with scikit-learn, Keras, TensorFlow, etc. packages helped a lot in designing hyperparameters tuning and outlier detections [16]. The hyperparameters that not only perform well with testing data but also offer the best result with cross-validation, should be taken as optimal hyperparameters. The developers need to avoid the local minima of the objective function. The transfer learning methodology with fine-tuning of the hyperparameters is a popular way of finding optimized results. Intel DAAL-induced gradient boosting can achieve 6.5 times faster results, in comparison with XGBoost library, on the same training data. The general gradient-boosting decision tree algorithm is computationally intensive and resource expensive [17] when it is dealing with large datasets and continuous features. Intel DAAL provides a highly tuned implementation of gradient boosting algorithm for classification and regression problem domains. It uses vectorization and parallelism in tree construction for optimal performance.

3 Proposed Work

In this paper, we have executed a basic linear regression model with scikit-learn and with DAAL library and compared the result in terms of execution time. The Intel DAAL is installed in the Anaconda framework. Also, we have compared the result of outlier detection with and without Intel DAAL. The basic methodology for API or software-based computational framework for deep machine learning-based scalable data processing is evolving around following steps [18, 19]:

- Select a class of model by importing the estimator class;
- Data arranged in a feature matrix and set the target vector class for desired output;
- The model hyperparameters are selected with the desired set of values;
- Split the data into sets as train, test, and validation;
- Fit the model with training data by the general fit() function of the instantiated model;
- Apply the model to test or new data:
  - For an unsupervised model, use the transfer() or predict() function to transfrom or infer properties;
  - For the supervised model, use predict() function to predict labels for unknown data;

The datasets may be real-life datasets or synthetic datasets for deriving the desired result as per our requirements. We have used synthetic datasets for performance analysis of the environments. The basic configurations, in terms of hardware and software, are as follows:

- Hardware: 2 x Intel Skylake Core i7-6500U@2.50 GHZ 4/8/12 cores, Max Turbo Frequency 3.10 GHz, 12 GB RAM, 24 GB Virtual Memory;
- Software: Anaconda framework, Intel-Python 2020 Update 2, MKL 2020 Update 1, Numpy 1.18.1, Intel DAAL 2020 Update 2/Intel DAAL 2019 Update 5 with daal4py 2019 update 5, Scikit-learn 0.23.1, Python 3.7, Conda 4.5.12, Pandas 1.0.5, Matplotlib 3.0.3 and other packages available in the distribution.

The dataset for outlier analysis consists of 5,000,000 rows with 03 columns to calculate the performance speed-up with and without the DAAL library. The dataset for the linear regression model consists of 10,000,000 rows with 6 columns. The model selection from the scikit-learn package is performed in the data processing without the DAAL environment scenarios. We used normalization and compute functions of the daal4py package to speed-up the computation in the DAAL environment. The environment is easy to integrate with different ML frameworks even for NCS based OpenVINO framework for edge computing. DAAL is helping in reducing irrelevant computations and predictive capabilities of the model so as to enable to speedup the computational results. It is helping to automate machine learning tasks to increase the efficiency of the model. Thus we have compared the computational time of the above two datasets in light of the DAAL framework to help us understand the speedup capability of the framework. This environment helps to easily troubleshoot and deploy different application models without configuring the development environments from scratch.

4 Results and Discussions

The tapping of the computing power to speedup the processing environment through DAAL with Anaconda framework is discussed and the results compared with non-DAAL environments as in Table 1.

Table 1 Computational speed up with and without DAAL framework

Model	Time					
	Computation time taken (in s)					
	Without DAAL			With DAAL		
	No of cores			No of cores		
	4	8	12	4	8	12
Linear model	37.040	30.387	26.405	24.878	22.497	18.576
Outlier analysis	28.974	25.897	22.704	19.428	13.258	11.441

The corresponding speedup comparisons are depicted in Figs. 1 and 2. It is evident

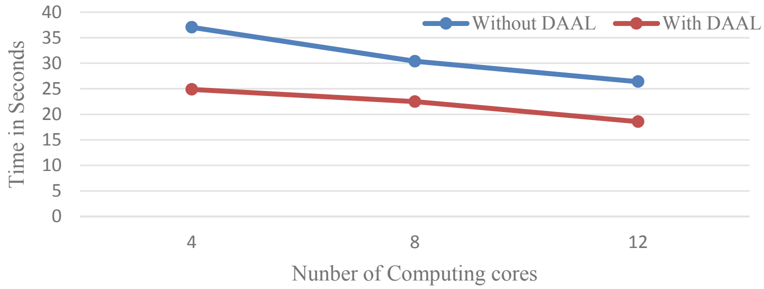


Fig. 1 Comparison of speedup performance of linear model

from the figures that the DAAL framework enabled the environment for maximum throughput in significantly less time. Intel optimizer DAAL framework improves scikit-learn efficiency in a significant manner for Intel Xeon series processors. It is also interesting to check whether all the popular machine learning algorithms are offering better speedup under different circumstances or not. The authors are presently working in this direction.

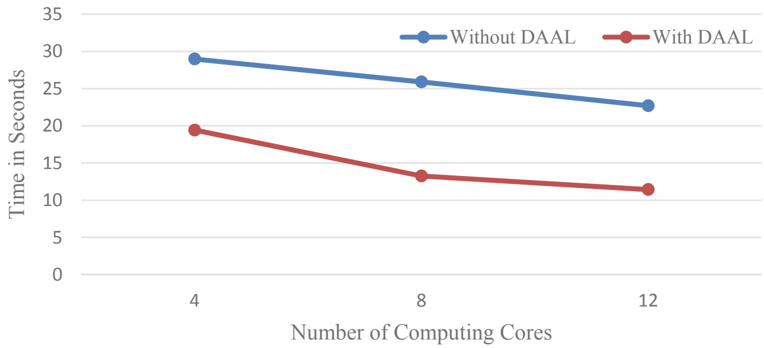


Fig. 2 Comparison of speedup performance of outlier analysis

References

1. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: Proceedings of the 20th VLDB conference Santiago, Chile
2. Mondal K (2015) Big data parallelism: issues in different X-information paradigms. Procedia Comput Sci 50. Special Issue on Big data, cloud and computing challenges, pp 395–400. http://doi.org/10.1016/j.procs.2015.04.028
3. Intel (2019) Machine learning based advanced analytics using Intel
4. Use Intel DAAL in Intel System Studio (2020) https://software.intel.com/content/www/us/en/develop/tools/system-studio.html
5. Wang P et al. Anaconda—guide to open-source tools & libraries for data science and machine. https://www.anaconda.com/library
6. Tiwari N, Mondal K (2019) NCS based ultra-low power optimized machine learning techniques for image classification. In: IEEE R10 symposium TENSYMP-2019, pp 562–565. ISBN: 978-1-7281-0296-2
7. Dennard RH, Gaensslen F, Yu H-N, Rideout L, Bassous E, LeBlanc A (1974) Design of ion-implanted MOSFET's with very small physical dimensions. IEEE J Solid-State Circuits SC-9(5)
8. Mondal K, Dutta P (2015) Big data parallelism: challenges in different computational paradigms. In: IEEE third international conference on computer, communication, control and information technology (C3IT, 2015). http://doi.org/10.1109/C3IT.2015.7060186
9. Patwary MdMA, Satish NR, Sundaram N, Liu J, Sadowski P, Racah E, Byna S, Tull C, Bhimji W, Prabhat, Dubey P (2016) PANDA: extreme scale parallel K-nearest neighbor on distributed architectures
10. Arthur D, Vassilvitskii S (2007) k-means++: the advantages of careful seeding. In: Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, pp 1027–1035
11. Hazan E, Duchi J, Singer Y (2011) Adaptive subgradient methods for online learning and stochastic optimization. J Mach Learn Res 12:2121–2159
12. Mohindru G, Mondal K, Banka H (2019) IoT and data analytics: a current review. WIRE's Data Min Knowl Discov 10(1002):1–27
13. Defazio A, Bach F, Lacoste-Julien S (2014) SAGA: a fast incremental gradient method with support for non-strongly convex composite objectives. In: Advances in neural information processing systems
14. Mondal K (2016) Application design and analysis of different hybrid intelligent techniques. Int J Hybrid Intell Tech 13(3):173–181. ISSN 1448-5869. http://doi.org/10.3233/HIS-160234

15. Friedman J, Hastie T, Tibshirani R (2017) The elements of statistical learning data mining, inference, and prediction. Springer, Berlin
16. Billor N, Hadib AS, Velleman PF (2000) BACON: blocked adaptive computationally efficient outlier nominators. Comput Stat Data Anal 34:279–298
17. Mondal K (2016) Design issues of big data parallelism. Springer AISC series. In: Third international conference on international system design and intelligent applications (INDIA 2016). ISSN 2194-5357
18. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: International conference on artificial intelligence and statistics
19. Szegedy C, Toshev A, Erhan D (2013) Scalable object detection using deep neural networks. In: Advances in neural information processing systems

Blind Descent: A Prequel to Gradient Descent

G. Akshat and N. R. Prasad

Abstract We describe an alternative learning method for neural networks, which we call Blind Descent. By design, Blind Descent does not face problems like exploding or vanishing gradients. In Blind Descent, gradients are not used to guide the learning process. In this paper, we present Blind Descent as a more fundamental learning process compared to gradient descent. We also show that gradient descent can be seen as a specific case of the Blind Descent algorithm. We also train two neural network architectures, a multilayer perceptron and a convolutional neural network, using the most general Blind Descent algorithm to demonstrate a proof of concept.

Keywords Blind Descent · Vanishing gradients · Neural networks · Gradients · Gradient descent

1 Introduction

We see Blind Descent as a prequel to gradient descent. By design, Blind Descent does not require calculation of gradients. Consequently, it does not face certain problems faced by gradient descent, including vanishing and exploding gradients. In gradient descent, there's also a need to store values from forward pass at each node for backpropagation. We do not need to do this for Blind Descent. However, not calculating gradients is what makes this method blind as we are not using gradients to guide the learning proces, hence the name Blind Descent.

In later sections, we describe the Blind Descent algorithm and its motivations. We also look at Blind Descent as a generalized learning process, which is more fundamental than gradient descent, which can be seen as a special case of Blind

G. Akshat (✉) · N. R. Prasad
Electrical and Computer Engineering, Carnegie Mellon University, Pittsburg, PA, USA
e-mail: akshatgu@andrew.cmu.edu

N. R. Prasad
e-mail: praghave@andrew.cmu.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_41

Descent. We also perform experiments on the MNIST dataset to demonstrate the working of the Blind Descent algorithm.

2 Prior Work

We did not find much prior works to build on. A remotely similar work is on Extreme Learning Machine [1]. However, the input-hidden layer appears to be randomly initialised and frozen and only the hidden-output layer appears to be minimised using pseudo-inverse solution. Even the incremental learning variant appears to do the same (but with convex optimisation for only one layer) [2]. Other methods use injected noise [3]. But, controlling the noise can be a difficult endeavor and moreover, the data itself can be said to have noise and the required information components. So, we decided to eliminate the noise present in the data instead of introducing additional noise. There are several variants of Extreme Learning Machine (ELM). But, as we can see in the survey, they are either too simplistic and reduce to linear regression or tend towards convex optimisation problems like SVM [4]. However, we did find random optimisation statistical papers of 1960s and one of the most influential papers that directly influenced our research is the backpropagation paper [5].

3 Blind Descent

The motivation behind Blind Descent was to try and overcome problems faced when gradients are used to guide the learning process in neural networks. An extreme step in that direction is to eliminate the use of gradients in the learning process. This leads us to Blind Descent. Calculating gradients do have their advantages, as they guide is the direction of steepest descent. Thus, using gradients, we not only know if we're going in the correct direction, but we also know that we're going in the best possible correct direction, locally speaking. Thus at first the idea of doing away with gradients seemed implausible, as we cannot in good conscience do away with the only guiding force we have in the learning process. But once we decided to let go of gradients, we discovered alternate methods that could be used to guide the learning process, which make up the Blind Descent algorithm.

In Blind Descent, we do not calculate gradients to guide the learning process. This is a one sentence elevator pitch for the algorithm repeated multiple times in this paper. But then how does the network learn? The weights at each node of the neural network are randomly initialised. Then on every iteration, we update the weights randomly. This randomness although, is guided. Only those updates to weights that lead to a reduction to the overall loss function are kept, other updates are discarded. This is the Blind Descent algorithm in its most basic form.

To define the random updates, we need to define the distribution from which the random numbers are picked. The distribution has a mean and a standard deviation, which defines the random updates. We center this distribution on the current value of the weights. The update rule in blind descent is given below:

$$x^{(t+1)} = \begin{cases} x^{(t)} + d(\mu = x^{(t)}, \sigma = \eta), & \text{if } \mathbf{L}(x^{(t+1)}) < \mathbf{L}(x^{(t)}) \\ x^{(t)} & \text{otherwise} \end{cases}$$

Here, $x^{(t)}$ is the weight at a node of the neural network at a certain time t. η is the learning rate, which here defines the standard deviation of the underlying distribution. $d(\mu, \sigma)$ is a distribution governing the random updates, centered around the value of the weight at the current time step. We call this distribution the *update distribution*. $\mathbf{L}$ is the loss function for the neural network under consideration.

The above definition of Blind Descent can have various variants. This is discussed in detail in Sect. 5. Different distributions can be used to update the weights. The standard deviations can become smaller and smaller as the loss decreases and we reach closer to a good solution, thus making standard deviation a variable quantity. This would be analogous to a learning schedule used while training neural networks using gradient descent frequently. We can also use concepts of momentum in Blind Descent, thus favoring the direction of previous motion in Blind Descent.

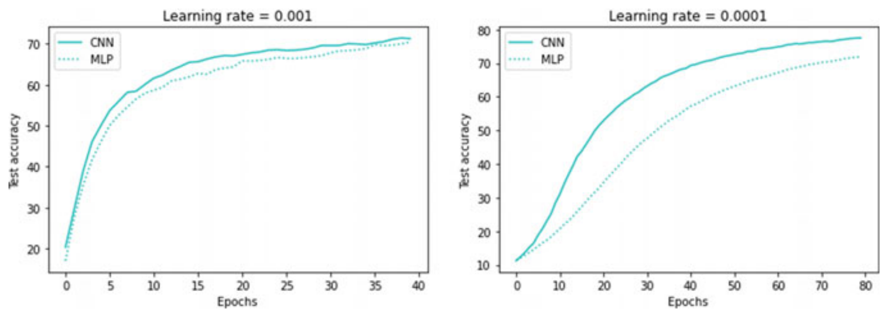
Blind Descent, as described above, can also be seen as a generalization to gradient descent. To obtain the equation of gradient descent from Blind Descent, we replace the *update distribution* $d(\mu, \sigma)$ with a deterministic function which is nothing but the gradient of the loss function. This observation shows that Blind Descent is a more fundamental learning process in neural networks compared to gradient descent. Gradient descent can be seen as a specific case of Blind Descent, with deterministic updates.

4 Experiments with Blind Descent

In this section we train a neural network using Blind Descent. We train a multilayer perceptron (MLP) and a convolutional neural network (CNN) on the MNIST dataset. We do not assume MNIST to be a standard dataset by any means, but it is a non-trivial dataset with 10 classes.

4.1 Training MLP with Blind Descent

We were able to achieve accuracies close to 70% on the MNIST dataset. The experiments performed are not optimum and various other techniques like momentum, standard deviation scheduler can be used to improve the results. Our focus is to present a proof of concept showing that Blind Descent works.



(a) Accuracy comparison between CNN and MLP architectures for learning rate 0.001 averaged over 10 sample runs.

(b) Accuracy comparison between CNN and MLP architectures for learning rate 0.0001 averaged over 10 sample runs.

Fig. 1 Blind Descent at work for MLP and CNN architectures

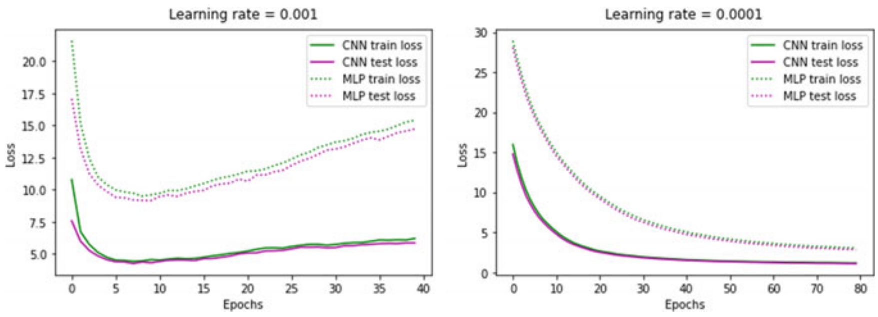
For training, we use a two layer deep neural network. The network architecture is [784, 256, 10], where 784 is the input dimensions of MNIST, the first layer has 256 neurons and the output layer has 10 neurons. We use cross entropy loss to measure the loss. While training, we do batch updates. Here are the steps followed during batch updates in Blind descent:

1. Calculate loss with current weights for entire batch
2. Do weight updates using Blind Descent
3. Calculate loss with new weights for entire batch
4. If new loss is less than previous loss, adopt new weights.

A unique feature we see in Blind Descent is that the training and testing loss keeps increasing along with the increasing accuracy (except the initial drastic decrease in loss). This can be seen in both Fig. 2. It might seem unintuitive as we are restricting updates only if the loss on update is lower than the loss before updates. Although this is true for one batch, the loss for the updated parameters for the next batch could be higher for the same parameters. The condition only compared the updated loss to the loss of the current batch. This is the main reason why the loss is increasing (excluding the initial decrease in loss). Even though that happens, the accuracy still increases. Notice that we never put in the condition for the accuracy to increase in our algorithm and it still increases!

4.2 *Training CNN with Blind Descent*

We demonstrate that even a single CNN layer can perform well with blind descent and achieve accuracy of about 73% even when we have not used momentum, learning rate schedule, automated hyper-parameter tuning etc. We choose a CNN architecture having one CNN layer consisting of 16 kernels, each of size 3×3. We have a max



(a) Training and Test loss for CNN and MLP architectures averaged over 10 sample runs for learning rate = 0.001.

(b) Training and Test loss for CNN and MLP architectures averaged over 10 sample runs for learning rate = 0.0001.

Fig. 2 Training and Tess Loss comparison between CNN and MLP architectures

pool layer after that with kernel size = 3, stride = 2 and padding = 1. The output of the max pooling layer is connected to a linear layer consisting of 128 neurons. The results for the CNN architecture are shown in Figs. 1 and 2.

5 Variants

This section highlights many variants of gradient descent which can help in faster Blind Descent.

- *Update Distributions*: Different *update distributions* can be used for the learning process using Blind Descent. In our work, we have used a normal distribution. Update distributions can in principle be stochastic functions and have both deterministic and non-deterministic components.
- *Standard Deviation Scheduling*: We can schedule the standard deviation based on the number of successful updates in Blind descent. As the loss decreases, it is likely that with a larger standard deviation, we're likely to miss an optimum update that reduces the loss. To counter that, we can reduce the standard deviation based as the number of successful updates decreases. We can also use the test accuracy for standard deviation scheduling. This is analogous to a learning rate schedule used with gradient descent [6].
- *Momemtum*: We can use the concept of momentum in Blind Descent as well (similar to gradient descent [7]). In the version of Blind Descent presented in this paper, the mean of the *update distribution* is at the current value of the weight. We can build up momentum with successful updates to the weights and shift the mean of the *update distribution* in direction of the building momentum. This moves the mean of the *update distribution* away from the current value of the weights

and towards the value suggested by the building momentum, thus accelerating the learning process.

- *Combining other learning techniques*: We can use Blind Descent with many other training techniques. Layer by layer training, first suggested in [8] can be used with Blind Descent. We can also incorporate Blind Descent while transfer learning. It is also possible to train a Blind Descent Network further using gradient descent and use Blind Descent training as an initialization.

6 Discussion

In this paper, we introduce the Blind Descent algorithm as a generalization to gradient descent. We trained two neural network architectures, namely a multilayer perceptron and a convolutional neural network to show that the algorithm works. Blind Descent has various advantages over gradient descent including not having to save forward pass values and calculating gradients at each node. This happens by design as we are not calculating gradients. Accordingly, Blind Descent does not suffer with limitations of gradient descent like vanishing and exploding gradients.

In Blind Descent we can choose to accept or reject an update, such a luxury is not available in case of gradient descent. This also remedies another disadvantage of gradient descent. For an instance of gradient descent, we are bound by the network initialization and update steps taken in the past, which are locally optimal but might be sub-optimal globally. Blind descent on the other hand gives us the opportunity to explore a larger solution space at every step, thus giving us a possible way out of a sub-optimal solution. We believe this feature of Blind Descent can be used to augment gradient descent, allowing us to explore a larger solution space.

Even though Blind Descent has some merits, without gradients, we do not know the direction in which we should proceed and thus we are blind. Thus we do not expect Blind Descent to compete with Gradient Descent in terms of performance.

Future work with Blind Descent includes implementing the many variants of Blind Descent as discussed in Sect. 5. Larger neural networks and other complex architectures can also be trained to test the Blind Descent algorithm. We also have a whole spectrum of *update distributions* to explore.

References

1. Huang G-B, Zhu Q-Y, Siew C-K (2005) Extreme learning machine: theory and applications. Neurocomputing 70:489–501
2. Huang G-B, Zhu Q-Y, Siew C-K (2006) Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans Neural Networks 17
3. Maryak JL, Chin DC (2001) Global random optimization by simultaneous perturbation stochastic approximation. In: Proceeding of the winter simulation conference. IEEE

4. Huang G-B, Wang DH, Lan Y (2011) Extreme learning machines: a survey. Int J Mach Learn Cybern 2:107–112
5. Rumelhart David E, Hinton Geoffrey E, Williams Ronald J (1986) Learning representations by back-propagating errors. Nature 323
6. Schaul T, Zhang S, LeCun Y (2013) No more pesky learning rates. In: International conference on machine learning, pp 343–351
7. Qian N (1999) On the momentum term in gradient descent learning algorithms. Neural Networks 12(1):145–151
8. Fahlman Scott E, Christian L (1990) The cascade-correlation learning architecture. Adv Neural Inform Process Syst

A Survey: Modelling Strategies for Predictive Maintenance

Vinayak Tyagi, Sanjay Silakari, Uday Chourasia, and Priyanka Dixit

Abstract This paper deals with various modelling Strategies to achieve predictive maintenance. Predictive Maintenance (PdM) is introduced into various industries and sectors. It helps to predict the condition of equipment of machines that are already in use this tells whether the maintenance is required or not. This technique ensures that cost-saving has done as compared to regular maintenance where unnecessary replacements have been done without proper utilization of resources. And this can be achieve using various modelling techniques as per the needs or requirements. The paper has four basic principles or strategies or ways to implement PdM as per the type of data available for business requirements. We have also covered the type of data that one should accumulate as per there business or user requirements. We have also tried to cover some of the basic questions that one commonly faces during the implementation of predictive modelling. The big importance of this technology is to predetermine the failure so, that proper timeframe is given for maintenance which simply means its cost-effective, lower the downtime, and better user experience. This approach is widely used in industries like oil, semiconductors, production, etc. (Susto et al. in IEEE Trans Ind Inf, [1]).

Keywords PdM (predictive maintenance) · Data mining · Machine learning · Classification algorithms · Remaining useful life (RUL)

1 Introduction

Organizational data is a valuable strategic asset because it represents the overall experience of an organization, the customer interaction data such as complaints, frauds, purchases, and responses of the customer about product provides an experience to an organization [2].

Types of maintenance are mainly divided into three main categories; these are as follows:

Reactive—Maintenance after the problem.

V. Tyagi (✉) · S. Silakari · U. Chourasia · P. Dixit
Computer Science Department, University Institute of Technology RGPV, Bhopal, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_42

Example: Replacement of car battery when it has a problem.
Problem: Unwilling failures can sometimes be dangerous as well as expensive too.
Scheduled—Regular interval maintenance.
Example: Change lubricant of a car in every 1000 km.
Problem: Incomplete utilization of resources leads to waste and sometimes may not eliminates all failures.
Predictive—Forecasting of problem arrival.
Example: Some cars can forecast problems with fuel pump, battery, and starter motor.
Problem: Difficult and complex to get an accurate forecast [3].

The data in this process play a very crucial role. Data in this technique mainly collected with the help of sensors that are placed on different parts of machinery. The data is noisy because of various disturbances in physical parameters.

2 Literature Review

Predictive maintenance models success mainly depends on three factors these are:

- Availability of the right data.
- Problem framing appropriately.
- Evaluation of prediction.

In this paper, we will elaborate on the first two points and give insights on how to choose the modelling technique that best fits the question you are trying to answer and the data you have at hand [4].

2.1 Data Collection

When collecting data to support a failure model, it is important to make an inventory of the following [5]:

- Which failure type can occur? And failure which we have to predict?
- A brief overview of "failure procedure" is its slow degradation or an acute failure.
- Which part of the machine/system is responsible for which type of failures?
- What can be measure from the parts of the machine/system that can reflect its working state?

The Life span of such a machine is in years, which means that amount of data collected is huge which helps to view the degradation process of the system [6].

Table 1 Questions table [6]

Analytics type	Question it answers	Down-time	Parts replacement cost	Parts replacement time
Descriptive	What happened?	What units/service area/customer had the most downtime?	What units/service area/customer had the highest parts replacement cost?	What units/service area/customer had the most unplanned parts replacement?
Diagnostic	Why did it happen?	Why?	Why?	Why?
Predictive	What will happen next?	When will downtime happen next?		When will the parts need replacement next?

2.2 *Problem Framing*

These are couples of questions which need to be kept in mind when thinking of framing a predictive maintenance model:

- What kind of output should the model give?
- Do we have enough historical data, or do we have just static data available?
- Do we have every measurement labelled that is which event corresponds to good function and which one corresponds to failure?
- In the case of machine failure, do we have any alert?
- What is the proportion of the number of failure events and events of well-functioning as and when labelled events are available?
- In case of failure, how long should you be able to indicate in advance?
- What are the performance targets that the model should be optimized for? High precision, high sensitivity/recall, high accuracy?
- What are the consequences of true and false predictions of the model? [6]

Some common questions which are answered by the predictive maintenance which are mainly according to the business perspective are (Table 1).

3 Dataset Description

The Dataset taken for this purpose is Damage Propagation Modelling for Aircraft Engine Run-to-Failure Simulation. NASA has created the Prognostics and Health Management PHM08 Challenge Data Set and is being made publicly available. the data set is used to predict the failures of jet engines over time. The data set was provided by the Prognostics Coe at NASA Ames.

The data set includes time-series measurements of various pressures, temperatures, and rotating equipment speeds for the jet engine. these measurements are typically measured in a commercial modern turbofan engine. All engines are similar but every engine has somewhat of different degrees of initial wear and difference in the manufacturing process which has not been given to users. There are 3 optional settings, which are used to tune the performance of each machine and each engine has the remaining 21 sensors which are used to collect different data at runtime [7].

For the creation for dataset the uses the simulation tool called C-MAPSS (Commercial Modular Aero-Propulsion System Simulation). Which is used to a large amount of realistic commercial turbofan engines dataset? With the help of various sensors, this environment has various editable inputs that allow users to further tune the parameter as per the need.

The simulated data generated were used as challenge data for the 1st Prognostics and Health Management (PHM) data competition at PHM'08 [7].

The sensor data has been fetched from 100 engines of the same model. For the task of predicting and fix failure before they arise. Earlier the maintenance was done after 125 flights, regardless of whether the equipment seemed to need it or not.

Sensor noise has been detected in the collected data. From sensor reading, it can observe that each engine develops a fault, overtime. the data stops for each engine when a failure has occurred for that particular engine. hence the actual RUL is known based on the length of the data.

4 Methodology

For Predictive maintenance, these four steps are very frequently used and multiple modelling methods have been used for PdM (Predictive Maintenance Modelling) as per the data. We have discussed some of the cases which should be kept in mind during implementation.

These Cases are described in detail as follows [8]:

- Prediction of RUL from the regression model.
- Time window failure prediction with classification model.
- Differentiating the odd/anomalous behaviours.
- Failure probability prediction by the survival models.

The above cases are used for the implementation of predictive maintenance. These cases have their respective purpose or output which can be achieved by certain specifications in the dataset only [8].

These cases are elaborated as follows:

CASE-I

Prediction of RUL from a Regression model.

Expected outcome

Predict numbers of cycles before the system fails.

Data features

- Labelled historical and static datasets.
- Along with every common event, events related to failures should also be labelled.

Assumptions/Requisite

- Based on the data features, RUL values are predicted, and this implies that both static and historical data is needed as well as the degradation process is smooth in the system.
- Available data should be labelled and measured during various moments of a lifetime.
- If multiple types of failures can take place within the system than a dedicated model should be made for each [9].

CASE-II

Time window failures prediction with classification models.

Description

Making the model which can predict precisely the long lifetimes of the machine can be a very difficult task so, mainly maintenance teams require to know whether the systems fail soon or not.

Expected outcome

Whether the machine fails or not in recent days of cycles (N days/ cycle)?

Data features

- Labelled historical and static datasets.
- Along with every common event, events related to failures should also be labelled.

Assumptions/Requisite

- Assumptions are very much the same as the above case of the regression model, but only differences are.
- Since we are dealing with time windows, so unlike the regression model, the degradation process is relaxed over here.
- The beauty of the Classification model is that one model can deal with multiple failures. For Example, Class-0: Represents no failures in N days. Class-1: Represents type-A failure within N days and so on [10].
- Failure labels in the dataset are available "enough" so, we can train and evaluate the model.
- In common words, if the model is applied to the system which has the failure which is absent in the data then the model will fail to predict such cases [11].

CASE-III

Differentiating the odd/anomalous behaviours.

Description

Odd over here means the anomalous behaviours or the behaviours out of leak it may also be termed as odd in evens.

Unlike the above two cases where the failure data is present, however, mainly in the real-world scenario which has a critical system, the failure data is absent. In that case, this case should be used [12].

Expected outcomes

Is this behaviour normal or odd?

Data features

Historical and static data is available, but many failure labels are unknows or only a few are known.

Assumptions/Requisite

- It is very much important to differentiate between the odd and normal behaviour because this leads to knowing about machine/system degradation which finally results in failures.
- In this case, mainly anomaly detection model is applied which has the biggest advantage as well as the pitfall too. It flags almost all the odds behaviour but it's not necessary that the given indication will always be right or right indication of failure will always be given hence if it does happen model will fail to give the timespan of failures.
- If few failures labels are available in the dataset it must be used to evaluate the model and, in another case, if no few failure labels are there than the model should be evaluated by the domain experts itself [13].

CASE-IV

Failure probability prediction by survival model.

Description

The above three approaches give a brief idea about maintenance before failure. In this case, we would mainly focus on the degradation process itself.

Expected outcome

How does failure risk change over time?

Data features

Static data is available, reports like dataset or dates and time are recorded when the machine is an unobservable state for failure.

Assumptions/Requisite

The probability of failure is predicted by survival models for each type of machine given the static features. It also estimates the groups of machines that are similar [8].

5 Conclusion

In this section, we can say that there are a few broad points which one can achieve after the successful implementation of PdM (Predictive Maintenance) in its operational environment. These are as follows along with its purpose of serving.

The Monitoring System which helps to detect the failure in the engine as well as it also predicts the time required for maintenance for the affected components of a machine.

Improvement Machine Availability after successful implementation of PdM, it ensures the reduction in downtime and improvement in on-time scheduling as well as it also provides the planed and optimizes maintenance approached which leads to reduce in engine out of service time.

This approach also reduces the unnecessary assistance required for the machine as well as it also limits the secondary damage to the machine. Which reduce the maintenance cost.

Acknowledgements I would like to express my special thanks of gratitude to my parents, friends, and professors, how helped me in doing this research work and I came to know about so many new things. Finally, I would like to thanks my guides and friends for all their support.

References

1. Susto GA, Schirru A, Pampuri S, McLoone S, Beghi A (2015) Machine learning for predictive maintenance: a multiple classifier approach. IEEE Trans Ind Inf
2. Siegel E Seven reasons you need predictive analytics today. PAW
3. Lombardi S Predictive maintenance with MATLAB: a prognostics case study
4. Matlab Expo 2017—Big data and machine learning for predictive maintenance
5. Predictive maintenance and monitoring using machine learning: demo and case study (Cloud Next 18)
6. How to optimize physical assets through ML-powered predictive maintenance
7. Saxena A, Goebel K, Simon D, Eklund N Damage propagation modelling for aircraft engine run-to-failure simulation
8. Machine learning for predictive maintenance: where to start? Blog by Big data republic
9. Getting-started-predictive-maintenance-models blog by Silicon valley groups for predictive maintenance
10. Abbas MA, Suhad MA (2017) Modelling the strength of lightweight foamed concrete using support vector machine (SVM). Case Stud Constr Mater
11. Bencheikh A, Cherkaoui A (2019) Chapter 25 transition model from articulatory domain to acoustic domain of phoneme using SVM for regression: towards a silent spoken communication. Springer Science and Business Media LLC

12. Anomaly detection with the normal distribution blog by anomaly.io groups
13. Choudhary P https://blogs.oracle.com/datascience/introduction-to-anomaly-detection

Abstract and Image Analysis of High-Temperature Materials from Scientific Journals Using Deep Learning and Rule-Based Machine Learning Approaches

Kavitha Jayaram, Prakash Gopalakrishnan, and Jayaram Vishakantaiah

Abstract In this digital world, many research papers have been getting added to the scientific journals in PDF format. It has become essential to the process and needs digitalization due to the exponential growth of newly published research papers. Survey of Abstract, Figures, and Captions convey the necessary information and extracted from scientific research journals. The section algorithm is used for extracting "Abstract" and "Result and Discussion" sections from the published research papers about high-temperature materials (HTM) by converting PDF to word documents. HTM got complete applications in aerospace engineering and electronic. After exposure to high-temperature test facilities, various characterization techniques are used to produce numerous high-resolution images. Abstract summarization, SEM figures, and caption extraction from HTM documents are performed using Deep learning, PDFFigure, and Machine learning (ML) techniques. MatLab is used for feature extraction and to classify SEM images by rule-based ML techniques. This summarization helps new researchers understand existing research datasets and proceed with new approaches to solve more complex problems. This research paper presents automatic extraction and summary of abstract, high-resolution figures, and captions from scientific journals on HTM; all these documents will be analyzed to categorize with a different tailor-made approach using deep learning and machine learning techniques.

Keywords Deep learning · Machine learning · Rule-based · PDF extraction · Summarization · Image classification · High-temperature materials

---

K. Jayaram (✉) · P. Gopalakrishnan
Dept. of Computer Science & Engineering, Amrita School of Engineering,, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: j_kavitha@blr.amrita.edu

P. Gopalakrishnan
e-mail: g_prakash@blr.amrita.edu

J. Vishakantaiah
Solid States and Structural Chemistry Unit, Indian Institute of Science, Bengaluru, India
e-mail: jayaram@iisc.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_43

1 Introduction

In the present digital world, scientific journals' regular and online publications are increasing with various types of images reported from the scientific instruments. Analysis of such scientific data quickly is especially crucial in the context of time analytics [1]. Data Scientists are extracting and analyzing tables or diagrams for discovering information from PDF documents. The term time to analytics refers to the total time that elapses between collecting the data and the time of analysis by the data analyst. In the context of enterprises, this concerns the time for collecting the data, storing them in a suitable data management system, and preparing them so that a data scientist can make use of this information [2]. Adam optimization is a stochastic gradient descent technique and used to update network weights more efficiently than a classical stochastic gradient descent method [3]. A concept-based extraction is introduced to improve text classification performance using a two-layer concept with better performance than the traditional one [4]. For reproducing biological vision in Artificial vision using machine learning computational models [5]. A NEXTWRAP project implemented low-level page segmentation techniques to extract and wrap pdf documents [6]. A new model is generated to parse PDF format information into semi-structured text and then pattern match with tree solution [7]. It is shown that the rule-based approach to extract titles from PDF gives better results [8]. Text summarization extracts a bigger context of larger documents into smaller sentences using deep learning [9]. A combination of Restricted Boltzmann machine and fuzzy logic for giving an overview of text using deep learning has been presented [10]. Extracting metadata from pdf using open source java framework stores extracted data in Oracle and XML (Extensible Markup language) files [11]. A content-based image retrieval system using primary features of images will check similarity for classification [12]. Detecting vehicles in traffic using advanced surveillance techniques using deep learning techniques reduces noise and provides better performance [13]. A new PDFigCapX is developed for the public dataset of computer science documents used to extract biomedical pictures and associated captions [14]. A novel and fast coding strategy was developed by reconstructing the classification dataset with reasonable accuracy [15]. Text summarization is broadly divided into two different approaches, such as Extractive Summarization and Abstractive Summarization, as shown in Fig. 1. Extractive summarization is to identify essential sentences from the original text and extract only those necessary without modification. In contrast, abstractive summarization generates new sentences after extracting sentences from the original text [9].

Literature studies reveal that "abstractive summarization" is the best method for analyzing abstract sections from the scientific research documents. The database consists of many research documents related to HTM materials, using section extraction from our previous paper need to extract abstracts and summarize the work. Understanding research being completed and what new research can be accomplished in HTM analysis. The scanning electron microscope (SEM) figures of SiO_2

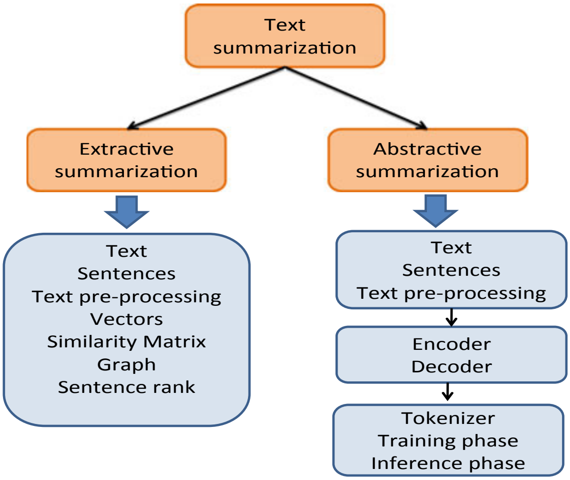


Fig. 1 Text summarization by extractive and abstractive approaches

are extracted from pdf research documents using PDFFigures[1] and then classify the images into different categories according to the pictures' quality (good, bad and worst images). The images can be classified according to their material properties using deep learning of pictures, which will be our future research work.

2 Implementation

The algorithm has been implemented using python nltk for abstract summarization and MatLab and python for image classifications.

The whole data (journal pdf papers) is trained into two networks where the first network is responsible for extracting abstract and in the next section is to extract parts labeled as "Figures/Fig.". Both systems are used to generate coordination marking for detecting location on these documents on a given page. A section extraction algorithm [16] extracts the title, authors, and abstract and stores them into a text document. The extracted abstract is used to generate a summary using abstractive deep learning. The PDFFigures functions has been used to extract SEM figures from the research papers and designed using a set of rules be categorized as "good, average, bad, and contradictory images" using the picture features such as intensity, visibility, contrast, entropy, and kurtosis. PDF document extraction of abstract using section algorithm and figures with its captions using python code is stored in separate word documents. The abstract file goes through the abstractive summarization method, which gives a summary of each abstract. The classification of figures is done on the rule-based approach into three categories, as shown in Fig. 2.

[1] http://pdffigures2.allenai.org.

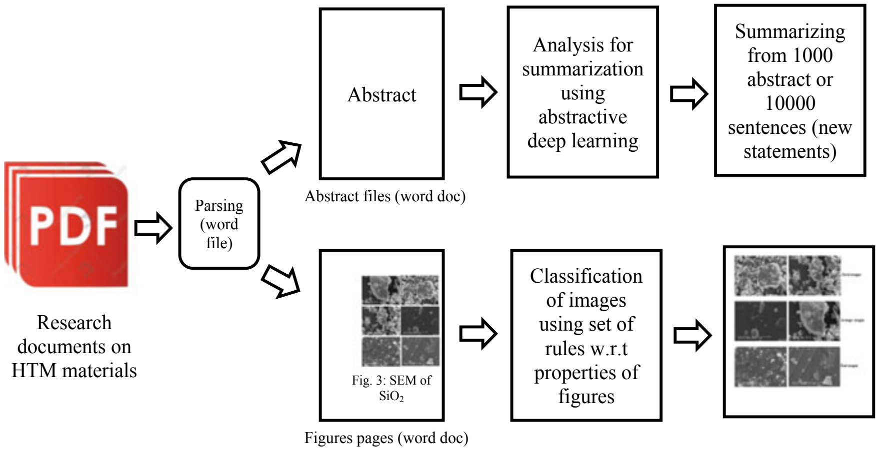


Fig. 2 Framework for summarization of abstract and classification of figures

2.1 Abstract Summarization

Applying normalization throughout the networks and Adam's [3] optimization (update of gradient descent method) algorithm is used to update each node's weights. The abstract section is extracted using a section extraction algorithm to summarize from the pdf research document in the on-going research. Analysis of text preprocessing is a good practice to reduce noise-free data, which helps clean the text data by removing punctuations, stop words, numbers, and special characters. "The bifurcated individual sentences form vector representations for each sentence used to store as a matrix for calculating the similarity matrix. This similarity matrix is converted into a graphical representation of vertices and similarity edges for ranking the top-ranked sentences from the final summary of an abstract".

The dataset has three columns known as article id (identification), text (abstract section), and source (the title of the paper). In the present work, we are interested in text columns, as it contains the abstract section of the research document. From the nltk library "sent_tokenize()" function is used to break text into individual sentences; this is used for forming vectors as bag-of-words (TF-IDF) representing the sentences. There are different methods for creating tf-idf but using "GloVe". The main disadvantage is ignoring the order of the words and number of features, hence using Wikipedia 2014 with the "Gigaword 5 GloVe" vector representation. The vectors of different terms, about one million, are stored in the dictionary. Generating consolidated vectors for text sentences by fetching vectors from words in sentences, then calculating mean or average for these vectors. Once vectors are formed for sentences, the cosine similarity (cosine_similarity) matrix is used to compute the similarity between a pair of sentences. This vector is converted into a graph to apply to the PageRank algorithm to rank the sentences. Now extract the top N sentences based on their rankings to summarize all the abstracts, calculating the percentage of sentences reduced concerning the main document. An abstractive summarization method outlines and summarizes the original text effectively.

"Input to text summarization is a long sequence of words, but the output has to be a shorter version of the input sequence. It uses a two-phase process, namely Encoder-Decoder, where variants of Gated Recurrent Neural Networks (RNNs) and Long Short Term Memory (LSTM) are encoder and decoder components, which helps us reduce the gradient of capturing from long terms. The task of the encoder network is to understand the input sequence for creating a reduced dimensional representation. The output is forwarded to the decoder network, which generates its sequence to portray the outcome. LSTM[2] reads the entire input sequence at each timestep; the encoder processes the information at every timestep and captures the input sequence's contextual information. The decoder is trained to predict the next word in the sequence given from the previous word". Condensing the important content from the sentences is a sensitive task in automatic text summarization. The LSTM produces new words (sentences) using the words fed.

2.2 Image Classification

The Matlab[3] coding has been implemented to retrieve the properties of the images like mean, standard deviation, contrast, entropy, and kurtosis, which frame the rules to classify figures. Convolutional Neural Networks (CNN) have advanced and improved architecture to model and classifies images by hierarchically training the data [17]. The images have been classified into four categories depending on a set of rules implemented in python, as shown below: the rules are classified into good, average, bad, and contradictory pictures.

Category 1: **Bad images**.

If ((139 $\leq$ **Kurtosis** $\geq$ 340)).

Category 2: **Average images**.

If ((15.5 $\leq$ **Contrast** $\geq$ 30) && (5 $\leq$ **Entropy** $\geq$ 6.8) && (2.4 $\leq$ $\mathbf{Entropy_c}$ $\geq$ 3.7)).

Then the Image is in **average Category**.

Few images had conflicts of interest to categorize with rules, hence considered contradictory pictures (failed to satisfy one rule).

---

[2] https://www.analyticsvidhya.com/blog/2019/06/comprehensive-guide-text-summarization-using-deep-learning-python/.

[3] Mathworks.

3 Results and Analysis

The present dataset consists of abstracts from journal papers selected from 2015 to 2020, including title, author names, keywords, and conclusion. The computational limitation set about "100,000 sentences to reduce the training time for our model". The encoder reduces the abstract dimension and feds to a decoder where it generates the probability of occurrences of a word in sequence. The decoder makes a decision using the greedy method to find the next word in the sequence. The model is used for fixing a maximum text length of 80 and a summary length of about ten sentences, which reduces the percentage of summarization. 90% of the dataset is split into training data and remaining 10% as testing data. Parameter setting for model building is: "Return sequences" = True (as LSTM needs to produce hidden state and cell state for every timestep), "Return state" = True (hidden state for last timestep only), "initial state" and stacked LSTM (to form multiple layers of LSTM to be stacked on top of each other). This model stops working once the validation loss increases (each step has been monitoring the validation loss) (val_loss) for more than ten epochs. There is a slight increase in validation loss after the 10th epoch; hence training has to be stopped after this epoch, as shown in Fig. 3. Around 96% of the abstracts have generated a good summary, whereas few abstracts were terminated without generating a summary.

The results obtained from the model, few abstract summary outputs are shown in Table 1. We have also demonstrated the percentage of summary with a reduced number of sentences (or words). Here summarizing and analyzing argumentative texts of the scientific research papers, which are well-structured analyses of overview. The images are classified using rules into three categories, the highlighted features of the image are used for classifying SEM images into three categories as shown in Table 2.

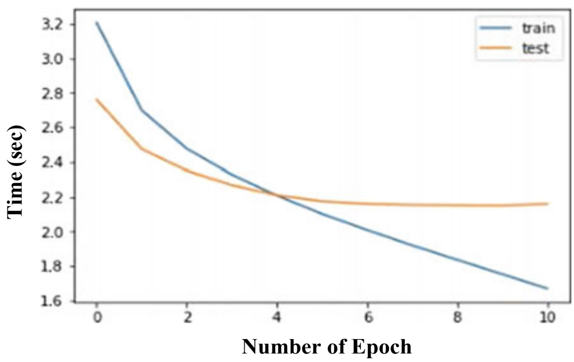


Fig. 3 Diagnostic plots to understand model behavior over time

4 Conclusion

We successfully summarized the abstracts extracted from journal papers using abstractive deep learning related to high-temperature materials with a reduced summary. The algorithm's efficiency was about 94% as few did not produce useful summary and few terminated without results. SEM image classification is implemented using rule-based machine learning techniques to categorize into three categories (good average and bad). It achieved a 90% classification success rate as 10% of the images could not be classified into any category as there were conflicts in

Table 1 Comparison of a few lists of selected abstract summaries

S. No.	Paper abstract with respective title	Summary generated
1	**"Characterization of anultra-high temperature ceramic composite (UHTCC) by ultrasonic testing"byAnimesh Talapatra, Openaccess NDT database** The ultra high temperature ceramic composites (UHTCC) are of interest for hypersonic vehicle leading edge. The ultrasonic testing is one of the widely used non-destructive testing (NDT) for the material characterization. The aim of investigations described in this article is measurement of the mechanical properties of two ceramic composite by pulse-echo ultrasonic testing. These two samples (ZrB_2–20%ZrC–20%SiC–5%Si3N4 and ZrB_2–20%SiC–5%Si_3N_4) were fabricated by hot pressing into 25 mm diameter and 4 mm thickness discs. The measurements were performed using pulse echo ultrasonic technique and ultrasonic transducers with the frequencies 5.0 MHz. It was determined that the velocity of ultrasonic longitudinal and shear waves varies when the frequency is changing. Ultra sound velocity measurements were used to evaluate mechanical properties of UHTCC. These composites were prepared from ZrB_2, SiC, ZrC and Si_3N_4 by ball milling and hot pressing	**(Reduced to 48%)** "Ultrasonic testing is one of the widely used non-destructive testings for material characterization. The investigations described in this article are measuring the mechanical properties of two ceramic composites by pulse-echo ultrasonic testing. It was determined that the velocity of ultrasonic longitudinal and shear waves varies when the frequency is changing"

(continued)

Table 1 (continued)

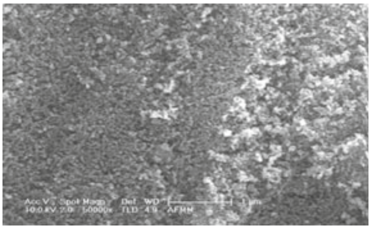
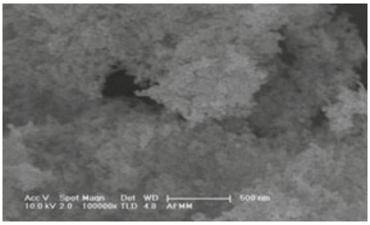
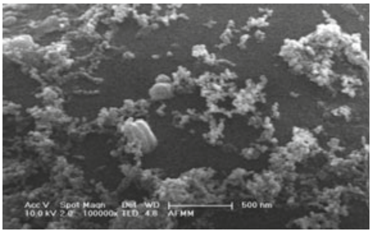
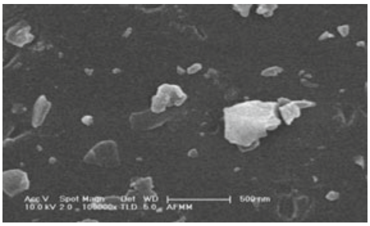
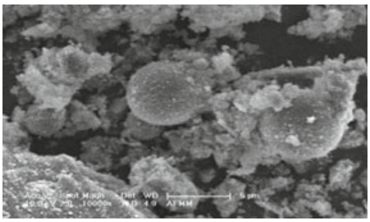
S. No.	Paper abstract with respective title	Summary generated
2	**"High performance oxide fibers for metal and ceramic composites", D. M. Wilson and L. R. Visser, Processing of fibers and composites conference, Italy (2000)** A family of oxide fibers. Nextel^{TM} 610 Ceramic Oxide Fiber, Nextel^{TM} 720 Ceramic Oxide Fiber and a new fiber, Nextel^{TM} 650 Ceramic Oxide Fiber, has been developed specifically for the reinforcement of metal and ceramic matrix composites. This paper summarizes room and high temperature properties for these fibers. The strength of single filaments and multi-filament rovings of Nextel 610, 650 and 720 fibers was determined between 25 and 250 mm gauge length. Weibull analysis was used to compare the statistical fracture distribution and gauge length dependence of strength. Fiber fracture statistics were in accord with Weibull theory; the effect of diameter variability on the statistical analysis was found to be small. Fractographic analysis on Nextel 610 fiber was used to identify primary fracture-causing defects; defect size was correlated with Griffith fracture predictions. High temperature single filament strength measurements were performed on Nextel 610, 650 and 720 fibers between 800 and 1400 °C. High temperature strength varied inversely with strain rate. In combination with tensile creep tests at 1100 and 1200 °C, these were used to compare the elevated temperature capability of each fiber and determine maximum use temperatures. The development of crystalline yttrium aluminum garnet (YAG) fibers that demonstrate further improvements in creep performance relative to Nextel 720 fibers is also discussed	**(Reduced to 29%)** "A family of oxide fibers, NextelTM 610 Ceramic Oxide Fiber, NextelTM 720 Ceramic Oxide Fiber, and a new fiber, NextelTM 650 Ceramic Oxide Fiber, has been developed specifically for the reinforcement of metal and ceramic matrix composites. In combination with tensile creep tests at 1100 and 1200 °C, these were used to compare each fiber's high-temperature capability and determine maximum use temperatures"

(continued)

Table 1 (continued)

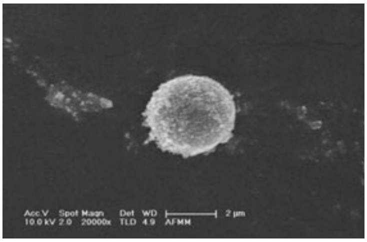
S. No.	Paper abstract with respective title	Summary generated
3	**"Ceramicmatrix composite (CMC) thermal protection systems (TPS) and hot structures for hypersonic vehicles", David E.Glass, 15th AIAA Space planes and hypersonic systemsand technologies conference, 2008** Thermal protection systems (TPS) and hot structures are required for a range of hypersonic vehicles ranging from ballistic reentry to hypersonic cruise vehicles, both within Earth's atmosphere and non-Earth atmospheres. The focus of this paper is on air breathing hypersonic vehicles in the Earth's atmosphere. This includes single-stage to orbit (SSTO), two-stage to orbit (TSTO) accelerators, access to space vehicles, and hypersonic cruise vehicles. This paper will start out with a brief discussion of aerodynamic heating and thermal management techniques to address the high heating, followed by an overview of TPS for rocket-launched and air-breathing vehicles. The argument is presented that as we move from rocket-based vehicles to air-breathing vehicles, we need to move away from the "insulated airplane" approach used on the Space Shuttle Orbiter to a wide range of TPS and hot structure approaches. The primary portion of the paper will discuss issues and design options for CMC TPS and hot structure components, including leading edges, acreage TPS, and control surfaces. The current state-of-the-art will be briefly discussed for some of the components. The two primary technical challenges impacting the use of CMC TPS and hot structures for hypersonic vehicles are environmental durability and fabrication, and will be discussed briefly	**Failed to summarize**

Table 2 Image classification of SEM micrograph taken for SiO_2 HTM

Images	Image intensity properties	Contrast properties
Bad images		
	(a) Brightness = 128.312 (b) Contrast = 7.98 (c) Entropy = 4.119 (d) Skewness = 3.082 **(e) Kurtosis = 139.524** (f) Visibility = 356.169	(a) Mean = 1.652 (b) Std Dev = 4.498 (c) Entropy = 1.888 (d) Skewness = 3.329 (e) Kurtosis = 132.75
	(a) Brightness = 112.361 (b) Contrast = 5.476 (c) Entropy = 3.628 (d) Skewness = 3.754 **(e) Kurtosis = 338.613** (f) Visibility = 296.385	(a) Mean = 1.086 (b) Std Dev = 3.459 (c) Entropy = 1.321 (d) Skewness = 3.856 (e) Kurtosis = 232.076
Average images		
	(a) Brightness = 133.816 **(b) Contrast = 22.767** **(c) Entropy = 6.463** (d) Skewness = 0.392 (e) Kurtosis = -0.164 (f) Visibility = 1493.471	(a) Mean = 2.592 (b) Std Dev = 3.993 **(c) Entropy = 2.678** (d) Skewness = 2.652 (e) Kurtosis = 65.05
	(a) Brightness = 137.433 **(b) Contrast = 15.629** **(c) Entropy = 5.511** (d) Skewness = 1.147 (e) Kurtosis = 4.791 (f) Visibility = 840.232	(a) Mean = 2.648 (b) Std Dev = 4.093 **(c) Entropy = 2.627** (d) Skewness = 2.618 (e) Kurtosis = 65.947
Good images		
	(a) Brightness = 68.922 (b) Contrast = 51.154 (c) Entropy = 6.822 (d) Skewness = 1.048 (e) Kurtosis = 0.723 (f) Visibility = 9430.149	(a) Mean = 10.506 (b) Std Dev = 12.068 **(c) Entropy = 4.845** **(d) Skewness = 1.368** **(e) Kurtosis = 4.363**

(continued)

Table 2 (continued)

Images	Image intensity properties	Contrast properties
	(a) Brightness = 113.152 **(b) Contrast = 38.140** (c) Entropy = 7.231 (d) Skewness = 0.664 (e) Kurtosis = 0.481 (f) **Visibility = 3240.813**	(a) Mean = 14.386 (b) Std Dev = 11.507 **(c) Entropy = 5.22** **(d) Skewness = 1.183** **(e) Kurtosis = 2.834**

The SEM images are classified into bad, average and good images (a rule based method is used (Machine learning technique) for classification). The feature which is made bold is the main rule to be satisfy the condition for classifying the image

rules. The future work is to improvise text summarization with evaluation method and image classification to achieve better results using deep learning technique.

References

1. Hansen M, Pomp A, Erki K, Meisen T (2019) Data-driven recognition and extraction of PDF document elements. Technologies. https://doi.org/10.3390/technologies7030065
2. Pomp A, Paulus A, Kirmse A, Kraus V, Meisen T (2018) Applying semantics to reduce the time analytics within complex heterogeneous infrastructures. Technologies. https://doi.org/10.3390/technologies6030086
3. Diederik PK, Jimmy B (2015) Adam: a method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, Conference Track Proceedings
4. Zhang Y, Gong L, Wang Y, Yin Z (2003) An effective concept extraction method for improving text classification performance. Geo-Spat Inf Sci 6(4):66–72. https://doi.org/10.1007/BF02826953
5. Das GP, Vance PJ et al (2019) Computational modeling of salamander Retinal ganglion cells using machine learning approaches. Neurocomputing. https://doi.org/10.1016/j.neucom.2018.10.004
6. Hassan T, Baumgartner R (2005) Intelligent text extraction from PDF documents. In: International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), pp 2–6. https://doi.org/10.1109/CIMCA.2005.1631436
7. Fang Y, Bo L (2005) A new method of information extraction from PDF files. In: 2005 International Conference on Machine Learning and Cybernetics, vol 3, pp 1738–1742. https://doi.org/10.1109/ICMLC.2005.1527225
8. Joeran B, Bela G, Ammar S, Nick F (2010) SicPlore Xtract: extracting titles from scientific PDF documents by analyzing style information (font size). In: 14th European conference on digital libraries, vol 6273. Springer, Berlin, pp 413–416
9. Kasimahanthi D, Kambala S, Baisetti S, Sankara Rao G (2020) Text summarization using deep learning. IRJET 7(5). e-ISSN: 2395-0056

10. Shirwandkar NS, Kulkarni S (2018) Extractive text summarization using deep learning. In: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), pp 1–5. https://doi.org/10.1109/ICCUBEA.2018.8697465
11. Azimjonov J, Alikhanov J (2018) Rule-based metadata extraction framework from academic articles. Cornell University. arXiv:1807.09009
12. Akshaya B, Sruthi Sri S et al (2019) Content-based image retrieval using hybrid feature extraction techniques. Springer Nature Switzerland, ISMAC-CVB
13. Haritha H, Thangavel SK (2019) A modified deep learning architecture for vehicle detection in traffic monitoring system. Int J Comput Appl. https://doi.org/10.1080/1206212X.2019.1662171
14. Li P, Jiang X, Shatkay H (2019) Figure and caption extraction from biomedical documents. Bioinformatics 35(21):4381–4388. https://doi.org/10.1093/bioinformatics/btz228
15. Huang Y, Huang K, Yinan Yu, Tan T (2011) Salient coding for image classification. IEEE Xplore. https://doi.org/10.1109/CVPR.2011.5995682
16. Jayaram K, Prakash G, Jayaram V (2020) Automatic extraction of rarely explored materials and methods sections from research journals using machine learning techniques. IJACSA 11(8): 447–456. https://doi.org/10.14569/IJACSA.2020.0110857
17. Neena A, Geetha M (2018) Image classification using an ensemble-based deep CNN. Adv Intell Syst Comput. https://doi.org/10.1007/978-981-10-8633-5_44

A Time Dependent Epidemiology Model for Hospital Resource Management in Usual Scenarios and Pandemic

Aditya Bora, Atharva Nirali, Chetana Chaudhari, Dhananjay Gavade, and Vikramdas Vaishnav

Abstract Medicine plays important role for patient's treatment and disease prevention in case of pandemic situations. A healthcare institutions such as pharmacies or hospitals must ensure the availability of medicines as well as other resources such as masks, beds, PPE kits, etc. Hospitals as a health service provider have to face several challenges. One of the challenges is unavailability of Medicines during crucial periods/outbreaks, also due to lack of centralized systems for inventory management results in poor treatment to patients in hospitals. Predictive analysis of the amount of medicine usage and resources can help in hospital asset management. But forecasting medicine usage based on only inventory data will not provide good results when an outbreak happens. As outbreaks cannot be time-series forecasted, it is not feasible to forecast medicine usage only using time-series forecasting models. We propose a forecasting model combination of time forecasting model and epidemiology SIR model to find different trends and seasonality during outbreaks.

Keywords Hospital asset management · Time series forecasting · Pandemic · Bidirectional LSTM · Epidemiology SIR model

1 Introduction

Inventory Management plays an essential role in Healthcare Sectors, especially when people's lives are endangered. Hospital administration faces many difficulties such as shortage or overstock of Pharmaceuticals, loss or theft of surgical equipment, creating sales reports, etc. A COVID-19 pandemic puts extensive pressure on demands of healthcare facilities of hospitals who already have their hands full.

A study conducted by Rajagopalan and Choutagunta [1] conveys asset management in government hospitals in India is fragile in pandemic situations. It shows an assessment of India's healthcare capacity by using countrywide and state-level metrics such as healthcare budget spending, bed capacity, doctors, nurses, and other

A. Bora · A. Nirali · C. Chaudhari (✉) · D. Gavade · V. Vaishnav
School of Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_44

healthcare personnel [2]. Addressed the shortage of PPE (personal protective equipment) in India and proposed methods for conservation and management of the same. To enhance the scarcity of resources, reduce shortages and expand capacity quickly, proper supply chain strategies and practices are needed.

The health authorities need to develop and adapt emergency response plans, in order to ensure continuity of clinical services critical to patients affected by a Pandemic as well as others suffering from critical ailments. Khichar et al. [3] discovered aspects of pandemic preparedness undertaken at a tertiary-care hospital at Rajasthan, India catering to a population of 12 million. Hospital was divided into five zones and involved strategies such as suspension of elective surgeries or routine outpatient services, deployment of residents and nursing personnel, etc.

However, there was no data provided to the hospital regarding the probable inflow of patients in near future which can help them decide their plans to accommodate the incoming patients as well as the Staff and Doctors required to treat them. Also a pre listed requirements of resources, like PPE kits, Ventilators, Masks if provided a week or month before would have proven very integral while treating patients. These plans didn't provide any kind of representation of nearby hospitals in case of shortage of any resources and probability of resource availability which can help them in a faster communication.

1.1 Contributions

The main contribution of this paper is to provide a module for evaluating and predicting the amount of medicines and resources required in hospitals. We have used a combination of two models i.e. LSTM based medicine requirement prediction using historic medicine inventory data and Epidemiology model (SIR) based medicine requirement prediction using ongoing disease outbreak data.

In normal conditions, time-series methods and LSTM alone perform well in forecasting medicine usage based on historical data. When Epidemic spreads, new diseases can have different trends and no historical data is available. Hence, we have selected a model where infection rate and mortality rate can be calculated and monitored. For pandemic situations, we have created a machine learning based Epidemiology model which can predict the infection spread and medicine requirement. Trend in the epidemic spread assumes different nature in different phases over time. A more appropriate method is to use SIR (Susceptible Infectious Recovered) model. Our modified version of SIR model uses sequence model to update parameters of SIR and then use it for prediction.

The following portion of the paper is structured as follows: In Sect. 2, an overview of the related work is provided. Section 3 focuses on methodology details. Section 4 illustrates the results of experimentation and implementation. Section 5 provides the conclusion of the paper.

2 Related Work

In order to understand Inventory control and management strategies used by hospitals [4], several studies have been conducted. Hospital mainly uses ABC Analysis and VED classification, which give only overview of inventory and doesn't consider scenarios like natural calamities, emergencies and pandemic [5]. Describes how big data plays role in healthcare industries. It facilitates healthcare through prediction of epidemics, early warnings of disease conditions, and helping in the discovery of novel biomarkers and intelligent therapeutic intervention strategies for an improved quality of life [6]. Proposed a medicine requirement forecasting system using ARIMA method, which uses time series data. Deep Care [7] uses deep learning approach to predict healthcare trajectories [8]. Uses Bidirectional LSTM for prediction purpose and able to achieve more accurate results.

Yang et al. [9], Indscicov.in. [10] have explained SIR model which are commonly used to study the number of people having an infectious disease in a population. Each entity in Population is classified into three categories i.e. Susceptible, Infectious and Recovered by a model. Ferguson et al. [11] describes an early stochastic mathematical model of COVID19 epidemic in India with the objective of determining its magnitude and assessing the impact on healthcare resources [12]. Is based on SEIR model they predicted that the cases will reach 36 crore by the mid of July and the deaths will be 15 lakh which is unrealistic. It estimated the effect of non-pharmacological interventions (NPIs) including social distancing, lockdown. However accurate data cannot be gathered because contact tracing is very difficult in such scenarios. Though SEIR model is suitable for COVID-19 epidemic, our paper tends to provide generalized solution that can be used in any scenario, hence SIR models can be used for such cases.

3 Methodology

This section focuses on models that are integrated together to build proposed system. Figure 1 shows system architecture of the prediction module, it explains how models work in normal as well as pandemic situations. It consist of two models as follows:

- Bidirectional LSTM for prediction of medicine requirements
- Epidemiology (SIR) model for disease outbreak.

In normal conditions, Model is trained on hospital data using only Bidirectional LSTM, which can be later used to forecast medicine requirements. Time series models like LSTM, ARIMA perform well in case of non-variational trends. When disease outbreak occurs, patient inflow in hospitals increases which is dependent on rate of spread of disease. Due to this varying trend, mentioned models tend to overfit over data and might not be useful. Hence, in pandemic situations, Differential

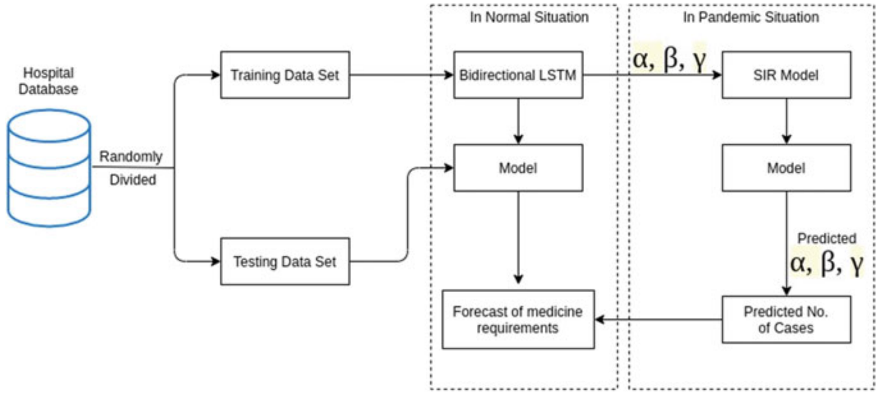


Fig. 1 System architecture of proposed system

model i.e. Susceptible Infectious Recovered (SIR) model is used to track the spread of infectious disease.

For Pandemic situations, Bidirectional LSTM is used to predict time varying mortality rate (α), transmission rate (β)and recovery rate (γ). These parameters are fed as input to a modified SIR model which can predict medical resources in epidemics. The output of the implemented system is a web application, where hospital administrators can forecast requirements of hospital resources and doctors and take necessary action. It also incorporates support modules such tracking medical resources in present and future, alerts in case of shortage of resources and communication to nearby hospitals. Main goal of the paper is to help hospitals manage medical resources in a better way in a pandemic situation.

3.1 Dataset

Model uses data which comes from mainly two resources i.e. state wise patient inflow data and medicine inventory data. Real time state wise data regarding number of deaths, confirmed cases and recovered patients during pandemics is collected from crowd sourced API [13]. In order to forecast medical resources for hospitals, healthcare resources data is also collected from different open government data platforms. In normal conditions, Medicine inventory data which comprises eight categories of having total 60,000 daily records divided into months and weeks [14, 15] provides Hospitals' geographical information and region wise medical resources information for Pune District. It contains over 196,000 records. Some preprocessing techniques are applied to remove data inconsistencies.

3.2 Bidirectional LSTM

This model takes historical data obtained and medical resources data as input and calculates alpha, beta and gamma i.e. time varying transmission rate, recovery rate and mortality rate using Bidirectional LSTM model.

Experimentation was conducted on different forecasting models such as ARIMA, Ridge regression, LSTM and Bidirectional LSTM. Bidirectional LSTM shows better performance as compared to other models. Bidirectional long short term memory is just putting two independent RNNs together. This structure allows the networks to have both backward and forward information about the sequence at every time step. Our model comprises two layers with 32 units. In order to improve working of the model, several techniques are used. Stochastic gradient descent is used as an optimizer. We have tuned parameters such as learning rate and momentum to reduce loss and get more accurate results.

3.3 SIR (Susceptible Infectious Recovered) Model

We have used the SIR model to forecast requirement of medical resources during epidemics. After predicting (t), $\beta(t)$ and $\gamma(t)$ using LSTM, system predicts number of Susceptible, Infected and Recovered cases by SIR model with predicted rates. SIR (Susceptible Infected Recovered) model divides population in three compartments i.e. Susceptible (individuals who does not have the disease yet), Infected (individual who currently has disease and can infect susceptible individual) and Recovered (individual who is recovered or dead from the disease and is no longer contagious). Traditional models ignores time-dependent nature of α, β and γ rates. We have developed a modified SIR model, in which the Recovered compartment is further divided into Recovered and Dead state. This model uses α, β and γ rate as a function of time i.e. (t), $\beta(t)$ and $\gamma\ (t)$ and can predict medical resources more accurately even in pandemic situations.

$$\mathrm{S}(\mathrm{t}+1) = \left(-\beta(\mathrm{t}+1) * \mathrm{S}(\mathrm{t}) * \frac{\mathrm{I}(\mathrm{t})}{\mathrm{n}}\right) + \mathrm{S}(\mathrm{t})$$
$$\mathrm{I}(\mathrm{t}+1) = \left(-\beta(\mathrm{t}+1) * \mathrm{S}(\mathrm{t}) * \frac{\mathrm{I}(\mathrm{t})}{\mathrm{n}}\right) - (\gamma(\mathrm{t}+1) * \mathrm{I}(\mathrm{t})) + \mathrm{I}(\mathrm{t})$$
$$\mathrm{R}(\mathrm{t}+1) = (\gamma(\mathrm{t}+1) * \mathrm{I}(\mathrm{t})) + \mathrm{I}(\mathrm{t}) + \mathrm{R}(\mathrm{t})$$

Mean absolute error is used as a metric. Three predicted rates in our system have three different MAEs i.e. MAE(α), $MAE(\beta)$ and $MAE(\gamma)$.

$$\mathrm{MAE}(\alpha) = \frac{1}{n}\sum \left|\alpha(t) - \hat{\alpha}(t)\right|$$

$$\mathrm{MAE}(\beta) = \frac{1}{n}\sum\left|\beta(t) - \hat{\beta}(t)\right|$$
$$\mathrm{MAE}(\gamma) = \frac{1}{n}\sum|\gamma(t) - \gamma(t)|$$

Total error for model is measured as

$$\mathrm{TotalError} \ = \mathrm{MAE}(S) + MAE(I) + MAE(R)$$

{S, I, R is predicted Susceptible, Infected and Recovered}.

Following algorithm shows working of our model:

Algorithm:

Input:*Historical data* $\{I, R, D$: *Infected*, *Recovered*, *Deceased*$\}$,

Available amount of medical resources $\boldsymbol{\Omega}$, **W***prediction window*

Output: $\left\{\hat{\alpha}, \beta, \gamma, \quad 0 \le \mathrm{t} \le \mathrm{T} - 2\right\}$,

$$\left\{\hat{\alpha}, \beta, \hat{\gamma}, \quad \mathrm{t} \le \mathrm{T} - 1\right\},$$

$$\left\{\hat{\alpha}, R, \hat{D}, \mathrm{t} \ge \mathrm{T}\right\}, \left\{\hat{\Omega}\right\}$$

Algorithm:

1. *Start*
2. *Calculate* $\{\alpha, \beta, \gamma, 0 \le t \le T\text{-}2\}$
3. *Train Bidirectional LSTM*
4. *Estimate* $\left\{\hat{\alpha}, \hat{\beta}, \hat{\gamma}\right\}$ *using trained model*
5. *Estimate the no. of infected* $\hat{I}$, *no. of recovered* $\hat{R}$ *and no. of deceased* $\hat{D}$.
6. *Estimate* $\hat{\Omega}\ (t) = \hat{I} - \Omega$
7. *While* $T \le t \le T + W$ *do*
8. *Estimate* $\hat{\alpha}(t)$, $\hat{\beta}\ (t)$ *and* $\hat{\gamma}(t)$
9. *Predict* $\hat{I}\ (t + 1)$, $\hat{R}\ (t + 1)$ *and* $\hat{D}\ (t + 1)$
10. $\hat{\Omega}\ (t + 1) = \hat{I}\ (t + 1) - \hat{\Omega}\ (t)$
11. *done*
12. *End.*

4 Results

Figure 2 and 3 shows that γ (Recovery rate) has an increasing trend and β (Transmission rate) has decreasing trend over time period of four months from March to July

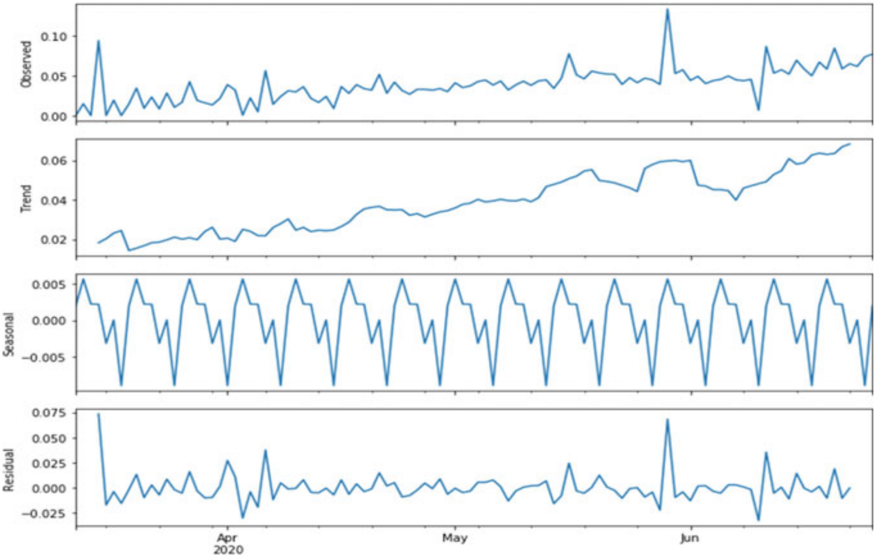


Fig. 2 Original data, trend, seasonality and residuals (from top to bottom) for γ (recovery rate)

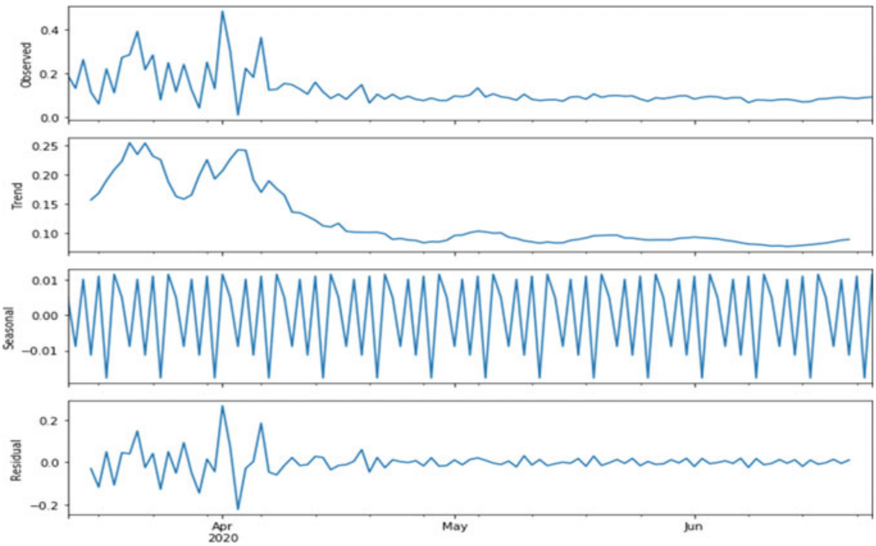


Fig. 3 Original data, trend, seasonality and residuals (from top to bottom) for β (transmission rate)

Table 1 Mean absolute error of different architectures for γ and β rates

Models used	MAE for γ rate	MAE for $\hat{\Omega}$ rate
Ridge regression	0.2169	0.3814
GRU (gated recurrent unit)	0.0807	0.1040
LSTM	0.0304	0.0810
Bidirectional LSTM	0.0140	0.0340

2020. It can be observed from results that transmission rate is higher at the start of disease outbreak and then gradually decreases with time. On the contrary, Recovery rate increases with respect to time. These rates show seasonality and trends which proves that rates have time-dependency.

Different time-series forecasting models were used experimentation. Results obtained i.e. Mean Absolute Error for different architectures is shown in Table 1. It can be observed that Bidirectional LSTM shows better performance in prediction of both γ (Recovery rate) and β (Transmission rate). Figure 4 shows decreasing Mean Absolute Error and Loss over number of epochs using Bidirectional LSTM.

Implemented system i.e. web application Fig. 5 uses these combined models to predict patient inflow and then predict requirement of hospital resources and doctors in daily basis. It can predict patient inflow and also forecast requirements of medical resources and has different modules such as prediction of medicine requirements for next month, medicine inventory, forecasting of number of patients in future, predicting medical resources like beds in near future, etc. It uses Flask a Python framework as backend and AWS RDS service as PostgreSQL database.

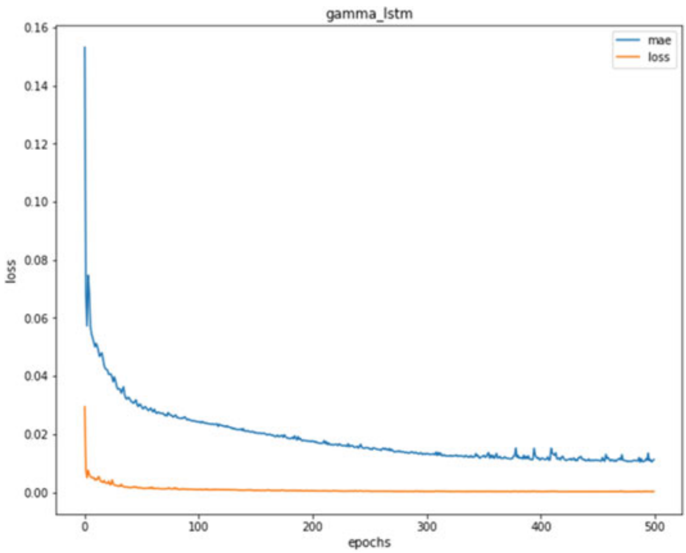


Fig. 4 Mean absolute error and loss of γ rate using bidirectional LSTM

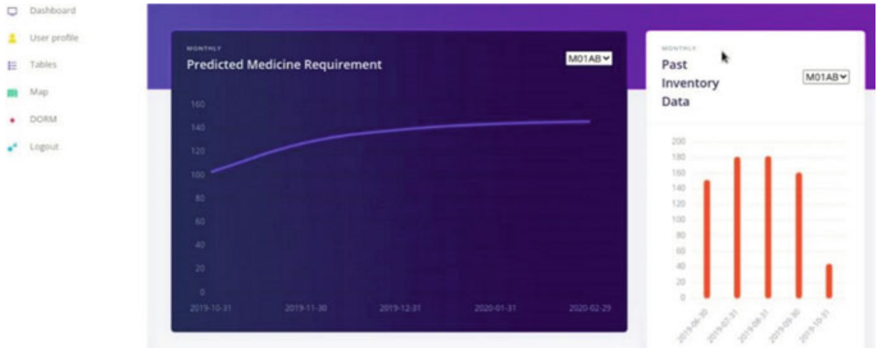


Fig. 5 Dashboard showing medicine inventory data and predicted medicine requirements

5 Conclusion

This paper attempts to make contribution in the healthcare industry and inventory management for government hospitals. The purpose behind this paper is to understand hospital administrator's needs and difficulties faced during critical situations like pandemics or natural calamities. Thus, we have designed a system that can adapt with the disease outbreaks and help in managing medical resources. It contains dashboard for same purpose and supporting modules. We have successfully developed a forecasting model using combination of Bidirectional LSTM and SIR model. By comparing predicted results with real-time patient inflow data, it can be concluded that accurate results are obtained. System can be upgraded in future which can provide people information regarding doctor and nearby hospital availability helping them to get quick and proper treatment. Such aspects can be useful in pandemic situations like COVID-19.

References

1. Rajagopalan S, Choutagunta A (2020) Assessing healthcare capacity in India. Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA
2. Bhattacharya S, Mahbub Hossain M, Singh A (2020) Addressing the shortage of personal protective equipment during the COVID-19 pandemic in India—a public health perspective. AIMS Public Health 7(2):223–227
3. Khichar S, Midha N, Bohra G, Kumar D, Gopalakrishanan M, Kumar B, Sakthivadivel V, Garg M (2020) Healthcare resource management and pandemic preparedness for COVID-19: a single centre experience from Jodhpur, India. Int J Health Policy Manag
4. Devnani M, Gupta A, Nigah R (2010) ABC and VED analysis of the pharmacy store of a tertiary care teaching, research and referral healthcare institute of India. J Young Pharm 2(2):201–205. https://doi.org/10.4103/0975-1483.63170
5. Dash S, Shakyawar SK, Sharma M et al (2019) Big data in healthcare: management, analysis and future prospects. J Big Data 6:54

6. Pratyaksa H, Permanasari AE, Fauziati S, Fitriana I (2016) ARIMA implementation to predict the amount of antiseptic medicine usage in veterinary hospital. In: 2016 1st international conference on biomedical engineering (IBIOMED), Yogyakarta, pp 1–4. https://doi.org/10.1109/IBIOMED.2016.7869815
7. Pham T, Tran T, Phung D, Venkatesh S (2017) Predicting healthcare trajectories from medical records: a deep learning approach. J Biomed Inf 69:218–229. ISSN 1532-0464
8. Althelaya KA, El-Alfy EM, Mohammed S (2018) Evaluation of bidirectional LSTM for short- and long-term stock market prediction. In: 2018 9th international conference on information and communication systems (ICICS), Irbid, pp 151–156. https://doi.org/10.1109/IACS.2018.8355458
9. Yang F, Liu H, Qi H, Liu X (2016) SEIR evolutionary simulation model of the infectious disease emergency. In: 2016 International conference on industrial informatics—computing technology, intelligent technology, industrial information integration (ICIICII), Wuhan, pp 315–318. https://doi.org/10.1109/ICIICII.2016.0082
10. Indscicov.in (2020) Explaining models of epidemic spreading [online]. Available at: https://indscicov.in/for-scientists-healthcare-professionals/mathematical-modelling/explaining-models
11. Ferguson NM, Laydon D, Nedjati-Gilani G et al (2020) Impact of nonpharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team, pp 20
12. Chatterjee K, Chatterjee K, Kumar A, Shankar S (2020) Healthcare impact of COVID-19 epidemic in India: a stochastic mathematical model. Med J Armed Forces India 76(2)
13. Covid19india.org. (2020) Coronavirus in India: latest map and case count [online]. Available at: https://www.covid19india.org
14. Opendata.punecorporation.org (2020) PMC open data store [online]. Available at: http://opendata.punecorporation.org
15. Open Government Data (OGD) Platform India (2020) Open government data (OGD) platform India [online]. Available at: http://data.gov.in

Load Forecasting Using LSTM and Its Comparison with Other Machine Learning Techniques and Their Implementation

Apoorva Mishra and J. N. Rai

Abstract This paper presents hourly electrical load forecasting as time series forecasting model using Multilayer deep learning and Long Short-Term Memory neural network Technique and their detailed comparative study with various machine learning techniques based on their mean Squared Error (MSE), mean absolute percentage error (MAPE), root mean squared error (RMSE) and training time. Load Forecasting has immense potential to help in modulating the generation and distribution potentials of our smart grids in accordance to the requirement so that optimum power is generated and supplied through various channels which would be effective in grid management and operations.

Keywords Hourly electrical load forecasting · Time series forecasting · Long short-term memory neural network · Smart grids · Deep learning · Machine learning · Long short term memory network (LSTM) · Recurrent neural networks (RNN) · Root mean square propagation (RMSProp)

1 Introduction

The optimum planning and management of energy distribution is a very crucial task when dealing with smart grids. There has been immense evolution in the smart grid, using the developments of information and communication technologies, and it is becoming a productive and durable system progressively. The smart grid systems are contrived so as to counter the problem of energy management and also to monitor, optimize and control the distribution of power. The decision for the flow or exchange of energy among all the utilities or devices connected to the grid is made after the governance and assessment of demand; which is an indispensable constituent of the energy management system of the electrical grid, thereby assuring the operational functionality, stability and dependency of the entire interconnected electrical system.

A. Mishra (✉) · J. N. Rai
Electrical Engineering Department, Delhi Technological University, Delhi, India

J. N. Rai
e-mail: jnrai@dce.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_45

This will be beneficial in Strategizing the supply according to the demand therefore the reliability of the grid is increased [1, 2]. It will also be economical as renewable energy sources are integrated with greater efficiency and lower costs. We will be using deep learning techniques for the predictions meanwhile comparing them with the other machine learning techniques [3].

2 Time Series Forecasting

Time series forecasting is done over successive time intervals using sequent data points. The measurements are arranged in a sequential manner and the forecasting can be univariate or multivariate in nature. In our problem statement we will be using univariate analysis with time as one variable and power the other. The LSTM Technique learns and tracks all the past dependencies of a variable in relation with the new observations. Normally in a neural network input is given to a hidden layer and output is obtained, but LSTM is based on the recurrent neural networks where input is given to a set of hidden layers which contain information of the previous inputs also and then output is obtained.

2.1 LSTM Description

It is an improvement of the recurrent neural network which can operate and keep a track of long term dependencies. In a RNN network output of the hidden layers is fed back again to the network itself unlike the normal feed forward network which does not work in loop. The RNNs have the same weights for all the time steps as they operate in a loop. The problem that recurrent neural networks face is that due to large scale dependencies on the previous data, if we want our model to learn better, as we increase our time steps, the problem of vanishing or exploding gradients becomes increasingly inevitable [4].

It consists of the forget gate, with the purpose of altering the portion of input or the previous stage which is to be passed to the next stage, deciding what percentage of the data of the previous stage we need to remember. It is made up of a sigmoid layer which outputs a number between 0 and 1 for each state depending on how much to keep and how much to discard. The output of this forgets gate can also be thought of as weights to the internal state for each input [5].

The inner state of the LSTM is defined by a variable s which is delayed by a one-time step and is used by the network to learn the relationship between the inputs and their sequence time thereby creating a recurrent loop. This filtering done by the forget gate also helps to reduce the effect of the vanishing gradient problem encountered in the recurrent neural network.

$$\mathrm{f} = (\mathrm{W}[\mathrm{h}(\mathrm{x} - 1), \mathrm{i}(\mathrm{x})] + \mathrm{b}) \tag{1}$$

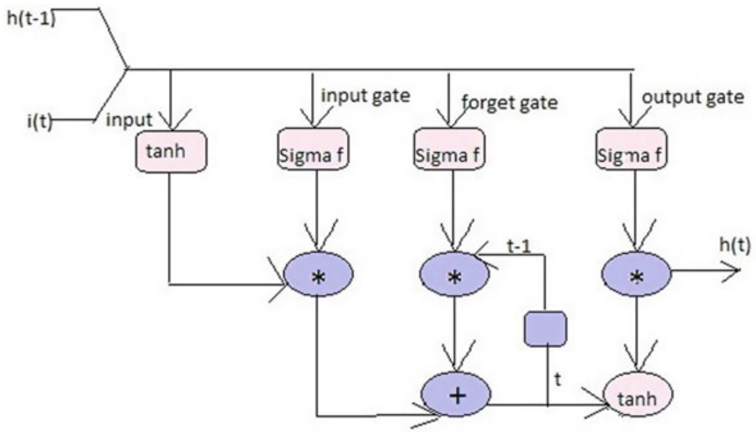


Fig. 1 LSTM structure

Then comes the input gate which decides which values are to be updates and what new information is added to the cell state and is also a sigmoid layer followed by a tanh layer the results of these two layers are multiplied together to get the new updated scaled value and is then added to the output from the forget gate to make a new cell state. This new state is again fed to tanh which gives values between − 1 and 1 to again decide which values are to be passed. Thus LSTM provides the advantage of adding or removing any information in the cell through gates [6]. Figure 1 shows the basic cell structure of a LSTM network.

2.2 Algorithm

After dividing the entire dataset into training and test data in the ratio 9:1 and reshaping the input in the form (number of samples, dimension of elements) we start building a sequential model.

Layer 1:

Add a LSTM layer with input dimension as 1 and output dimension as 50 with return sequence as True.

Add a Drop out layer to prevent overfitting.

Layer 2:

Add a LSTM layer with output dimension as 50 and return sequence as False.

Add a Dropout layer to prevent overfitting.

Layer 3:

Add a Dense neural network layer with output dimension as 1 and activation linear.

While compiling the model, the optimizer used is "rmsprop" and loss to be minimized is specified as "mse". The gradient descent optimizers are used to reach the global minima by finding the gradient and updating the weights and biases such that the cost functions can attain minimum value. However, sometimes non-convex functions arise due to which the optimizer is stuck in the local minima hence the global minimum value is not achieved by the cost function. RMSProp optimizer is used as it solves this problem. RMSProp restricts the oscillations in one direction, basically the vertical direction thereby converging faster to the solution as higher learning rates can be used [7].

BEGIN

Prediction values of the neural network are taken
The loss function is used to calculate the losses
The loss functions are differentiated and the gradients are obtained; these are also the partial derivatives.
These gradients update the values of weights and biases.

REPEAT

$$\mathrm{vdw} = .\mathrm{vdw} + (1-)\ .\mathrm{dw}^2 \tag{2}$$

$$\mathrm{vdb} = .\mathrm{vdb} + (1-)\ .\mathrm{Db} \tag{3}$$

$$\mathrm{W} = \mathrm{W} - \alpha \cdot \left[\mathrm{dw}/\left(\mathrm{v}^{1/2} + €\right)\right] \tag{4}$$

$$\mathrm{b} = \mathrm{b} - \alpha \cdot \left[\mathrm{db}/\left(\mathrm{v}^{1/2} + €\right)\right] \tag{5}$$

The following notations are used here:

W: weights, b: biases, dw: partial derivative of the weights, db: partial derivative of the biases, β: momentum value, α: learning rate, €: constant epsilon.

This method of updating the weights and biases while learning also comprises the backpropagation algorithm which is used in most of the neural networks.

3 Algorithm Used for Other Machine Learning Techniques

After dividing the entire dataset into training and testing in the ratio 9:1, we get properties of all the regressors Ridge, K nearest Neighbors, Random forest regressors, Gradient Boosting Regressors, MLP regressors and Extra trees Regressors. We fit the model, take out the training time, testing score and training score, RMSE, MAPE and MAE, store them in a dataframe for comparison purposes and plot each of them.

Description of the various kinds of regressors:

- Linear regression works similar to a linear relation in statistics. In machine learning since we have a larger number of dimensions of the inputs we call it a hyperplane instead of a line. This algorithm works on the assumption that the relationship between your input and output is linear in nature and does not have any noises in data. Also we need to make sure that we remove all the collinearity in data as those may overfit our model. For better predictions through linear regression we can normalize our input [8].
- Random Forest Algorithm works by constructing multiple layers of decision trees while training which outputs the mode in case of classification and the mean in case of regression problems. It constructs an ensemble of decision trees and then merges the outputs thereby increasing the accuracy of the prediction. This algorithm when searching through the features, while forming the trees, instead of selecting the most important feature, selects the best feature from the subset of features while splitting the node thereby adding some randomness [9]. The feature importance detection provided by random forest algorithms really contributes a lot in deciding which feature to drop depending upon its contribution to the model, thereby avoiding overfitting.
- The Gradient boosting works on the method of minimization of error. The prediction model given by a gradient boosting algorithm is generally the ensemble of many models arranged in the sequence which may produce the minimum error, mostly a set of decision trees. It works like the adagrad algorithm which trains a decision tree by assigning equal weights to all the observations [1]. Further as we proceed, after the first iteration and evaluation of the first decision tree, we assign higher weights to those data which are difficult to classify and lower weight to the data that is easy to classify so that we can improve upon the predictions from the first iteration. Repeating this for a particular number of iterations the model finally gives the weighted sum of the predictions.
- K-Nearest Neighbor works on the measured Euclidean distances between the query and the data points and then operating on a particular number (K) of data samples then finds the most suitable class for it in case of classification or takes the average of the points in case of regression. We train the dataset many times with different values of k and find those values of k at which the error is minimum when the algorithm is subjected to unseen data [10].
- MLP regressor is a multi-layer perceptron supervised learning, non-linear training technique that uses backpropagation to update weights. A basic perceptron network has 4 parts, an input layer, weights and biases, summation and an activation function. Perceptrons receives the input, pass them to the hidden layer which gives the weighted linear sum of the inputs, this is further passed on to an activation layer. The activation function which is non-linear in nature maps the input according to our requirements, which is passed to further layers. Adding many such layers makes the Multilayer Perceptron neural network [11].
- Extra Trees Regressor is also a variation of the Random Forest where instead of selecting an optimal feature we select a random feature for split [12].

- Ridge regressor is used in the cases when there is no unique solution available and the answer is to be calculated using approximation. Ridge regression operation is performed using L2 regularization which is the sum of squares of coefficients [13].

3.1 Evaluation Index Equations

$$\mathrm{MAPE} = (1/\mathrm{n}) * \Sigma\left[\mathrm{abs}(\mathrm{actual}-\mathrm{predicted})/\mathrm{actual}\right]$$

$$\mathrm{MAE} = \left[\Sigma\ \mathrm{abs}(\mathrm{actual}-\mathrm{predicted})\right]/\mathrm{n}$$

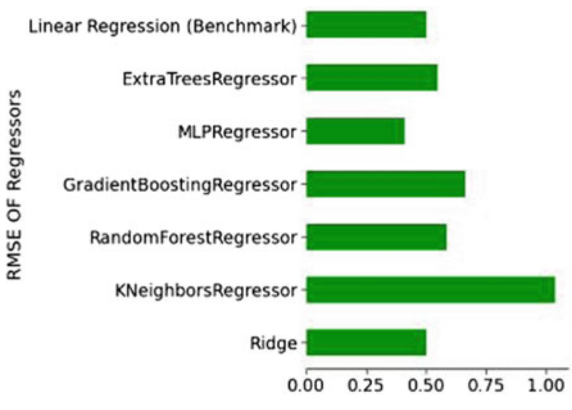
$$\mathrm{RMSE} = \left[\Sigma(\mathrm{predicted}-\mathrm{actual})^2/\mathrm{n}\right]^{1/2}$$

4 Results

Seaborn and matplotlib are used to plot the results. From the Figs. 2, 3 and 4 where we are comparing the RMSE, MAE and MAPE value of all the regressors it is clear that the lowest RMSE, MAPE and MAE value is for MLP regressor. Training score is not the highest but not too low for MLP regressor. Therefore, the MLP regressor is better than the other techniques which signify that a perceptron based network or a neural network would give the best performance on forecasting problems. Now we know LSTM is also composed of neural networks and is an improvement of the MLP networks. Comparing the MLP regressor to the LSTM network (Fig. 5):

LSTM results:

Fig. 2 RMSE values of all the regressors



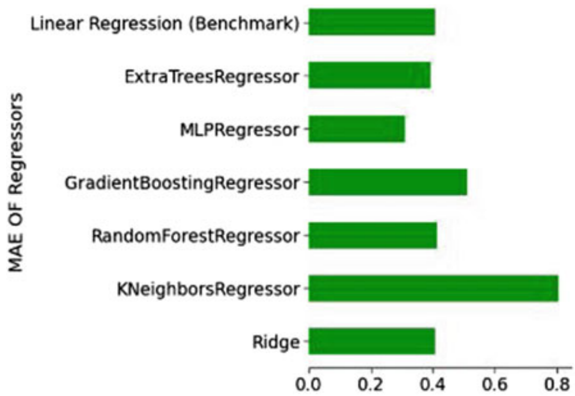


Fig. 3 MAE values of all the regressors

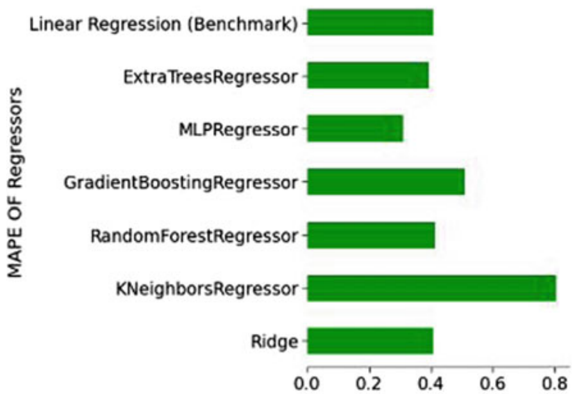


Fig. 4 MAPE of regressors

Fig. 5 Training time of regressors

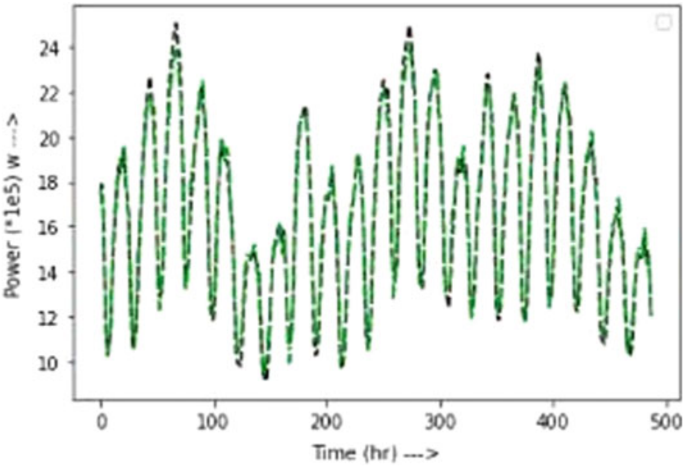


Fig. 6 Output of LSTM network predictions versus original power

TestMAPE: 0.41 MAPE

TestMSE: 0.27 MSE

TestRMSE: 0.51 RMSE.

MLP results:

TestMAPE: 0.311283 MAPE

TestMSE: 0.311283 MSE

TestRMSE: 0.40978 RMSE.

Though MAPE and RMSE values of MLP are better than LSTM. But, for the purpose of forecasting we would require the properties of learning and remembering information from the previous layers over a long chain, offered by LSTM. In Fig. 6 we can see that the dotted black line represents the original data for power and the green line represents the predicted power using the LSTM algorithm. Hence our model forecasts power correctly.

5 Conclusion

We can see from the data that MLP is always a better choice as compared to all other algorithms for forecasting as the key performance values of MLP method are better than the rest of the algorithms.

However, MLP alone does not support the operation of remembering long term data. Therefore we introduced a looping or recurrent loop method where all the weights and biases for all the layers were the same unlike MLP. Recurrent Neural Network was an improvement of the MLP network.

This Recurrent Neural Network However faces the problem of vanishing gradient when the data to be remembered becomes too large. To counter this problem LSTM was introduced which provided the benefits of MLP and Recurrent neural networks and is an improvised form of both the networks which support larger data memory and we can anytime use the forget gate and input gate to remove or add new data as per our requirements.

Hence we have successfully implemented LSTM neural networks to forecast load power studying and implementing all the other algorithms that can be used for forecasting. Concluding that LSTM is performing the best till now for load forecasting purposes.

References

1. Su F, Xu Y, Tang X (2018) Short and mid-term load forecasting using machine learning techniques. In: IEEE 2017 china international electrical and energy conference, 21 June 2018
2. Vantuch T, Vidal A, Ramallo-Gonzalez AP, Skarmeta AF (2018) Machine learning based electric load forecasting for short and long-term period. In: 2018 IEEE 4th world forum on internet of things (WF-IoT)
3. Hosein S, Hosein P (2017) Load forecasting using deep neural networks. In: 2017 IEEE power and energy society innovative smart grid technologies conference (ISGT)
4. He W (2017) Load forecasting via deep neural networks. Procedia Comput Sci 122:308–314
5. Tian C, Ma J, Zhang C, Zhan P (2018) A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. Energies 11:3493. https://doi.org/10.3390/en11123493
6. Kumar S, Hussain L, Banarjee S, Reza M (2018) Energy load forecasting using deep learning approach-LSTM and GRU in spark cluster. In: 2018 fifth international conference on emerging applications of information technology. https://doi.org/10.1109/EAIT.2018.8470406
7. Nguyen LC, Nguyen-Xuan H (2020) Deep learning for computational structure optimization. ISA Trans 103:177–191
8. Amral N, Ozveren CS, King D (2007) Short term load forecasting using multiple linear regression. In: 2007 42nd international universities power engineering conference
9. Dudek G (2014) Short-term load forecasting using random forests. In: Intelligent systems 2014, part of advances in intelligent systems and computing, pp 821–828
10. Lv X, Cheng X, Shuang Y, Tang Y-M (2018) Short-term power load forecasting based on balanced KNN. In: IOP conference series: materials science and engineering, vol 322, pp 072058.https://doi.org/10.1088/1757-899X/322/7/072058
11. Dudek G (2020) Multilayer perceptron for short-term load forecasting: from global to local approach. Neural Comput Appl 32:3695–3707
12. Ahmad MW, Mourshed M, Rezgui Y (2018) Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression. Energy 164:465–474
13. Cheung CM, Kannan R, Prasanna VK (2018) Temporal ensemble learning of univariate methods for short term load forecasting. In: 2018 IEEE power and energy society innovative smart grid technologies conference (ISGT)

Application of Machine Learning to Detect Neuroticism in Individuals Using Handwriting Analysis

Sheetal Thomas, Mridula Goel, Anmol Agarwal, and Asadali Abbas Hazariwala

Abstract A person's brain is more engaged while writing than in speaking or listening; emphasizing the importance of handwriting in describing a personality. Use of automated handwriting analysis can improve an organization's ability to select suitable candidates. The paper specifies the handwriting features related to neuroticism based on the Big Five Facets. It also proposes a model to automate identification of handwriting features related to neurotic traits, thereby reducing bias and time required to conduct a manual handwriting analysis. The paper establishes that machine learning techniques had a test accuracy of more than 95% in identifying handwriting features associated with neuroticism. The handwriting features like lack of symmetry in zones and heavy slashing strokes were identified with more than 99% test accuracy.

Keywords Handwriting analysis · Neuroticism · Machine learning · Deep learning · Data augmentation · Big five personality factor

1 Introduction

A person's brain is more engaged while writing than in speaking or listening; emphasizing the importance of handwriting in describing a personality [1]. Handwriting analysis can be used to understand an individual's personality. A trained graphologist is able to accurately identify the personality of an individual [2], but hiring graphologist's services may be expensive. Besides manual analysis is time consuming and

S. Thomas (✉) · M. Goel · A. Agarwal · A. A. Hazariwala
Birla Institute of Technology and Science, K. K. Birla Goa Campus, Pilani, Goa, India
e-mail: p20170420@goa.bits-pilani.ac.in

M. Goel
e-mail: mridula@goa.bits-pilani.ac.in

A. Agarwal
e-mail: f20170489@goa.bits-pilani.ac.in

A. A. Hazariwala
e-mail: f20170774@goa.bits-pilani.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_46

susceptible to subjective bias. Online tools are available for personality analysis and classification, but they are not effective in all situations. Designing an automated tool based on handwriting features can produce relatively more accurate insights. Such a tool can be useful for recruiters, educational and financial organizations. It can reduce risks in building business collaborations and other partnerships, especially in small group scenarios.

The Big Five Personality Factor Model is a widely accepted experiment-based model to identify personality traits [3, 4]. The Big Five Personality Factor questionnaire is used to classify human personality into Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness.

In the present study machine learning techniques are used to identify the trait of neuroticism from handwriting features. The accuracy of the test results thus obtained is found to be high. This will enable the development of an application to automate the process. A similar study using a non-invasive, neural network based system by Gavrilescu and Vizireanu on 128 handwritten samples showed an accuracy of 84% [5]. Another study by Anna Esposito and others using both handwriting and drawing patterns had also achieved a good level of accuracy [6].

2 Handwriting Analysis

A number of handwriting features like slant, size, pressure, and patterns visible in handwriting can give clues about the psychological characteristics of the writer [7]. Researchers find that innate writing formations do not significantly change over time and can be used as a valid method of analysis to gain insights into the individual's personality [8]. For example hurried irregularity in the letter slants of the writer would reveal moodiness, while the letter 't' crossed at the right of stem shows irritability. Study conducted by Eysenck to validate the assessment made by graphologists with respect to the writer's degree of neuroticism, concluded that there was a significant correlation between the graphological assessment and neurotic scores [9]. In the present work we have used graphology to identify handwriting features[1] that show neurotic traits of personality. These features were used as labels to feed our machine learning models.

3 Neuroticism

The personality trait of neuroticism captures intense negative emotions and feelings. In challenging situations where an emotionally stable individual would remain calm and resilient, neurotic individuals may experience anxiety, stress, and depression [10]. Anxiety moves them to take cautious financial decisions as they are sensitive to

[1] The first author is a trained graphologist.

financial losses. For this reason neurotic individuals prefer safer investment options. This behavior pattern may perplex financial advisors and they may not be able to offer products that satisfy the neurotic investor's emotional sensibilities [11]. Neurotic traits also cause individuals to take loans more often and make compulsive purchases [12], leading to financial distress [13]. Neurotic individuals may be less disciplined in repayment of loans.

An automated system of identification of individuals with high neurotic scores can help financial institutions to take calculated risks and reduce liabilities. Studies also suggest that neurotic individuals experience lower levels of contentment in jobs and relationships [14]. Neurotic individuals have a tendency to complain and are often over dependent on others. This can be cause of concern in the work setting as such individuals can add to complexities and conflict in teams. These persons are prone to anger and impulsiveness; they may find it difficult to sustain business commitments [15].

4 Machine Learning

In recent years researchers have started to explore the use of artificial intelligence in identifying personality traits. Studies have used Artificial Neural Networks (ANNs) to analyze individual personality traits using features such as baseline, letter slant, pen strain, letter 'i' and letter 'f'. These have been identified using polygonization method, grey-level threshold values and template matching [16]. Researchers have used Graphology, LSTMs, Optical Character Recognition and Convolutional Neural Networks, along with other Multi-modelling approaches, to identify handwriting features that can help to predict personality traits [17]. Connectome-based predictive modeling (CPM) a machine learning technique for extracting and pooling features from connectivity data has also been used to construct predictive models for personality recognition [18]. A recent study by M. K. Sharma and others shows that combining handwriting stroke features and single characters for personality recognition gave an impressive overall accuracy of 99.99% [19].

5 Experiment

5.1 Participants

The questionnaire design was based on the Big Five Personality Factor model. Convenience sampling was used to collect primary data from 302 participants. The questionnaire design consisted of twelve statements to gather behavioral patterns related to each of the five factors of the model. For the collection of sample handwriting texts we included three open ended questions related to behavior displayed in particular situations and a predetermined text of "The London Letter" [20]. Participants were sensitized prior to filling the questionnaire, about the motivation of the study and personal implications. This also, helped the participants to settle down into normal behavior patterns.

5.2 Procedure and Methods

Identification of Handwriting features. It is possible to infer handwriting features from a discursive handwritten paragraph or from few sentences, word, or even a single character [21]. An extensive literature review conducted during the study helped to gather handwriting correlates that show behavior associated with particular traits [22–24]. Table 1 shows the handwriting features that predict the handwriting correlates mapped to the Facets of Neuroticism as per the Big Five Personality Factor Model.

Data Pre-processing. Using the predetermined paragraph "the London letter" provided a handwritten text dataset controlling for any variations other than handwriting styles of individuals. However out of the 302 questionnaires, only 261 contained this predetermined handwritten text and were retained for further processing.

5.3 Data Processing and Augmentation

Data Processing. As deep learning models are 'Garbage in Garbage out' kind of model, it is necessary to perform the right data processing steps to extract features to be used for training our model. The images were read and processed using the OpenCV package in python. Non-Local Means Noise reduction methods available in OpenCV, which are based on Gaussian Blurring smoothing technique, were used to remove noise from the images. We have used different sets of values for window size and filter strengths to create various augmented versions of the dataset. The RGB images were converted to grayscale to normalize color variations. The grayscale images have 256 possible intensities of black color with 0 being black and 255 being white. The cropped images were resized to 1024×512 pixels per image to remove any variation in size that may have arisen due to cropping done to extract the predefined text (Fig. 1).

We used 5 different threshold values to augment our dataset, and applied a global thresholding technique to speed up the training process. Pixels with intensity greater than the threshold were set to maximum intensity i.e. 255 and intensity of all other pixels was reduced to 0. This thresholding helps to remove the various intensities of pixels; it thereby speeds up the training process dramatically and improves the soundness of the experiment.

Table 1 Neuroticism big five personality factor—big five facets—handwriting correlates—handwriting features

Big five personality factors	Big five facets	Handwriting correlates	Handwriting features
Neuroticism	Anxiety	Tense	• Too Angular, devoid of roundedness • Small size (middle zone)
	Angry hostility	Irritable	• T cross to right • Lack of Symmetry in zones
	Depression	Despondency	• Varying or erratic slant • Sudden change in speed • Sudden change in size and • Sudden change in pressure • Falling or descending bass lines
	Self-consciousness	Shy	• Capital Letter—Tall and Narrow • Arcade • Wide left margin
	Impulsiveness	Impulsive	• Extreme Right slant
		Moody	• Varying or erratic slant • Variable size
	Vulnerability to stress	Worrying	• 'M' lower loops-larger the loops more the worry

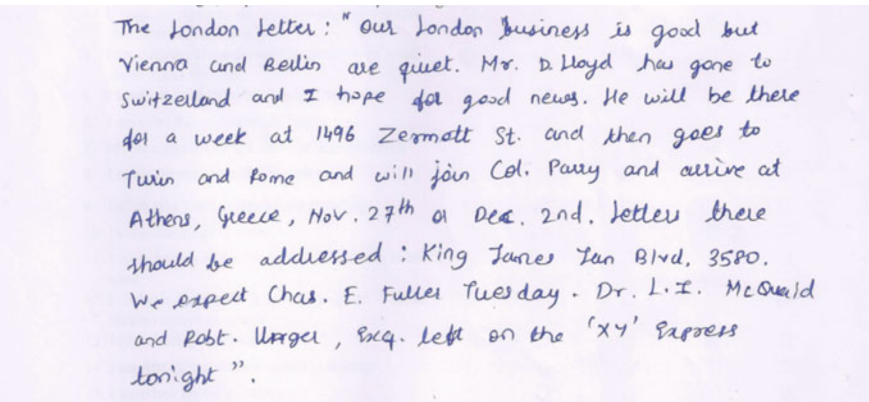
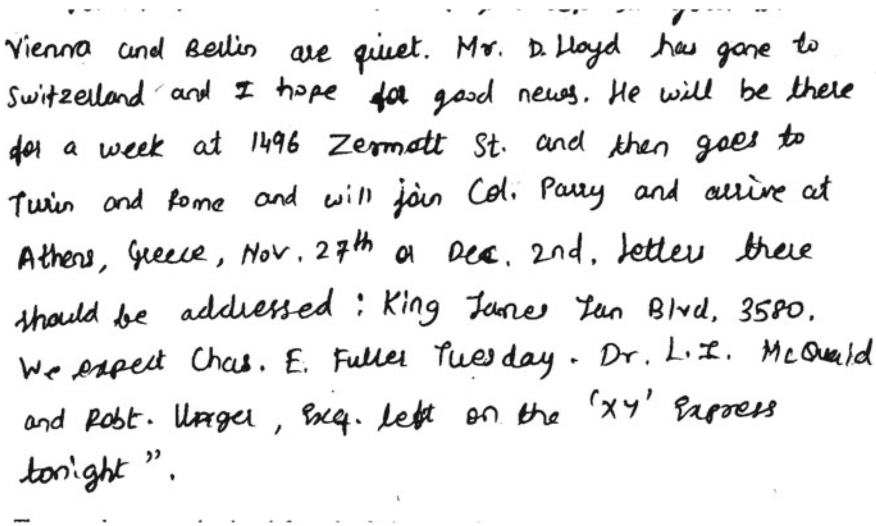

The London Letter: "Our London business is good but
Vienna and Berlin are quiet. Mr. D. Lloyd has gone to
Switzelland and I hope for good news. He will be there
for a week at 1496 Zermatt St. and then goes to
Turin and Rome and will join Col. Parry and arrive at
Athens, Greece, Nov. 27th or Dec. 2nd. Letters there
should be addressed: King James Jan Blvd, 3580.
We expect Chas. E. Fuller Tuesday. Dr. L. I. McQuald
and Robt. Unger, Esq. left on the 'XY' Express
tonight".

Fig. 1 Cropped and resized image



Vienna and Berlin are quiet. Mr. D. Lloyd has gone to Switzerland and I hope for good news. He will be there for a week at 1496 Zermatt St. and then goes to Turin and Rome and will join Col. Parry and arrive at Athens, Greece, Nov. 27th or Dec. 2nd. Letters there should be addressed: King James Blvd. 3580. We expect Chas. E. Fuller Tuesday. Dr. L. I. McQuald and Robt. Unger, Esq. left on the 'XY' Express tonight".

Fig. 2 Processed image after denoising, application of threshold and augmentation

Data Augmentation. Deep Learning models usually require huge data to train, but as our dataset was small, consisting of only 261 images, we used Image Augmentation techniques [25, 26]. This helped to increase the dataset size from 261 to 4176 images. This improved the model correctness and reduced overfitting [27, 28]. Before performing the data augmentations, the dataset was divided into training and testing sets to prevent any data leakage. The preprocessed images were then denoised and different thresholds were applied (Fig. 2). Further, the images were processed to obtain line segments. The colors were also inverted in some images to add variations (Fig. 3).

The row segmentation (Fig. 3) was done to compare the letter size in the handwriting samples, as this also an important handwriting feature. There were 16 classes of images created by varying the threshold values for pixel intensity and denoising filter strength. These images were then used to train the model.

5.4 Modeling a Neural Network

A Convolutional Neural Network (CNN) was used for the purpose of feature extraction and identification of a handwriting features in the given handwriting. The feature extraction was done by the convolutional and max pooling process.

The first layer of the model contained a Convolutional 2D layer consisting of 64 neurons, followed by a Max pooling layer followed by a dropout layer. These layers were added to reduce the dimensionality and overfitting of the model respectively. Another set of Convolutional 2D layer, Max pooling and dropout layer was repeated.

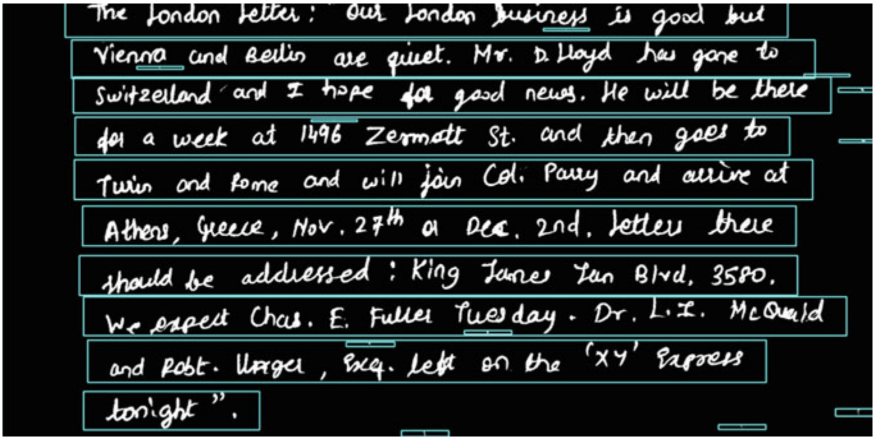


Fig. 3 Processed image with inverted colors

Subsequently, 4 fully connected layers were added with 2 dropout layers between them. The final layer consisted of one neuron which was used to predict the class of the handwriting feature.

This layer was activated through the sigmoid function which gives a value between 0 and 1. As per convention, we split the dataset into training and testing sets, with 80% images used for training and 20% for testing the correctness of the model. The training set was further divided in the ratio 80:20 to create a validation set to perform a K-Fold Cross validation while training the model. While splitting the data in train and test, it was made sure that the splits were exclusive and there was no data leakage. The model is represented in Table 2.

6 Results

Table 3 gives the percentage of trained accuracy and test accuracy obtained for the different features predicted through machine learning model applying handwriting analysis.

From the results we can observe that the model is able to predict the traits with a high accuracy. The best prediction was observed in the classes "Heavy slashing strokes" with a test accuracy 99.76% and "Lack of Symmetry in zones" with a test accuracy of 99.64%. It can be noted that most of the handwriting features had a test accuracy greater than 90%.

There remains some scope for improvement in the class "'M' lower loops- Larger the loops more the worry". This possibly occurred due to the few number of M letters present in the text. Other classes which could be improved included "Unfinished down stroke" and "Sudden change in speed, size, and pressure" which had test accuracy of 89.83% and 90.79% respectively.

Table 2 Layers used in he proposed model

Layer (type)	Output shape
conv2d_1 (Conv2D)	(None, 253, 509, 64)
Max_pooling2d_1 (MaxPooling2D)	(None, 63, 127, 64)
dropout_1(Dropout)	(None, 63, 127, 64)
Conv2d_2 (Conv2D)	(None, 61, 125, 32)
Max_pooling2d_1 (MaxPooling2D)	(None, 20, 41, 32)
Dropout_2 (Dropout)	(None, 20, 41, 32)
Flatten_1 (Flatten)	(None, 26,240)
dense (Dense)	(None, 256)
dropout_3 (Dropout)	(None, 256)
dense_1 (Dense)	(None, 128)
dropout_4 (Dropout)	(None, 128)
dense_2 (Dense)	(None, 16)
dropout_5 (Dropout)	(None, 16)
dense_3 (Dense)	(None, 1)

Table 3 Train and test accuracy of handwriting features

Handwriting features	Train accuracy (%)	Test accuracy (%)
Too angular, devoid of roundedness	99.55	98.44
Small size (middle zone)	98.11	97.73
Retraced letter d and t. Sharp and angular retracement	99.19	98.68
Tightly closed A and O	98.65	99.12
Lack of Symmetry in zones	96.77	99.64
Heavy slashing strokes	99.10	99.76
Sudden change in speed, size, and pressure	90.33	90.79
Unfinished down stroke	90.33	89.83
Arcade	98.93	98.92
Miniature size letters	93.05	93.30
Extreme right slant	98.86	98.80
Signature extreme right	95.72	95.93
'M' lower loops- larger the loops more the worry	90.90	90.43
Narrow left, wide right margins	98.05	98.21

7 Applications

The automated tool developed for handwriting analysis in this paper can have useful application in several organizational settings. It can also be used at an individual level to understand oneself.

Studies related to personality and financial decision making find that saving patterns are associated with extraversion and neuroticism; an extrovert individual may save less in the short term while a neurotic individual may have meager lifetime savings [29]. Financial institutions providing credit services would be interested in understanding the individual's financial habits and attributes [30]

Researchers have noted that higher extraversion and lower neuroticism are related to increased wellbeing [31, 32], suggesting that individuals may want to adopt behaviors associated with extraversion to improve their overall wellbeing. Neurotic individuals exhibit moodiness and lower emotional stability. The model developed in this study can make it easy and efficient for organizations to determine the personality of individuals. It can improve an organization's ability to select suitable candidates for various functions and positions [33].

8 Limitation

As the data set available for the study was limited in size, augmentation had to be done. Certain features were dropped due to lack of class variance because of this constraint. These included the handwriting features of Erratic slant, Depression valley, Capital Letter—Tall and Narrow, Wide left margin, i and j dots dashed and right to the stem, Wavy baselines, Encircled signature.

9 Conclusion and Future Scope

The robust results achieved during the experiment have shown the potential of the technique to identify neuroticism using machine learning. Training the tool on larger data sets will improve its applicability. It can help to popularize the use of handwriting in identifying personality traits. Several applications at both individual and organization level can be supported to make more informed choices and manage decision making risks.

The model architecture built in this paper can be used as a basis for developing such tools for other four of the Big Five Personality Factors.

References

1. van der Meer ALH, van der Weel FR (2017) (Ruud.: only three fingers write, but the whole brain works: a high-density EEG study showing advantages of drawing over typing for learning. Front Psychol 8:1–9. https://doi.org/10.3389/fpsyg.2017.00706
2. Giannini M, Pellegrini P, Gori A, Loscalzo Y (2019) Is graphology useful in assessing major depression? Psychol Rep 122:398–410. https://doi.org/10.1177/0033294118759667

3. DeYoung CG, Carey BE, Krueger RF, Ross SR (2016) Ten aspects of the big five in the personality inventory for DSM-5. Personal. Disord Theory Res Treat 7:113–123. https://doi.org/10.1037/per0000170
4. Digman JM (1990) Personality structure: emergence of the five-factor model. Annu Rev Psychol 41:417–440. https://doi.org/10.1146/annurev.ps.41.020190.002221
5. Gavrilescu M, Vizireanu N (2018) Predicting the big five personality traits from handwriting. Eurasip J Image Video Process. https://doi.org/10.1186/s13640-018-0297-3
6. Esposito A, Amorese T, Buonanno M, Cuciniello M, Esposito AM, Faundez-Zanuy M, Likforman-Sulem L, Riviello MT, Troncone A, Cordasco G (2019) Handwriting and drawing features for detecting personality traits. In: CogInfoCom Proceedings 10th IEEE international conference cognition Infocommunications, pp 79–84. https://doi.org/10.1109/CogInfoCom47531.2019.9089985
7. Chaudhari K, Thakkar A (2019) Survey on handwriting-based personality trait identification. Expert Syst Appl 124:282–308. https://doi.org/10.1016/j.eswa.2019.01.028
8. Klimoski RJ, Rafaeli A (1983) Inferring personal qualities through handwriting analysis. J Occup Psychol. https://doi.org/10.1111/j.2044-8325.1983.tb00127.x
9. Eysenck HJ (1948) "Neuroticism" and handwriting. J Abnorm Soc Psychol 43:94–96. https://doi.org/10.1037/h0053484
10. Soto CJ (2018) Big five personality traits. SAGE Encycl Lifesp Hum Dev 240–241. https://doi.org/10.4135/9781506307633.n93
11. Thomas S, Goel M, Agrawal D (2020) A framework for analyzing financial behavior using machine learning classification of personality through handwriting analysis. J Behav Exp Financ 26:100315 (2020). https://doi.org/10.1016/j.jbef.2020.100315
12. Brougham RR, Jacobs-Lawson JM, Hershey DA, Trujillo KM (2011) Who pays your debt? An important question for understanding compulsive buying among American college students. Int J Consum Stud 35:79–85. https://doi.org/10.1111/j.1470-6431.2010.00923.x
13. Xu Y, Beller AH, Roberts BW, Brown JR (2015) Personality and young adult financial distress. J Econ Psychol 51:90–100. https://doi.org/10.1016/j.joep.2015.08.010
14. Rukh G, Dang J, Olivo G, Ciuculete DM, Rask-Andersen M, Schiöth HB (2020) Personality, lifestyle and job satisfaction: causal association between neuroticism and job satisfaction using Mendelian randomisation in the UK biobank cohort. Transl Psych 10. https://doi.org/10.1038/s41398-020-0691-3
15. Cuncic BA How neuroticism affects your behavior prevalence of neuroticism. https://www.verywellmind.com/how-neuroticism-affects-your-behavior-4782188
16. Anasuya Devi HK (2006) Thresholding: a pixel-level image processing methodology preprocessing technique for an OCR system for the Brahmi script
17. Pathak AR, Raut A, Pawar S, Nangare M, Abbott HS, Chandak P (2020) Personality analysis through handwriting recognition. J Discret Math Sci Cryptogr 23:19–33. https://doi.org/10.1080/09720529.2020.1721856
18. Wang R, Huang S, Zhou Y, Cai ZG (2020) Chinese character handwriting: a large-scale behavioral study and a database. Behav Res Methods 52:82–96. https://doi.org/10.3758/s13428-019-01206-4
19. Kumar P, Gupta M, Gupta M, Sharma A (2020) Profession identification using handwritten text images 2:25–35. https://doi.org/10.1007/978-981-15-4018-9
20. Harralson HH, Miller LS (2018) Huber and Headrick's handwriting identification: facts and fundamentals. CRC Press
21. Sheikholeslami G, Srihari SN, Govindaraju V (1999) Computer aided graphology 1–7
22. Panagopoulos M, Rousopoulos P, Arabadjis D, Papaodysseus C (2009) Information system for graphological Identification. In: PCI 2009—13th Panhellenic conference on informatics
23. Punetha N, Pal AK, Singh Kushwaha G (2019) Characteristics and mood prediciton of human by signature and facial expression analysis. Int Res J Eng Technol 6:89–95
24. Raniman MA, Varghese D, Kumar GM (2013) HABIT: handwritten analysis based individualistic traits prediction. Int J Image Process 7:209–218

25. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6. https://doi.org/10.1186/s40537-019-0197-0
26. Gu S, Pednekar M, Slater R (2019) Improve image classification using data augmentation and neural networks. SMU Data Sci Rev 2:1–43
27. Mikołajczyk A, Grochowski M (2018) Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary Ph.D. work. IIPhDW, pp 117–122. https://doi.org/10.1109/IIPHDW.2018.8388338
28. Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning
29. Nyhus EK, Webley P (2001) The role of personality in household saving and borrowing behaviour. Eur J Pers 15:S85–S103. https://doi.org/10.1002/per.422
30. Alexandra L, Németh Erzsébet ZB (2017) Financial personality types and attitudes that affect financial indebtedness. Int J Soc Sci Econ Res 02:4687–4704
31. Steel P, Schmidt J, Shultz J (2008) Refining the relationship between personality and subjective well-being. Psychol Bull 134:138–161. https://doi.org/10.1037/0033-2909.134.1.138
32. Røysamb E, Nes RB, Czajkowski NO, Vassend O (2018) Genetics, personality and wellbeing. A twin study of traits, facets and life satisfaction. Sci Rep 8:1–13. https://doi.org/10.1038/s41598-018-29881-x
33. Anger S, Camehl G, Peter F (2017) Involuntary job loss and changes in personality traits. J Econ Psychol. https://doi.org/10.1016/j.joep.2017.01.007

Classification of Waste Objects Using Deep Convolutional Neural Networks

G. Rishma and R. Aarthi

Abstract Pollution induced by waste is one of the major problems across the globe for a long time. During the difficult natural calamities, it worsens the situation due to the improper way of disposing of the materials subsequently causes danger to human life. This also creates a risk for sanitation workers during the process of waste collection, sorting, and recycling. So there is a need for adapting the technology in real-time for proper waste segregation and management. In this paper, a model is proposed for the classification of waste objects according to their materials which can help in the automatic recycling process. The proposed approach uses the State-of-art deep learning architecture to construct the classifier. The architecture focusing on minimizing the structural complexity and maintain better classifier accuracy. The classifier uses the features extracted automatically from the Convolutional Neural Networks (CNN) to build a model. This model classifies the waste into plastic, paper, cardboard, trash, glass, and metal with an accuracy of 87%.

Keyword Waste classification · Convolutional neural networks · Deep learning

1 Introduction

The environment plays a vital role in the healthy living of people in the world. It's the home for many living things and provides all the needs like air, food, water, etc. Human's entire life depends on the well-being of all the environmental factors. It is the responsibility of humans to protect the environment and to ensure the maintenance of natural resources by reducing pollution in all possible ways. An unhygienic environment spreads disease easily by destroying natural gifts quickly. The effective way of recycling and reusing the waste would help to produce renewable energy

G. Rishma · R. Aarthi (✉)
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: r_aarthi@cb.amrita.edu

G. Rishma
e-mail: cb.en.p2cse19020@cb.students.amrita.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_47

sources. Though waste management is a self-responsibility of humans, the accumulation of waste is unavoidable. An automatic method of waste segregation will avoid the workforce required for management. The common materials like bottles, plastic/paper materials are widely used in the COVID pandemic situation to ensure the safety of individuals. A huge amount of waste products like bottles, plastic/paper cups, food wrappers, aluminum cans, etc. are used by humans in the field of medicine, hotels, and all common places. This increases the accumulation of waste day by day.

Every year a total of 2.01 billion of solid waste [1] is collected all over the world and are dumped in landfills which emit hazardous substance like carbon-di-oxide (CO_2) and methane. This hazardous substance is challenging in both developed and developing countries. This poorly managed waste is contaminating the world resources and transmitting diseases. If solid waste is not treated properly there will be a huge increase in the CO_2 emission. India is the largest garbage generator and generates more than the tenth parts of the world's waste [2]. India generated about 62 million tons of waste annually and some surveys predict that it can increase over 277 million tons by 2030 and 3.87 billion tons by 2050 [2]. An average Indian generates about 0.57 kg of waste per person per day [2]. The solid waste, especially plastic, paper, and cardboard waste generated is extremely high [2].

The solution for this waste management is to use technology in the right way to minimize waste. The proper segregation and remanufacturing and recycling waste into usable products would be a better solution. It is required to classify the materials into recyclable and non-recyclable material for further processing. Computer vision plays a major role in waste management. In the method of image processing and improvement in technology, the operating speed and efficiency of processing of images have improved. Deep learning approaches are being used in many applications like in computer vision, natural language processing, speech recognition, etc. Deep Learning models with CNN as the main model plays a major role in image classification. CNN model can learn the strong features on its own and is better than the traditional image processing where it is very complex to select the important features. Also, it is a better approach than the traditional approach because when human beings are involved in segregation they may be prone to diseases.

In this paper, the Convolutional neural networks are proposed for waste management. The major advantage of the CNN architecture is that it can learn and extract the features automatically. One major challenge that is addressed in our approach is to minimize structural elements of the network without compromising the accuracy. The experimental results show that the proposed CNN architecture outperforms the pre-trained networks such as DenseNet, ResNet, MobileNet, etc. The classifier can classify the waste material into different types such as plastic, paper, cardboard, trash, glass, and metal by maintaining better accuracy. The paper is organized as, in Sect. 2, brief on the works that are related with waste classification is summarized; in Sect. 3, explains about the proposed architecture of the work; in Sect. 4, the experiments on our proposed system to show how it outperforms other proposed CNN models. The final section discusses the conclusion and future work.

2 Related Works

Image classification is active research carried out over many years. Classification of objects in a natural environment with a complex background is a challenging problem. In the field of Waste Management, the shape of the object has been deteriorated by many factors and increases the complexity of classification. Few methods developed to Classify and detect waste objects help in automatic segregation and effective recycling. Many successful methods use a combination of computer vision with machine learning to classify the objects.

Lokesh et al. [3] used the combination of visual and physical features of waste bottles to segregate into various categories depending upon the material type they are made of to decide whether a bottle can be recycled or not. The classification has been done with Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Decision tree, and logistic regression where simple KNN shows better performance with the combination set of features. An approach followed by Witaya et al. [4] to classify the bottle images as PET (Polyethylene terephthalate) or Non-PET by using the correlation method applied to the histogram of RGB color of the image. The histogram of RGB colors coincide for PET bottles and shows an average correlation value near to 1 whereas for Non-PET bottles it's less than 0.5. Zhaokun et al. [5] developed a system that can automatically classify the plastic bottles according to different colors positioned on the conveyor belt for recycling. Firstly the bottles which are disjointed, adjacent, and overlapped are separated and then SVM is applied for color recognition. One major drawback of these algorithms is the selection of features for the object in the class which could represent them uniquely. Whereas deep learning models can extract the most important features by itself.

Deep neural networks are being used widely in image processing and computer vision. The convolution neural networks (CNN) models are preferred for waste classification because of their capabilities to learn strong features and the ability to classify the predetermined classes. In Lolith et al. [6] applied deep convolutional neural networks (DCNN) model with 22 convolution layers and 5 max-pooling layers for detecting the bottles in the indoor environment. A model was designed with 3 convolution layers and 2 fully connected layers by Sidharth et al. [7] for the trashnet dataset and achieved an accuracy of 76%. Also, Sachin et al. [8] developed a smart bin using this model by training it on trashnet dataset and obtained an accuracy between 80 and 86%. The DCNN models are also used in different IOT applications to improve the performance of the system. Nandhini et al. [9] and Sreelakshmi et al. [10] integrated the CNN model with robotic arm and conveyor belt applications respectively.

Deep Learning applications are also modeled using the pre-trained architecture so that the computation cost can be reduced rather than building a similar model from scratch. Models like DenseNet, ResNet, etc. are also used in waste management applications. Chutimet et al. [11] explored the VGG-16, ResNet-50, MobileNet-V2, and DenseNet-121 pre-trained networks for waste type and ResNet-50 showed good accuracy. Similarly in [12–19] used different pre-trained model such as Resnet [12], RecycleNet [13], DenseNet [14], VGG network [15], MobileNet [16, 17], AlexNet

[19], etc. for classification on the trash net dataset. All these pre-trained networks are very huge and show a maximum accuracy achieved ranging from 70 to 87%. Some pre-trained networks like AlexNet, Berkeley model, and GoogleNet are trained for classification but are not optimum for waste identification [20].

CNN based deep learning models are also integrated for waste classification applications like a reverse vending machine [21] and Automated Teller Dustbin [22] which is able to accept the waste product and classify them belonging to a particular item and credit the amount possible for the waste object deposited by the individual. Also, SmartBin [10] which is able to identify the waste and classify it and a website module integrated to display the results. As the size of the network increases the Computation cost and training time also increases. The work proposed in our paper focuses on designing a simplified deep neural network that is able to classify the waste objects with better accuracy. The model is compared with the pre-trained networks on various parameters.

3 Proposed Architecture

The objective of this work is to classify the objects based on the type of material like cardboard, glass, plastic, paper, trash, and metal. The overview of the CNN model is presented in Fig. 1, the sequence of training images is given as an input to the model. The CNN model is designed with a sequence of convolution layers, fully connected layers, and the output labels. Once the model is trained the learned parameters are used to predict and classify the material in the image.

Fig. 1 Overview of CNN model

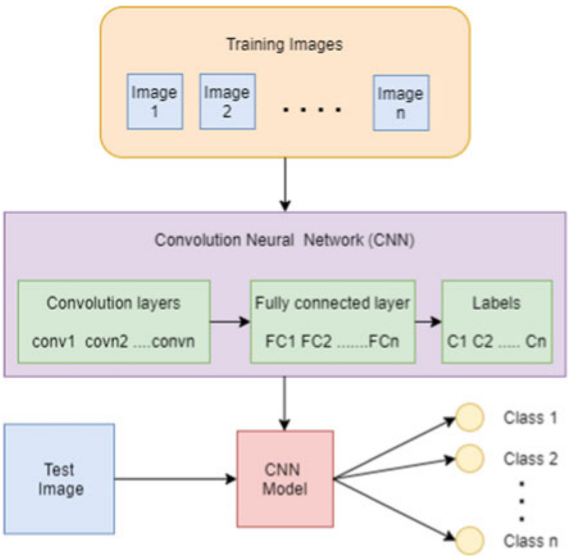


Table 1 Sample images and dataset description with total images of 2527

Image						
Class	Paper	Glass	Cardboard	Plastic	Trash	Metal
Count	594	501	403	482	137	410

3.1 Data Collection and Preprocessing

The Trashnet dataset proposed by Thung and Yang [23] is used in this work. The images in the dataset are divided into 6 classes: paper, glass, plastic, metal, cardboard, and trash. The images are of size 512 × 384 pixels. The Distribution of the dataset is shown in Table 1.

These images were captured by placing the object on the white background using either sun or room light. Every image has only one object placed in the frame. Sample images from the dataset are shown in Table 1. 80% of the images from the dataset are used for training and 20% for testing. The images in the dataset are resized to 100 × 100 and normalized to [0–1]. In normalization, each pixel of the image is scaled to a value between 0 and 1 by dividing the pixel value by 255. Data Augmentation has been applied to obtain better results by preventing overfitting. Some images from the dataset are rotated randomly by 10 degrees, zoomed randomly by a factor of 0.1, vertical or horizontal translation of random images by a factor of 0.2, and randomly flipping images horizontally.

3.2 Proposed CNN Model

The proposed CNN architecture is shown in Fig. 2. The network consists of seven layers. The first five layers are convolution layers, the next two layers are fully connected layers and the final is the output layer. The last layer is a 6-way softmax function which produces the distribution for all the 6 class labels from the dataset.

The image will be passed through a sequence of convolution layers and a max-pooling layer. The input image of size 100 × 100 × 3 is passed to the first convolution layer with 16 filters of size 3 × 3. The second and third convolution layers have 32 and 64 filters respectively each of size 3 × 3. The fourth and the fifth convolution layer have 128 and 256 filters respectively each of size 3 × 3. Each convolution layer is followed by a max-pooling layer which will perform the dimensionality reduction. All max-pooling layers in the network have a pool size of 2 × 2. The convolution layers have a stride of 1 pixel and for max-pooling layers, it is 2 pixels. After convolution layers, each neuron is connected to every other neuron. After processing through the convolution layers the flattening happens where the output

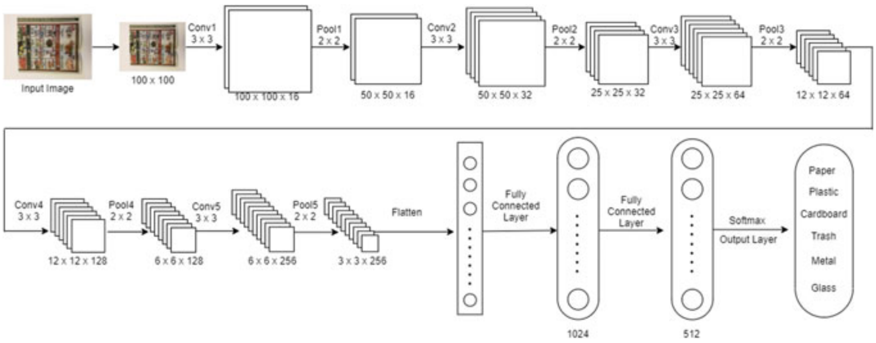


Fig. 2 Proposed CNN model

layer is converted to a single channel of neurons. This is followed by two fully connected layers each having 1024 and 512 neurons respectively. The last layer is the classification layer, a 6-way softmax function is used for classification. Relu activation function is used for the convolution layer and fully connected layer. To avoid the overfitting in the fully connected layers a dropout with a rate 0.5 is used.

In the Training phase, the model's learning rate is set as 0.001 with an Adam optimizer. The loss function used here is Categorical cross-entropy. The network was trained for a batch size of 50 and a set of epochs like 100, 150 and 200. The training time for the proposed network was about an hour for 100 epochs and the time increases with the increase of epochs. The details of experimentation and comparison are discussed in the next section.

4 Experimentation

The system is developed using python with Keras library and TensorFlow backend on the Google Colaboratory platform. An experiment was performed to find which CNN architecture will be able to provide better accuracy. The results shown in Table 2 were obtained by having a different number of convolution layers and two fully-connected layers with 100 epochs. From the results, it can be seen that the model with 5 convolution layers shows better training accuracy and testing accuracy than other architectures.

The architecture with 5 convolution layers, 2 fully-connected layers, and a softmax based output layer was concluded as the best architecture and further tested with different epochs like 100, 150, and 200. The results are shown in Table 3 the model with 150 epochs shows better results in and the training accuracy and testing accuracy. The corresponding accuracy and loss plot is shown in Table 4. By using hyperparameter the accuracy obtained is 86.56% and a loss of 0.8 units. The sample output from the proposed CNN model is shown in Table 5, it shows the sample correctly

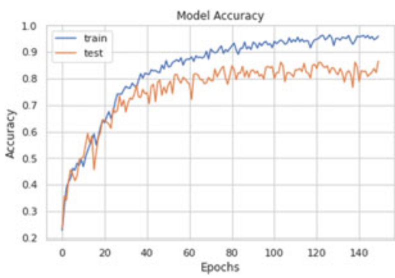
Table 2 Accuracy-varying convolution layer

Convolution layer	Training accuracy (%)	Testing accuracy (%)
1	76.94	70.15
2	89.11	78.45
3	95.15	80.63
4	92.42	80.23
5	97.27	83.20
6	90.49	79.25
7	93.46	78.06

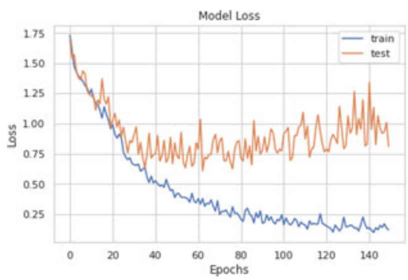
Table 3 Accuracy-varying epochs

Epoch	Training accuracy (%)	Testing accuracy (%)
100	97.27	83.20
150	98.07	86.56
200	97.08	83.39

Table 4 Accuracy and loss graph on proposed CNN



(a) Training and testing accuracy



(b) Training and testing loss

Table 5 Classified and misclassified images

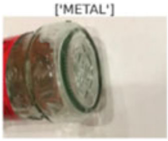
Class	Glass	Plastic	Metal
Correctly classified images	['GLASS'] 	['PLASTIC'] 	['METAL'] 
Misclassified images	['METAL'] 	['METAL'] 	['TRASH'] 

Table 6 Comparison results

Author	Method	Layer	Epoch	Accuracy
Sidhart et al. [7]	CNN model	5	100	76
Olugboja et al. [12]	ResNet-50	50	12	87
Cenk et al. [13]	RecycleNet	121	200	81
Rahmi et al. [15]	MobileNet	28	210	84
Sarah et al. [17]	CompostNet	159	18	77.3
Stephenn et al. [18]	MobileNet	28	500	87
Proposed	**CNN**	**7**	**150**	**87**

classified and misclassified images. For the prediction of each image, the classifier has taken 0.0132 s.

The comparison results are shown in Table 6 where the results of the other methods are considered from the papers available who have utilized the same trashnet dataset. This comparison result shows how our proposed method is better in the accuracy and number of layers used. In Table 6, the comparison is done with the other architectures like Resnet, RecycleNet, MobileNet, etc. are huge networks with 50, 121, 28 layers respectively, and having accuracy between 81 and 87%. But when comparing our approach with 7 layers (5 convolution and 2 fully connected layer) outperforms with an accuracy of 84% for 100 epochs and 87% for 150 epochs than any other known methods. Also our proposed network has the same number of layer as AlexNet, but AlexNet shows an accuracy of 74% with the same trashnet dataset.

The method is tested with images by combining 4 types of materials either of same or different into the same frame. The results obtained from out CNN model are shown in Table 7. The results shows that the CNN model failed to classify all the materials present in the image frame and it is sometimes able to detect either one of the material type. So the proposed CNN model will not work perfectly when different types of objects are present in the image or in a particular frame.

Table 7 Output results

Input image			

5 Conclusion and Future Work

In this paper, a simple and efficient CNN model is presented. The proposed architecture can be integrated into any hardware where it can classify waste objects according to their material type. It can conclude that this type of CNN model will work perfectly when there is a single object type present in the frame. This CNN model can be used in waste management applications like waste segregation over the conveyor belt, smart garbage Bin/Automated Teller Dustbin, reverse vending machine, etc. where only objects will be on the frame at a time. This CNN based model cannot be used when the waste is littered on the streets and roads. That is when the objects of different types are occluded or scattered in the same frame. This work can further be improved by choosing a better approach than the CNN model to resolve this kind of application where waste is littered in a complex environment.

References

1. The World Bank (IBRD & IDA) Trends in solid waste management. https://datatopics.worldbank.org/what-a-waste/trends_in_solid_waste_management.html
2. Times of India. In 30 yrs., India tipped to double the amount of waste it generates. https://timesofindia.indiatimes.com/india/in-30-years-india-tipped-to-double-the-amount-of-waste-it-generates/articleshow/74454382.cms
3. Kambam LR, Aarthi R (2019) Classification of plastic bottles based on visual and physical features for waste management. In: 2019 IEEE international conference on electrical, computer and communication technologies (ICECCT), Coimbatore, India, pp 1–6. https://doi.org/10.1109/ICECCT.2019.8869191
4. Srigul W, Inrawong P, Kupimai M (2016) Plastic classification base on correlation of RGB color. In: 2016 13th international conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON), Chiang Mai, pp 1–5. https://doi.org/10.1109/ECTICon.2016.7561304
5. Wang Z, Peng B, Huang Y, Sun G (2019) Classification for plastic bottles recycling based on image recognition. Elsevier Waste Manag 88:170–181. ISSN 0956-053X. https://doi.org/10.1016/j.wasman.2019.03.032
6. Gopan L, Aarthi R (2018) A vision based DCNN for identify bottle object in indoor environment. In: Hemanth D, Smys S (eds) Computational vision and bio inspired computing. Lecture notes in computational vision and biomechanics, vol 28. Springer, Cham. https://doi.org/10.1007/978-3-319-71767-8_37
7. Sidharth R, Rohit P, Vishagan S, Karthika R, Ganesan M (2020) Deep learning based smart garbage classifier for effective waste management. In: 2020 5th international conference on communication and electronics systems (ICCES), Coimbatore, India, pp 1086–1089. https://doi.org/10.1109/ICCES48766.2020.9137938
8. Hulyalkar S, Deshpande R, Makode K, Kajale S (2018) Implementation of smartbin using convolutional neural networks. Int Res J Eng Technol (IRJET) 05(04)
9. Nandhini S, Mrinal SS, Balachandran N, Suryanarayana K, Ram DSH (2019) Electronically assisted automatic waste segregation. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI), Tirunelveli, India, pp 846–850. https://doi.org/10.1109/ICOEI.2019.8862666

10. Sreelakshmi K, Akarsh S, Vinayakumar R, Soman KP (2019) Capsule neural networks and visualization for segregation of plastic and non-plastic wastes. In: 2019 5th international conference on advanced computing and communication systems (ICACCS), Coimbatore, India, pp 631–636. https://doi.org/10.1109/ICACCS.2019.8728405
11. Srinilta C, Kanharattanachai S (2019) Municipal solid waste segregation with CNN. In: 2019 5th international conference on engineering, applied sciences and technology (ICEAST), Luang Prabang, Laos, pp 1–4. https://doi.org/10.1109/ICEAST.2019.8802522
12. Adedeji O, Wang Z (2019) Intelligent waste classification system using deep learning convolutional neural network. Procedia Manuf 35:607–612
13. Bircanoğlu C, Atay M, Beşer F, Genç Ö, Kızrak MA (2018) RecycleNet: intelligent waste sorting using deep neural networks. In: 2018 innovations in intelligent systems and applications (INISTA), Thessaloniki, pp 1–7. https://doi.org/10.1109/INISTA.2018.8466276
14. Aral RA, Keskin SR, Kaya M, Hacıömeroğlu M (2018) Classification of TrashNet dataset based on deep learning models. In: 2018 IEEE international conference on big data (big data), Seattle, WA, USA, pp 2058–2062. https://doi.org/10.1109/BigData.2018.8622212
15. Hao W (2020) Garbage recognition and classification system based on convolutional neural network VGG16. In: 2020 3rd international conference on advanced electronic materials, computers and software engineering (AEMCSE), Shenzhen, China, pp 252–255. https://doi.org/10.1109/AEMCSE50948.2020.00061
16. Ahmad K, Khan K, Al-Fuqaha A (2020) Intelligent fusion of deep features for improved waste classification. In: IEEE access, vol 8, pp 96495–96504. https://doi.org/10.1109/ACCESS.2020.2995681
17. Frost S, Tor B, Agrawal R, Forbes AG (2019) CompostNet: an image classifier for meal waste. In: 2019 IEEE global humanitarian technology conference (GHTC), Seattle, WA, USA, pp 1–4. https://doi.org/10.1109/GHTC46095.2019.9033130
18. Rabano SL, Cabatuan MK, Sybingco E, Dadios EP, Calilung EJ (2018) Common garbage classification using MobileNet. In: 2018 IEEE 10th international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM), Baguio City, Philippines, pp 1–4. https://doi.org/10.1109/HNICEM.2018.8666300
19. Sakr GE, Mokbel M, Darwich A, Khneisser MN, Hadi A (2016) Comparing deep learning and support vector machines for autonomous waste sorting. In: 2016 IEEE international multidisciplinary conference on engineering technology (IMCET), Beirut, pp 207–212. https://doi.org/10.1109/IMCET.2016.7777453
20. Noorani S, Fernandes M (2017) Evaluation of convolutional neural networks for waste identification. In: 2017 international conference on computing methodologies and communication (ICCMC), Erode, pp 204–207. https://doi.org/10.1109/ICCMC.2017.8282675
21. Kokoulin AN, Tur AI, Yuzhakov AA (2018) Convolutional neural networks application in plastic waste recognition and sorting. In: 2018 IEEE conference of Russian young researchers in electrical and electronic engineering (EIConRus), Moscow, pp 1094–1098. https://doi.org/10.1109/EIConRus.2018.8317281
22. Sunny MSH, Dipta DR, Hossain S, Faruque HMR, Hossain E (2019) Design of a convolutional neural network based smart waste disposal system. In: 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT), Dhaka, Bangladesh, pp 1–5. https://doi.org/10.1109/ICASERT.2019.8934633
23. Thung G, Yang M (2016) Classification of trash for recyclability status, CS 229, Stanford University. Available: https://github.com/garythung/trashnet

Doodle Recognition Using Ensemble Learning

Harshit Gupta, Pratik Devnani, Kanishk Bhatia, and Shilpa Verma

Abstract The system uses machine learning techniques to identify the sketches drawn by people and recognize them correctly. The recognition system tries to consider nuances of size, shape, and orientation which might occur due to the individuality of human beings. Doodle and its recognition have recently drawn a lot of attention from the scientific community from using a sketch for search engines, psychological analysis, and many more, and it has initiated work in this domain. Here, the system has taken into consideration 30 different classes of objects that can be recognized. The method used here is ensemble learning, using an ensemble of CNN as well as LSTM algorithms to identify and categorize the doodle correctly.

Keywords CNN · LSTM · Doodle · Ensemble · Google · Bagging · Voting

1 Introduction

Image-based recognition has been present for a long time and it has various advantages and disadvantages. Although there are standard methods and techniques which give results, there has been a shift in the dynamic of this domain. Emphasis has largely been on a human's individuality, interpretation, and rendition of objects. Doodle recognition is the result of such renewed interest which aims to explore new possibilities in the recognition domain, to perfect them, and to standardize them into a format that is usable in real-life situations. "A doodle is a random drawing, or a collection of patterns made by a person while he is thinking about something else" is a very general definition that has been thrown around from the past two decades or so. Recently, doodles have been regarded as drawings that have a far deeper meaning and are considered to be extremely important than previously perceived.

Psychologists may use doodles for assessment purposes, but they are scientifically and professionally developed and tested. A set of assorted doodles would present a better understanding of an individual, particularly when it is coupled with alternative information like handwriting analysis. This technology can also provide suggestions

H. Gupta (✉) · P. Devnani · K. Bhatia · S. Verma
Department of Computer Engineering, Thadomal Shahani Engineering College, Mumbai, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_48

based on what the user fully or partially draws and enable roughly drawn sketches to be instantly converted into complete clip art.

The purpose of this work is to create a system that classifies doodles with the highest accuracy. We incorporated ensembles for this task, using Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) algorithms on the doodle dataset. The model predicts the object using separate instances of the algorithm and then takes an average based on a combination method to produce the output. The goal is to accurately categorize various objects, supported by the highest percent of the resemblance of the drawn sketch with the doodle of that object within the dataset.

Based on observation and experimentation, we chose to build an ensemble model using CNN. This method was able to achieve 98.32% training accuracy and 97.98% validation accuracy in the ensemble model using different instances of the CNN algorithm. We used the doodle dataset, which is made available by Google for its artificial intelligence research.

2 Related Work

2.1 Quick, Draw! Doodle Recognition [1]

Kristine and James experimented with variations of k-nearest neighbors and Convolutional Neural Network. However, this resulted in a computationally expensive model. The assumption that each category should be relatively similar, and limitation of this model were removed in KNN with K-means++. Moreover, they found that CNN outperformed KNN and its variations.

2.2 Drawing: A New Way to Search (Computer Vision) [2]

Keeping their emphasis on accuracy and efficiency, Jervis, Connie, and Nguyet applied machine learning and deep learning techniques to Google's quick draw dataset limited to 50 classes. They applied algorithms like logistic regression and support vector machine, but the results were not passable. They applied CNN and attained an accuracy of 81.83%. They also applied transfer learning with four different architectures out of which, VGG outperformed the rest of the architectures and had comparable performance to CNN.

2.3 Image Classifications Using LSTM and CNN Hybrid Neural Network [3]

Here, the authors have discussed the application of a hybrid system of CNN and LSTM for image recognition on the MNIST dataset. They proposed an architecture in which the output shape from the batch normalization is applied to LSTM architecture. The output generated is applied directly to the convoluted layer. They observed that the hybrid CNN-LTSM architecture outperformed CNN and LSTM with a significant margin where the hybrid model was able to attain 99.8% training accuracy and 98.2% validation accuracy. Therefore, we considered CNN and LSTM architecture for doodle recognition due to its performance.

2.4 Free-Hand Sketch Recognition Classification [4]

In this paper, which is presented by Wayne Lu and Elizabeth Tran, the authors explore the idea of using a CNN with a ResNet approach. The suggested technique uses modular residual units with varying architectures which include a stack of convolutional layers and a shortcut connection that carries the original input. Their experiments showed evidence of the intraclass overlap and interclass variation caused by variations in artistic interpretation. Their experiments also show that deeper networks provide higher classification accuracy with moderate dropout.

2.5 Ensemble of CNNs for Steganalysis [5]

In this paper, the authors study different ensemble methods that use CNNs as base learners for Steganalysis. Here, they designed various CNN models, each of which is trained on a random subset of the training dataset. The designed models used six groups of convolutional modules that were operating in unison and the output from this module was passed to the classification. Subsampling of the dataset caused some negative effects on the forensics of the image; however, the subsampling helped to remove irrelevant information. The collective working of the CNN overcomes difficulties of a single CNN and this made us gravitate towards ensemble learning.

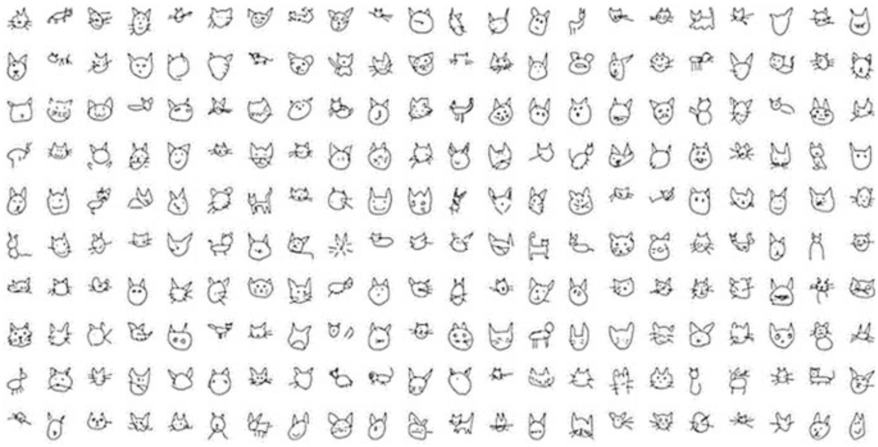


Fig. 1 Doodles of a cat in the database [6]

3 Proposed Model

3.1 Dataset

Google's QuickDraw is the world's largest doodling dataset, consisting of 50 million hand-drawn images across 345 categories. Some examples are given in Fig. 1. Each drawing consists of 28 by 28 raw pixel inputs which are NumPy images with values from 0 to 255. Each image has only two colours, black and white. To make training more tractable on modest computing resources, we elected to work with a subset of the data. The classes selected were chosen randomly from the overall pool of classes and were fixed throughout our experimentation. The number of examples per class is 10,000 and we selected 30 classes to classify [6].

3.2 Convolutional Neural Network

The Convolutional Neural Network, otherwise called the ConvNet, is a deep learning algorithm that takes an image as its input, calculates, and assigns learnable weights and biases to certain aspects of the image, and then identifies the image. Convolutional Neural Networks have two important properties [7]:

- The patterns learned are translation invariant. In simpler terms, if a certain pattern is detected and learned at the bottom right corner of the input image, a ConvNet can recognize it anywhere in the image. This results in ConvNet's being data-efficient when processing images.

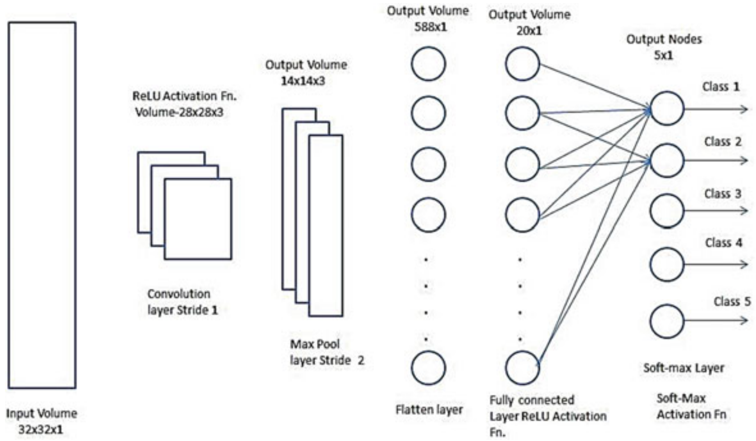


Fig. 2 CNN architecture [8]

- ConvNet can learn spatial hierarchies of patterns. The first layer focusses on learning comparatively smaller local patterns like edges while the subsequent layers will learn larger patterns, and this goes on increasing. This is one of the main reasons for a ConvNet to perform efficiently even when the input space is complex and visually abstract (Fig. 2).

The parameters that have been altered in different CNN instances:

- Number of layers
- Activation Function
- Number of epochs.

3.3 Long Short-Term Memory

LSTM is a permutation of the simple RNN, whose distinguishing feature is its capability to carry information across various timestamps. This is essentially how LSTM works; it keeps information that can be used at a later time which results in preventing older signals from gradually vanishing during processing [9].

There are a few advantages of using an LSTM model and they are as follows [10]:

- LSTM is capable of dealing with issues related to noise in the system, largely varied representations, and continuous-valued inputs efficiently.
- Parameter fine-tuning is not needed in an LSTM model. It performs efficiently over a broad spectrum of parameters such as learning rate, input, and output gate bias.

Different LSTM models have been used and the activation functions have been changed for each model.

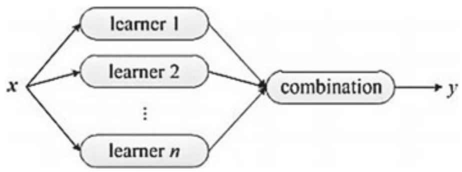


Fig. 3 Architecture of an ensemble model [11]

3.4 Ensemble Learning

"Ensemble methods use multiple learners to resolve a similar drawback collectively". This is contrasting to the classic learning approaches, where a single learner is constructed from the training data, whereas ensemble methods develop a batch of learners and use their combination to achieve results. An ensemble model has multiple learners known as base learners. Generally, these base learners are generated from the training data using a base learning algorithm. Ensemble learning methods are appealing since they are capable of turning weak classification algorithms into good classifiers with high classification accuracy as it combines different classifiers [11] (Fig. 3).

After generating a batch of base learners, ensemble methods integrate all the base learners using various methods and this combination is used for prediction. The method used for combination plays an important role in the accuracy of output.

Bagging: Bagging (also known as bootstrap aggregating) is a process in which homogenous base learners are used for creating an ensemble. These learners are trained in parallel and independently and their output is combined using some combination method. Now, in bagging, we fit individual models and average their predictions to get a good result with low variance. As our problem is a classification problem, we used the "voting" method for getting the final result from their ensemble model.

Voting: Amongst the present voting methods, majority voting is the most popular one. Every classifier has to vote for a single class label and the label that assimilates 50% of the total votes is the chosen output label. In the case of no voted majority, a rejection option is given and the combined classifier proceeds to make no prediction.

4 Experiment and Results

We experimented with different models varying based on the number of layers and activation functions. Here, we selected some classifiers and ran the model on the same dataset to attain the training and validation accuracy.

Figure 4 represents an example of a doodle that can be drawn by a user. Since, different user interprets things differently, they would draw the same object differently. However, the underlying representation or the structure of the object would

Fig. 4 Sample input image for recognition

Table 1 Accuracies comparison over doodle dataset

S. No.	Reported accuracies		
	Model	Training accuracy (%)	Validation accuracy (%)
1	CNN1 (4 layers + Relu)	98.09	97.85
2	CNN2 (3 layers + Relu)	98.13	98.06
3	CNN3 (3 layers + Selu)	98.07	97.37
4	LSTM1 (Relu)	91.72	90.28
5	LSTM2 (Softmax)	90.79	90.08
3	LSTM3 (Softplus)	91.66	89.95

more less be similar, which is converted into a single-dimensional matrix or a vector of size 256.

Table 1 displays the accuracy achieved by each of the different classifiers and their permutations which were tested. The analysis of the results exhibit that some permutations perform better compared to the others as there is a significant difference in percentage accuracies between the various models.

Table 2 shows the minimum and maximum accuracies achieved by singular runs of the CNN and LSTM models. It also shows the mean accuracy achieved by the same. The values clearly show that the CNN models outperform the LSTM models by a considerable margin, hence, we created our ensemble model using only CNN base learners.

Table 2 Accuracies comparison of models

Algorithm	Reported accuracies		
	Minimum accuracy (%)	Mean accuracy (%)	Maximum accuracy (%)
LSTM	89.97	90.20	91.72
CNN	97.71	97.91	98.13

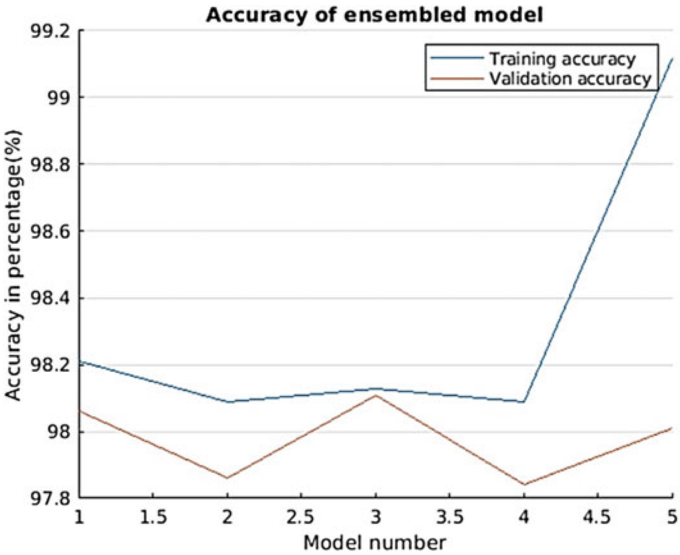


Fig. 5 Graph representing final accuracies of a CNN ensemble

Figure 5 represents the training and validation accuracy of an ensemble model of CNNs. We achieved an average of 98.32% training accuracy and 97.98% validation accuracy using this method.

5 Conclusion

In this paper, we developed various models of CNNs and LSTMs and finally decided to use ensemble strategies on CNN for their final classifier model. We were successful in compensating for the inaccuracies caused by the individual models that had lower accuracy and overall average it. To incorporate the advantages of our CNN models, we used the voting system for averaging the outputs from various base learners and this helped in overcoming the lower accuracy which may occur in some cases for some base learners. We managed to partially improve the results of some of these individual models. Therefore, we suggest that both, increasing the training set size and using ensemble results in better performance when compared to the use of a single independent model for classification. This reduces the probability of misclassification to some extent and increases the performance of the model. The proposed ensemble methods are generic and also applicable to most of the image recognition tasks even though we only tested our method on one dataset. As progress is being made in this domain, the future of web searching will eventually incorporate doodle-based searches. Another scope for doodles in the future is in the field of psychoanalysis.

A little bit of doodle analysis is present in this field, but it isn't as predominant as it could be. Graphology is also another field that can use doodles for analysis.

References

1. Kristine G, James W, Eric X (2018) Quick, draw! doodle recognition
2. Phu NM, Xiao C, Muindi J (2018) Drawing: a new way to search (computer vision)
3. Nagda M, Eswaran (2019) Poovammal image classification using a hybrid LSTM-CNN deep neural network. IJEAT 8(6)
4. Tran E (2017) Free-hand sketch recognition classification
5. Xu G, Wu H-Z, Yun QS (2016) Ensemble of CNNs for steganalysis: an empirical study. In: Proceedings of the 4th ACM workshop on information hiding and multimedia security (IH& MMSec '16). Association for Computing Machinery, New York, NY, USA, pp 103–107
6. Doodles of a Cat. https://quickdraw.withgoogle.com/data/cat
7. CS231n CNN for visual recognition. https://cs231n.github.io/convolutional-networks/
8. Convolutional Neural Network (CNN). https://www.edureka.co/blog/convolutional-neural-network/
9. Francois C Deep learning with python, pp 202–205
10. Illustrated guide to LSTM's and GRU's. https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation
11. Zhou Z-H Ensemble methods foundations and algorithm, pp 67–71

Training Optimization for a Hardware-Aware Approach to Deep Learning

Anjana Asok

Abstract Deep Convolutional Neural Networks (DCNNs) have been proven to be effective for various problems involving image, video and signal processing and classification. The size and complexity of the network are major concerns when it comes to real-life implementations of DCNNs. The increase in the number of parameters and a consequent increase in power and memory requirement have to be dealt with, for the hardware implementation of these networks. The quantization of parameters conceived as a method to reduce memory requirement has a trade-off on accuracy, and therefore various parameters have to be explored to get to a middle ground. The paper is divided in two parts - The first part explores the techniques to improve upon accuracy and to result in better generalization, for a deep convolutional neural network developed in MATLAB and the second section of the paper presents a comparison and usage of a scalable, efficient algorithm based on the Tensorflow library that does the feature extraction of the MNIST dataset as compared to the algorithm in MATLAB that performs the classification. The performance improvement is in terms of time-efficiency in training, an improved accuracy rate and a highly parallel algorithm that does effective classification.

Keywords CNNs · Hardware optimization · MNIST · Feature extraction

1 Introduction

Convolutional Neural Networks(CNNs) represent a class of multi-layer perceptrons that are composed of various layers of processing used to create complex computational models that can model high dimensional data to learn its representation with multiple levels of abstraction. CNNs have repetitive neuron blocks (that can be interpreted as 2-D convolutional kernels, in case of the MNIST dataset, that classifies handwritten digits), that are made to slide over patches of the image for repeated convolution to obtain feature maps which are taken to be compact comprehensive representations of the input data. The use of kernels or the weight-training of kernels

A. Asok (✉)
Birla Institute of Technology and Science, Pilani, Rajasthan 333031, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_49

is an important part of defining the network due to the spatial and temporal invariance criterion in recognition.

CNNs find a variety of applications in image, video and speech recognition as well as in natural language processing. They have been used in deep learning applications like image classification with excellent accuracy driven by the development of parallel algorithms and highly efficient optimization techniques. But the trade-off for such a network implementation that focuses exclusively on accuracy is the high power consumption, the increase in training time that extends to days, huge memory requirements and constraints imposed by limited processing power in embedded applications. Therefore, VLSI-based hardware implementations can be taken as a need of the hour for bringing the efficiency of CNNs to real-time applications with the constraints imposed by computing power, memory and time. The feasible techniques have to aim at a decrease in each of the requirements with a low impact on accuracy.

The major focus of this paper is in the performance improvement of an existing MLP algorithm by separating the feature extraction and classification parts and then the study of a novel algorithm that can potentially improve the feature extraction part to a considerate amount.

2 The MLP Network Structure

The initial model of feature extraction originally consists of a model that has the following architecture:

- Layer 1: Input Layer (28×28 Images)
- Layer 2: Convolutional Layer (6 kernels)
- Layer 3: Pooling Layer (Scale by 2)
- Layer 4: Convolutional Layer (8 kernels)
- Layer 5: Pooling Layer (Scale by 2)
- Fully Connected Output Layer: 10 neurons.

Our main focus was on the classifier part i.e. the fully connected layers. For testing purpose,we choose MNIST Handwritten Data Set. It is a database of 28×28 images.

- No. of training images = 60,000
- No. of testing images = 10,000.

The above architecture was used to obtain the features from the input data set as a vector of [60,000*128]. A separate classifier code takes care of the classification part where the modifications are made in terms of hardware optimization and network efficiency. The extracted features are fed into the classifier network.

The decrease in accuracy that comes about with quantization of weights, and outputs of each neuron in every layer is compensated by adding another fully connected layer of 60 neurons in between the output neurons and the feature layer.

3 Modifications and Optimization

3.1 Feature Extraction

3.1.1 Adding Another Fully Connected Layer

In order to model the effect of adding a fully connected layer in the classification model to make up for the loss in accuracy with the hardware optimization, a fully connected layer was added in the feature extraction part of the network with 60 neurons and its performance compared with that of adding two fully connected layers dividing the number of neurons into both. The objective of this was to see the trade-off in adding an extra layer to study its impact on efficiency in training the network as well as finding a lower bound on the number of neurons and the number of layers that the network might need for classification of the dataset without excessively compromising on the area of the implementation.

The literature survey and the modelling of effects are discussed in conjunction with the classification layer in the next section. The results show that there is a significant drop in accuracy upon using 2 fully connected layers of 30 neurons each as compared to using a single fully connected layer of 60 neurons. As the further implementations on-chip are expected to reduce the accuracy to a further low, this reduction in accuracy is not tolerable.

3.2 Classification Layer

The already available algorithm has dealt with the following issues with proposed techniques as:

1. Slow Convergence—Adaptive Learning Rate
2. Getting stuck in local minima—Adaptive Momentum
3. Reduced generalization—Dropout.

In order to prevent over-fitting during training, Dropout was used in which, neurons are randomly removed from the network along with its outgoing connections. This random dropping out of neurons during each training epoch ensures that the non-dominant neurons are given an equal or fair weightage as compared to the dominant ones.

3.2.1 DropConnect

An improvisation to the already available Dropout algorithm has been proposed in the form of DropConnect. DropConnect, is a generalization of Hinton's Dropout for regularizing large fully-connected layers within neural networks [1]. When training with Dropout, a randomly selected subset of activations are set to zero within each layer. DropConnect instead sets a randomly selected subset of weights within the network to zero. Each unit thus receives input from a random subset of units in the previous layer.

Training Network with Dropout: Each element of a layer's output is kept with probability p otherwise being set to 0 with probability $1 - p$. If we further assume neural activation function with $a(0) = 0$ for e.g. tanh and relu functions, then:

$$r = m * a(Wv) = a(m * Wv) \tag{1}$$

Training Network with DropConnect : Generalization of Dropout in which each connection, rather than each output unit, can be dropped with probability $1 - p$.

$$r = a((M * W)v) \tag{2}$$

where M is the weight mask, m is the node activation mask, W is the fully connected layer weights, v is the fully connected layer inputs, $*$ is used for element-wise multiplication and r represents the output layer values.

A theoretical justification as to how DropConnect helps regularize large neural networks is presented in [1]. Drop-Connect algorithm was chosen to be implemented to regularize the network at a better level to compensate the drop in accuracy due to overfitting in the case of a large number of training epochs. A combination of Dropout and DropConnect was applied on the 2 layers: the input layer and the hidden layer respectively to check for the best case scenario. The number of training epochs was set to 50 taking into account the delay and set as a benchmark in [2] for on-chip deep neural networks taking into account the delay.

The results obtained are shown in Table 1. Layer 1 is the input layer of 128 neurons and Layer 2 is the hidden fully connected layer having 60 neurons. The results show that the best case scenario in terms of accuracy is obtained when DropConnect is implemented in the weights between the input layer of 128 neurons and the fully connected layer of 60 neurons and DropOut for the fully connected layer of 60 neurons. This is expected as the mapping from the input layer to the hidden fully connected layer is prone to more over-fitting if the dominant weights are not nullified while training due to the fullyconnected layer mapping from 128 neurons to 60 neurons, thus reducing the feature information considerably. DropOut in the hidden layer of

Table 1 Quantized classification error percentages while implementing DropOut and DropConnect in fully connected layers

Layer1 | Layer2	DropOut (%)	DropConnect (%)
DropOut	7.37	19.65
DropConnect	6.37	19.42

60 neurons emphasizes the regularization already implemented and is supplemented by the random-weight-dropout applied in the previous layer.

Implementing DropOut and DropConnect for the fully connected hidden layer of 60 neurons however gave an error of 20% which is in line with the theory that random discarding of the neuron outputs and weights from the same layer can adversely affect training as the network won't be able to train efficiently due to an accumulated loss in information by the combined effect of weights and neuron information loss in the same layer.

3.2.2 Neurons in the Hidden Layer

The number of neurons in the hidden layer is important from a hardware perspective as the finding of a lower bound on the number of neurons that can give a reasonable accuracy can substantially reduce hardware in terms of the memory elements to be used, which has a direct impact on power consumption. In the original code, we started with taking the number of neurons to be 60 which is close to the mean of the number of input and output neurons, and guaranteed to give a decent performance as explained in [3].

However a rule of thumb in the selection of the number of neurons in the hidden layer that won't result in over-fitting is:

$$Nh = Ns/(\alpha \times (Ni + No)) \tag{3}$$

where Nh is the number of neurons in hidden layer, Ni the number of input neurons, No the number of output neurons, Ns the number of samples in training data set and α an arbitrary scaling factor usually in range of 2–10. Many papers recommend setting α to a value between 5 and 10. As emphasized in the [3], the number of free parameters in the model represented by its degree or number of non-zero weights has to be limited to a small portion of the degrees of freedom in the training data. The degrees of freedom is given by the number of training samples times degrees of freedom in each sample $= Ns \times (Ni + No)$ with the independence assumption of the degrees of freedom of each sample. α is an indication of the amount of generalization or the amount of over-fitting reduction that is to be needed in the network.

I have started with an alpha of 5 and gone upto 10 to increase the generalization of the network. The value of 10 theoretically gives an upper bound of 43 neurons in the hidden layer to prevent over-fitting with the values corresponding to the MNIST dataset and our network as 60000 training samples, 128 input neurons and 10 output neurons. The results obtained are shown in Table 2.

The experimental results verify that the best case accuracy is obtained with 39 neurons (nearer to 43) with an error percentage of 4.32% for the quantized phase of the training network.

Table 2 The error percentages for continuous and quantized phases of the classification layer with the change in the number of neurons in the hidden layer

No. of neurons	Stage-1 error (%)	Stage-2 error (%)
25	48.5	6.34
27	9.16	19.7
31	18.38	4.61
35	6.95	11.93
39	8.5	4.32
45	15.81	4.4

3.2.3 Permutation of Neurons in the Hidden Layers

Many methodologies have been discussed to fix the number of neurons in hidden layers in MLPs. A literature survey was done to go through the proposed techniques to come up with an estimate for out purposes. The optimization of several existing methods help in selecting the number of hidden layer neurons in DNNs. In a method proposed by Arai [4] in 1993, parallel hyperplane classifiers are used to estimate the optimum number and arrangement of neurons in hidden layer.

A coarse-to-fine search technique was proposed in 2010, to find the number of neurons in the hidden layer. Binary and sequential searches are used in this one. Lowest MSE on training is the criterion looked for and the optimum number of neurons is found by searching sequentially. A decision tree algorithm based on information entropy was proposed later on.

The major method used widely is a method proposed for 3 types of networks in MLP that eliminate trial-and-error and keep the generalization capacity. For our network, in order for simplicity of modelling on-chip, an objective was defined to see as to how the split in the number of neurons into 2 hidden layers from 1 hidden layer will impact the classification layer in terms of accuracy and training time. The experimental results obtained are enlisted in Table 3.

The method of geometric descent and arithmetic descent both mean that a large number of neurons in the first layer can effectively generalize the network than the

Table 3 The error percentages for continuous and quantized phases of the classification layer with the change in split of the number of neurons in the hidden layer for 50 epochs

HL1	HL2	Stage-1 error (%)	Stage-2 error (%)
20	40	45.11	35.05
40	20	36.99	23.14
10	50	89.9	74.25
50	10	71.43	36.63
40	50	25.24	8.7

other permutation with smaller number of neurons in the first hidden layer. This is reflected in the results obtained. However, the error percentages on adding one more layer in the feature extraction part is leading to lesser generalization and therefore, higher error percentages.

3.2.4 Output Quantization

In the available code, the output of the final fully connected layer was not quantized. The fully connected layer outputs were quantized to 3 bit values. This can result in an overlap of the values for the classifier. The algorithm developed picks the first index having the maximum value for the classifier as the label for the test input image. Upon testing, it was found that there is only a 1% drop in accuracy for the dataset for the quantized weights and neuron outputs as shown in Table 4. From further experimental results that gave only 100 overlaps from the training set of 60,000 images, it is concluded that the number of overlaps in the output quantized values are insignificant to affect the accuracy for the entire MNIST dataset.

3.2.5 Digit-Wise Separation of Training Data and Calculation of Toggle Factor for a Variation of Neuron Numbers from 60 to 2048

The first part involved the identification of a trend in the number of neurons in the hidden layer and the correlation with toggle factor.The toggle factor represents the number of bits that toggle from the absolute high for the three bit quantization of the value of neuron outputs from the hidden layer.The classification layer code was modified and a module written to convert output of neurons in the hidden layer into binary 3 bit values and for calculation of toggle factor. The results obtained for the 3-bit quantization, with a batch size of 1000 are shown in Table 5.

Table 4 Error values for output quantization with 60 neurons in hidden layer

No of epochs	Stage-1 error (%)	Stage-2 error (%)
50	6.03	6.99

Table 5 Toggle factor values for 3-bit quantization of hidden layer neuron outputs

No of neurons	Stage-2 error (%)	Toggle factor
60	4.45	0.5228
128	3.94	0.4092
256	3.12	0.4257
512	8.92	0.5795

Table 6 Toggle factor values for 1-bit quantization of hidden layer neuron outputs

No. of neurons	Toggle factor
60	0.6238
128	0.4826
256	0.6650
512	0.4594
1024	0.5024

Table 7 Toggle factor values digit-wise for 3-bit and 1-bit quantization of hidden layer neuron outputs

No of neurons	Toggle factor(3-bit)	Toggle factor(1-bit)
0	0.7330	0.1754
1	0.8150	0.3443
2	0.6819	0.2936
3	0.6709	0.3537
4	0.5522	0.2732
5	0.6575	0.2699
6	0.7529	0.4266
7	0.7444	0.5426
8	0.6453	0.4229
9	0.7853	0.5217

A 1-bit ADC was constructed and the results obtained for a run of 50 epochs with a batch size of 1000 are documented in Table 6.

The second part involved the separation of the MNIST training and testing data digit-wise between 0–9 and calculation of toggle factor for the individual digits. The classifier algorithm works perfectly for the digit-wise classification giving 100% accuracy in almost all cases.The results obtained are summarized in Table 7.

4 Tensorflow

The previous multi-layer perceptron model used has a serious drawback of having a very long-drawn training time for convergence (50 h for 2000 epochs)of the model to extract the features. Such a case is not tolerable in case of an on-chip model of a deep neural network where we cannot afford to spend that much time on feature extraction for a real-time application. Therefore, a new model developed using Tensorflow was considered with an aim to decrease the time required for training the network without compromising on the accuracy.

TensorFlow^{T}M is an open source software library for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them [5].

4.1 MLP Structure

- Layer 1: Input Layer (28x28 Images)
- Layer 2: Convolutional Layer (32 feature maps with 5 * 5 kernels)
- Layer 3: Pooling Layer (Sub-sampling by 2)
- Layer 4: Convolutional Layer (64 feature maps with 5 * 5 kernels)
- Layer 5: Pooling Layer (Sub-sampling by 2)
- Fully Connected Layer: The features obtained from Layer 5 are mapped to 1024 neurons
- Output Layer: 10 neurons.

The MNIST Handwritten Data Set [6] is used as having:

- No. of training images = 55,000
- No. of testing images = 10,000.

4.2 MNIST for Beginners

The code obtained from the tensorflow.org website for MNIST classification was run in its very raw version while feeding the MNIST data set consisting of 55,000 training examples and 10000 test examples. The MNIST for beginners is based on minimizing cross-entropy using the gradient descent algorithm with a 0.5 constant learning rate. Gradient descent is a fairly simple procedure, where TensorFlow simply shifts each variable a little bit in the direction that reduces the cost.

Here, new operations are added to the graph which implement back-propagation and gradient descent. Then it gives you back a single operation which, when run, does a step of gradient descent training, slightly tweaking your variables to reduce the loss.

Results
Accuracy on the MNIST data set using the beginner code was found to be 92%. This accuracy is pretty bad considering the already available networks that classify the MNIST data set have a reported accuracy of 99% or higher. The low accuracy is because of the fairly simple structure of the network and also that the optimizer used—the gradient descent optimizer while back-propagation is prone to overfitting and to be stuck in local minima. Therefore, we decided to go for the Deep MNIST bone-structure for our feature-extraction purpose and go for modifications on the same.

4.3 Deep MNIST

Although Deep MNIST code uses almost the same MLP structure as the beginner version, the major differences that this version has in comparison are:

4.3.1 Implementation of the Adam Optimizer Instead of the Steepest Gradient Descent Optimizer

Adam is an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments. The method is straightforward to implement, is computationally efficient, has little memory requirements, is invariant to diagonal re-scaling of the gradients, and is well suited for problems that are large in terms of data and/or parameters. The method is also appropriate for non-stationary objectives and problems with very noisy and/or sparse gradients [7].

This method differs in the manner that adaptive learning rates are calculated from first and second order moments of gradient estimates for the various variables. The Adam algorithm by means of adaptive moments takes care of two major problems in gradient descent algorithm:

(i) the problem of zero gradient and
(ii) the influence of initialization bias.

The problem of zero gradient is counteracted by initialization of moving averages to zeros, leading to zero biased moment estimates. The influence of initialization bias is counteracted by dividing the moments (say first and second order) by their expected value. In case of the second order moment, the estimate uses an exponential moving average of the squared gradient, with a particular decay rate as described by Diederik et al. [7].

4.3.2 Probability Association with DropOut

A probability is associated with the drop-out in case of the deep neural network as compared to that of the beginner network described in the Tensorflow network models. The association of probability p with drop out results in not entirely discarding the neurons in the hidden layer and weighing down the domination of the particular neurons. This results in more effective training of the network as the contribution of all neurons in the fully connected layer are taken into account.

4.4 Modifications

The following changes were made to the network from the already available algorithm:

(1) The number of neurons in the fully connected layer was reduced from 1024 to 128. The 128 neurons were mapped to the 10 output neurons and the algorithm was run for 20000 epochs and the accuracy was found to be 99.03%.
(2) Extraction of features in the form of the input feed to the MATLAB code for the classification layer to compare the accuracy obtained between the two features obtained as an output of 128 neurons. The results obtained are shown in Table 8.

It is noticed that the the tensorflow features when fed into the classifier code with output quantization results in a boost in accuracy by **3%** as well as the training time for the network is reduced by a factor of **96X**. The training takes only 30 min for 20,000 epochs for the whole MNIST training dataset as compared to 2.5 h for 100 epochs for the earlier feature extraction network based in MATLAB.

(3) Visualization of kernels: The difference between the two algorithms is mostly in terms of weight-training of the kernels used and how efficient the optimization process is using parallelization and use of tensors apart from the use of matrix multiplication in MATLAB that takes considerable computing power and memory. The obtained kernels for a single image from the MNIST dataset (Fig. 1) is shown in Fig. 2.

It is observed that the weight-training of kernels occur much more comprehensively in the Tensorflow network as compared to the MATLAB model. Almost all the feature maps preserve the spatial relation of the input pixels and in a much more cohesive manner. This is also a factor that hugely impacts the improved accuracy rate in the conceived model network.

(4) Change of function from ReLU to sigmoid: The implementation of the activation function in the deep MNIST network of the Tensorflow code is a ReLU function that maps the inputs using a rectified linear unit. Tensorflow offers a range of activation functions. The activation ops provide different types of non-linearities for use in neural networks. These include smooth non-linearities (sigmoid, tanh, elu, softplus, and softsign), continuous but not everywhere differentiable functions (relu, relu6, crelu and relu-x), and random regularization (dropout).

Table 8 The error percentages for continuous and quantized phases of the classification layer with the change in the number of epochs for the tensorflow features

No. of epochs	Stage-1 error (%)	Stage-2 error (%)
50	11.8	10.23
200	5.75	5.59

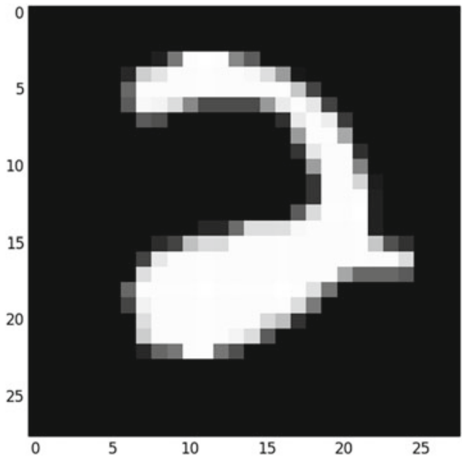


Fig. 1 The input image to the network

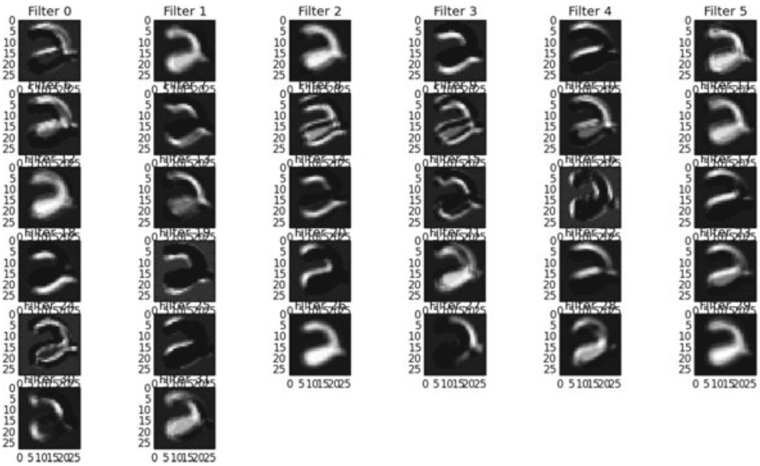


Fig. 2 The weight trained kernel output for the first convolutional layer in Tensorflow model for the input image

All activation operations apply component-wise, and produce a tensor of the same shape as the input tensor. The features that have to be input into the classifier code in MATLAB has to have an upper bound of 1. The rectified linear unit gave a greater upper bound due to which the sigmoid function was used for the final output layer. It was observed that the accuracy dropped in the case of sigmoid function being used for the outputs of all the convolutional layers and the final fully connected layer as well. Two major benefits of ReLUs are:

(i) sparsity
(ii) Less likelihood of gradient vanishing.

ReLU is defined as

$$h = \max(0, a) \tag{4}$$

where $a = Wx + b$.

This occurs in case of $a > 0$. The function being linear, here the gradient has a constant value. When the value of x increases, the gradient of the sigmoid function decreases and the incremental changes in gradient become too small to be detected. One of the reasons of the learning to occur faster in ReLUs is due to its constant gradient.

Sparsity is the other advantage of ReLUs. This is a case when $a <= 0$. The more of the ReLUs, the more the sparsity. The non-zero values in case of sigmoid function results in dense representations. Therefore, the more sigmoid functions are in the network, the more training error can occur as is theoretically established. To keep the outputs of the neurons bounded and also to not compromise on the training accuracy of the feature extraction, only the fully connected layer output was passed through the sigmoid function.

5 Conclusion

The initial part of the paper deals with performance improvement of an existing algorithm using regularization techniques like DropConnect, techniques and impact of adding another fully connected layer to the feature extraction code, studying the effect of permutations of neuron number in the arrangement of hidden layers, finding a lower bound on the number of neurons that can be incorporated in a particular hidden layer, the trends and correlation of toggle factor with the increase in number of neurons and finally quantization of output to reduce the memory requirements on-chip.

The main goal was to reduce the training time and quantize the parameters and obtain nearly the same accuracy. The weights were quantized to 4 bits, inputs and outputs were quantized to 3 bits without a significant drop in the accuracy. The goal was achieved by proposing certain techniques and keeping minimum hardware wherever possible. Also, the training speed was increased 96× by using stochastic gradient optimizers.

Various face, image and video recognition applications can be modelled using this work. The implementation of the feature extraction layer on chip is the next challenge with many new approaches being widely proposed.

Acknowledgements I would like to thank Prof. Mike S. W. Chen for the kind association with USC. The exploration helped me gain an in-depth understanding of deep learning and also helped me explore a new dimension in terms of a hardware aware approach to optimization. I am very thankful for his inspiration and mentoring which invoked my interest in the topic that was a great learning experience for me. I am also grateful to Mr. Rezwan A Rasul for his constant guidance and support throughout the course of the research. Finally, I have to thank my parents for the moral support and love they have extended throughout my life.

References

1. Li W, Matthew Z et al (2013) Regularization of neural networks using DropConnect. In: Proceedings of the 30th international conference on machine learning, Atlanta, Georgia, USA
2. Hasan R, Taha TM, Yakopcic C (2017) On-chip training of memristor based deep neural networks. In: IEEE international joint conference on neural networks (IJCNN), USA
3. Hagan MT (1996) Neural network design, 2nd ed
4. Arai M (1993) Bounds on the number of hidden units in binary-valued three-layer neural networks. Neural Netw 6(6):855–860
5. Documentation from https://www.tensorflow.org/
6. MNIST dataset. http://yann.lecun.com/exdb/mnist/
7. Kingma D, Ba J (2014) Adam: a method for stochastic optimization. In: International conference on learning representations

An End-to-End, Interactive Deep Learning Based Annotation System for Cursive and Print English Handwritten Text

Pranav Guruprasad, S. Sujith Kumar, C. Vigneswaran, and V. Srinivasa Chakravarthy

Abstract With the surging inclination towards carrying out tasks on computational devices and digital mediums, any method that converts a task that was previously carried out manually, to a digitized version, is always welcome. Irrespective of the various documentation tasks that can be done online today, there are still many applications and domains where handwritten text is inevitable, which makes the digitization of handwritten documents a very essential task. Over the past decades, there has been extensive research on offline handwritten text recognition. In the recent past, most of these attempts have shifted to Machine learning and Deep learning based approaches. In order to design more complex and deeper networks, and ensure stellar performances, it is essential to have larger quantities of annotated data. Most of the databases present for offline handwritten text recognition today, have either been manually annotated or semi automatically annotated with a lot of manual involvement. These processes are very time consuming and prone to human errors. To tackle this problem, we present an innovative, complete end-to-end pipeline, that annotates offline handwritten manuscripts written in both print and cursive English, using Deep Learning and User Interaction techniques. This novel method, which involves an architectural combination of a detection system built upon a state-of-the-art text detection model, and a custom made Deep Learning model for the recognition system, is combined with an easy-to-use interactive interface, aiming to improve the accuracy of the detection, segmentation, serialization and recognition phases, in order to ensure high quality annotated data with minimal human interaction.

Keywords Handwritten word detection · Handwritten text recognition · Automated text annotation

P. Guruprasad
Birla Institute of Technology and Science Pilani, K. K. Birla, Goa Campus, India
e-mail: f20171918@goa.bits-pilani.ac.in

S. Sujith Kumar · C. Vigneswaran · V. S. Chakravarthy (✉)
Bhupat and Jyoti Mehta School of Biosciences, Department of Biotechnology, Indian Institute of Technology, Madras, Chennai, Tamil Nadu, India
e-mail: schakra@ee.iitm.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_50

1 Introduction

Handwriting Recognition has been a field of extensive research, for the past few decades, and has evolved from approaches which employ the likes of Hidden Markov models [1], to Deep learning approaches [2, 3] in the recent past. Optical Character Recognition (OCR) systems that carry out the task of complete digitization, including text detection and recognition, have also been prevalent for a long time. Even though OCR systems have come a long way and produce excellent results on print text, their success has not carried over as much to cursive handwritten text digitization and face various challenges [4]. This can be attributed mainly to the variations in styles of handwriting, the spacing, lighting, geometrical orientations, noise, problems in segmentation, and much more. Even though many OCR systems of late are using Deep Learning approaches since the boom in the Deep Learning era, the results of these have been impressive only on handwritten documents with certain predefined structural formats, and not on simple, common handwritten documents and manuscripts. These drawbacks are because most OCR implementations are based on templating and feature extraction techniques. Thus, it is clear that there is still abundant room for improvements in the field of handwritten text recognition and digitization. To aid further research in these fields, we propose an annotation system for cursive and print handwritten English text, that provides fast and state-of-the-art quality annotated data while requiring hardly any human intervention or effort. Our annotation system offers a big advantage for users to create a comprehensive annotated dataset on any kind of handwritten documents of their choice, with no prerequisites for specified styles, formats, appearances or cumbersome preprocessing techniques. Upon passing mere photographs or scanned copies of the individual pages of a handwritten document into our pipeline, it converts them into a coherent annotated dataset that the user can use further for any desired tasks. Our pipeline consists of a word detection system built upon the state-of-the-art text detection model—EAST; an interactive user interface built using Python TKinter, that provides the user an opportunity to remove the common errors in the detection, segmentation and serialization phases which are committed by even the best text detectors; and finally a powerful Deep Learning model for the recognition phase, which boasts of a custom designed multi-dimensional LSTM (Long Short Term Memory units), Convolutional Neural Network (CNN) and a Connectionist Temporal Classifier(CTC). Even though we provide word level annotations, the recognition is implemented on a character level, thus allowing the recognition system to recognize words beyond its training data. In addition to this, meticulous and innovative data preprocessing techniques have been implemented on the images of the words that are detected and passed on to the recognition system, which helps bolster the robustness of the system and increase the accuracy of the annotations.

2 Related Works

The advent of Deep Learning and its rise in popularity, led to an increased need for annotated data for supervised learning tasks. Most of the early datasets have been annotated manually, and only in the recent past have there been attempts to automate the process of annotation and reduce the human efforts involved in it. However, it is important to keep in mind that in annotation, feedback from the user is indispensable in some cases in order to create the highest quality of annotated data that is nearly devoid of any errors.

There have been attempts to automate the annotation of data in various fields such as real time video feeds [5], object detection [6], and even the semantic web [7]. Similarly, with the evident need for annotated data for offline handwritten text, especially with the shift towards Deep Learning based approaches for offline handwriting text recognition, it is not surprising that there has been a lot of research in this area too. There have been quite a few works that have proposed a systematic arrangement of stages to create a complete annotation engine for handwritten text, comprising of varying levels of automation [8, 9]. However, a complete end-to-end pipeline to annotate handwritten text with very minimal human interaction is still considered to be a very challenging task as discussed in a part of a study by Ung et al. [10].

The two main components of our annotation pipeline are a word detection system and a handwriting recognition system. There has been extensive research in these fields for the past many years. There have been non-Machine Learning approaches, Machine Learning approaches, and most recently, Deep Learning approaches. Lavrenko et al. [11] presented a Hidden Markov Model based holistic word recognition approach, inspired by the results in cognitive psychology, for word recognition in handwritten historical documents. This method, which did not implement the segmentation of words into characters, gave a recognition accuracy of 65% which exceeded the results of the other systems during that time. Optical Character Recognition (OCR) systems have been around for a very long time, and there have been many attempts for Handwritten OCR, for various languages [12]. OCR techniques have evolved quantifiably over this period, and over the past few years with the advent of cloud computing, GPUs, and a better research community, have shifted towards some very impressive Deep Learning based models [13, 15]. However, as mentioned in the previous section, OCR techniques face various challenges [16] and have not been able to provide exciting results for cursive handwritten text documents that lack a predefined structure.

A work done by Shiedl et al. [17] implements a handwriting recognition system, with an architecture based on Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN) and a Connectionist Temporal Classifier (CTC), which provides impressive results for handwriting recognition. This was one of the main inspirations for the recognition architecture we present in this work. The 2D LSTM implemented in the recognition system of our work, is inspired by a work by Graves et al. [18].

Text detection has also been an extensively researched field, especially with the advancements in image processing, object detection and deep learning. Many works follow different categories of techniques for scene text detection and document text detection, like the sliding window techniques [19], single shot detection techniques [20] and region based text detection techniques [21]. Zhou et al. [22] proposed a simple yet robust deep learning based scene text detector known as EAST—Efficient and Accurate Scene Text Detector. In our work we build upon this EAST model as it not only outperforms a majority of the state-of-the-art text detectors in terms of accuracy and speed, but is easy to build upon, and works satisfactorily on different orientations of text words which provides a great advantage when dealing with handwritten text detection, given the variability in handwriting style from person to person.

3 Data and Preprocessing

This section mentions the datasets used for training and testing the detection and recognition architectures in our pipeline, and also discusses the various preprocessing techniques applied on the words, both before training the recognition system, and also before testing or implementing the recognition system on the unseen data, in order to help increase accuracy of recognition.

3.1 Data

The EAST model upon which our detection system is built, was trained on the ICDAR 2013 [23] and 2015 datasets [24]. The ICDAR 2015 dataset consists of a total of 1500 scene text images, out of which 1000 are training data and the remaining are test data. The ICDAR 2013 dataset contains 229 training images which were also used additionally for training. For the recognition system, our model was trained on the entire IAM Offline Handwriting Database [25]. This dataset contains words in both print and cursive handwritten English text that have been written by 657 writers. These pages have been scanned, automatically segmented, and then manually verified. Containing over 1500 pages of scanned text and over 1,00,000 labeled words, this dataset provides the massive volume of data that would be required to train the deep network of our recognition system, in order to perform well. Apart from this, the dataset offers a large diversity in the various features of offline handwritten text such as style, geometric orientation and spacing, thus ensuring the robustness of the handwriting recognition system. This enhances the performance of the system not only with respect to new, unseen data, but also transfer learning. Apart from the IAM offline dataset, we used the CVL dataset [26] to test the robustness of our recognition system, as the handwriting styles in this dataset make it very hard to recognize the words and are not considered to be very legible .

3.2 Preprocessing

3.2.1 For Training the Recognition System

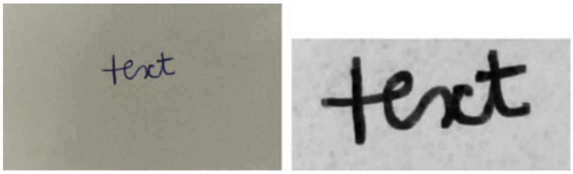
The IAM dataset consists of grayscale images of individual words. There is a variation in width and height of the images due to the lengths of different words and the heights of different characters. To maintain uniformity, and to make it easier to pass the images as inputs to the model, each image was resized such that it has the width or height of at least 1, and either a width of 128, or height of 32 at most . This resized input image was then copied into a complete white image which had a fixed width of 128 and height of 32 (128×32). Thus, every input had a uniform dimension of 128×32 without getting distorted in the process. Basic data augmentation techniques were used to increase the dataset size and make it robust to common variations that occur in the dataset. As a part of this, the input images were subject to random stretches, for which the stretch values were obtained by a random function set within a specified range. They were also shifted horizontally and vertically by small amounts so that even if parts of a character were cropped out, the recognition system would still be able to recognize the whole word. After these augmentations, the grayscale values of the images were also normalized, just to make the task easier for the network. Some of the images in the IAM dataset are damaged, and to take care of this problem, black images of the same dimensions were used in place of these damaged files. In addition to all these techniques, during the training process, the recognition system was additionally trained on the same images but with an added randomized Gaussian noise which helped reduce any chances of overfitting the IAM dataset, and enabled the recognition system to recognize the word to a good accuracy level irrespective of the gray values and contamination of the new image due to various natural causes.

3.2.2 Processing Applied to the New Unseen Data

Converting to the IAM format The pictures or scanned copies of the handwritten text documents that have to be annotated, may look very different from the images used to train the recognition system. The neural network not only learns how to read the words but also learns features like the contrast, the thickness of the words, the style and even learns the features of the surroundings.

To make sure that these factors do not affect the results, and to cater to a huge variety of handwritten text in documents, we process the word images (which are passed on in the pipeline to the recognition system) that do not resemble the images in the IAM dataset. To make sure they are very similar to the IAM images, three functions are carried out:

1. The contrast of the images is increased to a high contrast level
2. The word images are cropped to a very tight fit around the text
3. The thickness of the text is increased to make it resemble a bold font style.



(a) (b)

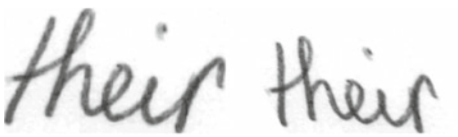
Fig. 1 Handwritten text image before and after the preprocessing steps to convert to IAM style text (**a**) Non IAM handwritten text image before preprocessing (**b**) Same handwritten text image as (**a**), but after the preprocessing step done to convert to IAM style image

Figure 1a shows a handwritten image that is initially not in the IAM data format, and Fig. 1b shows how the preprocessing step converts it to resemble IAM dataset style images.

This preprocessing step has been observed to increase the accuracy of recognition significantly for word images that look different from the training data.

Removing the slope and slant (Normalizing) The amount of slope and slant in cursive English words varies from writer to writer. Since this change in slope and slant affects the appearance and geometric orientation of the characters, this variation may have a significant effect on the performance of the recognition system.

To make sure that irrespective of the writer's style, the recognition system accurately recognizes the characters, while testing on the unseen data—we use a normalization technique to remove the slope and slant to a fair extent from all the words that come through the pipeline into the recognition system, so that they all have roughly the same amount of slope and slant. This makes sure that a cursive character written by two writers with completely different styles, would still look very similar to each other and thus prevent the recognition system from making any errors. When an input image is passed into the function that carries out this operation to remove the slant, the entire text part of the image is deslanted and made fairly upright, and the empty part of the image is filled with white colour. The algorithm we use for this slope and slant removal is based on a work by Vinciarelli et al. [27], and was implemented using OpenCV. Figure 2a shows an image of handwritten text with slant and slope, and Fig. 2b shows the same handwritten text after the slope and slant have been removed.



(a) (b)

Fig. 2 Handwritten text image before and after the slope and slant removal **a** a handwritten text image from the IAM dataset, **b** same handwritten text in (**a**), but after slant and slope removal

4 Methodology and Functioning

4.1 Phases

The functionality of the proposed pipeline can be divided into independent phases, with each phase contributing sequentially, towards the goal of completely annotating the document. In this section we discuss the functioning of the pipeline, phase by phase when the photo or scanned copy of a single page of the handwritten text document is passed as input. Figure 6 shows a simple overview of the sequence of phases in our pipeline.

4.2 Word Detection

The first phase in the pipeline is the detection of each of the words in the photograph of the page uploaded. Our detection model was built upon the pre-trained EAST model, which is a very robust text detector and achieves state-of-the-art text detection accuracies such as an F-score of 0.7820 on the ICDAR 2015 dataset. The page that is uploaded to our pipeline, is resized to a standard 720p image, before passing it into the detection system for detecting the individual words. For a page of handwritten text, our detection system which boasts of very impressive speed, produces its final set of bounding boxes around all the words it detects in the page, in just around an average of 0.5 s. In this network, one output layer gives the coordinates of all the initial bounding boxes that are predicted for the words detected in the page. In addition to this output layer, there is another output layer with a sigmoid activation function that outputs the probabilities signifying the presence of text in different regions of the image. After these values are obtained, a Non-Maximum Suppression technique is implemented to remove the weak and overlapping bounding boxes which are associated with lesser probability scores (which do not have a higher probability than a specified threshold). This then results in the final set of the bounding boxes predicted by the system, with their respective coordinates, for all the words it could detect in the page. After this function is complete, there are still some words that are undetected (which are usually one or two lettered words), some wrongly located bounding boxes, and some bounding boxes that do not cover the entire word vertically, or horizontally, or along both directions. These are common mistakes committed even by state-of-the-art object and text detection systems. Especially in our case, since it is handwritten text, the data is rife with noise and style variations, which only increases the chances of these detection imperfections. To make sure that we do not miss out on annotating every single word and symbol in the document, and to ensure that the accuracy levels of the recognized words which are used to annotate the document are kept high, the pipeline leads into the next phase, which is the interactive interface that allows the user to intervene and rectify these imperfections with minimal effort.

4.3 *Interactive Interface*

4.3.1 Editing, Adding and Deleting Bounding Boxes

As mentioned in the previous section, there are always certain cases where human intervention is inevitably required to create near flawless annotations, which is facilitated in our work by this interactive interface. The interface was completely built using Tkinter, which is a Python interface to the Tk GUI toolkit. Tkinter is widely used as a standard GUI for Python implementations, and works across all the popular operating systems.

This part of the pipeline first displays the entire image of the input handwritten page after being resized, on a Tkinter canvas of fixed size, for the user to see. Then the bounding box coordinates from the detection system of the pipeline are retrieved, and scaled according to the Tkinter canvas size. Once the coordinates are obtained with respect to the canvas size, these bounding boxes are displayed as interactive red rectangles at the coordinates where the detection system predicted the bounding boxes on the page. Before the serialization of words, this interface provides 4 main functionalities—Adding new bounding boxes, deleting bounding boxes, resizing bounding boxes and moving around of bounding boxes.

Adding Bounding Boxes: This feature can be used by the user when the detection system has not recognized certain small words or punctuation symbols. Usually text detection systems do not recognize one or two lettered words that are written so shabbily that they hardly look like text, or some punctuation symbols like full stops or commas, that are written too lightly and are barely recognizable as a written component. This can be corrected by drawing a bounding box over any symbol or word that has not been detected. The drawing of the bounding box is made very easy for the user and a left click and drag of the mouse anywhere on the canvas simply draws a bounding box, and this is logged into the system automatically as a new bounding box, along with its respective details such as coordinates.

Deleting Bounding Boxes: This feature comes in handy when the system has detected wrong spaces of the image as a text containing region, or if there are two bounding boxes over one hyphenated words, and many other scenarios like this. We allow the user to remove any bounding box on the canvas with the simple press of a key, after which the information stored regarding this bounding box is deleted in the system.

Resizing/moving Bounding Boxes: These features are useful when the bounding box coordinates detected by the detection system are not entirely accurate and do not cover the entire part of the word or symbol. This is a very common problem in the already existing methods and can affect the output of the recognition system adversely. The resizing functionality that we provide, allows the user to resize bounding boxes very easily, identical to how one would resize an image in a word document or a powerpoint presentation document, by dragging along the corners or edges of the bounding box. Moving the bounding boxes around as a whole, can be done by hovering the mouse over the desired bounding box and pressing any one of the arrow

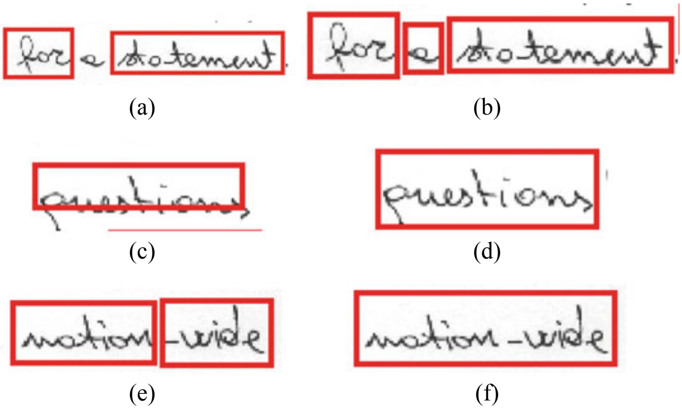


Fig. 3 **a** The detection system missed the detection of a one lettered word—"a". **b** The interface allows the user to draw a bounding box so that "a" is detected. **c** The detection system does not draw the bounding box accurately enough to enclose the whole word. **d** The interface allows the user to resize the bounding box so that the whole word is enclosed. **e** The detection system detects one hyphenated word as two separate Words. **f** The interface allows the user to delete the extra box, move around and resize the other box so that it is enclosed as one

keys according to the required direction. Once the bounding boxes have been shifted or resized, their respective modified coordinates are automatically stored. Once these editing operations are carried out and the user is satisfied with the positions of the bounding boxes around all the words and symbols in the page, the serialization of the words can be carried out. Figure 3 shows examples of when these operations are required and how our interface allows the user to fix them.

4.3.2 Serialization

The serialization phase is a very important phase as it determines the order of the detected words in the page of the document. Once the editing phase is complete and verified by the user, just a simple press of a specific key, automatically serializes all the words and symbols that have bounding boxes around them, and is represented on the canvas by straight lines between the bounding boxes to indicate their order, as it can be seen in Fig. 4. The serialization function was implemented by our own sorting algorithm that was based on the scaled canvas coordinates of the bounding boxes, the resolution of the image, the space between the lines of text and the space between the adjacent words in a line. Once this serialization is complete and is visible to the user, the user can further modify it in case they desire to change the order of any of the words. We have enabled a swapping feature to carry this task out, which we facilitated by constructing an elaborate dictionary data structure that contains information regarding the coordinates of each of the bounding boxes, their neighbours on either sides based on the serialized order, and their object tags in

Fig. 4 The black solid lines between the bounding boxes indicate the order in which the words have been serialized

the canvas. By just right clicking on two bounding boxes and pressing a key, the serialized order of the two bounding boxes are swapped. This change is not only shown on the interface, but also automatically logged in the system, and the new final serialized order is stored.

4.4 Recognition

Once the serialization phase is complete, the order of the words present in the document is now finalized. The images of these individual words, within each of the bounding boxes, are then extracted and stored using OpenCV. They are then passed on, one by one (in the same serialized order), as individual word inputs to the recognition system in the pipeline. Each input image that is passed into the recognition system goes through the three architectures of our recognition model: the Convolutional Neural Network (CNN), the multi-dimensional LSTM, and the Connectionist Temporal Classifier (CTC). The exact details and nature of these architectures are discussed in the next section.

For each input of an individual image to the recognition system, the output is a sequence of characters. The recognition system thus outputs a sequence of characters/individual character, for each of the detected, and serialized images passed into it from the previous phases of the pipeline. As the image passes through each of the layers of the CNN, the trained layers extract all the required features from that image. There are three main operations that are carried out on the image in the CNN, in each layer: the convolutional operation, a non-linear activation and a pooling function. Apart from these three operations, we add a Gaussian noise layer that adds standard Gaussian noise to the input, for reasons explained in Sect. 3.2.1. Then finally, after passing through two fully connected layers, a feature map is output.

The output feature map from the CNN is then passed as input to our 2D LSTM. An LSTM is used instead of a standard unidirectional or bidirectional RNN, as LSTMs prevent loss of information over long distances, and so is very helpful when dealing with long character sequences, which is very important for the task at hand. Our custom 2D LSTM was designed and implemented instead of using a standard one dimensional LSTM, because we felt that considering both the horizontal and vertical dimensions of handwritten text while recognizing it, would be much more effective than just working along one dimension. This is because the English cursive handwritten text has a myriad of variations along both the dimensions, and the system would be very robust if it could learn the features across both these dimensions. Moysset

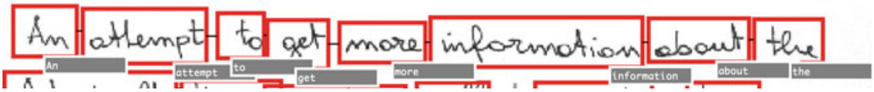


Fig. 5 The recognized words are displayed in editable text boxes below their corresponding detected words

et al. [28] show that 2D LSTMs give great results for recognition of handwritten text, and provide higher performances as compared to single dimensional RNNs or LSTMs even when used on complex, challenging and real life data.

The output sequence of the LSTM is mapped to a matrix which becomes the input to the final CTC layer. Connectionist Temporal classification, a work by Graves et al. [29] is a method that serves two purposes: it not only calculates the loss values required for training, but also decodes the matrix that is output from the LSTM, to obtain the final text that is present in the input image. During the training process, both the ground truth and the LSTM output matrix are fed to the CTC layer, and based on these, a loss value is calculated which is used to train the system to recognize the right sequence of characters. During the inference phase, only the LSTM output matrix is fed, and is decoded by the CTC layer, using a decoding algorithm, to get the text from the images. As the text from each input image is recognized by the recognition system, they are checked for any misspells, and are corrected using a state-of-the-art python spell checker known as Pyspellchecker. This spell checker which was developed and released very recently, even offers a feature where the users can add words of their choice to the dictionary, thus allowing them customize the dictionary to suit the task at hand. After this stage, all the recognized words and symbols are stored in the same serialized order. Then, all these stored texts are retrieved and displayed in text boxes on the Tkinter canvas, under each of their corresponding detected words/symbols, as the annotations for that word/symbol, as seen in Fig. 5. Then we offer a feature that allows the user to interact again with the pipeline before the final annotations are stored. The user is allowed to edit the recognized text present in each of the text boxes according to their desire. This ensures that any small mistakes made by the system are corrected completely before the annotations are finalized. After this stage, the user can click a button that shows up on the canvas, upon which the final modified annotations for all the words are stored and written into a text file in the same order as they appear in the original handwritten document, and are accessible to the user.

4.5 Details of the Recognition Model

The combination of a CNN, RNN and a CTC layer has been gaining popularity in the recent past, especially for text recognition tasks. We modify this architectural combination, and improve upon it by building our own 2 Dimensional LSTM to replace the standard RNN, which results in a significant increase in performance. Apart from this, the CNN in our system has also been designed in a way such that it is powerful and robust enough to deal with handwritten text images. The CNN model contains 10 layers, out of which each of the first 8 layers perform convolutional operations, non linear activations and pooling functions. The convolutional operations are carried out by filters of varying kernel sizes from 7×7 to 3×3, and a standard RELU non linear activation function is used. These 8 layers are followed by a Gaussian noise layer and 2 fully connected layers. The input to the CNN is the preprocessed image of dimensions 128×32, and the output is a feature map of size 32×512. The dimension of this 32×512 feature map that is passed as input to the LSTM, represents 512 features per time step, where each time step represents the position for the characters that may possibly be present in the word to be recognized. Each of these timesteps contain 512 relevant features extracted by the CNN layers. There are 32 timesteps because we set the maximum length of the character sequence that can be recognized, to 32. We found that for values greater than 32 the system performed worse and the loss values were considerably higher.

The 2D LSTM which was built using 256 hidden cells, processes this feature map further, by only carrying forward the relevant information. The output of the 2D LSTM is finally mapped to a matrix of dimension 32×80, where 80 is the number of possible characters that can be recognized. This is because, apart from the 79 different characters present in our training dataset, another extra character is required for CTC operations, known as the CTC blank. Therefore, this 32×80 matrix contains the probability scores with respect to the 80 different possible entries for each timestep.

This matrix is provided as input to the CTC layer, which during training is compared with the ground truth tensor to generate a CTC loss value. This CTC loss value was the error metric that was considered for training. We used an RMSProp optimizer with a decaying learning rate that was initialized to 0.01, and a batch size of 50 for the training process. During the inference phase, the text in the image is decoded by the CTC layer using a CTC beam search decoding algorithm, which is offered as a feature in the Neural Net module of Tensorflow. This algorithm was used instead of the standard greedy best path decoding algorithm, because of which we were able to improve the accuracy of the recognized words even further.

Figure 7 shows the summary of our recognition model architecture.

5 Results and Discussions

The IAM offline dataset which was preprocessed as discussed in Sect. 3.2, was used to train the recognition system, with a train-valid split ratio of 95:5. Therefore, a total of 115,320 words were used to train the model. In order to measure the performance of our recognition system, Character Error Rate (CER) was used as the error criterion, which is the standard among most works in this field. The CER was calculated based on the Levenshtein edit distance between the recognized word and the groundtruth word. This edit distance value between each ground truth and corre-

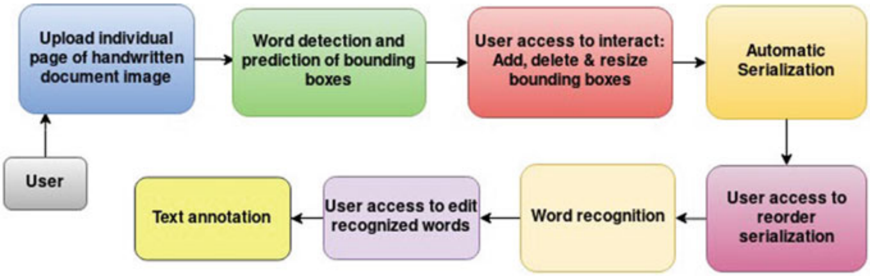


Fig. 6 An overview of the sequence of phases in our pipeline

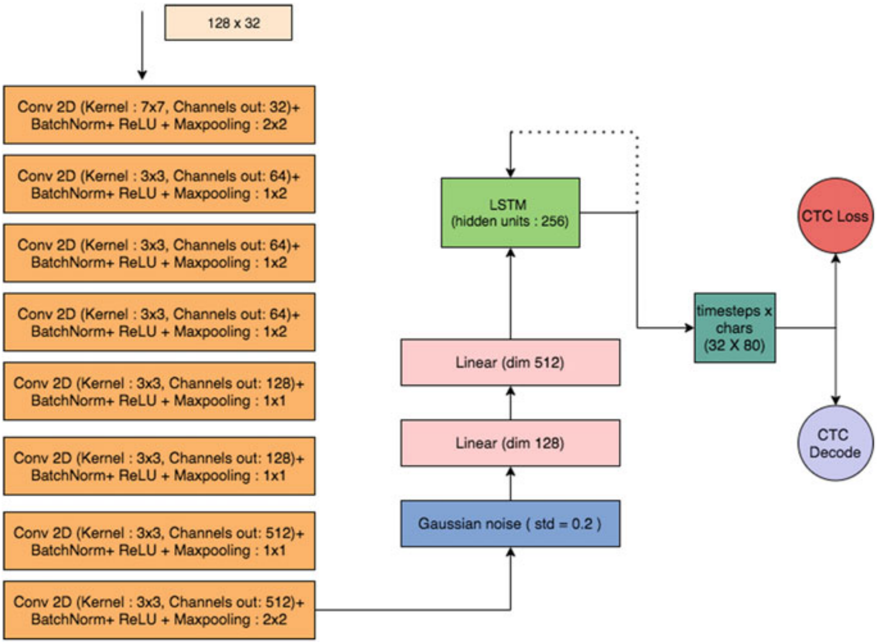


Fig. 7 Summary of our recognition system architecture

sponding recognized word, was summed for all the words in the epoch, and divided by the total number of characters in all the words in the epoch, to get the final CER value. Table 1 shows the CER obtained using the two different architectures that we tried for the recognition system: one that implemented a multi-dimensional LSTM, and the other that used a single dimensional bidirectional LSTM. The significant reduction in error rate upon using a 2D LSTM justifies our choice of building a custom multidimensional LSTM instead of using a standard unidimensional, bidirectional LSTM. Table 2 further compares the results of our work with other works

Table 1 Comparison of the two recognition models implementing different LSTM architectures (in terms of CER)

Model architecture	Character error rate (CER)	Epochs trained
CNN + bidirectional LSTM (single dimension) + CTC	12.36	110
CNN + multidimensional LSTM + CTC	9.3	100

Table 2 Comparison with other works that implemented different recognition architectures and were trained on the IAM offline dataset

Work	Model	CER
Ingle et al. [30]	Two Bidirectional LSTM layers	12.8
Ingle et al. [30]	1-D gated recurrent convolutional layers	14.1
Almazan et al. [31]	Kernelized Common Subspace Regression	11.27
Bluche [32]	Multi-layer perceptrons and Hidden Markov system	15.6
Our work	CNN + multidimensional LSTM + CTC	9.3

done in recent times that aimed to do similar tasks after being trained on the same datasets. As it can be observed, our proposed model performs better than all of them.

The decoding algorithm that we used during inference was a beam search decoding algorithm as mentioned in Sect. 4.5. Even though this gave much better results as compared to the greedy CTC decoder algorithm, we tried using an even better decoding algorithm known as word beam search decoding, in an attempt to further the recognition accuracy. This resulted in a small improvement in validation word accuracy. However, using a word beam search decoder limits the words recognized to those present in a dictionary/corpus that is created in the process. Thus, to make sure that our system is not limited by such constraints and performs well on words never seen before, we chose the beam search decoding algorithm over the word beam search decoding algorithm.

To further check the robustness of our recognition system, we checked it with handwritten text that have writing styles which are harder to interpret than the text in the IAM dataset. For this we used the CVL dataset. Figure 8 shows how the system was able to pass all the phases of the pipeline successfully, and recognize the words from the CVL dataset to great accuracy, even though these words are not very legible. It is clear from our results and methodology that the proposed pipeline not only provides very impressive results for automatically annotating the data, but also ensures that the user has to put in a negligible amount of effort to ensure a near flawless annotation.

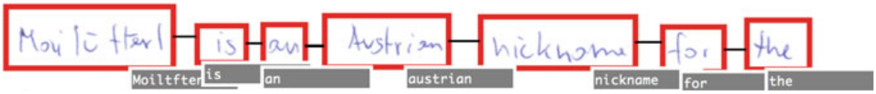


Fig. 8 Words from the CVL dataset recognized by our recognition system

6 Conclusion and Future Works

After observing the burgeoning need for annotated data in the field of handwritten text recognition, we have presented a robust and innovative pipeline that carries out annotation of cursive and print handwritten English text, with a great accuracy while ensuring the least amount of human effort required. We present an annotation pipeline that uses a word detection system based on a state-of-the-art model, and combine it with a powerful custom designed recognition system. To deal with the common errors that are committed by these systems, we provide a very intuitive interactive interface which ensures the removal of a majority of the flaws with minimum effort, resulting in very fast, high quality annotated data. The potential for this pipeline is very high and has various applications. It can be used to create large amounts of custom annotated data very easily, that can be used to train systems that focus on problems such as restoring old handwritten scriptures and manuscripts, evaluating exam answer sheets, designing literary based softwares, digitizing handwritten notes and much more.

This work has potential to further be improved, by using deeper networks and larger datasets in the case of availability of powerful computational systems and hardware, which we did not have access to. Better preprocessing techniques and more powerful CTC decoding algorithms are also aspects of the work that can be focussed on for significant improvements.

Future work can focus on extending our current work to regional languages, where there is a clear lack of significant amounts of annotated data. Designing annotation pipelines for regional languages may require more sophisticated segmentation techniques and better recognition systems, and is definitely something that requires much more meticulous research. Even though there are a large number of problems that can be solved by developing systems that aim to digitize or restore documents in regional languages, one of the main limiting factors for extensive research in this field is the dearth of high quality annotated data.

References

1. Gilloux M (1994) Hidden Markov models in handwriting recognition. In: Impedovo S (eds) Fundamentals in handwriting recognition. NATO ASI Series(Series F: Computer and Systems Sciences), vol 124. Springer, Berlin
2. Balci B, Saadati D, Shiferaw D (2017) Handwrittent Text recognition using deep learning. CS231n: Convolutional Neural Networks for Visual Recognition. Stanford Uni, Course Project Report

3. Ptucha R, Such FP, Pillai S, Brockler F, Singh V, Hutkowski P (2019) Intelligent character recognition using fully convolutional neural networks. Pattern Recogn 88:604–613
4. Sukanya R et al (2018) A study on handwriting analysis by OCR. Int J Sci Res Publ 8(1)
5. Manikandan NS, Ganesan K (2019) Deep learning based automatic video annotation tool for self-driving car (2019)
6. Kiyokawa T, Tomochika K, Takamatsu J, Ogasawara Ts (2019) Fully automated annotation with noise-masked visual markers for deep-learning-based object fetection. 4:1972–1977. https://doi.org/10.1109/LRA.2019.2899153
7. Tang J et al (2012) Automatic semantic annotation using machine learning. Mach Learn 535–578
8. Bhattacharya U, Banerjee R, Baral S, De R, Parui SK (2012) A semi-automatic annotation scheme for Bangla online mixed cursive handwriting samples. In: 2012 international conference on frontiers in handwriting recognition, Bari, pp 680–685
9. Stork L, Weber A, van den Herik J, Plaat A, Verbeek F, Wolstencroft K, (2019) Automated semantic annotation of species names in handwritten texts. In: Azzopardi L., Stein B., Fuhr N, Mayr P, Hauff C, Hiemstra D (eds) Advances in information retrieval. ECIR (2019) Lecture Notes in Computer Science, vol 11437. Springer, Cham
10. Ung HQ, Phan MK, Nguyen HT, Nakagawa M (2019) Strategy and tools for collecting and annotating handwritten descriptive answers for developing automatic and semi-automatic marking—an initial effort to math. In: 2019 international conference on document analysis and recognition workshops (ICDARW), Sydney, Australia, pp 13–18
11. Lavrenko V, Rath TM, Manmatha R (2004) Holistic word recognition for handwritten historical documents. In: First international workshop on document image analysis for libraries, Palo Alto, CA, USA, pp 278–287
12. Memon J et al (2020) Handwritten optical character recognition (OCR): a comprehensive systematic literature review (SLR). IEEE Access 8:142642–142668
13. Namysl M, Iuliu K (2019) Efficient, lexicon-free OCR using deep learning. In: 2019 international conference on document analysis and recognition (ICDAR)
14. Bartz C, Yang H, Meinel C (2017) STN-OCR: a single neural network for text detection and text recognition. arXiv preprint arXiv:1707.08831
15. Lee C-Y, Osindero S (2016) Recursive recurrent nets with attention modeling for OCR in the wild. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR)
16. Karez H, Mehmet K (2016) A detailed analysis of optical character recognition technology. Int J Appl Math Electron Computers 4:244. https://doi.org/10.18100/ijamec.270374
17. Scheidl H (2018) Thesis on handwritten text recognition in historical documents. Technische Universität Wien
18. Graves A et al (2007) Multi-dimensional recurrent neural networks. Lecture notes in computer science artificial neural networks (ICANN 2007), pp 549–558
19. Wang K et al (2011) End-to-end scene text recognition. In: 2011 International conference on computer vision
20. He P et al (2017) Single shot text detector with regional attention. In: 2017 IEEE international conference on computer vision (ICCV)
21. Huang Z et al (2014) Text extraction in natural scenes using region-based method. J Digital Inform Manage 12(4)
22. Zhou X et al (2017) EAST: an efficient and accurate scene text detector, in 2017 IEEE conference on computer vision and pattern recognition (CVPR). https://doi.org/10.1109/cvpr.2017.283
23. Karatzas D, Shafait F, Uchida S, Iwamura M, i Bigorda LG, Mestre SR, Mas J, Mota DF, Almazan JA, de las Heras LP (2013) Robust reading competition. In: Proceedings of ICDAR (ICDAR 2013)
24. Karatzas D, Gomez-Bigorda L, Nicolaou A, Ghosh S, Bagdanov A, Iwamura M, Matas J, Neumann L, Chandrasekhar VR, Lu S, Shafait F, Uchida S, Valveny E (2015) Competition on robust reading. In: Proceedings of ICDAR 2015

25. Marti U, Bunke H (2002) The IAM-database: an english sentence database for off-line handwriting recognition. Int J Document Anal Recogn 5:39–46
26. Kleber F, Fiel S, Diem M, Sablatnig R (2013) CVL-database: an off-line database for writer retrieval, writer identification and word spotting. In: Proceedings of the 12th international conference on document analysis and recognition (ICDAR), pp 560–564
27. Vinciarelli A, Luettin J (2001) A new normalization technique for cursive handwritten words. Pattern Recogn Lett 22(9):1043–1050
28. Moysset B, Messina R (2019) Are 2D-LSTM really dead for offline text recognition? Int J Document Anal Recogn (IJDAR) 22(3):193–208. https://doi.org/10.1007/s10032-019-00325-0
29. Graves A, Fernández S , Gomez F, Schmidhuber J (2006) Connectionist temporal classification. In: Proceedings of the 23rd international conference on machine learning (ICML ’06)
30. Ingle R, Fujii Y Deselaers T, Baccash J, Popat A (2019) A scalable handwritten text recognition system, pp 17–24. https://doi.org/10.1109/ICDAR.2019.00013
31. Almazan J, Gordo A, Fornes A, Valveny E (2014) Word spotting and recognition with embedded attributes. IEEE Trans Pattern Anal Mach Intell 12:2552–2566
32. Bluche T (2015) Deep neural networks for large vocabulary handwritten text recognition. Ph.D. dissertation, Université Paris Sud Paris XI

Enhancing the Classification Accuracy of Credit Default Using Extreme Gradient Boosting with Recursive Feature Selection

Reshma Thomas and E. R. Vimina

Abstract This research aims to propose an efficient model for the prediction of default of credit cards. The model is constructed using Extreme Gradient Boosting Ensemble technique on Taiwan based credit default dataset, which contains both financial and demographic attributes of credit card holders. The performance of the model is improved using Feature Selection and Hyperparameter Optimization methods. The result shows that the proposed model has more accuracy than many existing models for default prediction.

Keywords Credit default · Ensemble · XGBoosting · Feature selection · Hyperparameter optimization · Cross validation

1 Introduction

Financial institutions along with the economy is evolving dynamically and hence the way of assessing credit must also change. Institutions are keen on creating risk evaluation models and with plethora of data in hand, it is easier to create a sophisticated model for risk prediction thus benefitting the organization. This will give an edge for the institution to decide up on the credit limit to provide hence helping the credit card issuer to have a vivid understanding of their current customers and helps in the targeted marketing of credit services. Financial stability is crucial for the growth of an economy and monetary crisis can be minimized to a greater extend if there is a way to foresee the risks. One of the motives of predicting the risk is to make use of the economic status of the business as well as the customers to foresee the business performance and to analyse the behaviour of the customer in repaying the credit amount. This helps to reduce loss or damage and the business can use the extracted information for targeted marketing also. So, it is always helpful to have tools for predicting the customer's financial behaviour.

R. Thomas (✉) · E. R. Vimina
Amrita School of Arts and Sciences, Amritha Viswa Vidyapeetham, Kochi, Kerala, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_51

2 Literature Review

Umpteen studies were conducted in this domain dealing with the default of credit card clients. Different data mining algorithms were used for the prediction of default. However, machine learning models are observed to produce more accurate results than the traditional data mining algorithms. According to Hand and Henley, in the earlier days, statistical tools like Discriminant Analysis (DA), Logistic Regression, Bayesian Classifier, k-Nearest Neighbour algorithm was used to build models to predict the risk factor [1]. Koh and Chan [2] and Thomas [3] in their respective studies states that the evolution of Artificial Intelligence and Machine Learning enabled Artificial Neural Networks (ANN) and Classification Trees to predict credit default.

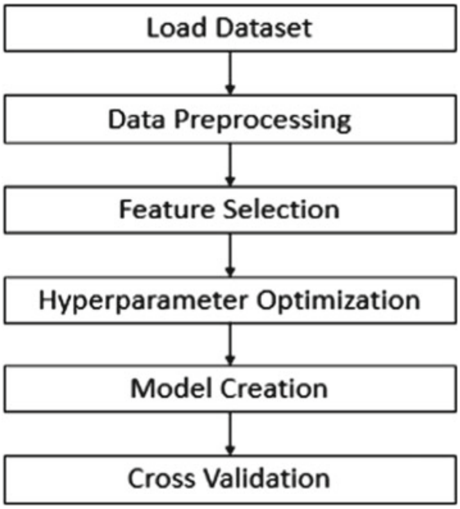
Many researchers tried predicting the probability to default, rather than focusing on classifying the clients to risky and non-risky. Forecasting the probability of default is a tedious task for practitioners and researchers and it requires more study [4]. A study conducted in 2009 used the Sorting Smoothing Method to deduct the actual probability of default. Among the six data mining techniques used, Artificial Neural Network is found to have the highest accuracy [5]. Li [6], in his study conducted in 2018 states that, in a comparative analysis among Logistic Regression, Support Vector Machine and Random Forest classifiers, Random Forest is found to have higher performance than the other two, especially when the dataset is huge. In recent times, Machine Learning algorithms are in popular demand for dealing with credit default. Also, the default made by a client may be due to many reasons. A study conducted by Kumar et al. [7] in 2018 suggests that, there are seven categories of variables that may impact customer credit default; there are relationship between consumer and creditors, annual income of consumer, debt-income ratio, consumer's occupation, home ownership, work duration and whether consumer possesses a saving/checking account or not. In [8], for the first-time bagging is explained as creation of n bootstrap replicates, where $n <$ sample space N whose majority vote determines the class label. This implies that bagging can yield good output with unstable procedure and even can diminish the efficiency of stable procedures.

This study builds a model for identifying potentially risky customers. The accuracy of the model is highly dependent on the features on which it is built. Hence in this paper we propose a methodology for extracting significant features and build the default prediction model using XGBoost classifier.

3 Proposed Method

To enhance the prediction accuracy, it is important to extract the right features from the dataset as curating the model with insignificant features can bring down the accuracy level. So, in this paper we suggest a methodology for building an accurate model for credit risk prediction by identifying the crucial features. The methodology

Fig. 1 The methodology adopted for model creation



comprises of pre-processing, feature selection and hyperparameter optimization (Fig. 1).

3.1 Dataset Description

The dataset chosen for the study is a Taiwan based dataset which contains 24 attributes which includes both the financial and demographic attributes of over 30,000 customers. This research used the binary variable (*default_payment*) as the response variable. The dataset contains demographic details like age, education, marital status etc. and financial attributes which includes the limit balance, repayment status and bill amount for six consecutive months. Train and Test data for the model is in 7:3 ratio. The publicly available dataset is obtained from UCI library [9].

3.2 Data Pre-processing

There were no missing values or null values in the dataset to be filled. However, the following pre-processing were done on the dataset: (i) created three new instances—*bill_amt_mean*, *pay_amt_mean* and *pay_pattern* which respectively shows the mean value of bill amount, payed amount and repayment status of six consecutive months. (ii) Some entries in the attribute *marriage* contained value *0* in the dataset, which we clubbed to *3* which represents the category *Others.*

3.3 Feature Selection

We used Recursive Feature Elimination to find out the significant features for model creation. This algorithm works by finding a subset of the attributes by eliminating the insignificant ones. The numbers of features can be restricted using the parameter *n_features_to_select.* The accuracy level is monitored for different values of this parameter and reached a maximum when *n_features_to_select* = *5.*

3.4 Hyperparameter Optimization

Setting the values of the hyperparameters to optimum is important to yield good results. The following hyperparameters are set to the values specified to achieve maximum results.

colsample_bylevel: 0.6, colsample_bytree: 0.7, gamma: 0.25, learning_rate: 0.1, max_depth: 10, min_child_weight: 5.0, n_estimators: 100, reg_lambda: 100.0, subsample: 1.0.

4 Findings and Discussion

The study and analysis are done on pre-processed database. For getting a more accurate model, the best parameters from the database is chosen for each model. Emil Richard Singh and Sivasankar [10] proposes a study on the same database where Information Gain is used to find the best features for model construction. This study initially applies the Recursive Feature Selection on the database to find the crucial and efficient attributes for the model. This is done by ranking the features based on their efficiency and importance while discarding those which are least significant. The process repeats until the desired number of attributes remain. The performance for each limit of feature is recorded and the best combination is chosen.

4.1 Performance Measure

Accuracy is a measure to identify the exactness of the prediction. It is evaluated by the proportion of correct number of predictions to the total number of instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$

Table 1 Accuracy (%) of models by different classifiers

Classifier	Accuracy (%)
Decision Tree	81.24
Random forest	81.7
XGBoosting	83.32

Table 2 Accuracy (%) of XGBoosting models

Exclusive of pre-processing and feature selection	80.81
Inclusive of pre-processing and feature selection	83.32

A comparison of the accuracies of models curated by Decision Tree, Random Forest and XGBoosting classifiers are given in Table 1 for a clearer picture. The accuracy level proves that application of pre-processing and feature selection elevates the accuracy of the models and Boosting techniques yield high-end models. Table 2 shows the difference in accuracies while applying XGBoosting classifier with and without feature selection. Figure 2 depicts the confusion matrix for the three algorithms.

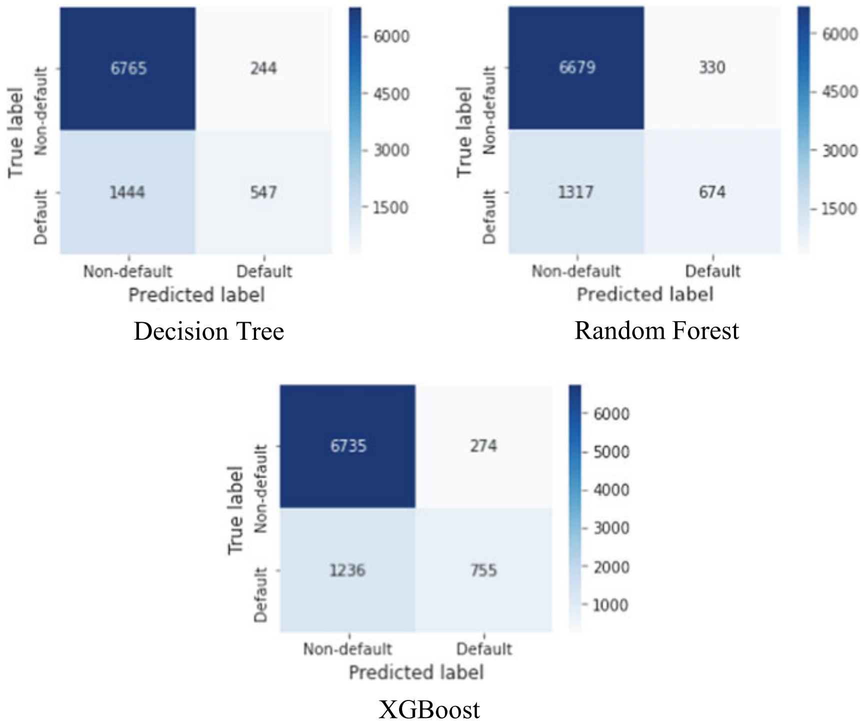


Fig. 2 Confusion matrix for decision tree, random forest and XGBoosting models

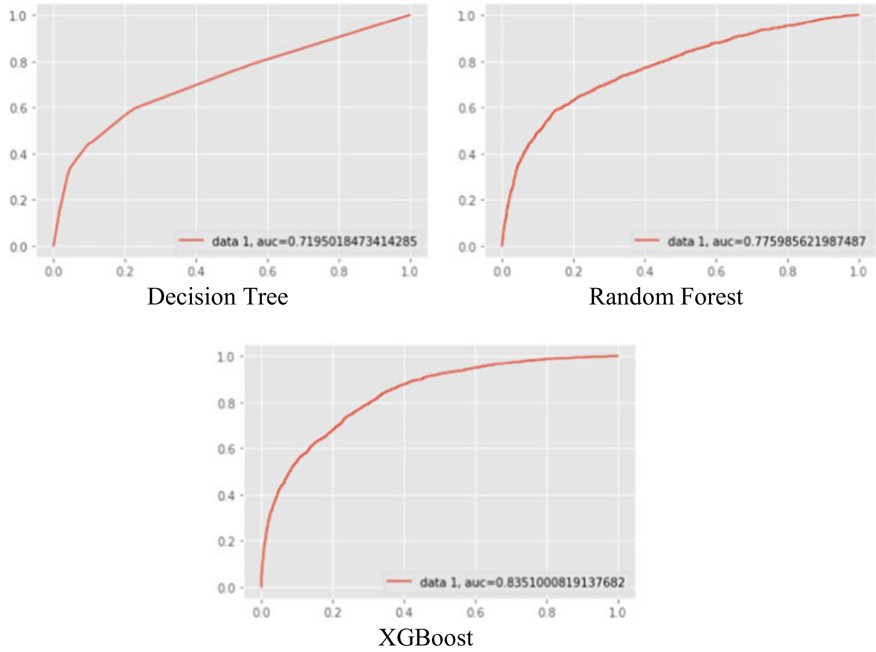


Fig. 3 Area under the curve plot for decision tree, random forest and XGBoosting models

Figure 2 depicts the AUC plot for Decision Tree, Random Forest and XGBoosting classifiers. The plot corresponding to XGBoost classifier occupies more area in comparison with rest of the models adhering to high efficiency (Fig. 3).

To analyse the curated model further, the feature importance plot for XGBoosting classifier is generated. Figure 4 implies the feature importance diagram of the classifier. The curated model recommends that the pay_pattern is the imperative factor determining default. This feature is created while pre-processing which is the average of repayment status of the customer. Along with this, the current limit balance of the customer is also a deciding factor. The demographic details are given less significance.

5 Conclusion

The study builds a model which outperforms the existing ones in predicting the default of credit card clients using the new methodology. The model relies more on the financial attributes of the client rather than the demographic details. Even though Random Forest and Decision Tree algorithms were found to have relatively good level of accuracy in classification, XGBoost classifier is found to have more performance when applied to a large dataset.

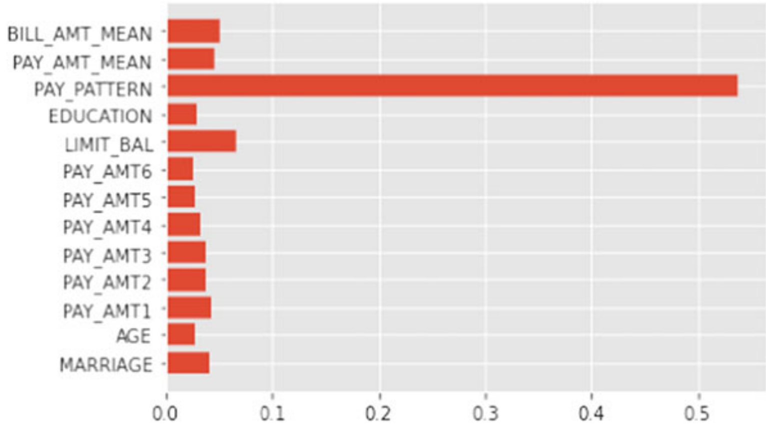


Fig. 4 Feature importance plot for the XGBoosting model curated

References

1. Hand DJ, Henley WE (1997) Statistical classification methods in consumer credit scoring: a review. J Royal Stat Soc Series A Stat Soc 160(3):523–541
2. Koh HC, Chan KL (2002) Data mining and customer relationship marketing in the banking industry. Singapore Manag Rev 24(2):1–27
3. Thomas LC (2000) A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers. Int J Forec 16:149–172
4. Baesens B, Setiono R, Mues C, Vanthienen J (2003) Using neural network rule extraction and decision tables for credit-risk evaluation. Manag Sci 49(3):312–329
5. Cheng YI, Lien C-H (2009) The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Exp Syst Appl 36:2473–2480
6. Li (2018) Research on bank credit default prediction based on data mining algorithm. Int J Soc Sci Human Invent 5(06):4820–4823. https://doi.org/10.18535/ijsshi/v5i6.09
7. Kumar M, Goel V, Jain T, Singhal S, Goel L (2018) Neural network approach to loan default prediction. Int Res J Eng Technol (IRJET) 5(4):4231–4234
8. Leo B (1996) Bagging predictors. Mach Learn 24(2):123–140
9. Credit default dataset. https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients
10. Emil Richard Singh B, Sivasankar E (2019) Enhancing prediction accuracy of default of credit using ensemble techniques. First international conference on artificial intelligence and cognitive computing, advances in intelligent systems and computing 815:427–436. https://doi.org/10.1007/978-981-13-1580-0_41

Deep Convolutional Neural Networks for Handwritten Kannada Numerals Recognition

G. Ramesh, W. Srihari, G. Srinidhi, and H. N. Champa

Abstract Kannada is an official language of India mostly spoken by people of Karnataka in Southwestern India. This paper describes the recognition/classification of Handwritten Kannada numerals using Deep Convolutional neural network(Deep-CNN). The occurrence of handwritten text is abundant. Convolutional neural networks (CNN) have been known for it's computationally efficient way of extracting features. For the extraction of features a series of convolution and pooling operations is performed. We have used the new Kannada MNIST dataset consisting of 60,000 samples of isolated handwritten numerals and a Handwritten dataset created by non-native users of the language called Dig-MNIST. The proposed system achieves an accuracy of 98.24% on the Kannada MNIST dataset and 86.85% on the Dig-MNIST dataset.

Keywords Computer vision · Convolutional neural network · Deep learning · Handwritten numerals recognition · Kannada numerals

1 Introduction

Deep learning has been widely used for the recognition of handwritten characters as well as numerals. For humans, the recognition of Handwritten digits is very easy but it is difficult for machines to recognize handwritten digits. Using Deep Learning and Machine Learning methods, this task can be achieved with the models having very high accuracy. Unlike traditional methods which involve different preprocessing steps, deep learning automatically identifies the features. Therefore, deep learning mostly depends on the data and hence can be applied to solve different kinds of problems, including handwritten characters and numerals recognition. There has been a huge progress in the field of Computer Vision with the application of deep learning in it. A Convolutional Neural Network (ConvNet/CNN) is one of the types of deep neural networks which is mainly used to analyze images. It takes images as input and extracts different features from the image. It contains one input layer, one output layer

G. Ramesh (✉) · W. Srihari · G. Srinidhi · H. N. Champa
UVCE, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_52

and one or many hidden layers. CNNs are very useful and feasible in dealing with patterns involving spatial arrangements, thus useful in recognition of handwritten characters and digits. One of the main advantages of a Convolutional Neural Network (CNN) when compared to the traditional machine learning techniques is that it consequently distinguishes the significant features in the image with no human oversight. The MNIST dataset is a standard dataset used worldwide consisting of 60,000 samples for training and 10,000 for testing. Kannada language is the official language of the state of Karnataka, India, spoken by over 50 million people all over the world. Similar to the MNIST dataset, the Kannada-MNIST dataset was introduced which contains the Kannada-MNIST dataset as well as Dig-MNIST dataset. The Dig-MNIST is a very challenging dataset when compared to the Kannada-MNIST as the dig-MNIST was created with the help of volunteers that were non-native users of the language. The aim of this paper is to provide a classifier that achieves a very high accuracy on the Kannada-MNIST as well as Dig-MNIST datasets without further preprocessing.

2 Related Work

The recognition of Offline Kannada Handwritten Numerals was implemented by Sharma et al. [1]. In this paper, the author has used a quadratic classifier based scheme for the recognition of offline handwritten Kannada numerals and achieved an accuracy of 90.34% with 2300 data samples. Here the author has given the idea of dimensional feature extraction using block segmentation.

Further down, the recognition of Handwritten Devanagiri Numerals using the SVM classifier was implemented by Jangid et al. [2]. In this paper, the author has used density and background directional distribution features for the zones, in which the numerals were already divided and have used SVM classifier with RBF kernel for classification.

Data Clustering is considered as an intuitive methodology for discovering similarities with regards to data and placing comparative information into gatherings. Here the author Mamatha et al. [3] has used a clustering algorithm for classification called the K-Means algorithm. The proposed method is experimented on 1,000 examples of Handwritten Kannada numerals and have achieved 96.00% accuracy.

A multi-layered neural network architecture was chosen for better accuracy by the authors Majumdar et al. [4]. For the feature selection, the authors use pixel and shape based features such as hole position, curve position, and block-wise corner position. The author has accomplished an accuracy of 97.20% on the dataset consisting of 10,500 images of Handwritten numerals.

Here the authors Gati et al. use the Skip CNN model for the classification of Kannada numerals [5]. The author without any additional preprocessing has accomplished an accuracy of 85.02% on the dig-MNIST and 97.53% on the Kannada-MNIST.

3 Problem Statement

This exertion is coordinated to settle the "Classification/Recognition of Handwritten Kannada numerals", by utilizing Convolutional neural networks, a dynamic procedure to yield dependable outcomes. The exertion is centred around the identification of these numerals from different wellsprings of writing. The work is likewise pointed towards breaking down the inadequacies in different conventional AI and OCR procedures that are right now existing as the answer for the difficult proclamation.

4 Network Architecture

The Convolutional neural network is used both as a feature extractor as well as a classifier. The ConvNet architecture used in this paper is a custom built network inspired by the AlexNet and VGGNet which are very deep network architectures and use several convolutional layers followed by one max pooling layer [6]. The size of the filter used in this architecture is 3 × 3, having a stride size of 1, and a max pooling layer of size 3 × 3. In Fig. 1, the architecture of the model is illustrated in detail. Zero Padding of size 1 × 1 is applied to the input of each convolution layer and after every two convolution layers the total number of filters for the convolution layers is increased by a factor of 2 going from 32 to 256. Dropout is applied after every two convolution layers to avoid over-fitting. It has a Fully Connected (FC) layer with 512 neurons and the activation function for every layer is the ReLU activation function except for the final output layer which has the softmax as its activation function.

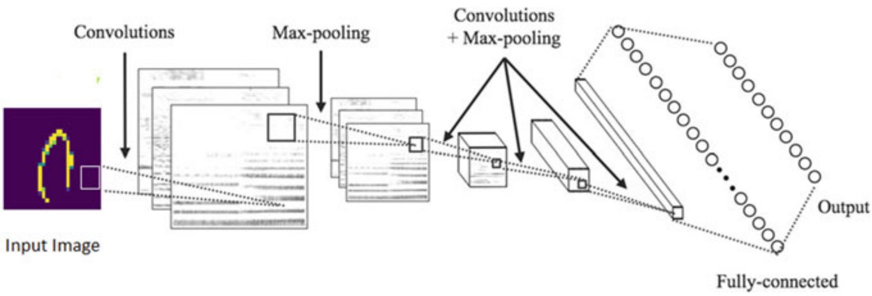


Fig. 1 Architecture diagram for the proposed model

5 Experimental Results and Discussion

5.1 Data Set

Our emphasis is precisely on the recognition of digits present in the Kannada-MNIST dataset as well as on the more challenging Dig-MNIST dataset. Both the datasets used in this paper are compiled by [7].

The main Kannada-MNIST dataset has a training set which comprises of 60,000 images converted to grayscale, each of the size 28 × 28 and a testing set which comprises of 10,000 images converted to grayscale with the same size and are divided equally among 10 classes.

The Dig-MNIST dataset contains 10,240 grayscale images of size 28 × 28 organized to give a more challenging dataset. This dataset was prepared with the help of people who were non-native users of the language and had trouble in writing these digits. Figure 2 represents the digits from 0 to 9 in Kannada.

5.2 Results

The Kannada MNIST dataset contains 60,000 images belonging to 10 classes for training and 10,000 images for testing. The Dig-MNIST dataset comprises of 10,240 images belonging to 10 classes over which the trained model is evaluated (Figs. 3 and 4). The grayscale images of size 28 × 28 was the input to our model as displayed in Fig. 1 and trained using the Adams optimizer having a learning rate(alpha) equal to 0.0001. The batch size of 64 was used for training. A dropout of 0.25 was added after every two convolutional layers and a dropout of 0.50 was added after the Fully Connected(FC) layer to avoid overfitting during training. The training is stopped after 50 epochs. In Fig. 5, the Accuracy versus Epochs graph of the model during training is shown and in Fig. 6, Loss Vs Epochs graph of the model during training is shown.

We accomplished an accuracy of 98.24% on the Kannada-MNIST dataset and an accuracy of 86.85% on the Dig-MNIST showing an improvement to [7] and [5]. The classification report is used to measure the quality of predictions from a classification algorithm. It shows the main classification metrics such as F1-Score, Precision and Recall. The classification report for the Kannada-MNIST is shown in Table 1 and the classification report for the Dig-MNIST is shown in Table 2.

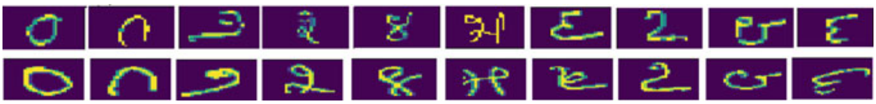


Fig. 2 Sample dataset: Kannada MNIST and dig MNIST respectively

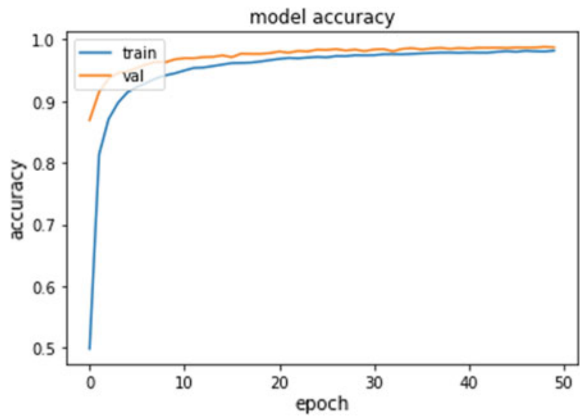
	0	1	2	3	4	5	6	7	8	9
0	0.0963	0.0032	0	0.0002	0	0	0	0	0.0003	0
1	0.0007	0.099	0	0.0002	0	0	0	0	0	0.0001
2	0.0003	0	0.0997	0	0	0	0	0	0	0
3	0.0003	0	0	0.0986	0	0.0001	0.0001	0.0009	0	0
4	0	0	0	0.0001	0.0995	0	0.0002	0.0001	0.0001	0
5	0	0	0.0002	0.0006	0.003	0.0956	0	0	0.0006	0
6	0	0	0	0	0	0	0.0978	0.0003	0.0001	0.0018
7	0.0002	0.0001	0.0003	0.0004	0.0001	0	0.0012	0.0973	0.0004	0
8	0	0	0	0	0	0.0001	0	0	0.0999	0
9	0	0	0	0	0.0001	0	0.0002	0	0.001	0.0987

Fig. 3 Normalized confusion-matrix : Kannada-MNIST

	0	1	2	3	4	5	6	7	8	9
0	0.084473	0.002051	0.002344	0	9.77E-05	0.000977	9.77E-05	0.001074	0.00146	0.00742
1	0.015625	0.077441	0.001758	0.00088	0	0.000488	0	0.002637	9.77E-05	0.00107
2	0.00166	0	0.094531	0.00068	9.77E-05	0.001074	0	0.001172	0.00049	0.00029
3	0.000293	0	0.001563	0.08857	0.000293	0.002734	0.000293	0.006152	9.77E-05	0
4	9.77E-05	0	0.002734	9.77E-05	0.084668	0.003711	0.001074	0.000488	0.00566	0.00146
5	0	0	9.77E-05	0.00029	0.000488	0.096582	0	0	0.00254	0
6	0.000391	0	0.001953	0.00088	0.000586	0.001074	0.075391	0.009766	0.00557	0.00439
7	0.000488	0.000195	0.001465	0.00029	0	0.000879	0.006543	0.087891	0.00186	0.00039
8	0.000293	9.77E-05	0.000586	0	0	0.002344	0.000879	0.000293	0.09365	0.00186
9	0	0.000195	0.000488	0	0	0.000195	0.01123	9.77E-05	0.00244	0.08535

Fig. 4 Normalized confusion-matrix: Dig-MNIST

Fig. 5 Accuracy versus epochs



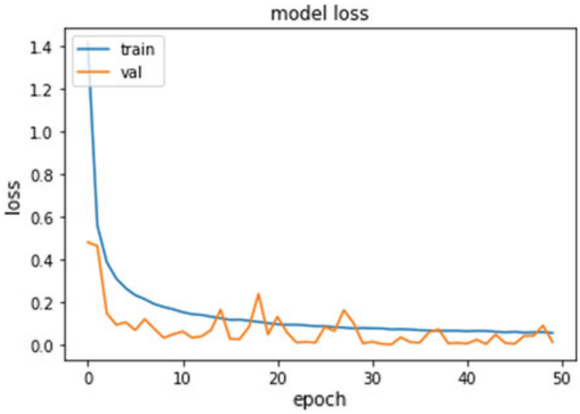


Fig. 6 Loss versus epochs

Table 1 Classification report of Kannada-MNIST dataset

Class	Precision	Recall	F1-Score
0	0.984663	0.963	0.973711
1	0.967742	0.99	0.978744
2	0.99501	0.997	0.996004
3	0.985015	0.986	0.985507
4	0.968841	0.995	0.981746
5	0.997912	0.956	0.976507
6	0.982915	0.978	0.980451
7	0.986815	0.973	0.979859
8	0.975586	0.999	0.987154
9	0.981113	0.987	0.984048
Accuracy	0.9824	0.9824	0.9824
Macro Avg	0.982561	0.9824	0.982373
Weighted Avg	0.982561	0.9824	0.982373

As we can see in the classification report(Dig-MNIST), the precision for class-5 (87.75%), class-6 (78.93%) and class-9 (83.47%) were higher when compared to [7] and [5]. The Normalized confusion matrix for the Dig-MNIST is shown in Fig. 3 and for the Kannada-MNIST is shown in Fig. 4. Table 3 shows the comparison of accuracy between the baseline accuracy, [5] and the proposed model when trained on the Kannada-MNIST dataset and evaluated on the Kannada-MNIST and Table 4 shows the comparison of accuracy between the baseline accuracy, [5] and the proposed model when evaluated on the Dig-MNIST and we have achieved better results in both the cases.

Table 2 Classification report of Dig-MNIST dataset

Class	Precision	Precision	F1-Score
0	0.81758	0.8447266	0.8309318
1	0.968254	0.7744141	0.8605534
2	0.879201	0.9453125	0.9110588
3	0.965921	0.8857422	0.9240958
4	0.98188	0.8466797	0.9092816
5	0.877551	0.9658203	0.9195723
6	0.789366	0.7539063	0.7712288
7	0.802139	0.8789063	0.8387698
8	0.82247	0.9365234	0.8757991
9	0.834766	0.8535156	0.8440367
Accuracy	0.868555	0.8685547	0.8685547
Macro Avg	0.873913	0.8685547	0.8685328
Weighted Avg	0.873913	0.8685547	0.8685328

Table 3 Comparison table: Kannada-MNIST

Author	Training set	Testing set	Accuracy (%)
Prabhu et al. [7] 2019	60,000	10,000	96.85
Gati et al. [5] 2019	60,000	10,000	97.53
Proposed method	60,000	10,000	98.24

Table 4 Comparison table: DIG-MNIST

Author	Testing set	Accuracy (%)
Prabhu et al. [7] 2019	10,240	76.17
Gati et al. [5] 2019	10,240	85.02
Proposed method	10,240	86.85

6 Conculsion

In this paper, Deep CNN was used to achieve a very high accuracy on the Kannada-MNIST dataset as well as the more difficult and challenging Dig-MNIST dataset. 98.24% on the Kannada MNIST dataset and 86.85% on the Dig-MNIST dataset is achieved without any additional preprocessing. This approach can be further used to achieve similar results on other languages which share the same characteristics of Kannada.

References

1. Sharma N, Pal U, Kimura F (2006) Recognition of handwritten Kannada numerals. In: 9th International conference on information technology (ICIT'06). IEEE
2. Jangid M, Dhir R, Rani R, Singh K (2011) SVM classifier for recognition of handwritten Devanagari numeral. In: 2011 International conference on image information processing. IEEE
3. Mamatha HR , Murthy KS, Veeksha AV, Vokuda PS, Lakshmi M (2011) Recognition of handwritten Kannada numerals using directional features and K-Means. In: 2011 international conference on computational intelligence and communication networks. IEEE
4. Majumdar A, Chaudhuri BB (2006) A MLP classifier for both printed and handwritten Bangla numeral recognition. In: Computer vision, graphics and image processing. Springer
5. Gati ES, Nimo BD, Asiamah EK (2019) Kannada-MNIST classification using skip CNN. In: 2019 16th international computer conference on wavelet active media technology and information processing. IEEE
6. Han Y, Kim J, Lee K (2016) Deep convolutional neural networks for predominant instrument recognition in polyphonic music. In: IEEE/ACM Trans Audio Speech Lang Process
7. Prabhu VU (2019) Kannada-MNIST: a new handwritten digits dataset for the Kannada language. In: arXiv preprint arXiv:1908.01242

Assessment of Phishing Email or URL Utilizing Machine Learning

Vidya Mhaske-Dhamdhere and Sandeep Vanjale

Abstract Phishing has become substantially bigger risk problem in online networking; and is usually performed by means of email mocking. In this paper we have focused mainly on the awareness on latest approach of real time phishing electronic mail type using system gaining knowledge of set of rules. We use machine leaning algorithm on spam base dataset random woodland set of rules work best which give true positive 97.2% and false negative is 0.88% and give efficiently class 94.82% and incorrectly type 5.17%.

Keywords Phishing email and URL · Cyber security · Privacy

1 Introduction

The present world is online computerized world. Each client's web for online correspondence and web based financial cycles for their everyday life for their own just as expert work. While utilizing on the web email or banking exchange administrations, clients may share their qualification or overlook security cautioning in view of occupied booked or ignorance of web safety efforts. These explanation causes phisher to trap to in phishing.

Phishing is one kind of cybercrime or phishing assault, which prompts budgetary misfortunes. There are number of methods to plan phishing assaults. The most mainstream thing is to make phishing site utilizing contains of genuine one sites, which seem as though an equivalent however their substance are supplanted by little character in space, sub domain, and number of specks, '@', '- ' like image substitution. It is exceptionally simple to supplant these substances.

V. Mhaske-Dhamdhere (✉) · S. Vanjale
Bharati Vidyapeeth Deemed to be University College of Engineering, Pune, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_53

2 Related Work

Simplicity in email communication has increased now days. Volume of Unsolicited bulk emails (UBE) [1] uses more internet and waste user time. UBE classification is done on basis of following 4 types as follows.

1. Document object model
2. Co-relation based feature selection
3. Wrapper based feature selection
4. Filter based feature Selection.

To evaluate the similarity of web pages layouts to identify phishing pages using learning based mechanism is used. Four Machine learning algorithms are used for extracting features of page layout. Learning based classification done on CSS features [1] to identify similar web pages. For experiment Phish Tank data set with 4000 data instances are used. Support vector Machine and Random Forest performed better accuracy of 95%.

Two level filter mechanisms [2] is used for detection of phishing sites using similarity based base approach, where first level filter is black list and second level filter is heuristic mechanism as second level filter is used. The feature extracted by five ways: feature based on attribute values, pathnames, filenames, plaintext, screenshot of data set Phish tank and Google search results.

Chanti et al. [3] machine learning is one of most important software detection approach, which play important role in implementing anti-phishing system. Author has considered phishing as classification problem, which can be solved using intelligent machine learning algorithm like neural network, dynamic self-structure. But software solutions are not providing permanent solution, to address this problem author has categories problem into two steps. First step is to develop and design anti-phishing techniques using users education on phishing websites and second step detection and identification of fake websites.

A stacking model is design for phishing feature detection of two parts that is URL and HTML. Stacking model works in multiple layers, which will show performance of system. Author has collected two real time databases 50K-PD with 49,947 database instances and 50K-IPD with 53,103 database instances classification on performance of parameter accuracy (97.20-20K-PD, 98.60-50K-IPD), missing rate (4.46%-20K-PD, 1.28%-50K-IPD), false rate (1.61%-20 K-PD, 1.28%-50 K-IPD). As this stacking model is using 2 types of features; URL related 8 features and HTML related 12 features are extracted. Total dataset is divided into training and testing dataset. Stacking model works in three layers. In first layer, original feature and predicted feature of training model will be combined and produced result. Result of combine feature of first layer will be given to second layer with stacking model. Lastly input feature of second layer and output feature will be combined and produced final result, which will be trained to predict final phishing Webpages.

3 Methodology

Heuristics is nothing but thumb rules through which URL is classified as weather it is Phishing or not. There are mainly 4 types of heuristics approaches are there. There are some Heuristics which is identifying after study some literature survey paper. They are as follows.

1. Short life of website
2. Replacement of top level domain
3. IP address as URL
4. URL Layout
5. Security Indicator of Browser
6. W3C recommendation and guidelines
7. URL link manipulation.

    i Link in email
    ii Link redirect
    iii Misspelled URL
    iv Number of Dots in URL
    v Presence of WWW other than first location in URL.

4 Results

Methodology used for designing email examples

According to phishing email categories, total emails and percentage is shown in below table. According to Fig. 1 only 49% legitimate emails are classified by participant which is quite less. According to Fig. 2, out of 149 participants 71 users click URL

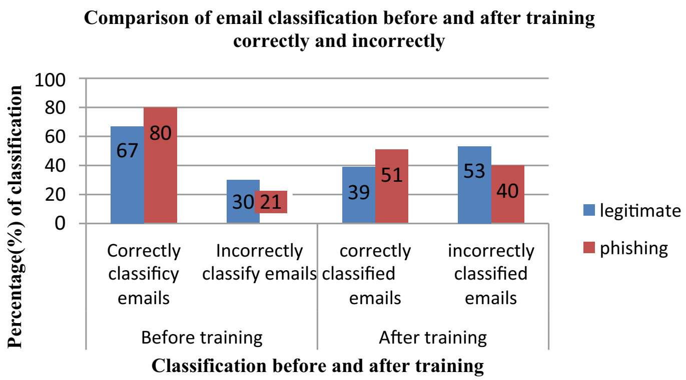


Fig. 1 Phishing email correctly and incorrectly classification before and after training

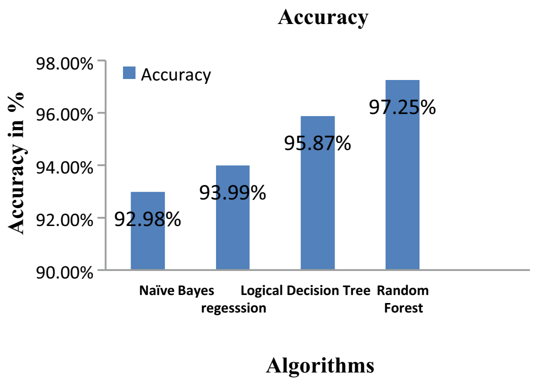


Fig. 2 Graphical representation of accuracy

once in a month and 53 users click URL once in a year so that we need to give training to these participants.

BCUD login notification, LIC policy benefit, ICICI credit card these email are repeated from before training as it is till ICICI bank credit card email is incorrectly classify major users. after asking users justification, users are saying that this email related to finance and it is very sensitive email.

University appointment work letter please click on email embedded link thats why 62 users are incorrectly classifies as phishing.

Overall comparison

See Table 1.

Finding of before and after training

1. Users required continuous training, they may be retained their knowledge.
2. Users ignore security warning messages.
3. 40% of uses are not able to identify phishing email after training.

To overcome this drawback, continuous training need to provide as well as this system required another approach which will classify real time email automatically using machine learning techniques (Table 2).

As we need to compare all algorithm accuracy which is shown in Fig. 4.4. Random forest has highest accuracy and Fig. 4.4 accuracy and Fig. 3 shows execution time of naïve Bayes is very less (Table 3).

Table 1 Before and after training email classification

Training mode	Before training		After training	
Classification of emails	Correctly classify emails	Incorrectly classify emails	Correctly classified emails	Incorrectly classified emails
Legitimate	67	30	39	53
Phishing	80	21	51	40

Table 2 Accuracy of algorithms

	Naïve Bayes	Logical regression	Decision tree	Random forest
Accuracy (%)	92.98	93.99	95.87	97.25

Fig. 3 Execution time in seconds

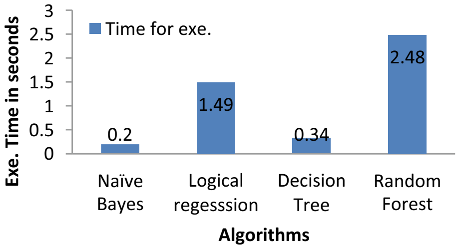


Table 3 Execution time of algorithms

	Naïve Bayes	Logical regression	Decision tree	Random forest
Time for exe. seconds	0.2	1.49	0.34	2.48

5 Conclusion

Phishing email is at present generally dangerous of patterns in net-work security threats. We have introduced a review of the security against these phishing email assaults. This review improves the understanding of the phishing messages issue, the current arrangement space, and the future degree to channel phishing messages. Approaches given in the writing despite everything have a lot of restriction on precision or performance, particularly with multi day phishing email assault.

Most classifiers used to distinguish phishing email depend on: supervised learning, for example they should learn before they can be utilized to detect another assault; unaided realizing, which is quicker, however has a low degree of exactness; or a cross breed (managed and solo) realizing, which is tedious and exorbitant.

References

1. Gangavarapu T, Jaidhar C (2020) Applicability of machine learning spam and phishing email filtering review and approaches. Artif Intel Rev
2. Mao J, Bian J (2019) Phishing page detection via learning classifiers from page layout feature EURASIP. J Wireless Commun Netw
3. Chanti S et al (2019) Classification of anti-phishing Solutions. SN Computer Science
4. Baadel S, Lu J (2019) Data analytics: intelligent anti-Phishing techniques based on machine learning. J Inf Knowl Manag 18(1)

5. Adebowale MA, Lwin KT, Sanchez E, Hossain MA (2019) Intelligent web-Phishing detection and protection scheme using integrated features of images, frames and text expert systems with applications
6. Mhaske Dhamdhere V, Vanjale SB (2020) Prediction and detection of Phishing emails by using multi-layer perceptron classifier. Int J Adv Sci Technol 29(9s):778–785
7. Dhamdhere V, Joeg P (2017) To study of Phishing attacks and user behavior. In: International conference on inventive computation technologies (ICICT 2017)
8. Dhamdhere V, Vanjale S (2019) A novel approach for phishing email real time classification using kmean algorithm. Int J Electr Comput Eng 2018:5326–5332
9. Dhamdhere V, Vanjale S (2019) PHISH safe guard-phishing detection: enhance anti-phishing system using machine learning algorithm. Int J Eng Adv Technol 1668–1671
10. Dhamdhere V, Vanjale S (2019) To enhance Phishing emails classification using machine learning algorithm. Int J Recent Technol Eng 2240–2242

Region Based Crop Prediction Using Artificial Neural Network

Taral Patel, Devkishan Patel, and Tanvi Patel

Abstract Agriculture is the backbone of economic growth of the country. We human beings totally depend on agriculture to satisfy all our daily food needs and sustain life on the earth. The crop rely on the different type of parameters like weather, soil moisture, sunlight, evaporation, temperature, cloud cover, humidity, and wind. There are three elements which are naturally available to the plant known as sunlight, air and water. For other elements plants entirely depend on soil minerals. Soil analysis is important to the farmer to know which type of crops can be cultivated in a particular soil characteristic condition. If a farmer knows which type of nutrients are in the soil like nitrogen (N), phosphorus (P), potassium (K). Soil pH is also very important for a crop. It will help the farmer, to which type of crop can be yield. So a farmer can improve the productivity of the next crop. Many researchers are focus on prediction of crop yield, weather condition prediction for crop, different soil suitable for crop and crop selection techniques. Machine learning algorithm is used to prediction most profitable crop for current location of soil and weather condition.

Keywords Machine learning · Prediction · Artificial neural network

T. Patel (✉) · D. Patel · T. Patel
Department of Computer Engineering, CGPIT, Uka Tarsadiya University, Bardoli, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_54

1 Introduction

India is an Agricultural country, as the traditional occupation of a country is Agriculture, the majority of the population is depending on it in one or other way. It is the major factor that decides the economic growth of our country. For the majority of farmers in rural area they are following the past tradition of their family. Follow the method of farming for crops. They continue to do their activities for living strategies. One generation passes their knowledge to another so, it helps to new generation which crop rotation help to improve their productivity and which crop is best for their particular land. Our agriculture sector face challenges like irrigation, fertilizers, less knowledge of new technology and not know which type of crop is suitable for his land. Now a day's government also focuses on Agriculture sector. Government facilitating farmers with many subsidies in fertilizers, tractor, cancellation of loan interest for loans etc.

Soil testing is an important part of nutrient management in agriculture. There are basically 18 components which define the fertility of the soil, in spite of them Nitrogen, phosphorus and potassium are major in soil nutrients. Analysis of the soil is very important because if the farmer knows the basic necessary components, it will provide that which type of fertility of his land has. As per the fertility it will helpful to the growth of the crop and also help to improve crop productivity [1].

2 Related Work

In [2], the WEKA (Waikato Environment for Knowledge Analysis) tool is used for Jrip algorithm implementation. Load data set in Weka preprocessing, generate the classification rule and compare the classification rule. Based on the nutrient rating of pH, EC, N, P and K parameter Jrip algorithm provide 99.71% accuracy. The result show that the soil is suitable for cultivation of many crops like paddy, ragi, maize, sugarcane, Green gram.

In [3], using the parameter of temperature, humidity, soil, pH level and using the Internet of thing device which take a real time data, then data will be store in agricultural cloud using the modified support vector regression algorithm and it will give the small plant predictions. In [4], using the parameter precipitation, minimum temperature, average temperature, maximum temperature and reference crop evapotranspiration, area, production and yield for rice crop prediction using different algorithm like Naïve Bayes, BayesNet and Multilayer Perception provide more accuracy, compared to SMO (Sequential Minimal Optimization) and SVM.

In naive Bayes classifier, the value of the dependent attribute is calculated by using the values of the independent attributes. Based on pH, N, P and K parameter naive Bayes provide 95.14% accuracy for soil is suitable for different crop or not [2]. By using the parameter climate data, soil it will provide rice crop prediction yield rate with good accuracy [4].

In agriculture J48 algorithm is used for analysis of soil condition based on pH value using the classification technique In [1], Used the soil parameter of pH, N, P, and K. Soil condition is classified as suitable soil or non-suitable soil based on the pH value of soil. The condition is if the pH value is greater than 8.5 then the soil is Non-suitable soil otherwise. It is suitable soil for crop cultivation. In [2], J48 (C4.5), Naïve Bayes, BayesNet classifiers, Jrip have been generated for the preprocessed soil dataset. Compare with other classifiers, the other classifiers gave the best result, but J48 provide 100% accuracy.

In [5], Prediction accuracy model for crop using the parameter of cloud cover, rainfall, minimum temperature, observation yield, predicting yield. Comparison of observation yield and prediction yield the check the prediction is accurate or not. It will give answer in prediction accurate yes otherwise no. Apply on different crop like soyabean, paddy, maize and wheat also check the average prediction accuracy. Crop with different accuracy like Soyabean—87% accuracy, paddy—85% accuracy, maize—76% accuracy wheat—80%. Used parameter climate data, soil it provides prediction crop yield rate.

In [6], Crop Selection Method (CSM) to achieve a net yield rate of crops over the season. The crop can be classified in four types: (a) the crops that can be planted during a season is known as seasonal crops. eg. Groundnut, cotton, wheat. (b) The crops that can be planted during the entire/whole year is known as whole year crops. Vegetable, Paddy, Toor, etc. (c) The crops that take short/less amount of time for the harvest is known as short time plantation crops. Potato, Vegetables. (d) These are the crops that takes a long time period for the harvesting is known as long time crops. Sugarcane, Canada.

As shown in Fig. 1, comparison of seasonal crops, whole year crops, short time plantation crop and longtime plantation crop. Short time plantation crop are provide more production compare to the long time crop and whole year crop because farmer can change the crop so fertility of soil is improved the it will provide more net yield rate. In this paper they create a crop sequence so it will help to farmer take the easily decision for cultivate next crop.

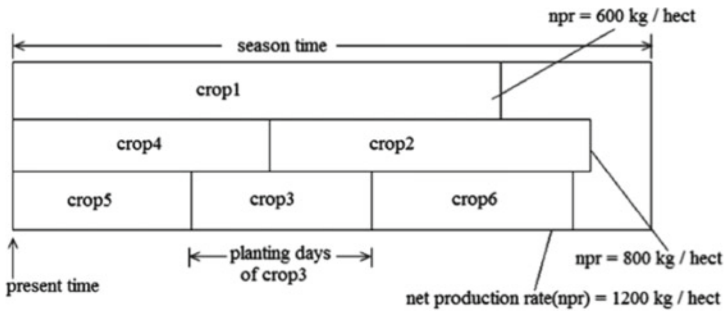


Fig. 1 Crop selection

In [7], using the parameters pH, phosphate, potassium, nitrogen, depth, temperature, rainfall. It was predicted, the single crop based on the parameter. In this paper, they also provide a fertilizer and single crop required nutrients. For easy to access they create an android application for farmers.

3 Proposed Work

The human brain is capable to make decision based on training. ANN is inspired by the human brain**.** ANN is effective in training the huge size of the dataset because of the parallel handling ability. In Ann three types of layer which are input layer, hidden layer, and the output layer [8, 9]. Feedforward Back Propagation is on the most popular algorithm in Artificial Neural Network. The neural process in the feedforward algorithm and back propagation network provide a training to the proposed neural. The input node carries the data related to the input variable, while the output nodes carries the data related to output variable. The hidden nodes which are depending on the specific problem so hidden node are easily extend based on the problem. The weights connections the data traverses between input layers to output layers. The propagation model weighted sum of inputs node calculated t is follows:

$$t_i = \sum_{j-1}^{n} W_{ij} X_j$$

Here

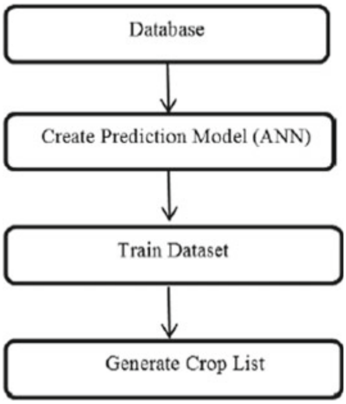
n—number of input node,

w—weight of the node i and j,

x—input node j.

As show in Fig. 2 first we collect a dataset of crop. After collecting of dataset we are going to apply prediction model of Artificial Neural Network (ANN) that is based on feedforword back propagation model. İn this model we train the data. After completed prediction model it train the database and given crop list. Which is based on the particular area of the location. It will provide suitable crop based on the parameter and give crop list to the farmer.

Fig. 2 Propose flow diagram



4 Result Analysis

```
THIS IS THE AREA: Shimla
 months 3
Bajra,171.0,14.55,6.5

 months 4
Banana,332.1,15.85,6.5

 months 4
Barley,332.1,15.85,6.5

 months 2
Bean,108.2,12.55,6.5

 months 2
Bottle Gourd,108.2,12.55,6.5

 months 3
Brinjal,171.0,14.55,6.5
```

All Simla Area crop

```
Cauliflower,Bottle Gourd,Potato,Barley,Guar seed,Bean,Groundnut
```

Final Result

Above result shows that the area of Simla region, it shows that different crop take different time for the harvesting of the crop. It also checks the average rainfall, average temperature and pH of the particular area and then it will provide a crop list to the farmers.

```
THIS IS THE AREA: Rajkot
 months 3
Bajra,6.5,29.77,6

 months 4
Banana,105.4,30.29,6

 months 4
Barley,105.4,30.29,6

 months 2
Bean,2.6,28.38,6

 months 2
Bottle Gourd,2.6,28.38,6
```

All Rajkot Area Crop

```
Drum Stick,Cauliflower,Bottle Gourd,Barley,Bean,Lentil,Orange
```

Final Result

Above result shows that the area of Rajkot region, it shows that different crop take different time for the harvesting of the crop. It also checks the average rainfall, average temperature and pH of the particular area and then it will provide a crop list to farmers.

5 Conclusions and Future Work

We can say that crop productivity is highly affected by soil and environmental condition. The combination of artificial neural network and agriculture will become a more interesting area of research in the near future. Artificial neural networks are one of the best solutions in search for a few agriculture problems, especially when it comes to predicting crop list.

The future work will be generate crop sequence based on soil nutrients and market price of the product. It will help to take the decision to a farmer which type of crop will be suitable for his land and which crop provide more profit.

References

1. Abhang K, Chaughule S, Chavan P, Ganjave S (2018) Soil analysis and crop fertility prediction. Int Res J Eng Technol (IRJET)
2. Hemageetha N, Nasira GM (2016) Analysis of soil condition based on pH value using classification techniques. IOSR J Comput Eng (IOSR-JCE) 18(6, Ver. III):50–54. e-ISSN:2278-0661, p-ISSN: 22788727
3. Radhika N (2018) Kind of crops and small plants prediction using IoT with machine learning. Int J Comput Math Sci (IJCMS)
4. Gandhi N, Armstrong LJ, Petkar O, Tripathy AK (2016) Rice crop yield prediction in India using support vector machines. In: 2016 13th international join conference on computer science and software engineering (JCSSE). IEEE, pp 1–5
5. Veenadhari S, Misra B, Singh CD (2014) Machine learning approach for forecasting crop yield based on climatic parameters. In: 2014 international conference on computer communication and informatics (ICCCI–2014), Coimbatore, India
6. Kumar R, Singh MP, Kumar P, Singh JP (2015) Crop selection method to maximize crop yield rate using machine learning technique. In: 2015 international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM)
7. Ravichandran G, Koteeshwari RS (2016) Agricultural crop predictor and advisor using ANN for smartphones. In: 2016 international conference on emerging trends in engineering, technology and science (ICETETS). IEEE, pp 1–6
8. Khairunniza-Bejo S, Mustaffha S, Ismail WIW (2014) Application of artificial neural network in predicting crop yield: a review. J Food Sci Eng 4(1)
9. Dahikar SS, Rode SV (2014) Agricultural crop yield prediction using artificial neural network approach. Int J Innov Res Electr Electron Instrum Control Eng 2(1):683–686

Smart Item Recommendation System for Offline Cloth Shop

Yash Ghanate, Aditya Mhaparle, Saurabh Yadav, and Snehal Mumbaikar

Abstract Recommendation System are widely designed for online e-commerce but the proposed recommendation system will be used by offline cloth shop. Offline shops can also use the recommendation system to increase their sales. Most of the shops recommend the same items to every customer irrespective of their interest either recommend items whenever some new stock arrives. Sometimes this may frustrate customer because customer may get some irrelevant recommendations or a customer may get recommendation very frequently. So the main aim of the proposed system is to control the product to be recommended and how frequently recommendation should be sent to the customer.

Keywords Recommender systems · Machine learning · Apriori algorithm · K means clustering · Hierarchical clustering · User-based collaborative filtering · Item-based collaborative filtering

1 Introduction

- The aim of a recommendation system is to produce meaningful recommendations for items or products which may be of interest to each customer.
- Recommendation system helps consumers find items of interest to them. Alternatively, they benefit the business by generating further sales. Recommenders programs are increasingly becoming a core tool in e-commerce on the Internet. Recommenders are constrained by the immense quantity of user data in current organizational databases and will be further stressed by the rising volume of user data available on the internet. New systems are required which can significantly improve the usability of recommender systems.
- Entire market is taken by online e-commerce website using recommendation system so recommendation systems should also be designed for offline shop to increase their sale. Each recommendation system follows a particular process to generate recommendations for the products. The approaches to recommendations

Y. Ghanate (✉) · A. Mhaparle · S. Yadav · S. Mumbaikar
Ramrao Adik Institute of Technology, Navi Mumbai, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_55

can be categorized according to the sources of information they use. You should recognize three potential sources of information as inputs for the recommendation process; the appropriate sources are the customer data (demographics), the product data (keywords, genres) and the customer-items (transaction data obtained).

2 Literature Survey

Various researchers implemented different recommendation systems which use User-User Collaborative, Item-Item Collaborative methods. There is various review based system as well as few RFM based recommendation system. The detailed study of these systems is given as follows.

Also some other recommendation systems for online as well as offline shops have been surveyed.

Following papers have been studied:-

- In Paper [1] authors Hyunwoo Hwangbo, Yang Sok Kim, Kyung Jin suggested a new approach for recommending products to customers by expanding the consumer to a collaborative filtering system to represent fashion product characteristics. By general, consumer preference for fashion items continues to fade over time. In the recommendation process, a decline function is implemented to represent this fact, which decreases the strength of interest over time. The item which the consumer wants to purchase is a product which substitutes the item preferred by the consumer beforehand.
- In Paper [2] authors Ruchika, Ajay Vikram Singh, Mayank Sharma show adaption of collaborative filtering in Apache Mahout Platform via Eclipse on a sample data set. Recommendation system described in this paper uses Item-based Collaborative Recommendation which finds the similar items and clusters them, the recommendation is provided from that clusters.
- In Paper [3] authors G. Krishna Kishore, D. Suresh Babu proposed a system which focuses to provide business advantage to retail industry. By moving the recommender technology to offline stores, a plethora of possibilities would be opened for entrepreneurs and retail stores. By serving their customers with personalized offers, they increase their customer loyalty and repeat customer transactions. This equips them to effectively combat the major online retailers like Flipkart, Snapdeal, and Amazon etc. Offline stores are an integral part of any economy and this work will help them to survive the battle with online retailers. K-prototypes algorithm (a variant of k-means algorithm to handle numerical as well as categorical attributes) is used to build recommender system.
- In Paper [4] authors Yingtong Dou, Hao Yang, Xiaolong Deng instituted the social recommender network research status broadly and collective filtering in particular. For the collaborative filtering, the paper shows the essential principles

and formulas of 2 basic approaches, the user-based cooperative filtering and item-based collaborative filtering. The paper contrasts the differences among cosine-based similarity, modified cosine-based similarity, and therefore mainly similarity based on the Pearson, for the measurement of consumer or object resemblance.

- In Paper [5] authors Alexander Tuzhilin, Gediminas Adomavicius argue that pertinent contextual data does matter in recommender systems and that it is important to take this information into account when providing recommendations. The paper introduces three different algorithmic paradigms contextual pre-filtering, post-filtering, and modeling for incorporating contextual data into the recommendation process.
- In Paper [6] authors Vasilis Aggelis, Dimitris Christodulakis give the scheme of how the RFM analysis can be used to cluster various customers. RFM analysis is a three-dimensional way to identify customers, or list them, to decide the top 20% or best customers. It is justified on the basis of 80/20 that 20% of customers produce 80% of sales. A consumer segmentation model recognized as the pyramid model is being used to classify consumers and conduct an analysis.

3 Traditional Recommendation System and Their Drawbacks

- In current existing recommendation system for offline shops, messages or sms regarding the offer is sent to almost each and every customers without considering whether the customer is actually interested in recommended item or not.
- Shopkeepers send “sms” or messages to customers regarding offer on any product, this sms is done in bulk and huge amount of customers receive this message. Whether the customer is actually interested in product or not this is not considered by shopkeepers and irrespective of that offer messages are sent to everyone.
- Most of the existing recommendation systems use rating of item for recommending the product to the user, but in offline shop ratings are not available.
- So the main motive of proposed recommendation system is to recommend only the interested product to the user. Even while sending some offer the users’ interest must be taken into account. The recommendation must be send to the user based on the pattern of its shopping.

4 Proposed Methodology

- This Item Recommendation System for Cloth Shop will be using combination of Collaborative Filtering (K means clustering and Hierarchical Clustering), Apriori Algorithm, RFM (Recency Frequency Monetory) analysis, Cosine Similarity. Design of the proposed system is shown in Fig. 1

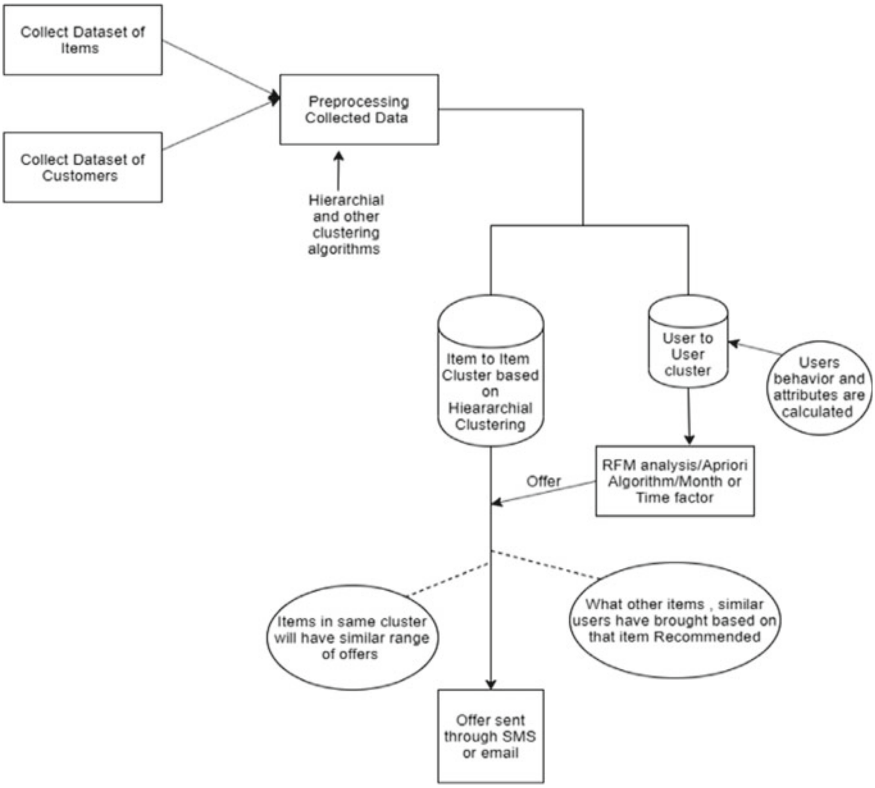


Fig. 1 System design

- The system consists of two different modules which are used together for recommendation. Similar items are grouped together using hierarchical clustering to form clusters of items. Similarity between different users is determined and based on that they are clustered together with a group of users most similar to them.
- Apriori algorithm is applied on the cluster of the customers for mining the most frequent itemset in the cluster. From this the item to be recommended is selected and then the item is searched in the clusters of the items and items similar to that are recommended to the customer.
- Whenever an offer is available on set of items, it is found whether the offered items are of customer's interest by comparing the items in the cluster of customers. If the item is present in the customer's cluster the offer is send to customer else that offer will not be send to that customer.
- RFM plays important role in recommending items to customers even if there is no offer on any item. RFM segments customers into different groups based on their RFM score. RFM helps in identifying the most valuable customers and how frequently recommendation should be send to each group of customers.

- Contextual behaviour of customers (time and amount) is considered to provide recommendation. There are two types of Collaborative Filtering, Item to Item and User to User Collaborative Filtering.
- In Item to Item Collaborative Filtering items with similar features are clustered together and in User to User Collaborative Filtering users with similar behavior are clustered together. Hence for Cloth recommendation system, collaborative filtering will be used to find similar set of items as well as similar set of customers.
- One of the notable advantages of CF is that, unlike content-based approaches that only produce similar products; they may consider profiles from other users to produce the required personalized recommendation. Some of the limitations in conventional collaborative filtering approach such as cold-starting were discussed in the proposed method.
- The main advantage of collaborative filtering recommendations is simply to implement, efficiency, and it provides a concise and justification for the computed predications. The key step of collaborative filtering approaches is to use the purchase history to classify the most appropriate user or object via the user-item purchase matrix.

4.1 K Means Clustering

K means clustering will be used to cluster similar users based on their behavioral attributes (Age, Frequency of visiting shop, Amount spent). Customers will also be clustered based on what items they have brought. K-means is applied on user-item purchase matrix. The proposed method of partitioning the entire dataset using K-means clustering decreases the time complexity and increases and data density.

4.2 Hierarchical Clustering

Hierarchical Clustering will be used to cluster similar items together (items here refer to cloths—t shirts, shirts, jeans, pants, sarees, dress etc.),attributes considered in Hierarchical clustering are gender (Male or Female), ethical wear or traditional wear, top wear, bottom wear, seasonal (winter and summer), top brands etc.

Men's clothes are jeans, shirts, t shirt etc. These are put in one cluster of Men. Further shirts can be clustered in Full Sleeves and Half Sleeves as these are types of shirts, similarly pants can also be divided into Formal pants, Chinos etc. Then again all these Men's clothes and Kids clothes will under one cluster of a brand. Hence in this way Hierarchical Clustering is used for Item to Item clustering.

4.3 Apriori Algorithm

Apriori algorithm is applied on clusters obtained through K-means clustering to find association rules for each cluster. It generates association rules for each customer based on their past transactions.

From association rules we get idea about what items customer may like. There can be various association rules suggesting various items but only top two most preferable items are recommended to the customers.

We aim to publish all proceedings papers in full-text xml. Our xml templates for La-TeX are based on CMR, our xml templates for Word are based on Times. We ask you to use the font according to the template used for your papers. Papers using other fonts will be converted by our typesetters.

4.4 Cosine Similarity

K-means is used to find cluster of similar customers. Since k means clustering is static, the total no. clusters is fixed at particular value (say 20 clusters), i.e. there are 20 clusters of customers. But if new customer with unique behavior (i.e. customer doesn't matches with the behavior of any existing cluster), then he should be assigned to a new cluster.

Hence no. of clusters should be now increased to 21, but this is not achieved by K-means. So the main aim of using cosine similarity is to change the number of clusters dynamically.

For calculating cosine similarity new transactions are compared with existing clusters and if the similarity is below the decided threshold then the no. of cluster is increased by 1 else no. of cluster remain same. This solves problem of Cold Start (issue of new customer).

$$similarity(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

Formula for calculating Cosine Similarity [7].

4.5 RFM Analysis

RFM stands for Recency, Frequency, Monetary amount—the three key elements in customer behavior that help to predict/identify customers who have higher response rates. Recency, frequency, monetary value is a marketing analysis tool used to identify a company's or an organization's best customers by using certain measures.

RFM analysis will help to analyze best customers of the cloth shop, which customers are most likely to respond quickly to recommendations, which customers are potentially losing out and how to retain their interest back. Ranking is given to each customer and based on this ranking recommendations are send to customers (i.e. if rank is 5 then after every 15 days recommendation are send to that customer and if rank is 4 then after every 30 days recommendations are send and so on).

5 Results

Aim of the proposed system is to send offer to only those customers who might be interested in the offer product. In traditional system offers are send in bulk irrespective of the customer interest. With the proposed system methodology offers can be sent to only interested and frequent user. As a experiment purpose one offer is selected and the result generated with traditional and proposed system is observed. It is found that offer on men product is send approximately to 500 customers in traditional but in our proposed system that particular offer is send to 150 customers over the 1000 customer's data. Figure 2 shows what product is recommended to the user when user visits shop by taking phone number of the user.

Figure 3 shows the Graph Analysis between Hybrid recommendation and Traditional recommendation system. Traditional Recommendation is sending offer message to all 500 customers, but hybrid recommendation will send only to those customers who might be interested in that item through recommendation algorithms. Hence sending bulk offer messages to customers irrespective of knowing their interest in item or not is avoided. Y axis shows different items and X axis shows no. of customers, for Item 1 Traditional Recommendation is send offer to 500 customers but Hybrid Recommendation will sent offer to only 100 customers.

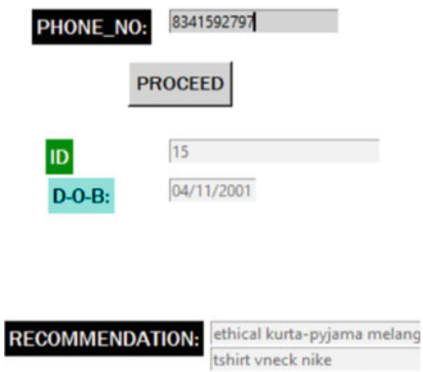


Fig. 2 Recommendation based on phone number

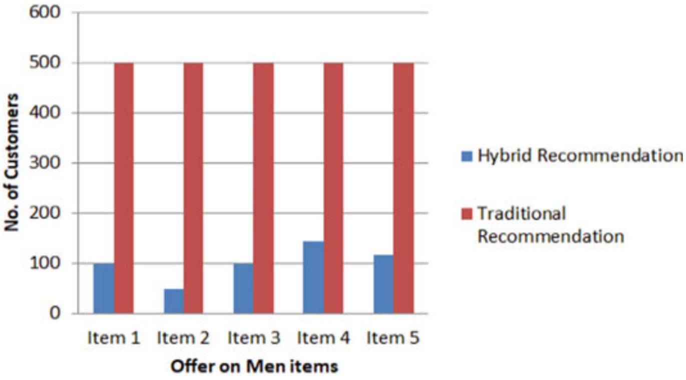


Fig. 3 Graph analysis between offers sent in traditional versus hybrid recommendation model for men items

6 Conclusion

Products are recommended to customer by retrieving the pattern how customer usually shops. Customers behavior is considered while recommendation. The system should be able to intelligently recommend items and give offers to customers using collaborative filtering. RFM helps to target the group of customers who can provide more benefit to the shop. The proposed system will have significant application in offline cloth shops. By recommending products of customers interest, customers interest towards that particular shop may increase and he/she can further share their good experience of shopping in that shop with their friends/family, thereby increasing net profit and sales of that cloth shop.

References

1. Hwangbo H, Kim YS, Jin K (2018) Recommendation system development for fashion retail e-commerce
2. Ruchika AVS, Sharma M (2017) Building an effective recommender system using machine learning based framework. In: International conference on infocom technologies and unmanned systems (trends and future directions) (ICTUS)
3. Krishna Kishore G, Suresh Babu D (2017) Recommender system based on customer behaviour for retail stores. www.iosrjournals.org
4. Dou Y, Yang H, Deng X (2016) A survey of collaborative filtering algorithms for social recommender systems. In: 12th international conference on semantics, knowledge and grids (SKG)
5. Tuzhilin A, Adomavicius G (2011) Context-aware recommender systems. Springer, Boston, MA
6. Aggelis V, Christodulakis D (2005) Customer clustering using RFM analysis. www.citeseerx.ist.psu.edu
7. Lahitani AR, Permanasari AE, Setiawan NA (2016) Cosine similarity to determine similarity measure: study case in online essay assessment

Breast Cancer Prediction Analysis Using Data Mining Techniques

Tanvi Patel, Devkishan Patel, and Taral Patel

Abstract All around the world cancer is a big issue of concern. In women, breast cancer represents the most prominent type of cancer. It develops when the development of cells in breast tissue get out of control. The repetitive uncontrolled division of cells construct up a mass of tissue called tumors. These tumors are classified into two types- benign tumors and malignant tumors. Data mining provides different classification strategies for the prediction of breast cancer. In the proposed work we have compared different classification such as an SVM, KNN, NB, RF, DT, ET and AdaBoost algorithms based on accuracy, execution time, kappa statistic error, mean absolute error, mean squared error, root mean squared error, true positive, true negative, false positive and false negative.

Keywords Data mining · Supervised learning algorithm · Classification · SVM · KNN · NB · RF · DT · ET · AdaBoost

1 Introduction

Data mining could be a method which is utilized to discover modern, covered up and valuable designs of information from huge database. It could be a portion of the information revelation process. It has widely used in many domains like healthcare, medical, banking, higher education, telecommunication etc. [1].

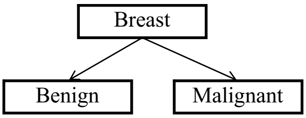
Classification may be directed machine learning procedure which allocates names or classes of distinctive objects or gather. It is a two-step process; the first step is constructing the model. This model is defined as the analysis of the training records of a database. The second step is utilized of show for classification the built demonstrate is utilized. A classification, strategy is utilized and makes a difference therapeutic specialist in their decision-making handle to discover out infections [2, 3].

Breast cancer is the over the world most common sorts of cancer among females. It may be a sort of cancer starting from the breast tissue, commonly from the inward lining of the drainage channels or labels providing the conduit with milk. Breast

T. Patel (✉) · D. Patel · T. Patel
Department of Computer Engineering, CGPIT, UkaTarsadiya University, Bardoli, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_56

Fig. 1 Type of breast cancer



cancer occurs in women as well as men, but in men, that has fewer chances to occur breast cancer [4].

Figure 1 shows the types of breast cancer. It is classified into two types:

1. Benign
2. Malignant.

1. **Benign Breast Cancer** [3–5]: This type of cancer does not spread to another area of the breast. The cancer cells most commonly develop in the milk ducts. These types of cancer can remove easily by using some treatment and there are fewer chances to grow back again and it is not harmful to women.
2. **Malignant Breast Cancer** [3–5]: This type of cancer is spread in another area in the breast. These types of cancer cannot easily remove and if it is removed, but it may be chances to grow back again and it is very harmful to women.

2 Related Works

A. **SVM** [6, 7]: SVM is a supervised machine learning algorithm. That can be used for classification and regression problems, but it is mostly used in classification problems. The SVM can be defined by a separating hyper plane. In that it has taken the input as a training data and the output is an optimal hyper plane which is categorized into a new example. In SVM if the data are unlabeled, then unsupervised machine learning approach is required.

B. **NB** [1, 8]: NB is a supervised machine learning classification algorithm. For classification issues which is based on Thomas Bayes's probability hypothesis. For content classification, Naïve Bayes is fundamentally utilized which includes high-dimensional training data sets. Spam filtration, opinion, analysis, and classifying news articles are a few cases of Naïve Bayes. It's calculated by the following conditions.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

P(A|B): Posterior Probability
P(B|A): Likelihood
P(A): Class Prior Probability
P(B): Prediction Prior Probability.

C. **KNN**: The KNN is the most basic classification algorithm in machine learning. It could be a Non-Parametric method which is utilized for classification and regression, but generally, the classification issue that is utilized in industry.

D. **DT**: Decision Tree is utilized for classification of nominal and numerical data. It is a tree like a chart or a show. It is an inverse tree since it has its root at the top and it grows to descend. The fundamental objective of making a decision tree to predict the value of the target attributes.

E. **RF**: RF is one of the fore most effective ensemble learning methods in patent recognition and machine learning techniques for high dimensional classification. This classification will handle lost values. It has also categorical values. In the medical field, this algorithm is used to recognize the proper combination of the components to validate the medicine. By utilizing patient's medical records this algorithm is supportive to recognize the disease.

F. **AdaBoost**: Adaboost is best utilized to boost the execution of decision trees on binary classification issues. More recently it may be alluded to as discrete AdaBoost since it is utilized for classification rather than regression. AdaBoost can be utilized to boost the execution of any machine learning algorithm. It is best utilized with frail learners. These are models that have finished accuracy, fair over random chance on a classification issue. The most suited and thus more common algorithms utilized with AdaBoost are decision trees with one level. Since these trees are so brief and only contain one decision for classification, they are frequently called decision stumps.

3 Proposed Work

Step 1: Breast Cancer Dataset

The data set is used from Wisconsin Hospital. It has total 569 instances where 62.74% is benign and 37.26% is malignant. It has been 32 patient attributes. That is include a patient ID (Fig. 2).

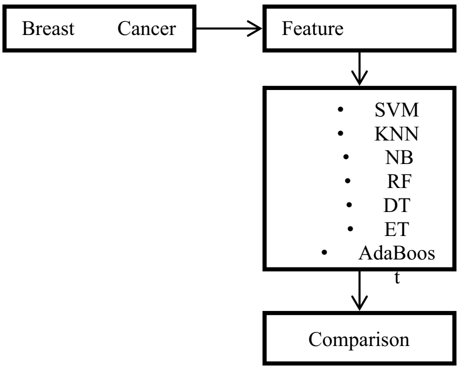


Fig. 2 Propose flow diagram

Step 2: Feature selection

In our proposed work we used uni-variate feature selection method for select the feature. The statistical test can be utilized to choose those features that have the most grounded relationships with the output variable. The scikit-learn library gives the Select K Best class that can be utilized with a suite of different statistical test to choose a particular number of features. We used 12 features for implementing.

Step 3: Classification

The different classification algorithms are used for classifying the data, such algorithms are SVM, KNN, NB, RF, DT, ET, and AdaBoost algorithms are applied on the dataset.

Extra Tree:

Extra tree algorithm is also known as Extremely Randomized Tree. It was proposed inside the primary objective of advance randomizing tree. The Extra-Tree strategy produces piece-wise multi direct approximations, instead of the piece-wise consistent ones of random forests [9].

Extra-Trees splitting algorithm (For numerical attributes).

- **Split_a_node(S)**
- **Input**: the local learning subset S corresponding to the node we want to split
- **Output**: a split [a < ac] or nothing
- **If Stop_split(*S*)** is **TRUE** then return nothing
  - Otherwise select **K** attributes $\{a_1, \ldots, a_k\}$ among all non constant (in S candidate attributes);
  - Draw *K* splits $\{s_1, \ldots, s_k\}$, where s_i = **Pick_a_random_split**(S, ai),∀ i = 1, …, K;
  - Return a split $s*$ such that Score($s*$,S) = $\max_{i=1,\ldots,k}$ Score(si, S).

- **Pick a random split(*S*, a)**
- **Inputs**: a subset*S* and an attribute a
- **Output**: a split
  - Let $a^S{}_{max}$ and $a^S{}_{min}$ denote the maximal and minimal value of a in *S*;
  - Draw a random cut-point a_c uniformly in [$a^S{}_{max}$, $a^S{}_{min}$];
  - Return the split [a < a_c].

- **Stopsplit(*S*)**
- **Input**: a subset *S*
- **Output**: a Boolean
  - If |*S*| <n_{min}, then return **TRUE**;
  - If all attributes are constant in *S*, then return **TRUE**;
  - If the output is constant in *S*, then return **TRUE**;
  - Otherwise, return **FALSE**.

The Extra-Trees part method for numerical attributes. It has two parameters: K the number of properties arbitrarily chosen at each node and nmin, the minimum test estimate for part a node.

Comparison:

All classifiers are compared to each other based on accuracy, execution time, KSE, MAE, RMSE, TP, TN, FP, FN.

4 Result Analysis

As described in the proposed work we will get the following results with the help of Extra Tree. An Extra tree algorithm was checked on the breast cancer data set and compare with another classifier.

Result Evaluation Parameters

Accuracy: Accuracy is calculated as the number of instances predicted positively divides by the total number of instances.

$$\text{Accuracy} = \frac{\text{Number of correct prediction}}{\text{Total of all cases to be predicated}}$$

Kappa Statistic Error (KSE): It is a metric that compares observed accuracy with an expected accuracy (Random change). It is used to evaluate single classifier amongst themselves.

$$KSE = \frac{P_o - P_e}{1 - P_e}$$

where

P_e is identical accuracy

P_0 the hypothetical probability of chance agreement.

Mean Absolute Error (MAE): It is a measure of different two contiguous variables [9, 10].

$$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

where

x_i is true value

y_i is prediction value.

Root Mean Squared Error (RMSE): It represents the sample standard derivation of the difference between prediction value and observed values [9].

$$RMSErrors = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2}{n}}$$

where

$Yi' - yi$ in yi is the observed value of the it's an observation.

Table 1 Comparison of all individual classifiers

Algorithms	Accuracy (%)	KSE	MAE	RMSE	TP	TN	FP	FN	Time (S)
SVM	62.28	0.00	3.27	61.42	0	71	0	43	0.63
KNN	95.61	90.45	4.39	20.94	38	71	0	5	0.15
NB	97.37	94.32	2.63	16.22	40	71	0	3	0.08
RF	95.61	90.53	4.39	20.94	39	70	1	4	0.45
DT	92.98	85.06	7.02	24.49	39	67	4	4	0.24
ET	98.25	94.37	2.63	16.22	41	70	1	2	0.27
Adaboost	97.37	94.37	2.62	16.22	41	70	1	2	3.98

Result of Proposed Work

Table 1 shows the comparison of different classification algorithms. For every different classification algorithms the data set is trained and tested. From our proposed work, first the train and test the dataset with the individual classification algorithms. Then after the feature selection method is used to select the best feature of a given data set. Then applied individual classifier algorithms and compared all of these algorithms and it will be observed that the extra tree gives the highest accuracy 98.25% compare to all other classification algorithms and give the best result

Figure 3 shows that the comparison of different classification algorithm such as SVM, KNN, NB, RF, DT, ET and Adaboost. Extra Tree classifier gives the highest accuracy compared to other algorithms.

Comparison

Table 2 shows the comparisons by the existing method and other previous work on different classification algorithms. It includes different instances and attributes which are compared with algorithms to calculate accuracy.

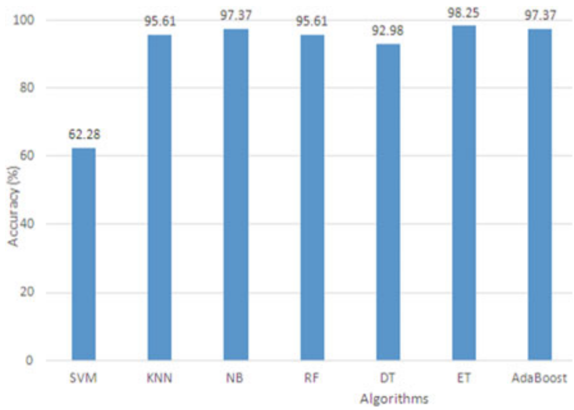


Fig. 3 Graphical representation of different classification algorithms

Table 2 Performance comparison by the existing method and proposed method

S. No.	Author/year	No instances	No attributes	Algorithms	Accuracy (%)
1	Delen (2005) [10]	699	11	ID3	92.99
				C4.5	95.57
				CART	92.42
				NB	97.42
2	Shajahaan (2013) [1]	699	11	J48	94.56
3	Senturk (2014) [6]	699	11	DT	97.89
				ANN	87.11
				Logistic Regression	93.78
				SVM	96.23
				NB	97.78
				KNN	93.78
4	Sumbaly (2004) [4]	569	11	J48	94.48
				SVM	96.70
				NB	96.13
5	Umesh (2015) [7]	683	11	NB	97.36
				RBF	96.77
				J48	93.41
6	Saranya (2016) [8]	699	9	NN	95.61
7	Kumar (2017) [11]	699	9	NB	94.19
8	Proposed work	569	12	**ET**	**98.25**

Source: Bold represents comparition of different algorithms

5 Conclusions and Future Work

Several data mining techniques are used for the classification of the breast cancer prediction. On these the supervise learning classification algorithms such as SVM, KNN, NB, RF, DT, ET and AdaBoost algorithms and compare to each other based on accuracy, execution time, kappa statistic error, mean absolute error, root mean squared error, true positive, true negative, false positive and false negative. Univariant selection method is used to select the features; we used 12 features for the implementation. After observing all classification algorithms the extra tree gives higher accuracy that is the 98.25% comparable to all other algorithms.

In future work, we will add more feature selection methods, different dataset and also work on voting ensemble techniques, it may give the best accuracy of breast cancer prediction.

References

1. Shajahaan, SS, Shanthi S, ManoChitra V (2013) Application of data mining techniques to model breast cancer data. Int J Emerg Technol Adv Eng 3(11):362–369
2. Shrivastava SS, Sant A, Aharwal RP (2013) An overview on data mining approach on breast cancer data. Int J Adv Comput Res 3(4):256–262
3. Venkatesan E, Velmurugan T (2015) Performance analysis of decision tree algorithms for breast cancer classification. Indian J Sci Technol 8(29):3–8
4. Sumbaly R, Vishnusri N, Jeyalatha S (2014) Diagnosis of breast cancer using decision tree data mining technique. Int J Comput Appl 98(10):16–24
5. Sivakami K, Saraswathi N (2015) Mining big data: breast cancer prediction using DT-SVM hybrid model. Int J Sci Eng Appl Sci 1(5):418–429
6. Senturk ZK, Kara R (2014) Breast cancer diagnosis via data mining: performance analysis of seven different algorithms. Comput Sci Eng
7. Umesh DR, Ramachandra B (2015) Association rule mining based predicting breast cancer recurrence on SEER breast cancer data. In: Emerging research in electronics, computer science and technology, international conference
8. Saranya P, Satheeskumar B (2016) A survey on feature selection of cancer disease using data mining techniques. Int J Comput Sci Mob Comput 5(5):713–719
9. Kumar V, Mishra BK, Mazzara M, Thanh D, Verma A (2019) Prediction of malignant and benign breast cancer: a data mining approach in healthcare applications, vol 4
10. Delen D, Walker G, Kadam A (2005) Predicting breast cancer survivability: a comparison of three data mining methods. Artif Intel Med
11. Kumar UK, Sai Nikhil MB, Sumangali K (2017) Prediction of breast cancer using voting classifier technique. In: Smart technologies and management for computing, communication, controls, energy and materials international conference

Unsupervised Extractive Text Summarization Using Distance-Based Clustering Algorithms

S. Divya, N. Sripriya, S. Mohanavalli, and S. Poornima

Abstract Due to existence of numerous data in huge data sources, organizing and managing those is essential to extract required information according to user preference. Summarization plays a vital role in providing meaningful and precise information from long term document(s). Embedding of sentences is performed using BERT model, which makes the system capable for further processing. The vectors those lie closer in the n-dimensional space are grouped together. The number of clusters to be generated is estimated using gap statistics method. A novel approach is proposed in the sentences ranking and selection, which decides the quality of the summary. The clustering of sentences is carried out using three different clustering algorithms with three different distance metrics. The comparative performance analysis of the system is done for manually created dataset and the results are recorded.

Keywords Automatic summarization · Distance based clustering · Distance estimation method · Summary

1 Introduction

According to the survey performed by the International Data Corporation (IDC), circulation of digital data around the world would grow annually from 4.4 zettabytes in 2013 to 180 zettabytes in 2025. Given long document(s) as input, Text Summarization technique aims at generating a well-organized and continuous summary of text, which comprises highly informative points present in input document(s). In Natural

S. Divya (✉) · N. Sripriya · S. Mohanavalli · S. Poornima
Department of Information Technology, SSN College of Engineering, Kelambakkam, Chennai, India

N. Sripriya
e-mail: sripriyan@ssn.edu.in

S. Mohanavalli
e-mail: mohanas@ssn.edu.in

S. Poornima
e-mail: poornimas@ssn.edu.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_57

Language Processing, there are two main approaches of automatic text summarization. (i) Extractive Summarization selects the informative phrases to generate the summary. (ii) Abstractive Summarization generates the condensed version of the input document(s). The terms in the document must be embedded by converting text into vectors of real numbers. These n-dimensional vectors can be grouped together using many clustering algorithms.

In this paper, an unsupervised text summarization is performed with various clustering algorithms and the results were compared. Sentences from the input document(s) were represented as vectors which are clustered using three different clustering algorithms by applying three different similarity metrics and the results are evaluated.

The organization of the paper follows the below sections. Section 2 describes the literature survey. Section 3 presents the BERT embedding method. Sections 4 and 5 describes various clustering algorithm and the various similarity metrics applied to cluster the input. Section 6 explains the proposed system and performance evaluation metrics of various clustering techniques with different distance calculation metric to select sentences. Section 7 presents the conclusion and future work.

2 Literature Survey

Text summarization is a well-known problem in literature right from 1950s. The first solution coined for this problem [1] by using frequently available words from the document and ranks the sentences. A powerful statistical technique named Latent Semantic Analysis (LSA) [2], which works based on two steps, (i) Construction of term-document matrix (ii) singular value decomposition.

A shallow embedding model named Word2Vec [3] that learns low-dimensional vectors from the beginning. Prediction is based on the context and performed using one of two models: CBOW and Skip-Gram. Continuous Bag of Words (CBOW) detects current words based on its context and Skip-Gram takes the target word as input and delivers its context as output. Global Vector coined as Glove [4] depends on the frequency of words and co-occurring with other words. Attention based [5–8] encoder decoder models were proposed to generate summary based on informative words or sentences from the document. The latest successful language modeling for NLP task is Bidirectional Encoder Representation from Transformers (BERT) [9]. Contextualized word embedding are presented using BERT which helps to get a better performance in processing.

Clustering of sentences in a high-dimensional space can be applied on text summarization [10–12]. These systems use Term Frequency-Inverse Document Frequency (TF-IDF), which is applicable only on multi-document summarization. Another method uses Latent Semantic Indexing [13], which detects sentences that represents the latent hypothesis in the document [14, 15]. A graph-based approach that models the document as a similarity graph, holds sentences as nodes [16] to lexical chains [17, 18].

Analysis of different clustering algorithms is stated in [19, 20]. The most common clustering technique is to estimate the distance between the centroid and data point and group the data points closer to the centroid. This distance estimation is performed using various metrics. Our focus is to perform various distance-based clustering algorithms in which, different distance measures are applied to find the similarity between text objects.

3 Bidirectional Encoder Representation from Transformers

BERT works with the base of transformer architecture and is pre-trained with a suitable dataset for an application. Transformer model performs sequence-to- sequence processing and it utilizes attention mechanism which is efficient as it could be used as encoder or decoder for language modelling.

BERT is comprised of two models: (i) BERT base—12 layers (transformer blocks), 12 attention heads and 110 million parameters. (ii) BERT large—24 layers, 16 attention heads and 340 million parameters. This deals with hundreds of millions or billions of parameters. This may lead to memory over utilization, which led to the motivation of developing techniques for parameter reduction. A Lite BERT for Self-Supervised Learning of Language Representations (ALBERT) uses three characteristics to reduce parameters. (i) Factorized embedding parameterization, (ii) Cross-layer parameter sharing and (iii) Inter-sentence coherence loss. ALBERT model performs position embedding which pads input on the right instead of left. While compared with BERT-base model which has 110 M parameters, ALBERT model has only 31 M parameters with the utilization of same number of layers and hidden states. The computational cost remains same as that of BERT architecture. In this paper, ALBERT-base-v2 model is used for embedding.

4 Clustering Algorithms

Text clustering is the process of grouping text under certain number of clusters. The focus of clustering is that the intra-cluster similarity must be maximum whereas the inter-cluster similarity must be minimum. Few clustering algorithms are explained.

4.1 K-Means Clustering Algorithm

This is an elementary procedure where the number of cluster k is determined in prior, which represents the number of centroids. Each data point is picked up iteratively and is associated to its nearest centroid. Once the first level of clustering is done, re-calculation of centroid value and re-binding of data point is iterated until the clusters are fixed at a point where there is no change in the centroid value or when the data point in each cluster do not move any more.

The goal of this algorithm is at minimizing an objective function known as squared error function which is stated as:

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\left\| x_i - v_j \right\| \right)^2 \tag{1}$$

where, $\left\| x_i - v_j \right\|$ is the Euclidean distance between two data points, x_i and v_j, c_i is the number of data points in *ith* cluster and c is the number of centroids.

4.2 K-means++ Algorithm

K-means++ varies from K-means only in the centroid initialization. First centroid is randomly selected and the distance between data point and centroid is calculated. Next centroid is selected with maximum probable distance. This is iterated until k-centroids are chosen. The other clustering steps remain the same as that of K-means Clustering.

4.3 GMM Algorithm

Naturally, modelling of huge dataset is performed by Gaussian distribution (Univariate or Multivariate). The Probability Density Function (PDF) of a Gaussian Distribution for one dimension is given by,

$$G(X|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2}$$

where, μ and σ^2 are mean and variance of the distribution. For n-dimensional Gaussian distribution, the PDF is given by,

$$G(X|\mu, \Sigma) = \frac{1}{\sqrt{2\pi}|\Sigma|} \exp^{\left(-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)\right)} \tag{3}$$

where μ is a n-dimensional vector denoting the mean of the distribution and $\sum$ is the $d \times d$ covariance matrix.

Assuming that there are k clusters, μ and $\sum$ is estimated for each k using Expectation–Maximization algorithm. This selects some random values and recursively tries to estimate the better data until the value gets optimized.

5 Similarity Estimation Metrics

In order to calculate the distance between the centroid and data points, following distance calculation metrics were applied.

Euclidean Distance: This is the straight-line distance between two points. Consider the two points $A(x_1, y_1)$ and $B(x_2, y_2)$. The Euclidean distance is calculated as,

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

This formula is obtained by simply applying Pythagoras theorem.

Manhattan Distance: This is the distance between two points measured along axes at right angles. The distance two points $A(x_1, y_1)$ and $B(x_2, y_2)$ is calculated as,

$$ManhattanDistance = |x_2 - x_1| + |y_2 - y_1| \tag{5}$$

Cosine Similarity: This is the measure of similarity of an inner product space among two vectors.

$$\cos(\theta) = \frac{A.B}{\|A\|\|B\|} \tag{6}$$

6 Proposed System and Performance Evaluation

Given a large text document, an automatic text summarization technique generates summary in an understandable sequence of sentences. The vector representation for the given input text is generated using ALBERT-base-v2 model. This model increases the training speed by splitting the embedding matrix into two matrices and by using repeated layers split among groups.

Clustering of similar feature vectors are performed using three different distance-based clustering algorithms. K-means, K-means++ and Gaussian Mixture Model are used for clustering the vectors, which requires in determining the number of clusters. The sequence in which the proposed system works is stated in the below diagram (Fig. 1).

The number of clusters is defined using Gap Statistics method. Once the number of clusters was fixed, the centroids were selected and iterated to optimize the centroid selection process. Based on the distance between the centroid and each data point, the data point is associated with the cluster which has minimum distance.

Once the clustering is done, desired number of sentences must be selected from each cluster to generate a summary of the input document. Ratio of the sentences to be selected from the input document is pre-defined. In this work, 50% of the

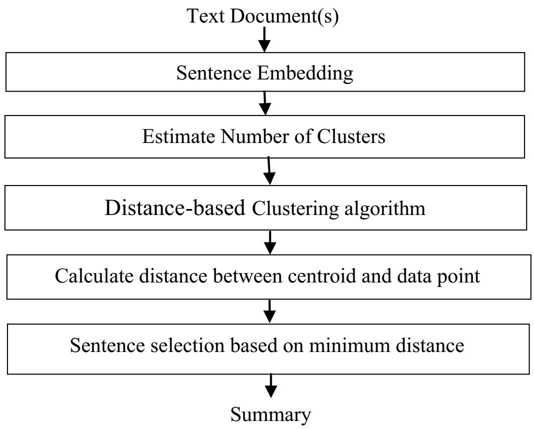


Fig. 1 Unsupervised extractive summarization

input document is estimated to be selected for the summary. Initially one sentence is selected from each cluster based on the minimum distance between the centroid and each data point.

Now, the ratio of the selected sentence is checked with the required ratio of the sentence. If the ratio is not satisfied, selection of sentence is repeated until the number of sentences selected is greater than or equal to the defined ratio. The ratio is adjusted in such a way that the number of sentences selected must be greater than the ratio and the same number of sentences is selected from each cluster. The performance is evaluated based on the following metrics.

Precision: It is computation of the number of sentences existing in both relevant summary and the retrieved summary by the number of retrieved sentences.

$$Precision = \frac{|\{relevant\ summary\} \cap \{retrieved\ summary\}|}{|\{retrieved\ summary\}|}$$

Recall: It is the fraction of the relevant summaries that are retrieved.

$$Recall = \frac{|\{relevant\ summary\} \cap \{retrieved\ summary\}|}{|\{relevant\ summary\}|}$$

F-Score: It is the average of Precision and Recall.

$$F - Score = \frac{2.P.R}{P + R}$$

The proposed system is processed with human generated dataset, which comprises of information collected from documents of four different contexts. The extractive summarization system works with three different clustering algorithm by varying three different distance calculation methods and the performance is observed and recorded in Table 1.

Table 1 Comparison of clustering algorithm with different distance calculation methods

Clustering algorithms	Precision	Recall	F-score
K-means clustering algorithm			
Euclidean distance	0.840	0.782	0.809
Manhattan distance	0.910	0.83	0.867
Cosine distance	0.820	0.733	0.773
K-Means++ clustering algorithm			
Euclidean distance	0.810	0.772	0.790
Manhattan distance	0.91	0.808	0.855
Cosine distance	0.850	0.713	0.775
Gaussian mixture model clustering algorithm			
Euclidean distance	0.94	0.86	0.897
Manhattan distance	0.91	0.89	0.899
Cosine distance	0.89	0.76	0.819

This concludes that, among three different distance-based clustering algorithms with variant in distance calculation method, Gaussian Mixture Model with Manhattan-distance calculation methods works comparatively better than the other two clustering algorithms.

7 Conclusion and Future Work

While comparing the summary performance of various clustering algorithms with various distance metrics, the summary generated by Gaussian Mixture Model provides informative and clear summary. Text embedding improves the understandability of input text, which helps in generating meaningful summary. As a future work, different hierarchical clustering algorithm planned to be implemented to generate the summary and performance to be evaluated.

References

1. Luhn HP (1958) The automatic creation of literature abstracts. IBM J Res Dev 2(2):159–165
2. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R (1990) Indexing by latent semantic analysis. Am Soc Inf Sci 41:391–407
3. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representation in vector space. In :Proceedings of workshop at ICLR, arXiv, pp 1301–3781
4. Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. EMNLP 14:1532–1543
5. Rush AM, Chopra S, Weston J (2015) A neural attention model for sentence summarization. EMNLP

6. Lopyrev K (2015) Generating news headlines with recurrent neural networks. arXiv:1512.01712
7. Nallapati R, Zhou B, dos Santos CN, Gulehre A, Xiang B (2016) Abstractive text summarization using sequenceto-sequence rnns and beyond. arXiv:1602.06023
8. Li J, Luong M-T, Jurafsky D (2015) A hierarchical neural autoencoder for paragraphs and documents. arXiv preprint arXiv:1506.01057
9. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN (2017) Attention is all you need. Adv Neural Inf Process Syst 5998–6008
10. McKeown K, Radev DR (1995) Generating summaries of multiple news articles. In: Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval, SIGIR ’95, New York, NY, USA. ACM, pp 74–82
11. McKeown KR, Klavans JL, Hatzivassiloglou V, Barzilay R, Eskin E (1999) Towards multi-document summarization by reformulation: progress and prospects. In: Proceedings of the sixteenth national conference on artificial intelligence and the eleventh innovative applications of artificial intelligence conference innovative applications of artificial intelligence, AAAI ’99/IAAI ’99, Menlo Park, CA, USA, 1999. American Association for Artificial Intelligence, pp 453–460
12. Bookstein A, Klein ST, Raita T (1995) Detecting content bearing words by serial clustering—extended abstract. In: Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval, SIGIR ’95, New York, NY, USA. ACM, pp 319–327
13. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R (1990) Indexing by latent semantic analysis. J Am Soc Inform Sci 41(6):391–407
14. Steinberger J, Jeek K (2004) Using latent semantic analysis in text summarization and summary evaluation. Proc ISIM 04:93–100
15. Wang M, Wang X, Xu C (2005) An approach to concept oriented text summarization. In: Proceedings of ISCIT05, ISCIT05, pp 1290–1293
16. Kruengkrai C, Jaruskulchai C (2003) Generic text summarization using local and global properties of sentences. In: Proceedings of the 2003 IEEE/WIC international conference on web intelligence, WI ’03, Washington, DC, USA, 2003. IEEE Computer Society, p 201
17. Barzilay R, Elhadad M (1999) Using lexical chains for text summarization. Adv Autom Text Summarization, pp 111–121
18. Silber HG, McCoy KF (2002) Efficiently computed lexical chains as an intermediate representation for automatic text summarization. Comput Linguist 28(4):487–496
19. Steinbach M, Karypis G, Kumar V (2000) A comparison of document clustering techniques, KDD workshop on text mining
20. Zhao Y, Karypis G (2004) Empirical and theoretical comparisons of selected criterion functions for document clustering. Mach Learn 55(3):311–331

Performance Evaluation of TCP Variants for IoT Built on Visible Light Communication

B. R. Vatsala and C. Vidya Raj

Abstract Internet of Things is the most prominent technology in the field of research. IoT protocols have to be carefully designed to meet the limited resource constraints of the system. One of the limitations in IoT environment is restriction in data storage, hence large data flows generated from IoT applications must be transferred quickly with suitable flow control and reliability. TCP is the preferred transport layer protocol for this. We build IoT on Visible Light Communication physical medium which is a promising technology with high bandwidth and evaluate performance of TCP variants with reference to goodput, packet delivery ratio and congestion window size in both error free and error prone scenarios.

Keywords TCP variants · Internet of things · Visible light communication · Flow control

1 Introduction

IoT applications such as smart cities, digital diagnosis, remote patient monitoring systems, digital supply chain, surveillance cameras, smart education systems etc., generate large data. Since IoT nodes have limited storage area these data have to be transferred to the destinations as quickly as possible there by raising a need for wireless network with large bandwidth. Visible Light Communication (VLC) is a potential solution for this with bandwidth in terms of Giga hertz [1]. It is harmless compared to radio waves and assures security.

Flow control is vital to large data flows in IoT network with VLC physical medium. TCP was designed to be a prominent flow control protocol with in-order packet delivery and reliability, but it was not popular in IoT paradigm as it requires significant

B. R. Vatsala (✉) · C. Vidya Raj
The National Institute of Engineering, Mysuru, India
e-mail: vatsalabr@nie.ac.in

C. Vidya Raj
e-mail: vidyarajc@nie.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_58

processing. IoT devices exhibit constraints regarding processing and power leading to usage of UDP at the transport layer.

Constrained Application Protocol (CoAP), lightweight RESTful applications are deployed using UDP since it is free from timeouts and retransmissions, however many network infrastructures that use middle boxes like firewall does not allow UDP packets to flow through them this motivated to employ TCP in application protocols like CoAP [2] in accordance with guidelines from IETF. IoT application layer protocols for instance, Message Queue Telemetry Transport (MQTT) [3] and Advanced Message Queuing Protocol (AMQP) [4] are deployed using TCP and are not finding any problem in IoT market place. HTTP which a prominent application layer protocol in internet is coming with version HTTP/2 for IoT context is also based on TCP. Moreover UDP is not suitable for QoS requirement like guaranteed delivery for applications such as critical health care and military Surveillance systems. Hence TCP can be a promising transport layer protocol for IoT with guaranteed delivery and flow control. This encouraged us to use TCP as transport layer protocol in our work.

Several variants of TCP such as NewReno, Vegas, Veno, Bic, Scalable, Hybla, HTCP, Highspeed, Illinois, Yeah, Westwood and Westwoodplus exist which are to be evaluated for deployment in IoT environment. Our work focuses on building an IoT infrastructure with VLC as physical medium and evaluating the performance of these TCP variants with reference to goodput, packet delivery ratio and congestion window size.

The paper is organized into 5 sections which focuses on: TCP-Variants in Sect. 2, building IoT over VLC in Sect. 3, performance evaluation of TCP variants in Sect. 4 and conclusion in Sect. 5.

2 TCP-Variants

2.1 TCP-NewReno

TCP NewReno [5] employs congestion control in two phases: Slow Start and Additive Increase Multiplicative Decrease (AIMD). At the beginning it enters Slow Start where the congestion window size (cwnd) is set to 1 MSS (Maximum Segment Size) and raises it by 1 MSS for every ACK received. It gets into AIMD phase when cwnd reaches ssthresh, a predefined threshold and increments cwnd by 1/cwnd for each ACK received, until congestion window reaches the receiver's advertized window, or 3 DUP ACKs are received. In the latter case it decrements the cwnd to half of its current value, it continues in the same phase until ACKs for all outstanding data segments are received. The major disadvantage of NewReno is that it always reduces window size on encountering packet loss predicting congestion and ignores losses due to network error there by underutilizing the available network bandwidth.

2.2 TCP-Vegas

TCP Vegas [6] estimates the network congestion level before it occurs and attempts to avoid it. It maintains two packet rate thresholds a and b. It computes the difference between expected flow and actual flow rate the sender. If the difference is less than a, it increments window size by one segment. If it is more than b, it decreases window size by one segment. Otherwise does not alter the window size. Vegas makes the effective utilization of the available network bandwidth.

2.3 TCP-Veno

TCP Veno [7] combines the featured of both Reno and Vegas. It keeps AIMD algorithm for congestion avoidance, and uses RTT measurements to estimate network congestion level. Upon encountering packet loss, it first decides whether the packet losses are due to congestion or error in network. If it is due to congestion, Veno halves the window size otherwise it reduces window size to $4/5^{th}$ of its current value.

2.4 TCP-Bic

TCP Bic [8] uses binary search technique and additive increase technique for congestion control. At the beginning it employs SS and maintains two variables min which is the current window size for which no packet loss occurs, and max which is the minimum window size for which packet loss occurs. It uses binary search technique where congestion window is updated with the midpoint value between min and max, thus congestion window increases logarithmically. This process is repeated until the difference between min and midpoint falls below a preset threshold, Smin. If the distance between the min and midpoint value is larger than preset threshold, Smax then additive increase technique is employed. Here current congestion window value is added with Smax. TCP Bic is more efficient in case of high speed networks as the congestion window grows more rapidly.

2.5 TCP-Westwood and TCP-Westwoodplus

TCP Westwood [9] uses Additive Increase Adaptive Decrease (AIAD) congestion control mechanism which is efficient for heterogeneous networks. During congestion it samples the current network bandwidth for every ACK it receives and estimates available bandwidth. It adjusts the congestion window based on estimation, instead of halving the window. TCP Westwoodplus [10] differs from westwood in sampling event. İt samples current network bandwidth for every RTT.

2.6 TCP-HTCP

HTCP [11] is a protocol designed especially for long distance networks with high speed. Here modification to TCP is made by adjusting the sender packet rate based on current condition of the network. It increases sender data rate as a function of the elapsed time since the last packet drop occurred at the source. It makes best use of the available bandwidth.

2.7 TCP-Highspeed

TCP Highspeed [12] is a protocol designed for network with large bandwidth delay product, it defines two functions a(w) and b(w) where w is congestion window size. Initially they are set to 1 and 0.5 respectively. During congestion avoidance phase it increases cwnd by a(w)/w when it receives ACK and decreases window size by (1-b(w))*w on observing packet losses due to duplicate acknowledgements. It performs well for large cwnd size.

2.8 TCP-Hybla

TCP Hybla [13] is optimized for long distance network (large RTT) such as satellite networks. It is designed to overcome the inefficiency of original TCP by mistaking packet loss due to network error and large RTT as congestion and thus maintaining small window size. It tries to have same data transmission rate for wireless network as that of referenced wired network.

2.9 TCP-Illinois

TCP-Illinois [14] is designed with emphasis on high-speed, long-distance networks. It uses packet loss as the primary signal for congestion along with queuing delay as secondary sign of congestion. During congestion avoidance phase it increases cwnd by α/w when it receives ACK and decreases window size by βw on observing packet losses. Where α and β are calculated based on queuing delay.

2.10 TCP-Scalable

TCP-Scalable [15] focuses on providing high throughput and scalability. It modifies the conventional congestion control algorithm by decreasing window size to 1/8th of the current window size on encountering packet losses and continues in the same phase until no packet losses occur. During recovery phase it increases window size by 1 packet for every 100 acknowledgements and thus making window size to grow faster.

2.11 TCP-Yeah

TCP Yeah [16] uses a mixed loss/delay approach to calculate congestion windows. Its main target is high efficiency, fairness and less RTT. It detects congestion by observing queue occupancy and decides congestion to be low when queue occupancy is low and congestion is high when queue occupancy is high. It operates in two modes fast and slow, when congestion is low it enters fast mode where it increments congestion window as in TCP Scalable and when congestion is high it enters slow mode where it behaves like Reno.

3 Building IoT Over VLC

IoT system is built over Visible Light Communication in NS3 by using IoT protocol stack and VLC module [17]. It consists of two source nodes (A, C) two sink nodes (B, D) connected via two gateway nodes (G1, G2) in dumbbell shape topology as shown in Fig. 1. 6LoWPAN [18], which is the network layer protocol of IoT is installed on all nodes with IPV6 addressing scheme.

The source nodes are connected to gateways through VLC channel. Transmitter and Receiver devices are attached to this channel. The two gateway nodes are

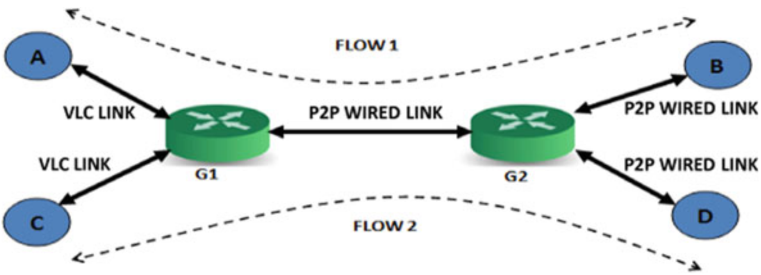


Fig. 1 IoT topology

Table 1 Network parameters

Parameters	Value
Bandwidth and delay of channel between source and gateway	10 Mbps, 0.01 ms
Bandwidth and delay of channel between gateways	2 Mbps, 45 ms
Bandwidth and delay of channel between sink and gateway	2 Mbps, 45 ms
Receiver error value (Rerror)	0.0 and 0.1
Packet size	1024 Bytes
Socket buffer size	4 MB

connected using point to point wired Link. The Sink nodes are also connected to gateways through point to point wired links. The network parameter values used are as shown in Table 1. BulkSend applications (which use TCP at the transport layer) namely flow1 and flow2 are run between end to end nodes A, B and C, D respectively. The channels are set with error model such that receiver error value (Rerror) is 0.0 in first case and 0.1 in second case.

4 Performance Evaluation of TCP Variants

The simulation is run for 100 s in both cases. Figure 2 shows the plot of goodput values for TCP- variants considered in our research work. It is clear that goodput decreases with increase in receiver error as more packets are lost during transmission

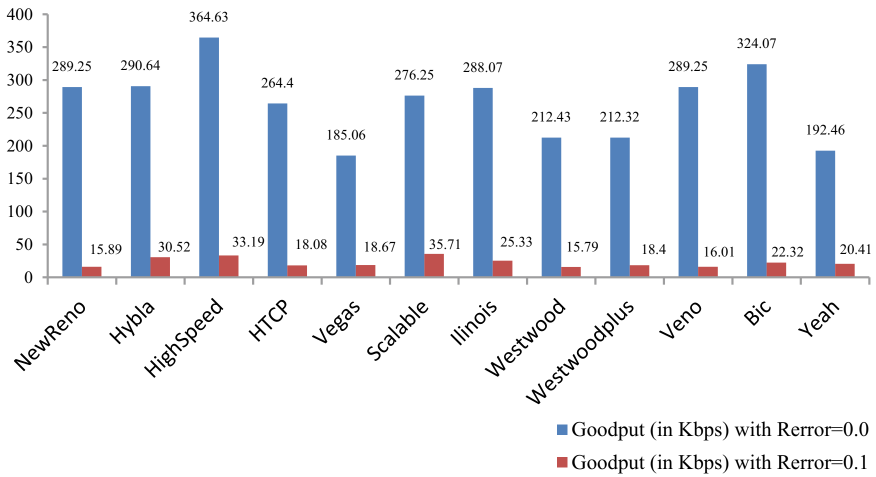


Fig. 2 Goodput of TCP variants in error free and error prone network

in second case. It is evident from the resulting graph that TCP-HighSpeed exhibits highest goodput in first case, since it is basically designed to increase throughput in high bandwidth network like VLC. Bic also exhibits fairly well as it utilizes window size effectively. All other TCP-variants except Vegas and Yeah have goodput with more than 200 kbps.

Scalable TCP has highest goodput in second case compared to other variants since it decreases window size to 1/8th of the original instead of halving when packet loss occurs. Also TCP-Highspeed performs well as discussed earlier for case 1. TCP Reno exhibits least goodput since it decreases window size to half on encountering packet loss and remains in the congestion control phase until acknowledgement packets for all outstanding segments are received.

Figure 3 depicts the plot of packet delivery ratio, TCP Vegas and Yeah exhibit 100% PDR in first case, Vegas carefully chooses window size based current sender data rate and expected data rate thereby avoiding packet drops also the goal of Yeah is high efficiency and chooses window size based on both loss and delay.

In second case HTCP makes best utilization of network bandwidth and adjusts window size to avoid packet losses and has highest packet delivery ratio.

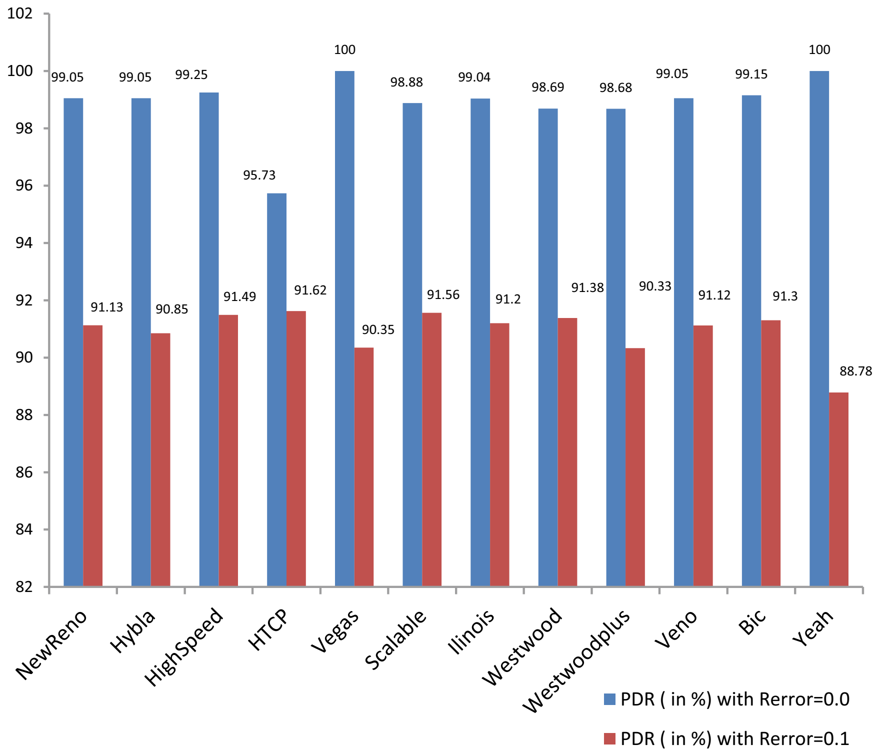


Fig. 3 Packet delivery ratio of TCP variants in error free and error prone network

Table 2 Average CWND for source node A

Tcp variants	Average CWND in bytes (Rerror = 0.0)	Average CWND in bytes (Rerror = 0.1)
Bic	176,460.35	1477.98
Highspeed	173,771.57	1695.55
Htcp	179,389.15	1285.64
Hybla	150,753.26	1878.16
Illinois	150,382.96	1873.53
New Reno	172,160.27	1086.51
Scalable	204,635.95	2459.71
Vegas	2118.53	1098.58
Veno	172,160.27	1141.28
Westwood	52,105.41	1092.72
Westwoodplus	52,132.52	1227.93
Yeah	26,270.34	1377.31

We also notice that 10% error prone network is not feasible for Yeah as it enters slow mode when packet loss occurs and decreases window size by half. All TCP variants achieve reasonably good packet delivery ratio as TCP retransmits the lost packets.

Table 2 represents the average cwnd size for source node A. It is obvious that cwnd size is large in case 1 due to receiver error 0.

TCP Scalable has highest average window size in both cases as it decreases window size to 1/8th on encountering packet loss and recovers window size faster. Vegas and Veno have least average window size in case 1 and case 2 respectively. Vegas decreases window size by 1 when sender rate is more than expected rate. Venos halves window on encountering packet loss due to congestion. Figures 4 and 5 represent the congestion window plot for source node A in case1 and case 2 respectively.

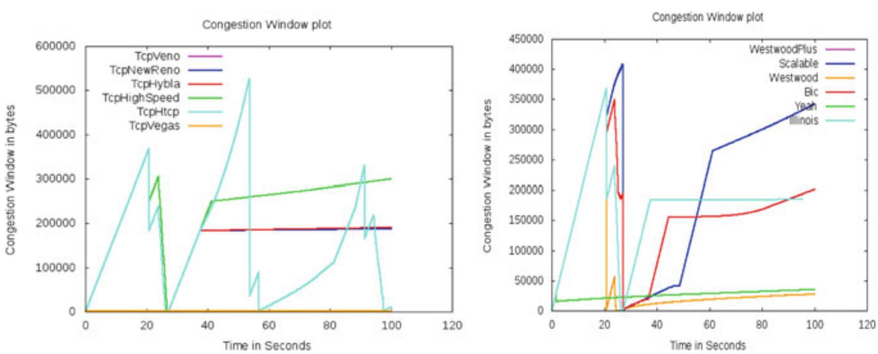


Fig. 4 Congestion window plot for source node A with error value 0.0

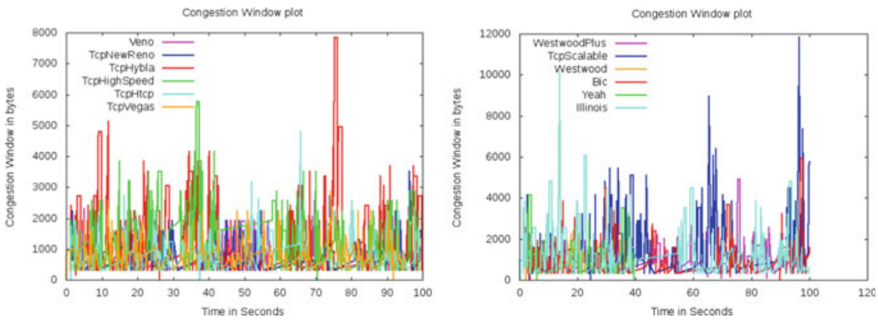


Fig. 5 Congestion window plot for source node A with error value 0.1

The main intention of building IoT system over VLC (which has high bandwidth and less network error) is to achieve high goodput and packet delivery ratio thus we can conclude that TCP variants Highspeed, Bic, Vegas, Yeah, HTCP and Scalable are more suitable.

5 Conclusion

IoT is built over Visible Light Communication medium to meet the large bandwidth requirement of some IoT applications. To cater the need of reliability and flow control, TCP is employed as the transport layer protocol in IoT network stack. TCP Variants are evaluated with reference to the performance parameters goodput, packet delivery ratio and cwnd size in both error free and 10% error prone cases. It was found that in error free scenario TCP high speed achieves highest goodput of 364.63 Kbps, TCP Vegas and Yeah achieve 100% packet delivery ratio and TCP Scalable exhibits maximum average cwnd of 204,635.95 Bytes. In 10% error prone condition TCP Scalabe turns out to be best with 35.71kbps goodput, HTCP achieves highest packet delivery ratio of 91.62% and TCP Scalable has maximum average cwnd size of 2459.71 Bytes. The performance evaluation TCP variants gives an idea about proper choice of these variants based on the requirement of the system.

Acknowledgements We are grateful to The National Institute of Engineering in completion of this work and the authors whose papers have been referred.

References

1. Haas H (2011) Wireless data from every light bulb. TED Global. Edinburgh, Scotland
2. Bormann C, Lemay S, Tschofenig H, Hartke K, Silverajan B, Raymor B (2016) CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets. IETF Internet Draft

3. Information technology—Message Queuing Telemetry Transport (MQTT) v3.1.1. ISO/IEC 20922:2016 (2016)
4. Information technology—Advanced Message Queuing Protocol (AMQP) v1.0 specification. ISO/IEC 19464 (2014)
5. Floyd S, Henderson T, Gurtov A (2004) The NewReno modification to TCP's fast recovery algorithm. Request for Comments: 3782, Network Working Group
6. Brakmo LS, Peterson LL (1995) TCP Vegas: end to end congestion avoidance on a global internet. IEEE J Sel Areas Commun 13(8):1465–1480
7. Fu CP, Liew SC (2003) TCP Veno: TCP enhancement for transmission over wireless access networks. IEEE J Sel Areas Commun 21(2):216–228
8. Xu L, Harfoush K, Rhee I (2004) Binary increase congestion control (BIC) for fast long-distance networks. In: Proceedings of the IEEE conference on computer communications—23rd annual joint conference of the IEEE computer and communications societies (INFOCOM '04), pp 2514–2524
9. Mascolo S, Casetti C (2001) TCP Westwood: bandwidth estimation for enhanced transport over wireless links. J ACM Mobicom
10. Grieco LA, Mascol S (2004) Performance evaluation and comparison of Westwood+, New Reno and Vegas TCP congestion control. ACM Computer Commun Rev 34(2)
11. Leith D, Shorten R (2004) H-TCP: TCP for high-speed and long-distance networks. Procceedings of PFLDnet
12. Floyd S (2003) RFC 3649—high speed TCP for large congestion window
13. Caini C, Firrincieli R (2004) TCP Hybla: a TCP enhancement for heterogeneous networks. Int J Satell Commun Network 22(5):547–566
14. Liu S, Basar T, Srikant R (2006) TCP-ıllinois: a loss and delay-based congestion control algorithm for high-speed networks. In: Proceeding of first ınternational conference on performance evaluation methodologies and tools
15. Kelly T (2003) Scalable TCP: improving performance in highspeed wide area networks. Computer Commun Rev 32(2)
16. Baiocchi A, Castellani AP, Vacirca F (2007) YeAH-TCP: yet another highspeed TCP. Proc PFLDnet
17. Aldalbahi A, Rahaim M, Khreishah A, Ayyash M, Little TDC (2017) Visible light communication module: an open source extension to the ns3 network simulator with real system validation. IEEE Access 5:22144–22158
18. Culler D, Chakrabarti S (2009) 6lowpan: incorporating IEEE 802.15. 4 into the IP architecture, IPSO Alliance.White Paper

AI Powered Smart Traffic Control System for Emergency Vehicles

Vedant Kumar, Siddhant Kumar, L. Sreekar, Pradhuman Singh, Pratik Pai, Shivani Nimbre, and Surendra Singh Rathod

Abstract Vehicular traffic is endlessly increasing everywhere within the world and may cause terrible traffic jams at intersections. Traffic congestion and tidal flow are major facts that cause delays to emergency vehicles. Fire brigade officials, ambulances, and police officers often get delayed due to such congestion and traffic. With the use of the right technology, like Artificial Intelligence and real-time monitoring of the traffic, such a predicament can be moderated thereby saving the lives of the needy. The solution presented allows emergency Vehicle (EV) drivers to select the route of commute on the mobile application. Modules on that route selected are activated for the adaptive traffic control system. The driver's GPS is dynamically updated on the cloud and fetched by the Raspberry Pi module. As the vehicle enters a given radius, the module checks for other EVs in the vicinity for priority assignment. The module starts traffic density detection and changes the traffic light states to clear the traffic if the density is above a certain threshold. After the EV comes closer to the signal, the light turns green irrespective of the traffic density. The camera cross validates if

V. Kumar · L. Sreekar (✉) · P. Singh · P. Pai · S. Nimbre
Department of Electronics and Telecommunication Engineering, Bharatiya Vidya Bhavans Sardar Patel Institute of Technology, Mumbai, India
e-mail: sreekar.l@spit.ac.in

V. Kumar
e-mail: vedant.kumar@spit.ac.in

P. Singh
e-mail: pradhuman.singh@spit.ac.in

P. Pai
e-mail: pratik.pai@spit.ac.in

S. Nimbre
e-mail: shivani.nimbre@spit.ac.in

S. Kumar · S. S. Rathod
Department of Electronics Engineering, Bharatiya Vidya Bhavans Sardar Patel Institute of Technology, Mumbai, India
e-mail: siddhant.kumar@spit.ac.in

S. S. Rathod
e-mail: surendra_rathod@spit.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_59

the EV has passed. Once it does, traffic signals switch to regular operation. This will allow the emergency vehicles to reach the destination on time and save the lives of those in need without delay. The entire solution is robust and reliable and can manage traffic efficiently keeping in mind the imperativeness of the emergency vehicles.

Keywords MobileNet · MobSE · Imagenet · Raspberry Pi · MQTT · SE blocks

1 Introduction

We propose a holistic approach to the Traffic Management and clearance system by incorporating traffic density detection using Artificial Intelligence. The system also includes a user-friendly mobile application which shall be used by the driver of the emergency vehicle in order to select the route of commute. The system does real-time GPS monitoring of the emergency vehicle with the help of GPS of the driver's mobile phone. This eliminates the need of installation of separate GPS modules or any kind of transmitter on the emergency vehicle itself. The android application will send an activation signal to all the modules that lie on the driver's selected route of commute. The modules will be interfaced with the pre-installed CCTV cameras on the traffic signals for traffic density detection. For further cross validation with the GPS monitoring, the system performs emergency vehicle detection thus improving the accuracy of the system. All emergency vehicles can be monitored and managed during operation efficiently as the GPS of the driver gets recorded in the central database through the Android application. The system incorporates a parallel control system model and toggles between normal and smart mode of operation. The modules will be interfaced with the current traffic signals in such a way that the modules will only be activated if the traffic signal lies in the route of commute of the emergency vehicle. This is called Smart mode of operation. The signals will be operated by the pre-existing system after the emergency vehicle leaves the traffic post.

2 Literature Survey

After extensive research from various sources, it has been discovered that the death toll and injury rate due to road accidents is on the rise. In India, the death toll per hour is as high as 15, and the injury rate per hour is as high as 53. According to the statistics, in India, this results in loss of life every 4 min. In this life- threatening situation, emergency vehicles play a very significant role. Because of traffic jams, more than 20% of the patients lose their lives in the ambulance itself. If the patient's condition is critical the chances of dying increase to many fold [1]. Comparative insights bolster the condition of fire-related crises, crime-related emergencies, and a lot more domains in India. According to the official report given by a part of the Indian Council of Medical Research (ICMR) viz. Management of Acute Coronary

Event (MACE) Registry in 2017, Around 50% of people who experience a heart attack in India arrive late at the hospital. According to the doctors, a patient should reach the hospital within 30 min. However, in reality, every second person arrives at the hospital roughly 400 min later after a heart attack has occurred. This example proves that the delay is caused by the increasing traffic jams on the Indian roads. The structure to lessen vehicular traffic has been proposed in a wide range of ways yet those are arrangements that would require years of implementation and moreover, they are not very efficient. The main aim of this project is to save lives. Hence our paper explores the effectiveness, feasibility and technical scope of an advanced and integrated traffic management system, based largely on minimizing response time solely to promote faster response times for priority vehicle services. The goal of this paper is to add value to the healthcare system, the traffic system, the research communities and other partners that are creating approaches to solve such problems.

In metropolitan cities, the rise in traffic jams is an inevitable issue. To tackle this problem, there is a need for a smart traffic management system that incorporates advanced techniques like computer vision [2, 3]. In this proposed solution, the movement of vehicles is detected using computer vision and video analytics [4, 5]. The solution is designed to monitor the traffic signals alongside the driving route of the emergency vehicle. The position of emergency vehicles is tracked with respect to the traffic signals lying in its path. The intervals of traffic signals are controlled on the basis of the following: (1) The proposed approach helps drivers of emergency vehicles to pick the commuting path on the smartphone device. Modules are enabled for the adaptive traffic control network on the designated road. GPS of the emergency vehicle is continuously updated on the cloud and fetched by a module. If the vehicle enters within the defined radius, other emergency vehicles within the vicinity are detected. After that, a traffic clearance algorithm is started if traffic density crosses a certain threshold. (2) A number of deep neural networks like YOLO [6], ResNet [7], MobileNet [8], VGGNet [9] etc. have been trained on large data sets like PASCAL VOC, COCO dataset, ImageNet and are used for object detection purposes. In an object detection task, the number of bounding boxes of a particular class is equal to its density. Furthermore, with transfer learning, a highly accurate object detection model can be presented for classes on which the model hasn't been trained previously [10, 11]. Transfer learning utilizes features extracted by the deep neural network along with tuning of the fully connected layers for accurate detections and predictions [12] (Fig. 1).

3 Methodology

The proposed solution depends on the vehicle density as an important metric in turning the signals green. If the density detected by the model deployed is less than the threshold, we wait for the EV to enter the inner zone in which the lights are turned green irrespective of the Density. We are using the Haversine formula which determines the distance between two points on a sphere given their respective

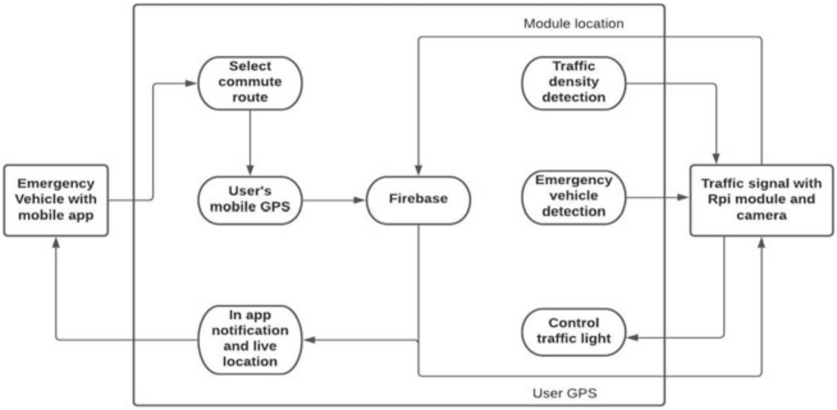


Fig. 1 Use case of the system

longitude and latitude [13].

$$d = 2r \sin^{-1}\left(\sqrt{\sin(\frac{\phi_2 - \phi_1}{2}) + \cos(\phi_1)\cos(\phi_2)\sin(\frac{\lambda_2 - \lambda_1}{2})}\right)$$

where,

ϕ_1, ϕ_2 are the latitude of point 1 and latitude of point 2 (in radians),
λ_1, λ_2 are the longitude of point 1 and longitude of point 2 (in radians).

3.1 Emergency Vehicle Detection

The Deep learning approach is used for detection of the vehicle density as well as the Emergency vehicle (EV). Detection of Emergency vehicle (EV) is to validate if it has crossed the approaching signal or not as GPS can be misleading at times. The model used for counting vehicle density and detecting emergency vehicle is a modification of MobileNet V1. In MobileNet, computationally expensive convolution layers are replaced by a convolution layer that performs depth wise separable convolution, which makes use of a 3 × 3 depth wise convolution layer followed by a 1 × 1 convolutional layer. This reduces the number of parameters and does a similar thing like that of a regular convolution layer. In the proposed solution, 3 × 3 dilated convolution with a dilation rate of 2 which has the receptive field of a 5 × 5 matrix has been used. It can deliver a broader field of view at the same computational cost. The Fig. 2 shows the architecture of a single block of the proposed model which is similar to the MobileNet architecture.

MobileNet V1 consists of 13 of these blocks in a row. The proposed solution has introduced SE blocks between the blocks of the MobileNet (after blocks 1, 3, and 5).

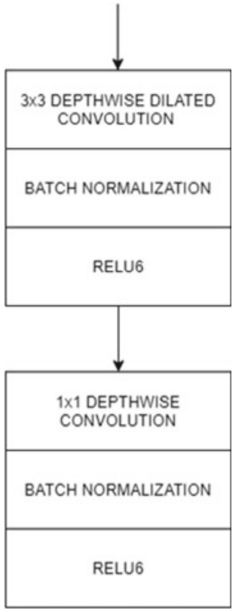


Fig. 2 Architecture of a block of MobSE

The name given to the model is MobSE, as the architecture of the presented model is similar to MobileNet and makes use of SE blocks between the base layer blocks. These SE blocks used in squeeze and excitation networks improve the inter dependencies of channels and is a form of self-attention on channels. The computational complexity of SE blocks is almost negligible. The main use of SE is to selectively enhance useful features and suppress less useful ones. In transfer learning, SE can be thought of as giving importance to the class one needs as output over others. SE blocks can also be used to improve predictions in the object detection task.

SE blocks recalibrates the feature maps by making the use of a 'squeeze' operation which is done by global average pooling, followed by a dense layer. ReLU adds non-linearity and output channel complexity is reduced by a ratio 'r'. At last, 'excitation' operation is performed that makes use of two fully connected layers and the scale operation reweights the feature maps. Similarly, the fire module used in the squeeze net used 1×1 convolution blocks to squeeze the feature maps and 1×1 convolution followed by 3×3 convolution was used in the end. The addition of the fire module however increased the parameters involved substantially which is not the case in SE block. The hyperparameter involved in SE block is reduction ratio 'r'. With a bigger value of r, the intermediate representation is smaller. Reduction in size of the representation to C/r and then expanding its backup to C is to put a limit on the model complexity and provide generalization. SE block implementation can be used in any CNN architecture and require minimal computational overhead. The Fig. 3 shows the addition of SE block between the base block of the MobSE model.

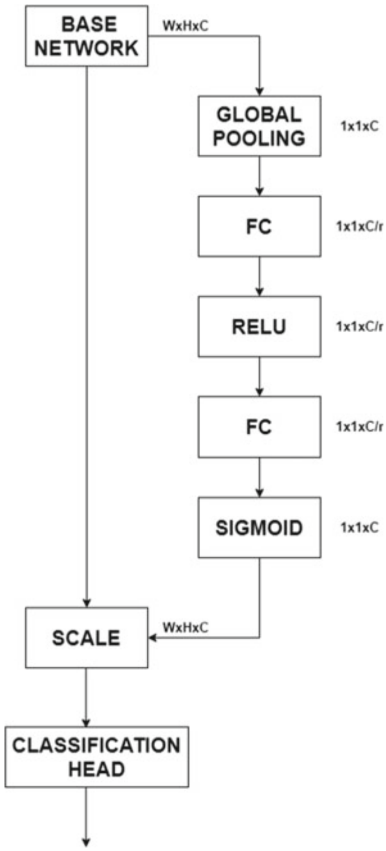
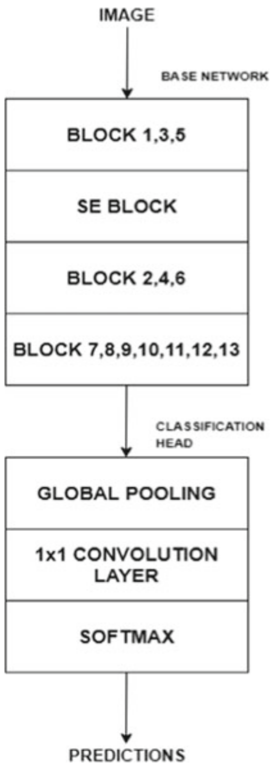


Fig. 3 SE block implementation in MobSE

For Vehicle detection, the model was trained on ImageNet dataset and the evaluation metrics were taken as Top-1 accuracy and Top-5 accuracy. The accuracy values were compared with MobileNet V1 paper.

For emergency vehicle detection, transfer learning technique was used to train the model in which the old network head was replaced by a set of fully connected heads which was placed after the SE block. During backpropagation, head layer weights were updated whereas base model weights were unchanged. The model was trained on images of ambulances that were scraped from the internet and the evaluation metrics considered were Precision and Recall. This model was compared with the MobileNet V1 model so as to evaluate the performance of the model. The Fig. 4 shows the complete architecture of the proposed model MobSE.

Fig. 4 Architecture of the MobSE model



3.2 *Evaluation Metrics*

$$\text{Precision} : \frac{\text{True Positive}}{\text{True positive} + \text{False positive}}$$

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

Top-1 accuracy: The percentage of time that the classifier gave the correct class the highest probability score.

Top-5 accuracy: The percentage of test examples for which the classifier involved the correct class among the top 5 probabilities.

Top-5 error rate is the percentage of test examples for which the correct class was not in the top 5 predicted classes

Top-1 error rate is the percentage of test examples for which the correct class was not given the highest probability score.

3.3 Mobile Application

In order to provide a complete holistic system, and to loop in our end user, our system leverages the usage of an android mobile application. The major functionality of the mobile application is to act as a conduit of interface between the user and the system. Using the mobile application, the system constantly fetches and gathers data pertaining to the location of the user in order to cross verify whether the vehicle has entered the vicinity of the traffic signal. Apart from that the mobile application provides a route from the source to the destination, along with all the intermediate traffic signals, where the hardware module is installed.

Flow of events for mobile application:

As the user opens the mobile application, the application requests permission to access the location of the mobile phone. This information is stored in a database (Firebase) and is constantly updated as and when the mobile's location changes. Once the user grants access to this, he is required to enter the destination. Once the destination is entered the mobile application finds the route from the source to the destination with all the intermediary traffic signals where the hardware module is installed.

As soon as the user enters the destination, the application has 2 sets of coordinates: the source and the destination. A list of coordinates, at specific intervals, along the route from the source to destination are obtained. The polyline function, used to display a route, is used to connect these coordinates and form a complete route. The coordinates of the hardware module stationed at the traffic signals is fetched from the database and is stored as a list. These coordinates encompassed in the list are then translated to the map and are displayed as markers.

The Fig. 5 depicts the operations involved when an emergence vehicle arrives at an intersection.

3.4 Hardware

The communication of the software application to the module and module to signal control happens via MQTT protocol. MQTT is a publish/subscribe communication protocol that utilizes TCP/IP sockets or WebSockets. The software application will analyze the route selected and it will send an activation notification for all the modules that need to be activated lying on that route. The activation notification is published via MQTT protocol (Here, the software application acts as a publisher node). The Raspberry Pi module (s) acts as a subscriber node and it receives the activation notification. Once the required modules are activated, the modules will then start video capturing and processing. The frames are then published to the cloud (publisher node = Raspberry Pi, subscriber node = cloud) with a fixed frames per second (FPS) rate.

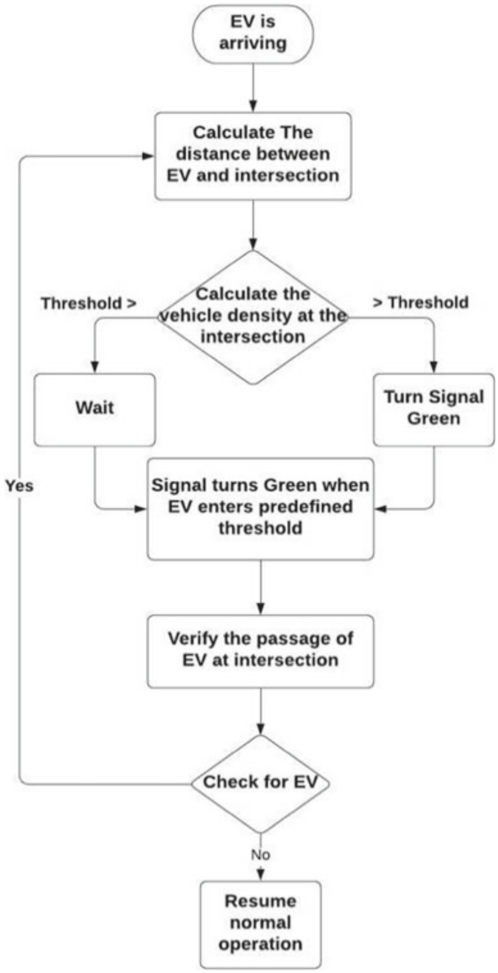


Fig. 5 Algorithm flowchart

Machine Learning and Image processing algorithms will be then applied in the cloud to analyze the frames and then the traffic signal timings will be altered accordingly. At this stage, the cloud acts as a publisher node and the Raspberry Pi module which alters the signal timings by controlling the relay acts as a subscriber node. The Fig. 6 shows the connection diagram used for the prototype of the solution along with Table 1 that contains the components along with their functionalities. The Fig. 7 shows the circuit diagram of the modules that need to be interfaced for our solution to be implemented at the traffic signals.

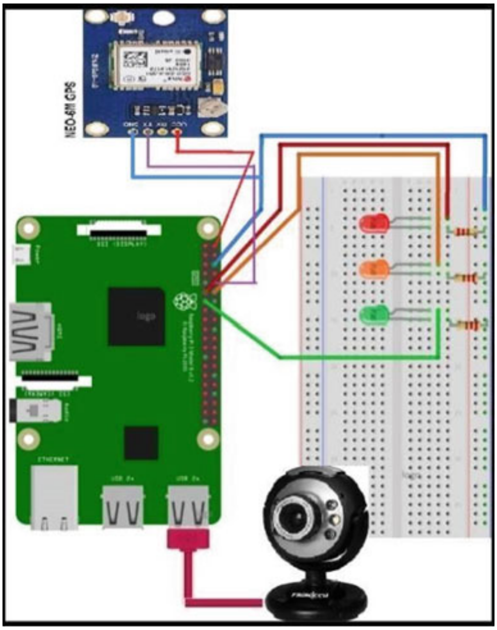


Fig. 6 Connection diagram

4 Results and Discussion

The proposed MobSE model achieved 3.7 and 3.6% better results than the standard MobileNet model for Top-1 and Top-5 accuracy respectively. For ambulance detection, the MobSE model achieved 3.1 and 4.7% better results than MobileNet for recall and precision respectively. After multiple tuning of the reduction ratio (r), the value of 16 achieves a good balance between complexity and accuracy. The Fig. 8 depicts the output of our machine learning model on a video frame (Tables 2, 3, and 4).

5 Conclusion and Future Scope

In this paper we presented a MobSE model for enhanced object detection tasks. The overall architecture makes it easy for use in devices supporting limited computations like Raspberry Pi. The SE block acted as a self-attention module thereby improving the over-all performance. Appropriate reduction ratio along with the use of dilated convolution yielded better results. The complete solution provided, makes sure that the emergency vehicles are given utmost priority over other vehicles. The app-based solution along with hardware components enables drivers to provide the service in a prompt manner thereby saving lives of the people.

Table 1 Specifications and functionality of the components used

S. No.	Component	Specifications	Functionality
1	Raspberry Pi 3B	Single Board Computer 1.4 GHz processor Input voltage: 5 V Input current: 2.5A	1. Acts as a subscriber node 2. Enabled Wi-Fi to be installed on traffic signals for controlling the traffic lights
2	Ublox Neo6M GPS	Operating Voltage: 2.7–3.6 V Operating Current: 45 mA	1. Interface with Raspberry Pi module for obtaining the module location
3	T5875DV	Input voltage: 90–264 VAC, Input Frequency: 47–63 Hz, Input current 0.5 A (max), Output Voltage + 5.1Vdc, Nominal Load Current: 2.5 A	1. Power supply for the raspberry pi module. 2. Installed at the main power supply unit of the traffic signals
4	Easy power micro SMPS	AC-DC Dual Output Input: 230 V Output: 5–3.3 V	1. Step down the 230 V AC supply to power the raspberry pi 2. Functions as step-down buck converter
5	74HC157N	4 channels 2:1 analog multiplexer Input voltage: 3.15 V (Minimum high-level voltage)	1. Act as a switch between raspberry pi and the existing circuitry in the traffic lights
6	IC 7805	Input voltage: 7 V (min) Input voltage: 25 V (max) Operating current (IQ): 5 mA. Output: 5 V	2. Voltage regulator 3. Used as a supply for the LED indicating the current working mode of raspberry pi

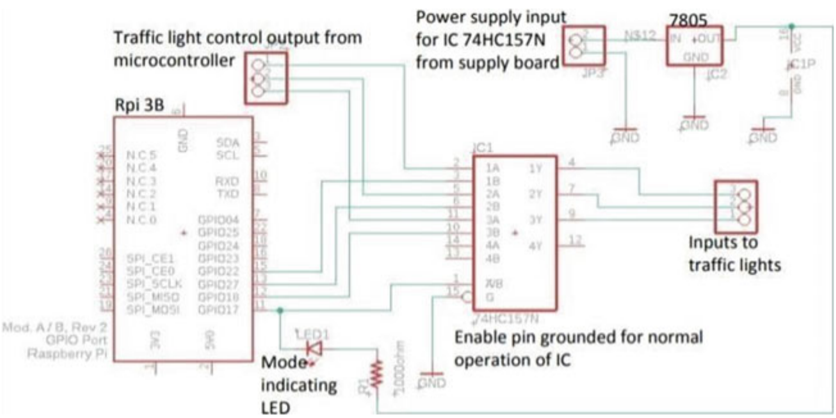


Fig. 7 Module to control the traffic lights via Raspberry Pi module

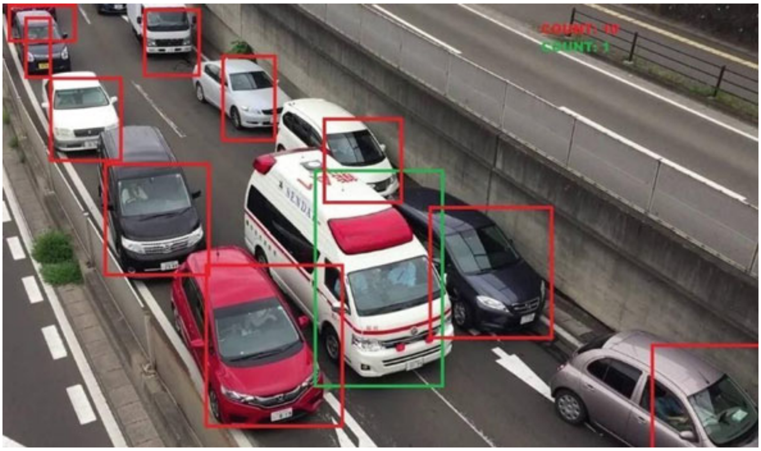


Fig. 8 Sample output of ambulance detection with the overall vehicle density

Table 2 Accuracy comparison of MobileNet and MobSE

Model	Accuracy TOP-1	Accuracy TOP-5
MobileNet	70.6	87.9
MobSE	74.3	91.5

Table 3 Values of error for various values of reduction ratio ‘r’

Ratio ‘r’	Error TOP-1	Error TOP-5
4	26.2	8.9
8	25.93	8.47
16	25.7	8.5
32	25.85	8.33

Table 4 Recall and precision values of the two model for ambulance detection

Model	Transfer learning	Recall	Precision
MobileNet	Imagenet	0.81	0.83
MobSE	Imagenet	0.841	0.877

In the further scope of our project, in order to make it more efficient as we aim to reduce power consumption and operational complexity, the hardware modules that are in operational condition continuously could be introduced to a power saving mode, which focuses on the hardware module (the microprocessor and the sensors connected) to be in power off state and when needed be switched to an operational condition using a trigger pulse. In order to do so we need to activate only the modules that lie along a particular route. From our understanding of the application and upon research we have devised a rough algorithm to achieve this. Since we have the list of coordinates where the hardware modules have been installed and we have the list of coordinates along the route, at specific intervals; using these two lists and comparing if a point (Latitude, Longitude) lies in a very close vicinity of any of the points along the route, we need to send a trigger pulse to activate that. However, this is completely theoretical but could be extended and deployed in order to enhance the functionalities of the mobile application.

We plan to expand the scope of solution to detect more emergency vehicles such as police vans, fire fighting vehicles etc. apart from ambulances which was the primary focus of the paper. We will also take into consideration the severity of the incident thereby giving priority to a more emergent situation when the two emergency vehicles approach the intersection at the same time.

References

1. Kanzaria HK, Probst MA, Hsia RY (1997) Emergency department death rates dropped by nearly 50 percent. Health Aff 35(7):1303–1308
2. Zaid A, Suhweil Y, Yaman MA (2017) Smart controlling for traffic light time. Proc IEEE Jordan Conf 1–5
3. Djahel S, Salehie M, Tal I, Jamshidi P (2013) Adaptive traffic management for secure and efficient emergency services in smart cities. In: 2013 IEEE International Conference on Pervasive Computer Communication Workshops (PERCOM Workshops), pp 340–343
4. Chung J, Sohn K (2018) Image-based learning to measure traffic density using a deep convolutional neural network. IEEE Trans Intell Transp Syst 19(5):1670–1675. [Online]. Available: https://doi.org/10.1109/tits.2017.2732029
5. Kanungo A, Sharma A, Singla C (2014) Smart traffic lights switching and traffic density calculation using video processing. Proc Recent Adv 1–6
6. Masurekar O, Jadhav O, Kulkarni P, Patil S (2020) Real time object detection using YOLOv3. Int Res J Eng Technol (IRJET)
7. Szegedy C, Ioffe S, Vanhoucke V (2016) Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv preprint arXiv:1602.07261
8. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T et al (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. CoRR
9. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
10. Pan SJ, Yang Q (2010) A survey on transfer learning. IEEE Trans Knowl Data Eng 22(10):1345–1359
11. Yao Y, Doretto G (2010) Boosting for transfer learning with multiple sources. In: 2010 IEEE conference on Computer vision and pattern recognition (CVPR). IEEE, pp 1855–1862
12. Xu Y, Pan SJ, Xiong H, Wu Q, Luo R, Min H, Song H (2017) A unified framework for metric transfer learning. IEEE Trans Knowl Data Eng 29(6):1158–1171
13. Mr. Reid "Shortest distance between two points on earth" http://wordpress.mrreid.org/haversine-formula/

Retweet Prediction for Large Datasets of Random Tweets

Saurabh Sharma and Vishal Gupta

Abstract In the information age, social media plays an important role in the life of online users. The ever-increasing use of social media for various purposes opens up a broad research area. The activities performed by users on social media can provide insights into their interests, opinions, likes, dislikes, emotions, etc. This article is an attempt to understand a general trend of information diffusion through retweets. The explicit features of a tweet are positively correlated with retweet count. The results of the regression analysis of five different algorithms have pointed out a strong linear correlation between retweet count and explicit features of a Tweet. The contribution of this article is to provide a simple and fast way of analysis of very large tweet datasets, using a minimum number of features without the complexity of feature generation and implementing complex algorithms.

Keywords Twitter · Retweet prediction · Information diffusion · Correlation analysis · Regression analysis · Big data

1 Introduction

Online Social Networks (OSNs) are the networks of users connected through the internet. These digital identities i.e. OSNs are a rich source of data about user behavior, activities, and interests. Twitter has proved to be an open place for interacting, sharing, discussing, and collaborating with other users from different walks of life. Twitter has seen enormous growth in terms of the number of active users per year. Recent studies have shown that the Twitter universe has a lot to offer for research fields of social sciences, human behavior, network evolution, information propagation, data mining, information overload, digital surveys, product branding, marketing, and so on.

The power of OSN such as Twitter comes from the fact that any news or information can reach a global audience in a few minutes. To spread any information,

S. Sharma · V. Gupta (✉)
Department of Computer Science and Engineering, UIET, Panjab University, Chandigarh, India
e-mail: vishal@pu.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_60

the users can post content on their timeline for the consumption of other users. The content can be posted by two common methods: first by creating a new tweet, and second, by reposting a tweet of some other user. The first method comes under the "active action" by creating new content and the second method comes under a "passive action" by not generating new content. The act of retweeting is very interesting for understanding the information flow in the network. There can be several factors responsible for this behavior; for example, news content in the tweet, the influence of friends, interest in a topic, emotion/sentiment of the tweet, etc. [1–5]. In this study, some basic statistical features, related to the content of the tweet, have been investigated for developing a language-independent, fast, and reliable model for predicting retweets.

The rest of the article is divided into the following sections. Section 2 covers the related work of other researchers on the study of retweets; Sect. 3 explains the motivation for this study. Section 4 describes the methodology and algorithms used and Sect. 5 discusses the results and findings. The last section concludes this article.

2 Related Work

The research field of information propagation is widely studied and always remains active due to the evolution of social networks with time. The methods of sharing information become more interesting with the use of images, videos, and links to external resources.

To study the ever-growing and evolving research area of information diffusion, it is important to study the retweeting patterns on Twitter. Recent studies have dealt with this research problem from many different directions. The three main categories to study retweets are:

i. Classification of tweets to determine which tweet will get retweets?
ii. Classification of users to determine which user will retweet a specific tweet?
iii. Regression analysis to determine why few tweets receive more retweets than others?

All these three main branches are very active and still have lots of scope for further research. This research article focuses on the third category where retweet prediction is studied with the help of regression analysis.

The complete process of retweet prediction can be summarized in four parts. These are collecting data, feature selection, proposing a model based on feature vector, and in the last step, interpreting, evaluating, and/or comparing the results. A very similar approach is used for tweet recommender systems [6–10].

To collect data, few researchers used publically available datasets and some researchers produced their fresh dataset by collecting data using Twitter API. The limitation of using public datasets is that sometimes few features are not available as per requirement. However, the benefit of creating a new dataset is that it can be customized according to the specific needs of the research domain. Both these methods are popular and acceptable for research work [1, 2, 4, 11, 12].

The example of features ranging from temporal analysis, social interaction, language-specific features, location-based features, user profiling, topic-based, interestingness and news factors, sentiment and emotion-based features, etc. [3, 13–16]. All these features used in retweet prediction are commonly extracted based on three factors. These are tweet content, author of the tweet, and user of the retweet. The popular way of using features includes extracting new features and creating a combination of features from different factors. The benefit of using various combinations leads to better prediction. However, the drawback of this approach is it is not feasible for domain-independent, language-independent, topic independent scenarios.

The models used for retweet prediction can be broadly divided into two categories: classification and regression. There are multiple combinations of these techniques with various combinations of features. It ranges from classical machine learning algorithms to deep learning models [3, 6, 8, 9].

To evaluate the results, there are very few standard measures for classification and regression algorithms. Some of the common measures are precision, recall, F-measure, r-squared, RMSE, and MSE [2, 3, 5, 8, 15].

3 Motivation

Recent studies [5, 8, 13, 17] have found that retweet prediction can be studied using different types of features such as text features, user profile features, network-based features, and a combination of these features. The features, related only to tweet content, can be categorized into two broad categories: explicit and implicit features as shown in Fig. 1. The explicit features are those which are provided directly in API response as key-value pairs or can easily be derived from them as measurable

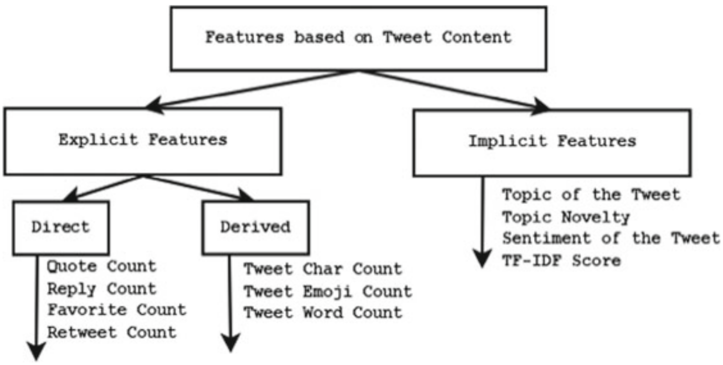


Fig. 1 Explicit and implicit features of a Tweet

Table 1 Description of direct explicit features

	Quote count	Reply count	Favorite count	Retweet count
Data type	int64	float64	float64	float64
Count	101,681,675	101,681,675	101,681,675	101,681,675
Mean	163.68	215.76	4792.55	2341.21
Std	1652.06	3122.62	41,456.3	18,934.1
Min	0	0	0	0
Max	414,331	846,494	4,554,490	3,614,140

quantities. The implicit features are those features that can be generated/calculated using some specific formula/equation or algorithm.

For this study, the authors tried to answer the following research questions:

- Which explicit features are easily available from random feeds of tweets using standard streaming Twitter API?
- Which limited explicit features can be used for retweet prediction for a very huge random dataset or live streaming data?

4 Methodology

4.1 Data Collection

Twitter data for this study is collected from the open-access Twitter archives of the year 2018 for August and September [18, 19]. This archive contains an unfiltered raw feed of tweets collected using standard search Twitter API. The dataset contains a random sample of 100 million tweets. The unit of data analysis is a single tweet where each tweet has a unique identity.

Whether a tweet received any retweet or not is checked by looking at the retweet count of that tweet. No two tweets are related to each other in any way i.e. every tweet is considered as a random tweet. The author of the tweet, user of the retweet, the text of the tweet, and any other information (which can relate any two tweets)s is not extracted for this study. The description of the dataset is given in two tables. Table 1 shows the basic statistics of direct explicit features and Table 2 shows the basic statistics of derived explicit features of tweets.

4.2 Correlation Analysis

The first step of the regression analysis is to find the correlation among the various features of the dataset. Figure 2 has shown the correlation matrix for all the

Table 2 Description of derived explicit features

	Tweet char count	Tweet emojis count	Tweet word count
Data type	int64	int64	int64
Count	101,681,675	101,681,675	101,681,675
Mean	85.23	0.4632	11.32
Std	44.5093	1.9523	7.96556
Min	1	0	1
Max	494	140	70

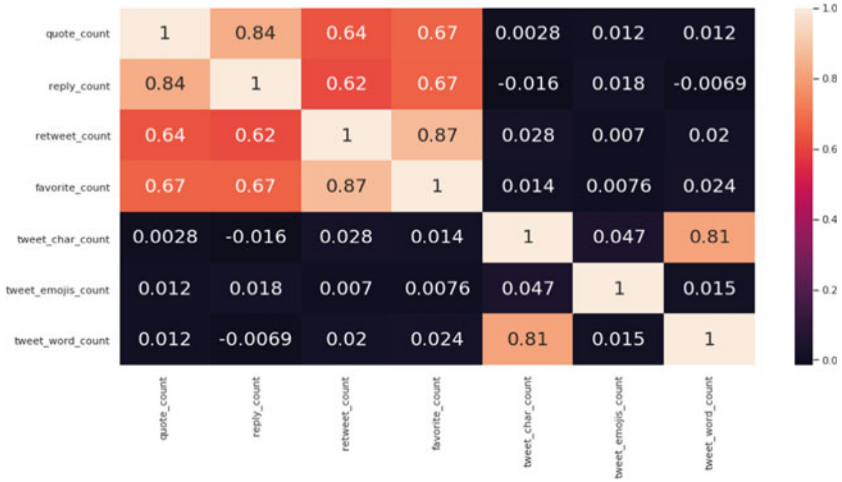


Fig. 2 Correlation matrix of explicit features of a Tweet

explicit features. The correlation among explicit features of the tweets shows significant positive values with retweet count. The recent studies have used implicit features for their studies, however, for this study, the implicit features do not provide any significant correlation either positive or negative with retweet count. Hence, it is safe to assume that these implicit features are not useful for the regression analysis of a huge dataset of random tweets.

Figure 3 has shown the different steps of the methodology. The correlation analysis has helped to filter useful features for regression analysis. This is a very important step because removing non-correlated features reduces the computational needs which leads to faster execution and efficient performance of algorithms on very huge datasets. The explicit features quote count, reply count, and favorite counts used for regression analysis.

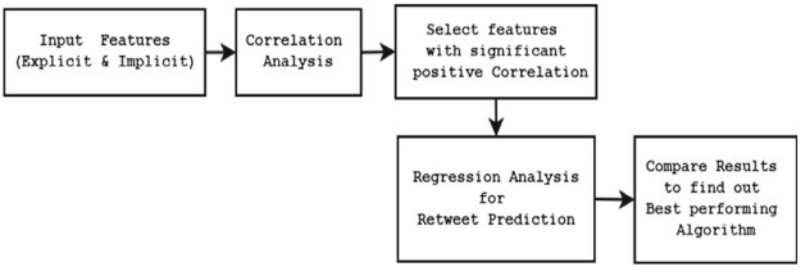


Fig. 3 Methodology

4.3 Regression Analysis

To establish a linear correlation between one dependent variable and the independent variable(s), the standard approach is to perform regression analysis. All features used for this study fall under the category of discrete variables. To study the relationship between the retweet count and various independent variables, five regression algorithms are used. For regression analysis, the dataset is divided into train and test datasets where 70% data (70 million tweets) is used for training and 30% (30 million tweets) data is used for testing.

The very basic algorithm for regression analysis is linear regression. This algorithm is used to find any linear relationship between the feature vector and the dependent variable. The second algorithm used in this study is a linear regression with an elastic net. This algorithm is used to overcome the limitations of basic linear regression and it also combines linearly L1 and L2 penalties from Lasso and Ridge regressions. The other three algorithms are gradient boosted tree regression, random forest regression, and decision tree regression.

The purpose of this analysis is to provide a simple, fast, and general-purpose model for big data analysis of random tweet feeds. The implementation was done on the Apache Spark cluster using pySpark interface, Apache Spark uses distributed parallel computing for fast and efficient execution. Apache spark generates a standard output for various regression algorithms in the form of Root Mean Square Error (RMSE).

5 Results

The result from five different algorithms, in the form of RMSE value, provides a comparison of their performance. The linear models performed better as compared to other regression models.

The RMSE is one of the standard measures to understand the absolute fit of the model. The value of RMSE can be understood as the standard deviation of the unexplained variance of the model as shown in Eq. 1. The unit of RMSE is the same as the dependent variable. In this study, the unit of RMSE is the number of retweets.

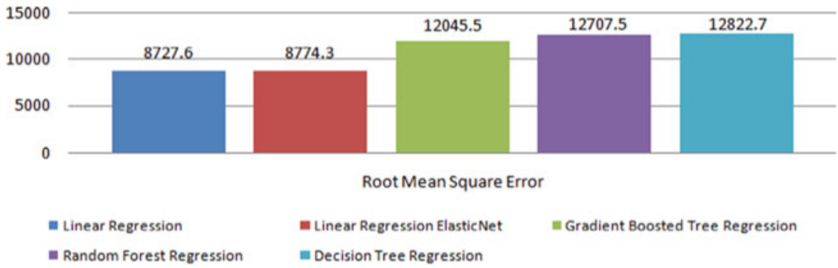


Fig. 4 Root mean square error of algorithms for regression analysis

$$RMSE_{fo} = \sqrt{\frac{\sum_{i=1}^{N}\left(v_{fi} - v_{oi}\right)^2}{N}} \quad (1)$$

where, N = sample size, v_{fi} = forcast value for ith data item, v_{oi} = observed value for ith data item.

The acceptance range for the RMSE is less than the standard deviation of the dependent variable. Table 1 shows that the standard deviation is 18,934 for the retweet count. All five algorithms are performed within an acceptable range as shown in Fig. 4. This indicates that regression models are best suited for retweet prediction with a few input features on the datasets as huge as 100 million random tweets.

The linear regression models performed better than other models. The R-squared is another measure to explain the performance of linear regression models. It explains the percentage of variance explained by the model. The basic linear regression model has a slightly better performance than the elastic net model as given in Table 3. This can be interpreted as the linear correlation between the retweet count and input vector follows an absolute linear pattern with no or minimum regularization needed. In other words, for a tweet, the number of retweets is directly proportional to the quote count, reply count, favorite count of the tweet. This finding encourages future researches to explore this pattern with other combinations of dependent and independent variables (Table 3).

Table 3 Comparison of linear regression algorithms

	R-squared	Root mean square error
Linear regressions	0.77899	8727.682583
Linear regression ElasticNet	0.77738	8774.302461

6 Conclusion

In this study, the authors proposed a simple way of analyzing very huge datasets of general topic tweets using very limited features provided in the standard API response. The implicit and explicit features of a tweet showed two different patterns in the correlation matrix. The explicit features are found to be positively correlated whereas implicit features did not show any direct correlation. The feature vector of only three variables can explain approximately 2/3 of the variance of retweet count. The linear regression model found out to be the most efficient algorithm in capturing the correlation as compared to other algorithms.

Recent studies have shown that for analyzing Twitter, a large number of features and complex algorithms are needed for accurate prediction of retweets. The main drawback of those studies was that those techniques needed a huge amount of computational power for scale-up, feature engineering, and in some cases manual help for data preparation. This study is an attempt to provide a solution to the above-mentioned problems and can be used for many new scenarios that focus on general-purpose pattern mining on Twitter.

The future work shall be aimed at using various combinations of tweet features including user profiling, user activities, and sentiment analysis. The main focus will remain on the pattern mining from the random data streams to understand a global trend irrespective of geolocation, communities, domains, and sections of users.

References

1. Jiang B, Yi F, Wu J, Lu Z (2019, Aug) Retweet prediction using context-aware coupled matrix-tensor factorization. In: International conference on knowledge science, engineering and management. Springer, Cham, pp 185–196. https://doi.org/10.1007/978-3-030-29551-6_17
2. Jain DK, Kumar A, Sharma V (2020) Tweet recommender model using adaptive neuro-fuzzy inference system. Futur Gener Comput Syst. https://doi.org/10.1016/j.future.2020.04.001
3. Fan C, Jiang Y, Yang Y, Zhang C, Mostafavi A (2020) Crowd or hubs: information diffusion patterns in online social networks in disasters. Int J Disaster Risk Reduction 46:101498. https://doi.org/10.1016/j.ijdrr.2020.101498
4. Jalali NY, Papatla P (2019) Composing tweets to increase retweets. Int J Res Mark 36(4):647–668. https://doi.org/10.1016/j.ijresmar.2019.05.001
5. Firdaus SN, Ding C, Sadeghian A (2018) Retweet: a popular information diffusion mechanism—a survey paper. Online Soc Netw Media 6:26–40. https://doi.org/10.1016/j.osnem.2018.04.001
6. Cao Y, Zhang J, Ma Y, Shao P (2018) A study on predicting the microblog retweet based the random walk model. In: International conference on management science and engineering management. Springer, Cham, pp 527–541. https://doi.org/10.1007/978-3-319-93351-1_42
7. Chen G, Kong Q, Xu N, Mao W (2019) NPP: a neural popularity prediction model for social media content. Neurocomputing 333:221–230. https://doi.org/10.1016/j.neucom.2018.12.039
8. Liu G, Shi C, Chen Q, Wu B, Qi J (2014) A two-phase model for retweet number prediction. In: Li F, Li G, Hwang S, Yao B, Zhang Z (eds) Web-age information management. WAIM 2014. Lecture notes in computer science, vol 8485. Springer, Cham. https://doi.org/10.1007/978-3-319-08010-9_84

9. Mao H, Xiao Y, Wang Y, Wang J, Xiao Z (2018) Topic-specific retweet count ranking for Weibo. In: Phung D, Tseng V, Webb G, Ho B, Ganji M, Rashidi L (eds) Advances in knowledge discovery and data mining. PAKDD 2018. Lecture notes in computer science, vol 10939. Springer, Cham. https://doi.org/10.1007/978-3-319-93040-4_49
10. Oliveira N, Costa J, Silva C, Ribeiro B (2020) retweet predictive model for predicting the popularity of tweets. In: Madureira A, Abraham A, Gandhi N, Silva C, Antunes M (eds) Proceedings of the tenth international conference on soft computing and pattern recognition (SoCPaR 2018). SoCPaR 2018. Advances in intelligent systems and computing, vol 942. Springer, Cham. https://doi.org/10.1007/978-3-030-17065-3_19
11. Nesi P, Pantaleo G, Paoli I, Zaza I (2018) Assessing the reTweet proneness of tweets: predictive models for retweeting. Multimed Tools Appl 77(20):26371–26396. https://doi.org/10.1007/s11042-018-5865-0
12. Son J, Lee J, Oh O, Lee HK, Woo J (2020) Using a heuristic-systematic model to assess the Twitter user profile's impact on disaster tweet credibility. Int J Inf Manage 54:102176. https://doi.org/10.1016/j.ijinfomgt.2020.102176
13. Firdaus SN, Ding C, Sadeghian A (2019) Topic specific emotion detection for retweet prediction. Int J Mach Learn Cybern 10(8):2071–2083. https://doi.org/10.1007/s13042-018-0798-5
14. Liu Y, Zhao J, Xiao Y (2018) C-RBFNN: a user retweet behavior prediction method for hotspot topics based on improved RBF neural network. Neurocomputing 275:733–746. https://doi.org/10.1016/j.neucom.2017.09.015
15. Jiang B, Lu Z, Li N, Wu J, Jiang Z (2018, June) Retweet prediction using social-aware probabilistic matrix factorization. In: International conference on computational science. Springer, Cham, pp 316–327. https://doi.org/10.1007/978-3-319-93698-7_24
16. Guoheng R, Wei W, Liyong C (2019) Effective prediction of microblog user retweet behaviors based on markov random field. In: Deng K, Yu Z, Patnaik S, Wang J (eds) Recent developments in mechatronics and intelligent robotics. ICMIR 2018. Advances in intelligent systems and computing, vol 856. Springer, Cham. https://doi.org/10.1007/978-3-030-00214-5_1
17. Daga I, Gupta A, Vardhan R, Mukherjee P (2020) Prediction of likes and retweets using text information retrieval. Procedia Comput Sci 168:123–128. https://doi.org/10.1016/j.procs.2020.02.273
18. https://archive.org/details/archiveteam-twitter-stream-2018-08
19. https://archive.org/details/archiveteam-twitter-stream-2018-09

To Predict Employability of Student by Using Artificial Neural Network

Manjushree D. Laddha, Arvind W. Kiwelekar, Laxman D. Netak, and Prasad C. Mahajan

Abstract In the higher education institute or universities employability of the student is the key factor that decides the quality of the education. One of the outcomes of education acquired is employability according to students' perspective. To sustain in this competitive world a higher education along with different hard and soft skills are required. Being a computer engineer many skills are required for employability. This study aimed to determine the employment status depending upon the competencies related to the courses that were learned during the engineering education are considered. As well as some factors related to the students' competitive exams, some factors related to their primary, secondary education, location and student profile were also examined. For this study, the survey was conducted with 14 different parameters. From these parameters, it is predicted that whether the students will be employed or not? These parameters are considered as a very high contribution for placement. To predict whether the student will get a placement or not we have applied Artificial Neural Network (ANN) and Logistic Regression. Our Finding shows that 87.5% accuracy, we got by using Artificial Neural Network (ANN) and 62.5% by Logistic Regression.

Keywords Employability · Computer engineering · Parameters

1 Introduction

Employability is the major key factor for the students who are learning technical courses. This engineering degree aimed that students' are fully equipped with practical as well as theoretical knowledge. If it is possible for a student to know whether

M. D. Laddha (✉) · A. W. Kiwelekar · L. D. Netak · P. C. Mahajan
Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, India
e-mail: mdladdha@dbatu.ac.in

A. W. Kiwelekar
e-mail: awk@dbatu.ac.in

L. D. Netak
e-mail: ldnetak@dbatu.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_61

he or she is placed in the company or not then this will be a great deal. If not selected they can work hard to acquire skills required for employability. This information helps instructor as well as Career Development Cell (CDC) to take major action to improve their students knowledge and skills to acquire a job in reputed company.

For predication of placement of students [5] Naive Bayes classification by using Weka and Datamere tools were done. These predictions were done on 60 students, gave 71.66% accuracy. For outcome based education system a placement predication model prepared [6]. In this paper, different categories (dream companies, super dream companies and mass recruiter companies) are designed to predict in which category student will be selected by considering their past academic performance.

The objectives of this paper [7] is to apply the clustering techniques for the student placement data set and comparing performance of the clustering techniques. The output information of the clusters is added to the student dataset which is used as input to the Naive Bayes technique in classifying and predicting the placement chance of students.

A method is proposed to [4, 8] help students prepare for the job recruitment process. The courses are identified for a specific job. Tests are conducted in this course. They also proposed refined assessment to improve students' scores. From this students can focus on their weak concepts and receive timely feedback for their improvement.

Many different placement prediction models for pre-final year engineering graduate students are available in this paper. The survey on these predication models with different techniques are presented [2].

Investigation of this [3] paper shows that if student have low competence then extra in house training is required with the industry experts.

Many different machine learning models are used to predict student placement [1] like linear regression model, K-neighbor regression model, decision tree regression model, XGBoost regression model, gradient boost regression model, light GBM regression model and random tree classifier model.

We predict employability based on the major parameters which are required for computer engineering. The parameters are categories as communication skills, competitive exams scores and courses of computer engineering and we get accuracy 87.5% by ANN.

2 Data Collection

2.1 Method

The main instrument used for this study is surveying questionnaire. Actually 33 different parameters are considered in this survey out of this we have considered 14 parameters in this paper. The questions which we have used for this survey are related to past learned courses, related to their Secondary School Certificate (SSC), Higher

Secondary Certificate (HSC) scores, English Language, Competitive programming platforms, locations, course knowledge etc. Our focus is to give thrust on important aspects which are related to their employability.

2.2 Participants

Total population of 68 students of computer engineering from the academic years from 2019–2020 is considered. These students have completed all the four year courses of B.Tech Computer Engineering and only their last semester is remained. They were exposed to theoretical as well as practical aspects of graduate courses. They were acknowledged with competitive platforms and competitive exams also.

3 Data Analysis

3.1 Data Analysis and Pre-processing

The dataset used in this research consists of 68 records containing 14 variables as shown in Table 1. We have used R programming to analyse the dataset and construction of models. Table 1 describes basic descriptive statistics for 14 variables. For the sake of simplicity, we rename these 14 variables as shown in the second column of Table 1 for further use. Before building logistic regression model and ANN model over the dataset, we have performed preprocessing over the data. In the dataset, we have two categorical fields namely SSC from and HSC from. These two fields tell us from where the student has completed his/her education either from a rural or urban area. We have encoded these categorical variables using the "factor" method in R. Again, if we are looking at Table 1, we came to know that the values are not properly scaled. Hence we performed feature scaling over the dataset using the "scale" method of R. This takes all our features on the same scale.

Once the dataset is ready, we have found out correlation matrix over the dataset. The correlation of all thirteen independent variables with the dependent variable is shown in Fig. 1. From this, we have selected 6 independent variables having higher correlations. Those are 'x1', 'x7', 'x8', 'x9', 'x11' and 'x13' along with dependent variable 'y'. The correlation of these six independent variables with the dependent variable is shown in Fig. 2. Once the dataset is preprocessed it is ready for constructing a logistic regression model and ANN model. The "sample.split" helps to divide the dataset into two sets, training set and test set. While splitting the dataset into two sets, the split ratio is set to 0.89. Hence there are 60 records in the training set and 8 records in the test set.

Table 1 Basic descriptive statistics for variables

Name of variable	Symbol	1st Qu.	3rd Qu.	Max.	Mean	Median	Min.
HSC English marks	x1	63	78	95	69	70	35
HSC maths marks	x2	66	83	93	74.47	76	35
SSC English marks	x3	77	91	94	81.52	85	44
SSC maths marks	x4	88	97	128	90.77	92	54
HSC from	x5	1	2	2	1.441	1	1
SSC from	x6	1	2	2	1.522	2	1
GATE score	x7	7	19	492	20.35	15	0
CET/JEE score/diploma percentage	x8	85	114	133	97.95	102	7.99
C programming marks	x9	50	70	95	61	60	40
Data structures marks	x10	50	70	95	61.57	60	40
Rating for knowledge of machine learning	x11	1	3	5	2.435	2	1
Rating for knowledge of mobile application development	x12	2	3	4	2.391	2	1
Rating for knowledge of web application development	x13	3	4	5	3.217	3	1
Placed	y	0	1	1	0.4058	0	0

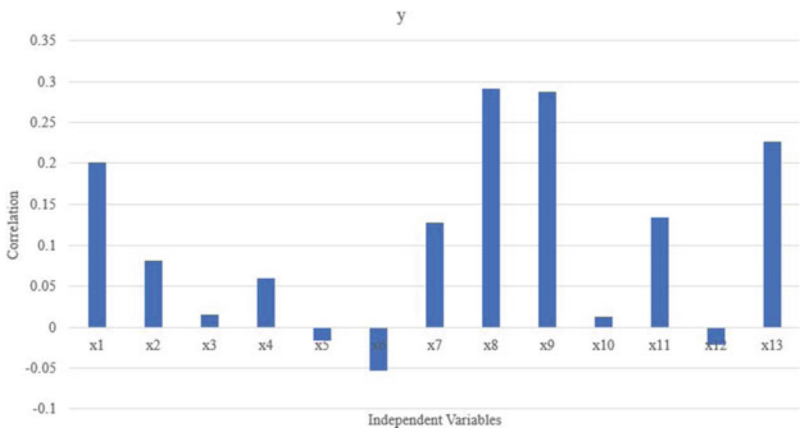


Fig. 1 Correlations of thirteen independent variables with dependent variable

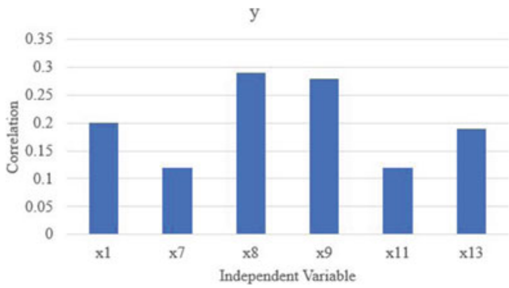


Fig. 2 Correlations of selected six independent variables with dependent variable

4 Model Construction

4.1 Logistic Regression Model

The logistic regression model is constructed using the "glm" method. While constructing the model we have specified binomial family, which tells glm to perform the logical operation. Once the model is constructed, the "predict" method is used to predict values over the test set. There are three arguments to the "predict" method. The first argument is the fitted object of a class inheriting from "glm". The second argument is the type of prediction required. As we are constructing logistic regression we have set the type as "response" which gives predicted probabilities. And the last argument is the test data without including the dependent variable.

4.2 Artificial Neural Network Model

While building an ANN model, the question arises like how to decide the number of layers and numbers of neurons in the hidden layers. So, unfortunately, there is no thumb rule to choose an optimal number of hidden layers and optimal numbers of neurons in the hidden layers. We know that number of nodes in the input layer is equal to the number of independent variables. So as we have 6 independent variables, the number of input nodes in the input layer is 6. Here we are trying to predict whether a student will be placed or not as binary output we have one output node in the output layer. Therefore, we choose the number of neurons in the hidden layer as the mean of input nodes and output nodes as 6 plus 1 divided by two that is three point five. Hence we will take round number 4 as neurons in the hidden layer.

To be honest, we are not working with any complex data set like for example image just to find some patterns in the pixels. Here we have a simple data set with some independent variables and one dependent variable and there is no spatial structure. So to be honest, we don't need many hidden layers. The Model would work very well with only one hidden layer. But since we are using ANN we have gone for two hidden layers. The "h2o.deeplearning" method from the "h2o" package helps to construct the ANN model. As we know there is an activation function in the ANN, the method "h2o.deeplearning" helps to set the activation function. "Tanh", "TanhWithDropout", "Rectifier", "RectifierWithDropout", "Maxout", "MaxoutWithDropout" are some activation functions used in ANN. Here "Rectifier" is used. The method "h2o.deeplearning" gives functionality to set the number of hidden layers and the number of neurons in the hidden layers. As discussed above, there are two hidden layers and 4 neurons in each hidden layer. The "hidden" argument of the method "h2o.deeplearning" is set to a vector in R as "c(4, 4)". The number of elements in the vector indicates the number of hidden layers and the values indicate the number of neurons in each hidden layer. After specifying the required arguments, the ANN model is constructed. The model is used to predict output over the test set using "h2o.predict" method. There are two arguments to the method. The first argument is the model and the second argument is the test set.

5 Result and Discussion

To describe the performance of a classification model, the confusion matrix is used. Here the confusion matrix is used to compare the predicted values with actual values. Table 2 is the confusion matrix for logistic regression whereas Table 3 is the confusion matrix for the ANN.

There are two possible predicted classes: "1" and "0" (dichotomous). We are predicting whether a student will get a placement or not, for example, "1" would mean the student will be placed, and "0" would mean the student will not be placed. The classifier made a total of 8 predictions (e.g., 8 students are being tested for the checking will get placement or not). Out of those 8 cases, the logistic classifier predicted "1" 4 times, and "0" 4 times whereas the ANN classifier predicted "1" 2

Table 2 Confusion matrix for logistic regression

n = 8		Predicted placement	
		0	1
Actual placement	0	3	2
	1	1	2

Table 3 Confusion matrix for ANN

n = 8		Predicted placement	
		0	1
Actual placement	0	5	0
	1	1	2

Table 4 Evaluation of confusion matrix

	Logistic regression	ANN
Accuracy	0.625	0.875
Mis-classification rate	0.375	0.125
True positive rate	0.667	0.667
False positive rate	0.4	0
True negative rate	0.6	1
Precision	0.5	1
Prevalence	0.375	0.375
F-score	0.5714	0.8
Cohen kappa coefficient (k)	0.25	0.713

times, and “0” 6 times. In reality, 3 students in the sample are placed, and 5 students do not.

For binary classifier, Table 4 gives a list of rates that are often computed from the confusion matrix along with F-score and Cohen Kappa Coefficient (k).

From Table 4, it is observed that the accuracy of the ANN is more than logistic regression. The second row of the table tells us how often the model goes wrong. Precision gives an idea about how frequently a model is correct when it predicts output as 1. Prevalence tells how many times “1” actually occurs in the sample dataset.

F-score is the measure of accuracy in the statistical analysis of binary classifier. F-score has a possible value in between 0 and 1. For logistic regression, we got an F-score value as 0.5714. For ANN, it is 0.8. F-score tells the perfection of precision and recall.

Cohen Kappa coefficient values in Table 4, indicate the actual values are much closer to predicted results by ANN than that of the predicted result of logistic regression.

6 Conclusion and Future Scope

In this paper, we met our objective of predicting employability of student from their past performance data. As we know that employability of students’ is not only beneficial to them, but also for the progress of the university. Predition of employability is useful for instructors to help students and give timely feedback to improve their knowledge. Due to this prediction Career development center will also be benefited as more student’s will be employed in reputed company and inturn more companies will come to university for placement.

In our paper, we proposed two models to predict students’ employability by using logistic regression and ANN and we got accuracy 62.5% and 87.5% respectively.

This prediction can enlighten students to improve their lacuna and cope up with current requirements of the industry. This model also helps to university to make policy to improve employability for the future years.

In our work we have considered data of computer engineering students, in the future scope for the competency of employability, we may consider more parameters as well as data from other engineering streams also. For large scale of student data, we can use other machine learning languages, techniques and tools.

References

1. Aravind T, Reddy BS, Avinash S, Jeyakumar G (2019) A comparative study on machine learning algorithms for predicting the placement information of under graduate students. In: 2019 third international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). IEEE, pp 542–546
2. Joy LC, Raj A (2019) A review on student placement chance prediction, pp 542–545
3. Lundberg GM, Krogstie BR, Krogstie J (2020) Becoming fully operational: employability and the need for training of computer science graduates, pp 644–651
4. Marling C, Juedes D (2016) CS0 for computer science majors at Ohio University, pp 138–143
5. Mavani U, Lobo VB, Pednekar A, Pereira NC, Mishra R, Ansari N (2020) Naïve Bayes classification on student placement data: a comparative study of data mining tools. Information and communication technology for sustainable development. Springer, Germany, pp 363–372
6. Rao AS, Aruna Kumar S, Jogi P, Chinthan Bhat K, Kuladeep Kumar B, Gouda P (2019) Student placement prediction model: a data mining perspective for outcome-based education system. Int J Recent Technol Eng (IJRTE) 8:2497–2507
7. Shukla M, Malviya AK (2019) Modified classification and prediction model for improving accuracy of student placement prediction
8. Soumya M, Sugathan T, Bijlani K (2017) Improve student placement using job competency modeling and personalized feedback, pp 1751–1755

MNIST Image: Color to Sound Conversion and Classification Using SVM

S. Srivalli Devi and A. Geetha

Abstract Sound is a longitude wave. Light, transverse one. Frequency is the common term in between these two. Using the visible light spectrum and audible sound ranges, the color (light) in an image is converted into the sound for that specific frequency. MNIST dataset containing the handwritten images of digits from zero to nine is used for the research and those images are converted from image (light) to sound (audio). The image is mapped to piano musical audio notes. Then the audio is classified as belonging to disease category image or not, using SVM. The accuracy we got is nearly 83.35% using audio classification.

Keywords SVM · Light · Sound · Classification · MNIST · Light-to-sound conversion

1 Introduction

MNIST dataset consist of handwritten numbers stored digitally as images. These images are used in many image classification algorithms. There exists a relationship between light and sound. Using this relationship, in this paper we propose a novel idea to classify sound data converted from image data. Image data is wrangled to sound data and the classification is done on sound data using SVM. Here is a list of ideas discussed in the paper:

- Image to sound conversion.
- Classification of the sound which has been converted from image.
- Classification done by using SVM for audio.
- Audio classification done by using pyaudioanalysis package in Python.
- The sounds that are used in this research are Western musical notes.
- The algorithm is done on MNIST dataset and the classes for classification are divided into two zero and non-zero.

S. Srivalli Devi (✉)
Government Arts and Science College, Palladam, India

A. Geetha
Chikkanna Government Arts College, Tirupur, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_62

Table 1 Color values table

Color	Frequency (THz)	Wavelength (nm)
Violet	668–789	380–450
Blue	631–668	450–475
Cyan	606–630	476–495
Green	526–606	495–570
Yellow	508–526	570–590
Orange	484–508	590–620
Red	400–484	620–750

- Zero class means the image is zero, else otherwise.

2 Relationship Between Light and Sound

Light. The main source of light to our Earth is Sun. In electromagnetic spectrum, visible light is a small portion. Cameras capture the light and thus the image is produced for later viewing. This image can be viewed on display devices. Light is a form of transverse wave, even in vacuum it travels. Vibrations of light waves are perpendicular to its travel direction.

Sound. In vacuum, sound can't travel. Vibrations of sound waves are in same direction of its travel. Sound can be recorded using microphones and can be heard using speakers. Sound is a longitudinal wave. Sound can be transmitted to solids, liquids and gases [1]

Visible Spectrum of light. The only region if the spectrum which is observable by the human eye. The range is about 400–700 nm.

Wavelength and frequency. Relationship between Frequency and Wavelength can be expressed as

$$\text{Wavelength} = \text{Veocity}/\text{Frequency} \tag{1}$$

Colour chart that connects these two terms is given below.
The color conversion values is obtained from the Table 1 as given in [2]
The electromagnetic spectrum is shown in Fig. 1 from an article in Wikipedia [3]

3 Conversion from RGB Pixel to Musical Note

RGB Color model. A photopgraph (image) stored in a computer consists of pixels. These pixels store the color information. The RGB color model consists of three

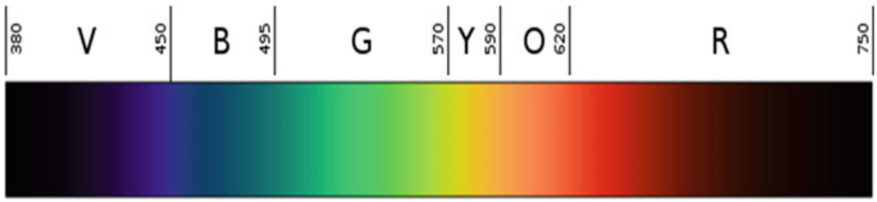


Fig. 1 Light spectrum and its wavelength

primary colors Red, Green, Blue. When these colors mixed togather in different proportions they produce different other colors. Wavelength and frequency differs for each and every color.

Sound. Similar to light, sound also has a sound spectrum and there is an audible range in sound spectrum within the frequencies which the sound can be heard by humans. The audible range for humans ranges from 20 to 20,000 Hz [4, 5].

Musical notes. Musical sound is symbolically denoted by note. There are twelve notes whose frequencies are fixed are used, in Western music, which are mathematically related to each and defined just about the central note A4. Pitch is 440 Hz.Pitch classes are denoted by A, B, C, D, E, F and G.

The colour of sound. From the given conversion chart in the Fig. 2 the values for each color and their corresponding sound musical note is obtained. This is done by a series of if else conditions.

4 Methodology

Dataset used. The dataset used here is MNIST dataset. The dataset consist of handwritten digits from 0 to 9. MNIST (Mixed National Institute of Standards and Technology) database is dataset for handwritten digits, distributed by Yann Lecun's [6] website. The dataset consists of pair, "handwritten digit image" and "label". Digit ranges from 0 to 9, meaning 10 patterns in total, handwritten digit image: This is gray scale image with size 28 × 28 pixel. Label: This is actual digit number this handwritten digit image represents. It is either 0–9. MNIST dataset is widely used for "classification", "image recognition" task.

Procedure

Initially the color to note conversion is done for MNIST images for training images. And are used in classification using SVM. For example, Let us assume that we are going to convert an MNIST image for the number 1 to corresponding note file. This generated note file is then converted to music by reading each line which contains notes information and creates a soundfile and appends the notes in sound form for

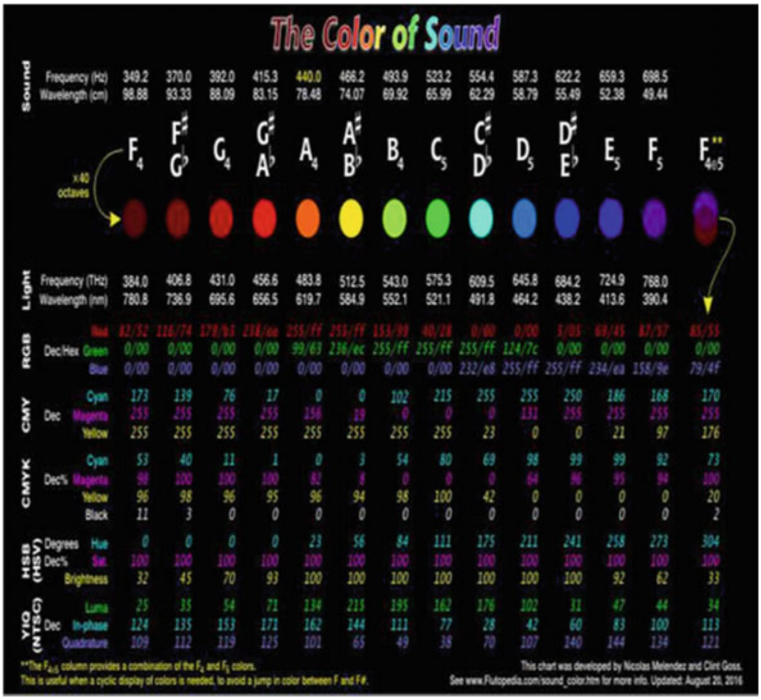


Fig. 2 Colour of sound

every note read line by line from the text file. Now this sound file is used in classification task. In this paper the classification is done for two classes: zero and non-zero, if the pixel has got the values RGB (2552371) then the note A♯B♭ is stored in the first line of the file for that particular image. Similarly, second pixel's RGB values are extracted and checked for the corresponding musical note and this note is written in the second line of the file and procedure continues on till the last pixel is read. Procedure for conversion of pixel values to written notes. These notes are written in a file.

Step 1 Store the color to note information in a text file line by line.
Step 2 for each pixel do.
Step 2.1 Extract the Red Green Blue values.
Step 2.2 check and write the Note in the file, each note in a single line, line by line.

The else part (last one) is used for colors that are not in the range within the colour to sound spectrum (Table 2).

```
Else
if r== 0: PrintF5♯G♭note in the file
if r in range(1,5): Print C♯D♭note in the file if r in
range(5,40):Print D♯E♭note in the file if r in range(40,69):Print
C5 note in the file
```

Table 2 Corresponding color value's note to be printed on file

if	Red in range		Green in range		Blue in range		Note to be printed in file ♭
	255	256	236	256	0	256	A♯B
Else if	255	256	99	236	9	256	A4
Else if	238	256	0	100	0	9	G♯A♭
Else if	179	238	255	256	0	256	G4
Else if	153	179	255	256	0	256	B4
Else if	116	154	0	255	0	256	F♯G♭
Else if	87	116	0	256	158	256	F5
Else if	82	88	0	1	0	158	F4
Else if	69	82	0	256	234	256	E5
Else if	40	70	255	0	0	234	C5
Else if	5	41	0	255	255	256	D♯E♭
Else if	0	5	255	0	232	255	C♯D♭
Else if	0	1	124	256	255	229	D5
Else if	0	1	0	125	0	255	C5

```
if r in range(69,82):Print E5 note in the file
if r in range(82,87):Print F4 note in the file
if r in range(87,116):Print F5 note in the file
if r in range(116,153):Print F♯G♭note in the file if r in
range(153,179):Print B4 note in the file
if r in range(179,238):Print G4 note in the file
if r in range(238,255):Print G♯A♭note in the file if r== 255:Print
A4 note in the file
```

Procedure for conversion of written notes on the file to music

Step 1 the notes written in file are read line by line.
Step 2 each note is then matched with the corresponding note sound from sound file.
Step 3 each note is then appended one after the other till the last pixel's note is appended in the sound file for that image.
Step 4 similarly for all the training, testing and classification required images the sound file is generated All the musical files are stored in MP3 format.

Classification: Classification is the grouping of elements into predefined groups. The metrics used here for evaluating are precision, recall, accuracy and f1 score. The confusion matrix as in Table 3.

Precision:

Precision is calculated using the below formula

Precision = (true positive)/(true positive + false positive)

Table 3 Confusion matrix

Actual	Predicted		
		Positive	Negative
	Positive	True positive	False negative
	Negative	False negative	True negative

Recall:

Recall is calculated using the formula
Recall = (true positive)/(true positive + false negative)

F1-score:

The f1-score is computed using the formula
F1 = 2([Precision*recall]/[Precision + recall])*

Pyaudioanalysis package:

The MP3 files are stored in a folder of two: ZEROSONGS and NONZERO_SONGS (each folder representing a different class) and then this trained classifier is used to classify an unknown audio MP3 file. The pyaudioanalysis [7] package written in python is used to achieve this audio classification.

5 Result

The parameters (c) and precision (PRE), recall (REC) and f1 scores for ZEROSONGS and NON-ZERO SONGS and their overall accuracy and f1 scores are displayed in the table. Confusion matrix is generated and the values obtained are shown above. The parameter selected is 0.01000 because the best f1 and best accuracy is obtained on the selection of that parameter. Accuracy of about 83.35% is obtained (Fig. 3).

6 Conclusion

The novel approach used for the classification of images by converting them to sound (music) has achieved about an accuracy of about 83% which is a good score. This approach can be applied in multiple areas wherever image data is used.

```
          ZEROSONGS                      NONZERO_SONGS          OVERALL
          C       PRE    REC    f1     PRE    REC    f1     ACC    f1
          0.001   81.5   90.9   86.0   74.9   56.7   64.5   79.9   75.2
          0.010   82.0   91.0   86.3   75.4   58.2   65.7   80.4   76.0   best f1    best Acc
          0.500   81.4   89.0   85.0   71.3   57.4   63.6   78.8   74.3
          1.000   80.4   87.6   83.8   67.9   55.0   60.8   77.1   72.3
          5.000   81.8   86.9   84.2   68.3   59.4   63.5   78.0   73.9
          10.000  81.6   85.8   83.7   66.6   59.5   62.9   77.3   73.3
          20.000  82.2   85.9   84.0   67.3   61.0   64.0   77.9   74.0
Confusion Matrix:
          ZER     NON
ZER       61.62   6.12
NON       13.48   18.77
Selected params: 0.01000

Out[2]: (0.0, array([0.83352057, 0.16647943]), ['ZEROSONGS', 'NONZERO_SONGS'])
```

Fig. 3 Screenshot of the results obtained

Appendix

MNIST Modified National Institute of Standards and Technology database.

SVM Support Vector Machine.

References

1. Kuehni RG (2011) On the relationship between wavelength and perceived hue. Color Res Appl 37(6):424–428. https://doi.org/10.1002/col.20701
2. Bruno TJ, Svoronos PDN (2006) CRC handbook of fundamental spectroscopic correlation charts. Published by CRC Press, 225 p. ISBN 0-8493-3250-8 (978-0-8493-3250-0).
3. https://en.wikipedia.org/wiki/File:Linear_visible_spectrum.svg
4. Rosen S (2011). Signals and systems for speech and hearing, 2nd ed. BRILL, p 163
5. Rossing T (2007) Springer handbook of acoustics. Springer, pp 747, 748. ISBN 978-0387304465
6. http://yann.lecun.com/exdb/mnist/
7. Giannakopoulos T (2015) Pyaudioanalysis: an open-source python library for audio signal analysis. PloS one 10(15) Public Library of Science

Secure Image Transmission Using Style Transfer

Dheeraj Komandur, Yash Shekhadar, Hrishikesh Mahajan, and Shebin Silvister

Abstract In this paper we introduce a novel method for secure transmission of images using style transfer and image cryptography. Style transfer is a process of modifying the style of an image while still preserving its content. In our system the secret image is encoded with a key image and the encoded image is sent over an untrusted network. The key image is converted into a styled image by Cycle GAN and this styled image is transmitted to the receiver. Cycle GAN is used on the receiver end to convert the style image to the key image. The receiver decodes the encoded image and gets the secret image. The proposed novel architecture provides opportunity for a new scope of research and development in the field of image transmission using style transfer.

Keywords Style transfer · Deep learning · Image cryptography · Generative adversarial neural networks

1 Introduction

The objective of this paper is to propose a novel architecture for secure image transmission using style transfer and image cryptography. The proposed architecture uses style transfer to convert the key image (K) to a styled image and transmit it through the untrusted network. Once the receiver gets this styled image, it is converted back to the original key image (K′). This process is done by Cycle GAN's G and F generators [1]. The novel approach is indicative of the fact that further analysis and studies can be performed in this domain and open up research in secure transmission of image using style transfer.

D. Komandur (✉) · Y. Shekhadar · H. Mahajan · S. Silvister
School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_63

2 Related Work

In "A Neural Algorithm of Artistic Style" [2], Leon A. Gatys et al. implemented a system of natural style transfer which is based on deep neural network methods. The objective of the paper is to transfer characteristics of a picture, like artistic masterpieces to regular pictures. This transfer of characteristics makes the current image look like the picture was painted by the artist himself. Neural Style transfer is efficient and fast for one-way conversion of an image to its style transferred form, but the image that has been transformed cannot be reverse engineered to its original form by any means. Also, artistic transformation of characteristics is spread throughout the image uniformly with random distribution. Style transfer is not specific to the main object of the picture.

In "Image-to-Image Translation with Conditional Adversarial Networks" [3], Phillip Isola and et al. implemented conditional adversarial networks for image-to-image translation for style transfer. These networks learn the mapping from input image to output image. This algorithm, named "pix2pix" is effective at synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images, among other tasks. Though pix2pix has extremely high accuracy and is used by many for style transfer of images which transfers only the characteristics of the main object, it does not provide any support for style transfer of unpaired images. This disadvantage is solved by Cycle-GAN based style transfer mentioned below.

In "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks" [4], Jun-Yan Zhu et al. developed an Image-to-image translation with mapping between an input image and an output image using a training set of aligned image pairs. They present an approach for learning to translate an image from a source domain X to a target domain Y in the absence of paired examples. The algorithm learned mapping G: X → Y such that the distribution of images from G(X) was indistinguishable from the distribution Y using an adversarial loss.

Figure 1 explains the architecture of CycleGAN. It contains two adversarial generator networks represented by letters G and F. X and Y are datasets of two unpaired image classes. There are two adversarial discriminators Dx and Dy which provide binary output for classification of its respective class. The loss is calculated over both the domains and hence is named cycle-consistency loss.

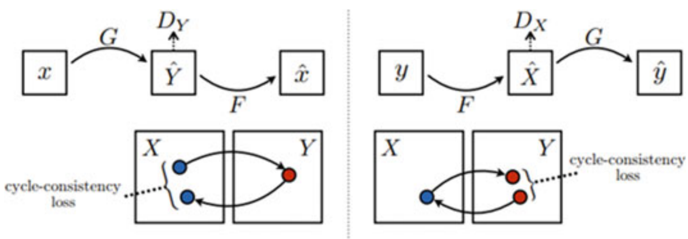


Fig. 1 Cycle generative adversarial neural network architecture

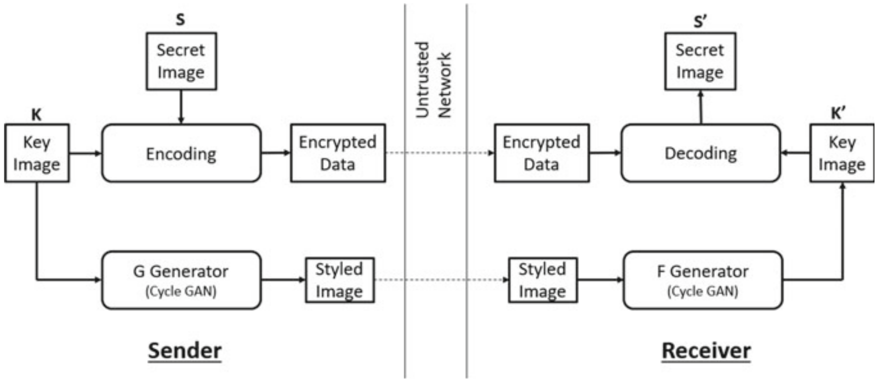


Fig. 2 Architecture diagram of simplex transmission

3 Overview

In this section the overview of our novel architecture is explained. The transmission process involves two parties: the sender and the receiver. A Cycle GAN network is trained on a certain training dataset till a good accuracy is achieved. Once a Cycle GAN network is trained sufficiently its weights and checkpoints are stored (Fig. 2).

The sender and the receiver load the frozen weights to their Cycle GAN architecture. The sender uses the G generator while the receiver uses the F generator. While it is not necessary for sender to load weights of F generator and for the receiver to load weights of G generator, but to make it a duplex transmission model in which both parties communicate by switching the role from sender to receiver and vice versa, both the generator weights can be loaded.

4 Methodology

4.1 Cycle GAN Working

Image to Image translation is the process of converting a given image into a new synthetic image with a specific detail inside the image being changed, for example converting a summer landscape to a winter landscape or converting an apple into an orange.

Cycle-GAN does not require paired image datasets for translation. Hence, ensuring that there is no need to have any domain related similarity between the style and the content images.

Cycle-GAN contains two adversarial generator networks represented by letters G and F. X and Y are datasets of two unpaired image classes. There are two adversarial discriminators Dx and Dy which provide binary output for classification of its respective class. The loss is calculated over the entire network in both directions i.e. x to y translation and y to x translation, hence the loss is called cycle-consistency loss and the network is named Cycle-GAN.

To ensure that the network does not generate the same image as the translated image, cycle consistency loss needs to be introduced into the architecture. For each image x from domain X, the image translation cycle should be able to bring x back to the original image, i.e.,

$$x \rightarrow G(x) \rightarrow F(G(x)) \approx x.$$

We call this forward cycle consistency. For each image y from domain Y, G and F should also satisfy backward cycle consistency:

$$y \rightarrow F(y) \rightarrow G(F(y)) \approx y.$$

4.2 Encoding Working

Encoding is done at the sender side. Consider two images, the secret image which will contain the data we want to hide and the key image which will contain the front image. In the first stage, the secret image and the key image will be superimposed over one another hence giving each pixel a value ranging from 0 to 510. In the second stage, the key image will be translated into another image form using the A-to-B part of the Cycle-GAN. This transformed data will be sent over the network where even if intercepted it will remain indiscernible as the key image has been transformed and it will be impossible to recover the secret from the superimposed image.

Multiple strategies can be employed for mixing the key image and the secret image, we will be using the superimposing technique which will obscure the secret to a large extent but shuffling the image pixels and then superimposing can also be done (Fig. 3).

4.3 Decoding Working

The receiver receives the encoded data from the sender via the untrusted network and decodes it. At the decoding stage the translated key image will be converted back to its original form using the B-to-A part of the Cycle-GAN. The B-to-A part of the network has been trained along with the A-to-B part of the network on the same dataset and hence will be the only network capable of doing so. This recovered key image will be subtracted from the superimposed image sent over the network to finally retrieve the secret image. The loss of the retrieved secret image will be equivalent to the cyclic loss of the Cycle-GAN (Fig. 4).

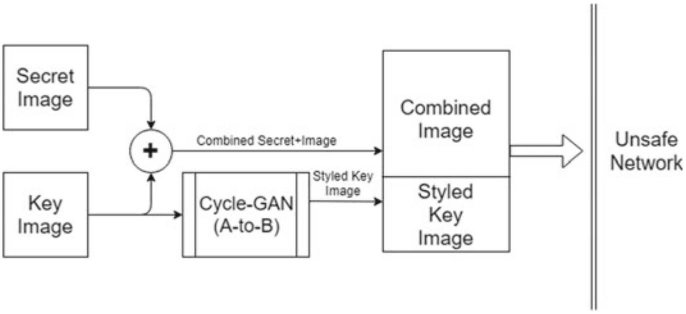


Fig. 3 Encoding process

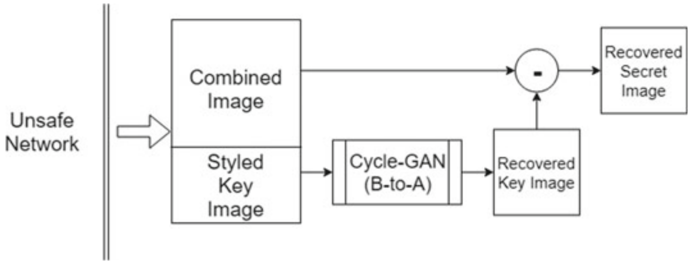


Fig. 4 Decoding process

5 Security and Future Scope

This method makes use of the Cycle-GAN model which makes it secure. Any attacker will need to have the access to the entire model architecture, data used to train the architecture and also the exact same model weights in order to successfully intercept the data and then decode the secret image. In order to make the communication extra secure, custom datasets can be generated and the models can be trained on these datasets.

If the attacker does intercept and try to extract the secret data, only distorted representations can be reproduced (Fig. 5).

6 Result

Figure 6 shows the result of our architecture's input and output with images generated in between the sender and the receiver. The peak signal to noise ratio, a very common image reconstruction metric, comes out to be 30.3 dB which is a minor loss equivalent to that of a general compression algorithm with minor changes to the image with regards to contrast and color hues [5].

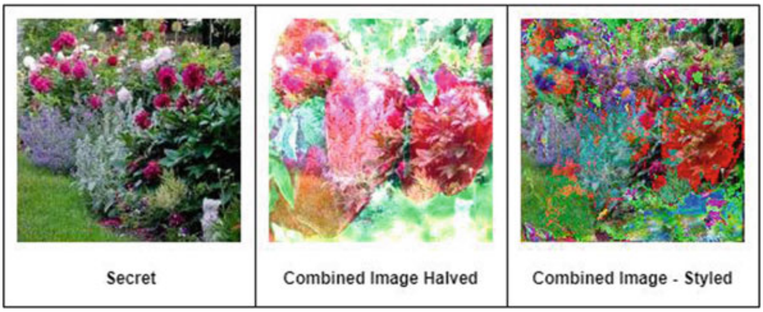


Fig. 5 Attacker failed attempt

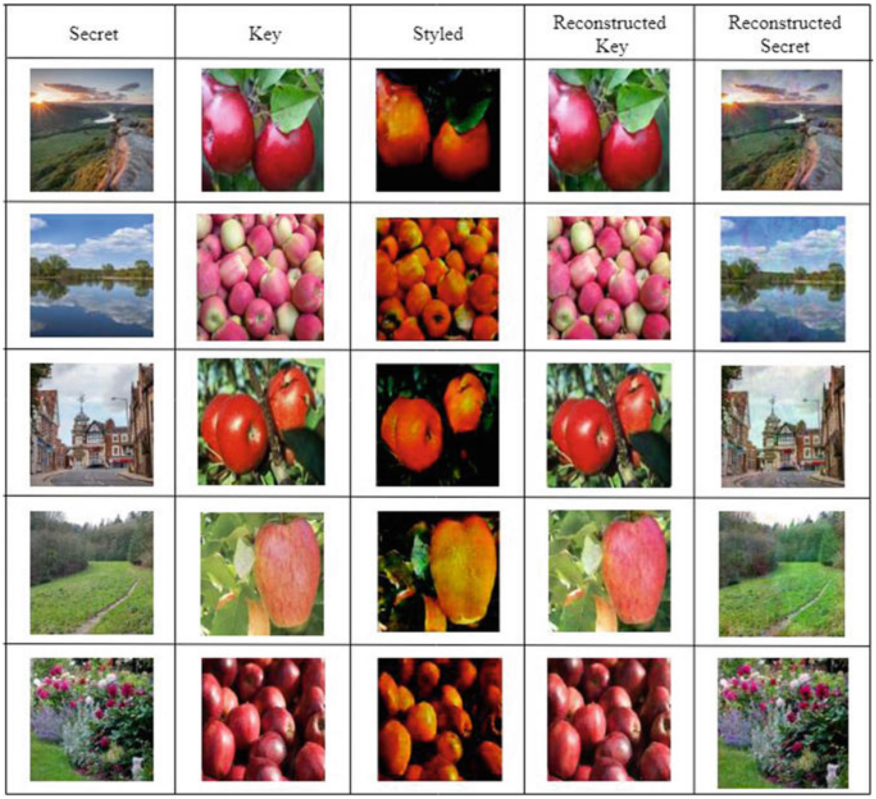


Fig. 6 Results

7 Conclusion

In this paper a style transfer based image cryptography algorithm was discussed. The results produced, obscure the images beyond recognition with the images merged just by superimposing each other. Using advanced methods to merge the secret and the key shall further improve the security of the implementation. Techniques like shuffling used by Prarthana Madan Modak et al. in “A Comprehensive Survey on Image Scrambling Techniques” [6] and advanced algorithms like AES for image encryption [7] can be used to merge the secret and key images, thus enhancing the security, such that, when the encoded data is tampered with, it becomes impossible to retrieve the secret image from the data without the model.

The proposed method for transmission of images via an untrusted network using style transfer and cryptography is secure and novel, thus providing an opportunity for research in the field of secure image transmission using style transfer technique.

References

1. Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks, arxiv 1703.10593
2. Gatys L, Ecker A, Bethge M (2015) A neural algorithm of artistic style. arXiv https://doi.org/10.1167/16.12.326
3. Isola P, Zhu J, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, pp 5967–5976. https://doi.org/10.1109/CVPR.2017.632
4. Zhu J, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE international conference on computer vision (ICCV), Venice, pp 2242–2251. https://doi.org/10.1109/ICCV.2017.244
5. Pradhan A, Sahu AK, Swain G, Sekhar KR (2016) Performance evaluation parameters of image steganography techniques. In: 2016 international conference on research advances in integrated navigation systems (RAINS), Bangalore, pp. 1–8. https://doi.org/10.1109/RAINS.2016.7764399
6. Modak PM et al (2015) A comprehensive survey on image scrambling techniques
7. Arab A, Rostami MJ, Ghavami B (2019) An image encryption method based on chaos system and AES algorithm. J Supercomput 75:6663–6682. https://doi.org/10.1007/s11227-019-02878-7

A Framework to Detect Hibiscus Flower Using YOLOV3 and SSD MobileNet

M. Mahesh, R. Rohan, V. Padmapriya, and D. N. Sujatha

Abstract The advancement in technology and intense usage of various devices like laptop, tablet, mobile phone etc. have reduced the existence of natural pollinators likes bees. This decline of bees is also attributed to various factors like high temperature due to climate change, extensive usage of fertilizers and insecticide to enhance the crop yield. The yield, type and the quality of crop depends to a great extent on the pollination In this context, the probable solution is to use technology to mimic pollinators. The identification of the Hibiscus flower is the very first step in deciding whether the flower is ready for pollination. In this paper, we have designed a framework for detecting Hibiscus flower and its parts through TensorFlow implementations for YOLOV3 and SSD MobileNet. The simulation study proves that the data set is appropriately trained in YOLOV3 and SSD MobileNet with the loss reducing in successive iterations. Finally, both training and validation sets converge justifying that the training is appropriate. The results also shows that the accuracy of both YOLOV3 and SSD MobileNet increases with the increase in the epochs.

Keywords Detection · Hibiscus · Image · SSD MobileNet · YOLOV3

1 Introduction

In today's agricultural world, the decline in the bee population due to harmful pesticides and fertilizers has reduced crop production and growth. Natural pollinators

M. Mahesh · R. Rohan · V. Padmapriya · D. N. Sujatha (✉)
Innovation Cluster Lab, Department of Computer Applications, B. M. S. College of Engineering, Bengaluru, India
e-mail: dns.mca@bmsce.ac.in

M. Mahesh
e-mail: maheshm.mca18@bmsce.ac.in

R. Rohan
e-mail: rohanr.mca18@bmsce.ac.in

V. Padmapriya
e-mail: padmapriyav.mca@bmsce.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_64

such as bees tend to pollinate flowers, thereby helping in large scale crop production. However, with their decline, skilled labour in farms have to be used. Manual work, however, consumes time and increases production cost [1]. Hand pollination also tends to be error-prone and can be difficult to accurately judge whether the flower is ready to be pollinated or not. To combat this, automation in agriculture or more commonly called Intelligent Agriculture has become popular in recent times. Identification and classification of the parts of the flower is the first step to understand whether a flower is ready to be pollinated. In this paper, we have used two popular object detection algorithms You Only Look Once (YOLOV3) [2, 3] and Single Shot Multibox Detector (SSD) [4] to identify the parts of a Hibiscus flower. In this work, we propose to build two models using YOLOV3 and SSD MobileNet to detect the parts of a Hibiscus flower predominantly the stigma and stamen which help in determining whether the flower is ready for pollination or not and we study the tradeoff between the algorithms in terms of loss and accuracy.

The reminder of the paper is organised as follows: Sect. 2 presents literature survey in the area of image detection. The proposed system model to detect hibiscus flower and its parts is discussed in Sect. 3. We elucidate the usefulness of our approach in Sect. 4 through simulation and performance analysis. Finally, we provide concluding remarks in Sect. 5.

2 Literature Survey

You Only Look Once is a state-of-the-art algorithm [5] where prediction of the objects is based on looking at the image only once and running a convolutional neural network on the same. YOLO is extremely fast and reliable. It predicts less false positives. However, YOLO struggles to detect small objects and small clusters accurately. The paper [6] discusses the algorithms like R-CNN, Fast R-CNN and YOLO. In real-time, it is found that YOLO provides a trade-off between speed and accuracy. YOLO has a better generalizing capability of representing objects which makes it ideal for applications that rely on robust object detection. The proposed work [7] deals with training the YOLO network with degraded images to improve the precision of object detection. The YOLO network which is trained with the degraded images extracts more features and has a better object detection accuracy. Ertam et al. [8] uses Tensorflow to classify MNIST dataset. The ReLu activation function achieves an classification accuracy of 98.43% on the test data. Nitin R Gavai et al. [9] shows the experimental performance of MobileNets model on Tensorflow platform to retrain the flower category datasets. It was found that the proposed model compromises on the accuracy. Gogul and Kumar [10] proposed a deep learning approach using Convolutional Neural Networks (CNN) to recognize flower species with high accuracy. A machine learning classifier such as Logistic Regression or Random Forest is used on top of it to yield a higher accuracy rate.

3 System Model

Figure 1 depicts the Proposed System Model for Detecting Hibiscus Flower. It consists of three stages: *Image Pre-processing and Data Cleaning*, *Image Annotation & Training*. The details of each phase is described below:

1. *Image Pre-processing and Data Cleaning*: This consists of two sub phases: Image data acquisition and Image data cleaning
   a. *Image Data Acquisition*: For this study, image data acquisition was reproduced by using a camera to obtain photos of Hibiscus flowers, while the rest of the images were obtained by scraping free stock sites such as Pexels and Unsplash. The total image database obtained during this phase comprises of 1230 images.
   b. *Image Data Cleaning*: The images were scraped from websites, sometimes contained other images not pertaining to the required class. Hence, data pre-processing techniques were used to remove these outliers. The images obtained through photos were cleaned to remove extreme distortions, blurriness and redundancy to reduce overfitting during training. The final image database was reduced to 1152 after pre-processing and the data cleaning phase.
2. *Image Annotation*: SSD algorithms are annotated using the Pascal VOC format whereas YOLOV3 uses the YOLO format. However, since we are using TensorFlow implementations of the algorithms, the images are annotated using the

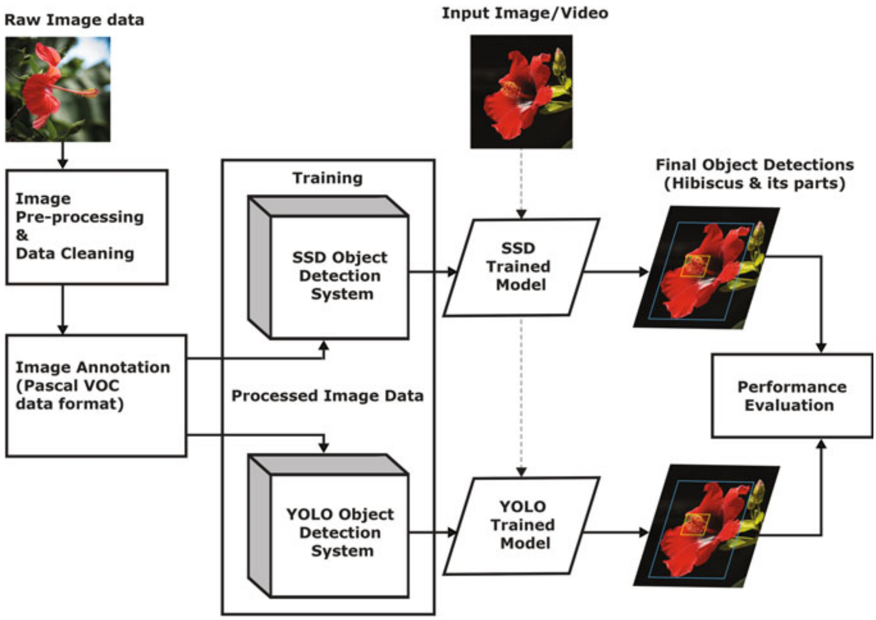


Fig. 1 Proposed system model for detecting Hibiscus flower

Pascal VOC format only. The Pascal VOC data format is an image labelling format where the image class is stored in terms of its height, width and depth as an XML file.

3. *Training*: The image data consists of a single class labelled as Hibiscus. Once the annotated images are obtained, the dataset is divided into 80% training and 20% validation set for training and testing the object detection models respectively. The training dataset is fed into the SSD MobileNet and YOLOV3 object detection systems where the training takes place. Once the training is complete, we obtain the SSD MobileNet and YOLOV3 trained TensorFlow models which can be served or exported using TensorFlow serving for production environments.

4 Simulation Study and Performance Analysis

In this paper, we demonstrate a framework for detecting Hibiscus flowers along with its parts specifically the stigma and the stamen through TensorFlow implementations for YOLOV3 and SSD MobileNet. The system is implemented based on Fast R-CNN for SSD MobileNet and YOLOV3 architecture.

Figures 2 and 3 show the loss of the Training and Validation Parameters using YOLOV3 and SSD MobileNet. This graph is a plot of number of iterations versus the percentage of loss. At 1000th iteration, the loss of both training and validation data sets was found to be 500% and at 14,000th iteration, the loss was as low as 10% for YOLOV3. Similarly for SSD MobileNet, at 2000th iteration, the loss of both

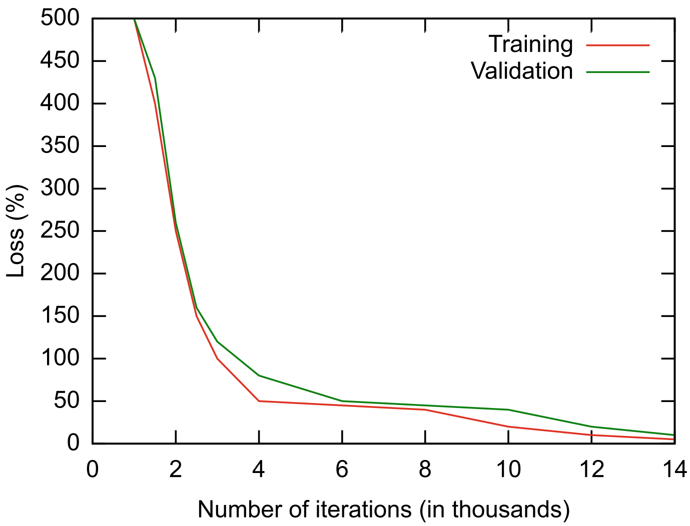


Fig. 2 Loss of the training and validation parameters using YOLOV3

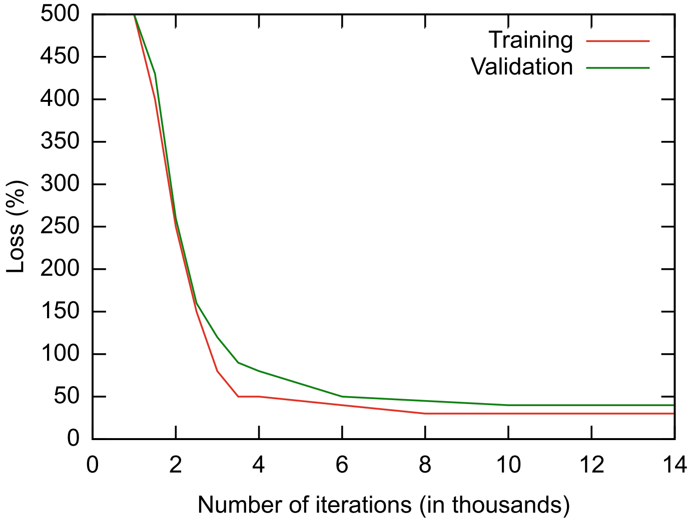


Fig. 3 Loss of the training and validation parameters using SSD MobileNet

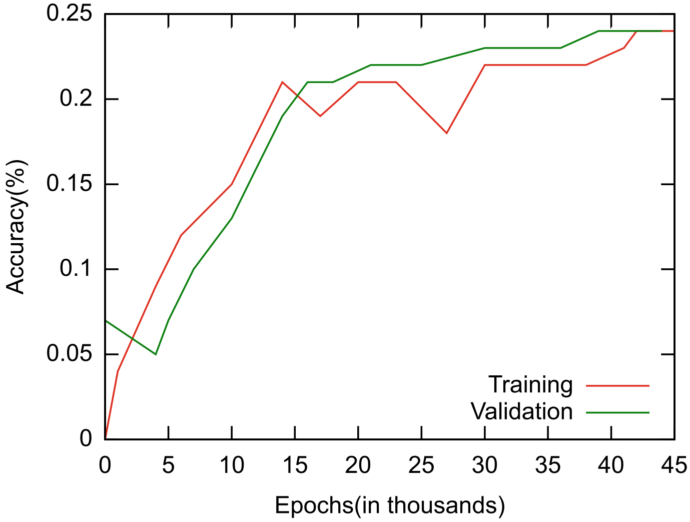


Fig. 4 Accuracy of the training and validation parameters using YOLOV3

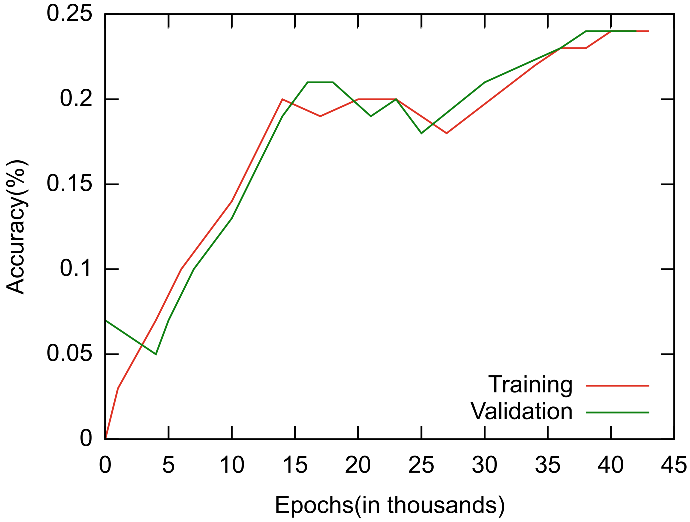


Fig. 5 Accuracy of the training and validation parameters using SSD MobileNet

training and validation was found to be 500%. However, at 14,000th iteration, the loss decreased to 30%. It is evident from the graph that as the number of iteration increase, the percentage of loss decreases. We can also see both the training and validation sets converge justifying that the data sets have been correctly trained using YOLOV3 and SSD MobileNet.

Figures 4 and 5 show the accuracy of the Training and Validation Parameters using YOLOV3 and SSD MobileNet. This graph is a plot of number of epochs versus the percentage of accuracy. Initially the accuracy of both training and validation parameters is 0% and 0.07% respectively. However, it is evident from the graph that the accuracy of both YOLOV3 and SSD MobileNet increases with the increase in epochs and converge from 45th epoch onwards for YOLOV3 and from 40th epoch for SSDMobileNet.

5 Conclusions

In this paper, we have a proposed a framework for detecting Hibiscus flower and its parts through TensorFlow implementations for YOLOV3 and SSD MobileNet. The simulation study indicates that the data set is appropriately trained in YOLOV3 and SSD MobileNet with the loss being reduced in successive iterations. Finally both training and validation sets converge justifying that the training is appropriate. We have also observed that the accuracy of both YOLOV3 and SSD MobileNet increases with the increase in the epochs.

References

1. Tyagi AC (2016) Towards a second green revolution. In: Conference on irrigation drainage. Wiley Online Library, pp 388–389. https://wileyonlinelibrary.com
2. Redmon J, Farhadi A (2017) YOLO9000: Better, Faster, Stronger. In: Computer Vision and Pattern Recognition, pp 6517–6525
3. Redmom J, Farhadi A (2018) YOLOv3: An Incremental Improvement. Computer Vision and Pattern Recognition
4. Wei L, Dragomir A (2016) SSD: Single Shot multiBox Detector
5. Redmon J, Divvala S, Girshick R, Farhadi A (2015) You Only Look Once: Unified, realtime object detection
6. Du J (2018) Understanding of object detection based on CNN family and YOLO. In: IOP Conf Ser: J Phys. 2nd International Conference on Machine, Vision and Information Technology
7. Liu C, Tao Y, Liang J, Li K, Chen Y (2018) Object detection based on YOLO network. In: IEEE 4th Information Technology and Mechatronics Engineering Conference, China, pp 799–803
8. Ertam F, Aydin G (2017) Data classification with deep learning using Tensorflow. In: 2nd International Conference on Computer Science and Engineering, Turkey, pp 755–758
9. Gavai NR, Jakhade YA, Tribhuvan SA, Bhattad R (2017) MobileNets for flower classification using Tensorflow. In: International Conference on Big Data, IoT and Data Science, pp 154–158
10. Gogul I, Kumar VS (2017) Flower species recognition system using convolution neural networks and transfer learning. In: Fourth International Conference on Signal Processing Communication and Networking, India, pp 1–6

Cricket Highlights Generation Using Video Information Retrieval and Feature Selection

Saicharan Gadamshetti and Rajeswari Sridhar

Abstract In multimedia domain, video processing consumes more time. In this paper, we dealt with improving video feature extraction with optimal resources for cricket highlight generation. These days, both domestic and worldwide broadcasting sports channels are day in and out streaming sports events. To watch full game alongside hectic work hours is tedious but to follow-up the game using highlights daily is something everyone is aligning towards. So, our project focuses on simulating latest technologies to generate cricket highlights. Our model considers both excitement and incident driven features to acknowledge and clip essential events in an exceedingly consistent fashion. Replays, scoreboard, bowler runner up sequence and audio intensity are samples of suggestions accustomed in seizing such events. To test our model, we administered a survey from user reviews to compare highlights of our clips with broadcasted clips. Projecting user approval rate and overlap between both highlight videos adds value of our framework in real time automated highlights generation domain.

Keywords Feature · Highlights · Replays · Audio · Scoreboard

1 Introduction

Continuous occurrence of global sports events every day, keeping up with every news is impossible for a sports fan. This fuels necessity of building systems that are efficient enough to summarize videos continuously, generating highlights domain has been solved to the most extent leveraging contemporary developments of computer vision, deep learning and machine learning. Producing highlights automatically can be either incident driven where rules of the game are taken into account or excitement driven where reactions from audience, commentators or even players are used. Incident-based features need to be used to extract 4 important events namely boundaries, milestones, wickets and sixes. In contradiction, excitement-based features extract

S. Gadamshetti (✉) · R. Sridhar
Department of Computer Science and Engineering, National Institute of Technology Tiruchirappalli, Tiruchirappalli, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_65

audio intensities of the match to classify. Incident based features can be extracted using replay detection and text extraction in a video, to contrast excitement driven features leverages on audio classifiers.

Motivation: To generate Cricket Highlights Generation with desired output in optimal time with maximum precision and minimal resources. Our framework focuses on excitement and incidents features to summarize video. Our primary motivation for this problem statement is that there are many features present in cricket video which might be extracted to get better highlights instead of training on high dimensional CNN and RNN.

Contribution: This project is being done in two phases, in the first phase, identification of the topic, research work on the topic, dataset preparation and a simulation is being carried out. In the second phase the implementation part of the project is carried out. The overall system can further be used in various cricket videos to generate highlights.

Organization: This paper is organized as follows. Section 2 provides glimpse of the related research work. Section 3 presents the Proposed System Architecture. Section 4 contains Implementation, Performance Evaluation and results. Conclusion of paper is done in Sect. 5.

2 Related Work

Automated video summarization uses following features namely audio intensity, camera motion, replay detection, contextual info like segmenting shots and positions of the cricket ball. In [7], Tang proposed multi hierarchical model that relies on HOG (Histogram oriented gradients of histogram) and CH (Histogram colours) to detect crucial events in cricket match. These detected events were classified using HMM (Hidden Marcov Model). In [8], Namudhri introduced MPEG-7 to summarize videos. Entire cricket match was divided into video clips (shots that belongs to same camera take). Elicited key-frames were clustered into 5 different classes with respect to field view. Hidden markov model is used to generate complete event using sequences of above generated key-frames. Drawback of this approach is not segmenting video automatically but diving manually and continuity of event is missing in final video. In [9] Kolekar proposed model based on audio intensity levels to extract crucial events. Hypothesis here is that important events are generally followed by crowd cheer (Increment in audio energy). This approach fails when crowd cheers leading to false highlight generation. In [10], Sengupta formed highlights from a match by viewing low-level concepts and high-level features using apriori radical algo. But here they focused on wicket fell concept-mining. With comparison to the literature, our framework dealt with more generalized aspects of the match considering all events. As each approach did not arrive at prompt results, we took hybrid approach to enhance the results.

3 Proposed System

Our methodology as described in Fig. 1, involves 2 levels. In first level video segmentation is done using audio cues. Second phase extracts low level features in cricket video namely replay detection, Scoreboard extraction, eliminating ground and crowd view. Initially large video is skewed to excited clips removing unwanted clips and low-level features are extracted from each video clip and assigned a label (event recognition) which are later used to classify it as highlight or not.

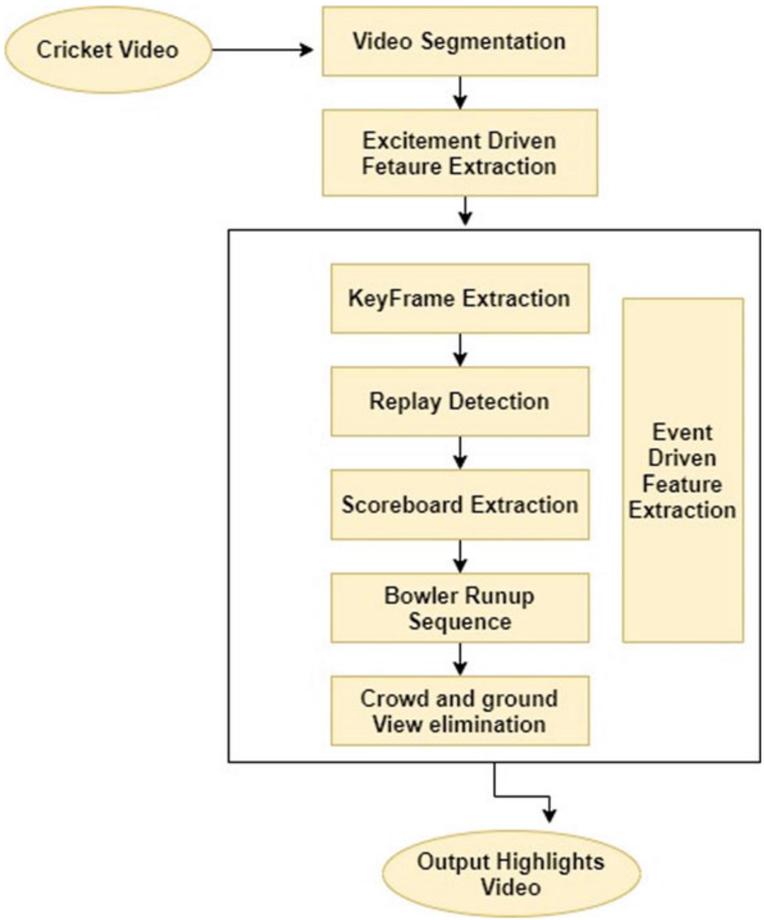


Fig. 1 General overview of the system

Fig. 2 Start and End time of each event after merging audio chunks having energy value greater than threshold

	energy	start	end
0	232.247	140	160
1	350.412	375	395
2	268.751	980	995
3	185.992	1050	1070
4	200.013	1105	1115
5	210.945	1140	1155

Table 1 Proposed algorithm for video shot segmentation

Input: Full Cricket Match Video
Begin
Step 1: Extract audio from input video using librosa library
Step 2: Divide audio into 5 second chunks
Step 3: Compute short-time energy of each chunk
Step 4: Classify each chunk as excited or not based on threshold
Step 5: Merge consecutive excited chunks into clips
Step 6: As audio energy raises only after batsman hits. So to cover up bowler run up include 5 seconds to start time of each clip
Step 7: Extract video subclips from video based on start time and end time using subclip function in moviepy editor
End
Output: Excited video clips

3.1 Video Shot Segmentation Using Audio Cues

This phase involves segmenting video into video clips and classify each video clip as excited or not. We can also divide video using pyscene detect library but obtained clips do not represent entire sequence (from bowler runup to batsman hitting). Using trial and error method, we got optimised threshold value as 180 (Fig. 2 and Table 1).

3.2 Keyframe Extraction

After segmenting the videos, key-frame extraction is identifying frames that have complete representation of a video shot preserving salient features while removing redundant frames. Using Katna library we have extracted keyframes from each video clip. Each video clip generates roughly around 8–10 frames. So instead of processing each frame to extract low level features, it is sufficient to process only keyframes. Keyframes are generated for each video clip obtained from the excitement feature

Table 2 Proposed algorithm for replay detection

Input: 400 images dataset of both classes (replay and on-going), keyframes of each clip ***Begin*** **Step 1**: Load pre-trained VGG19 model using Keras.applications and remove last two fully connected layers **Step 2**: Feed each image into VGG19 and get vectors for each frame **Step 3**: Train SVM using obtained vectors with 70% dataset for training and 30% for testing **Step 4**: Save the trained model using joblib **Step 5**: load the model and predict each keyframe as replay or not ***End*** ***Output***: Label each keyframe as replay or not

extraction module. Keyframes directory is built where each clipname is folder name which contains keyframes of that clip.

3.3 Replay Detection

Replay in a sports video is referred as a footage which is played more than once generally after important event. It plays important role in predicting key events Replays can be detected by observing that the absences of scoreboard in all portions of the video. Transfer learning is inculcated to determine a frame as current event or replay/advertisement using VGG19 and Support Vector Machine (SVM) framework. Last two fully connected layers of VGG19 is removed and vectors obtained from it are fed to a SVM which is trained using 400 images of two classes. Offset for training is set to zero (Table 2).

Replay frames are eliminated from above obtained clips. To eliminate replay frames in a clip, say 'clip1'. if we find at least three successive keyframes as replay, then eliminate corresponding frames between those keyframes. After this module we get replay free excited video clips (Fig. 3).

3.4 Scoreboard Extraction

Scoreboard is used in predicting boundaries, wickets and sixes. Accuracy is most important here to find key-events We built a CNN model consisting of four layers and trained it using MNIST dataset. This model is used to predict digits when we give digit image as an input. Figure 4 shows step by step output of scoreboard extraction (Table 3).

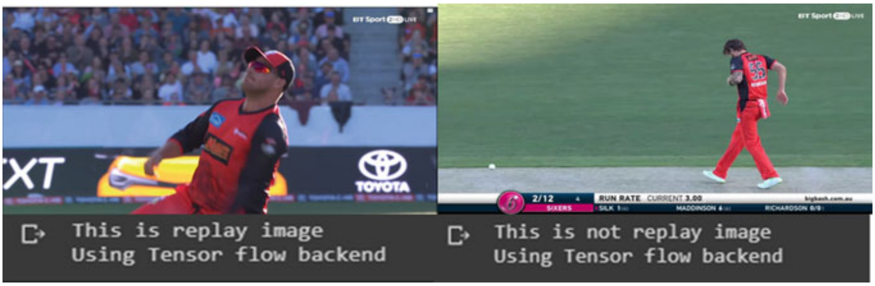


Fig. 3 Results of correctly labelling a replay and a non-replay image

Fig. 4 Results of a self-built CNN model for extracting the score and wickets

Table 3 Proposed algorithm for scoreboard extraction

Input: cricket image
Begin
Step 1: Crop scoreboard
Step 2: Denoising image using cv.fastNlMeansDenoisingColored
Step 3: Gray Scale conversion
Step 4: Ostu binary thresholding
Step 5: Separate score and wicket image using Numpy arrays
Step 6: Feed digit images to self built CNN model and predict score and wickets
End
Output: Extract Score and Wickets

3.5 Player Runup Detection

We used the Yolov3 model and achieved significant results. This module takes a frame as an input and recognises it as bowler runup or not. Steps Involved:

(1) Feed input frame to Yolov3 model
(2) It outputs a frame which contains detected objects (name, id, probability)
(3) If number of persons detected in any frame is in between 4 and 6 then we can classify it as bowler runup sequence.

Fig. 5 Detecting the number of players in a given frame using Yolov3

This module is useful to check whether a video clip contains bowler runup sequence or not to ensure video clip begins with bowler bowling the ball. So, it removes clips which started from the middle of the event (Fig. 5).

4 Crowd View Detection

Edge Pixels percentage (PEP) can be used to classify frame as audience or close-up. As pixels that consists of audience have more edges. Canny edge detector is applied and calculated the ratio of edge pixels to the total number of pixels. If video clips contain any crowd view, then it must be removed since it is not important event. Algorithm for crowd detection:

(1) RGB image is converted into YCbCr.
(2) Canny edge detector is applied to find edge pixels and compute PEP.
(3) If PEP > 0.8 then frame belongs to class crowd Else frame belongs to class close-up

4.1 Ground View Detection

The frames of RGB format are converted to HSV format. The hue component of the converted HSV images plotted. Grass pixel ratio is found if hue segment lies in range of 30 to 60 for an individual pixel. If GPR > 0.9 then it can be classified as field view, so it must be removed.

4.2 Highlights Generation

All the keyframes generated from the clips that are produced after eliminating replays are taken and labelled in a database as shown in Fig. 6.

1	ClipName	FrameNumber	Score	Wickets	PlayerCour	GroundView	CrowdView
2	clip1	frame1	3	3	3	0	0
3	clip1	frame2	3	3	4	0	0
4	clip1	frame3		3	7	0	0
5	clip1	frame4	2	0	7	0	0
6	clip1	frame5	2	0	7	1	0

Fig. 6 Frames labelled with mentioned attributes

Finally, the database is processed to get the highlights by comparing the score and wickets field if there is a difference of important events like four, six or wicket. The player count field is analysed and if it contains more than 5 players it refers to starting of the ball and it should be included in the highlights since highlights start with beginning of every ball. Finally, crowd view and ground view frames must be removed from video clips and thus the highlights are generated.

5 Performance Analysis

To measure the performance improvement, after enhancing the model the following evaluation metrics are used.

a. Official video highlights comparison with generated video
   Here we compared generated highlights of our framework with broadcasted highlights. We used metrics like overall accuracy, precision and recall.
b. Event Recognition
   Here we also tested performance of our framework to check whether important events are detected or not. 4 events namely wickets, sixes, boundaries and milestones of players were assessed and mean average precision for every event is calculated. A confusion matrix is computed to evaluate the accuracy of classification as highlight or not and is given in Tables 4 and 5 for 2 matches.

From Table 4, the following is the calculation of Precision, Recall, accuracy and F-score for match 1.

Precision $=$ **91%**
Recall $=$ **86%**
Accuracy $=$ **80%**

Table 4 Confusion matrix for boundaries prediction in match 1

Actual	Predicted	
	Positive	Negative
Positive	19	2
Negative	3	Nil

Table 5 Confusion matrix for boundaries prediction in match 2

Actual	Predicted	
	Positive	Negative
Positive	17	1
Negative	2	Nil

Table 6 Mean average precision of each event

EVENT	MAP (%)
Boundaries	93
Sixes	95
Wickets	92

f1-score = **89%**

Match 2 contains 18 boundaries out of which 17 are included in our highlights and 1 is not recognised. Other 2 clips are included in our highlights which are not actually boundaries.

Precision = **95**%
Recall = **90**%
Accuracy = **85**%
f1-score = **93%.**

Therefore, Mean average precision (MAP) of boundaries for both the matches = **93%**

Similarly, we computed MAP for events like sixes and wickets and is given in Table 6.

Summarization Effectiveness for Match 1:

Input video size: 52 min.
Output Video size: 4 min.

Recognised 19 boundaries, 3 sixes and 3 wickets out of 21 boundaries, 3 sixes and 4 wickets with Mean average precision of 93% for boundaries, 95% for sixes and 92% for wickets.

6 Conclusions and Future Work

Scoreboard Extraction using OCR accuracy is low as it is trained on wide set of images, so we built our own CNN model and trained with MNIST dataset. Unlike other models we built iterative algorithm that extracts very possible event and excitement driven information. Replay detection is done using transfer learning (VGG19 and SVM) whose results are accurate and time complexity is less comparatively.

Most of the sports video is garbage, which is eliminated in initial phases using audio cues, so video processing complexity decreases largely. Ball by Ball video segmentation is done based on bowler run-up sequence using yolov3. All in all, our model is compatible in terms of results and computation power. Future work includes addition of important events like cricketer celebrations, reaching milestone, fan's emotional moments etc. Gathering more data sets for testing purposes for comparisons between the original ones and those generated by our model. Ongoing work includes deploying same model to other sports and trying to improve results.

References

1. Nasir M, Javed A, Irtaza A, Malik H, Mahmood M (2018) Event detection and summarization of cricket videos. J Image Graphics 6:27–32
2. Shih HC (2018) A survey of content-aware video analysis for sports. IEEE Trans Circuits Syst Video Technol 28(5):1212–1231
3. Chakraborty S, Thounaojam D (2019) A novel shot boundary detection system using hybrid optimization technique. Appl Intell 49:03
4. Goyani M, Dutta S, Gohil G, Naik S (2011) Wicket fall concept mining from cricket video using a-priori algorithm. Proc Int J Multimedia
5. Zhang K, Chao W-L, Sha F (2016) Video summarization with long short-term memory. In: European conference on computer vision. Springer, pp 766–782
6. Soomro K, Zamir AR (2014) Action recognition in realistic sports videos. In: Computer vision in sports. Springer, pp 181–208
7. Tang H, Kwatra V, Sargin ME, Gargi U (2011) Detecting highlights in sports videos. In: Multimedia and Expo (ICME). IEEE, pp 1–6
8. Namuduri K (2009) Automatic extraction of highlights from a cricket video using mpeg-7 descriptors. In: Communication systems and networks and workshops, 2009. COMSNETS 2009. First International. IEEE, pp 1–3
9. Kolekar MH, Sengupta S (2008) Caption content analysis based automated cricket highlight generation. In: Communications conference, pp 461–465
10. Kolekar MH, Palaniappan K, Sengupta S, Seetharaman G (2009) Semantic concept mining based on hierarchical event detection for soccer video. JMM 4(5)

Design of Secure Biometric System Using Cancelable Techniques

Aarti Laxman Gilbile and Pramod D. Ganjewar

Abstract Individual identification can be accurately done by measuring biological parameters termed as bio-metrics. These have been proved as an unusual tool for identity verification. Recognition of biometrics and applications based on it is increased tremendously, so the privacy protection monitors are raising the privacy concerns. For reducing the privacy threats, the research work has increased to find methods to protect the biometric data. Security of biometric template is that the most challenging aspect of biometric identification system. The biometric data are stored as it is within the database which increases the rate of compromising it. This can also cause serious threat or misuse of the individual's identity. This paper proposes simple and a unique approach to store a biometric sample within the sort of template by using Cancelable Biometrics.

Keywords Cancelable biometrics · Revocable · Biometric sensor · Multi-biometrics · Biometric security · Template protection

1 Introduction

Authentication system based on biometrics has a many advantages than regular authentication systems with token which uses PINs, passwords, identification cards and user IDs. Authentication system based on biometrics uses different biometric traits such as fingerprints, palm prints, face, iris, voice, keystroke, signature, etc. Every user has their unique and permanent biometric characteristics. These traits can never be forgotten or never get lost unlike password or identification cards. However, if database of biometric authentication system is stolen then recovery of system is not possible due to above mentioned characteristics of biometrics. Hence,

A. L. Gilbile (✉) · P. D. Ganjewar
MIT, Academy of Engineering, Alandi, Pune, India
e-mail: algilbile@mitaoe.ac.in

P. D. Ganjewar
e-mail: pdganjewar@comp.maepune.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_66

for providing protection to the biometric data, biometric template protection is important. In this, biometric characteristics are arranged with a systematic adjustment for securing actual biometric information. In case, if transformed biometric template is stolen, then by changing the transformation parameter, the new template for user is formed and this newly generated template restores the stolen template.

The primary goals of cancelable biometrics are:

1. *Non-reversibility*: Even after performing large number of computations, it should not be possible to regenerate the actual biometrics from the biometric template.
2. *Accuracy*: All transformations made during template generation will not affect the quality of fingerprint recognition.
3. *Diversity*: Different applications should not use single biometric template. This will trigger revocability, from which new template may be introduced when compromise occurs.

2 Related Work

Kaur et al. [1] addressed security concerns during the outstanding use of biometric systems. Cancelable biometric is a technique where the user's real biometric data is converted into a pseudo-biometric data. This pseudo-biometric data is stored and it is used while verifying the user also. Usage of pseudo-identity decreases the loss of privacy and in the event of compromise it allows revocability. They proposed a technique called as 'Random Distance Method' which generates the pseudo biometrics and reduces its size to 50%. They executed this method on unimodal and multimodal data generated from biometric data like palm-vein, palm-print, face for analyzing recognition and protection performance. Here, Euclidean distance is calculated between the feature points and some random point and generates a new transformed vector which is considered as transformed template. For the template produced with this method the performance is improved.

Wu et al. [2] proposed many cancelable biometric techniques for electrocardiograms (ECG) to decrease the threats. The inherent biological and their intrinsic life signs of ECGs make them very difficult to forge or steal. R peaks of the ECGs are detected and they are considered as feature. Different biometric templates associated with the ECGs of a person can be built using the principle of "signal subspace collapse" so that a compromised template like password can be revoked. To recognize the identity of any person just with his/her ECGs, the popular multiple signal classification technique is used. Unlike other cancelable techniques, the identification can be carried out without understanding the transformation of the distortion. Because of this, now the recovery of original ECGs from their template is very difficult. They have tested various experiments on real ECGs of 285 subjects.

Jin et al. [3], for template protection, proposed a hashing based on ranking which is motivated by two-factors one is cancelable biometrics and other is dubbed "Index-of-Max" (IoM) hashing. This hashing transforms an individual vector function of biometric into a hashed code with discrete index using some parameters which are randomly generated. From this hashing notion, they demonstrated two realizations based on hashing schemes, Uniformly Random Permutation and Gaussian Random Projection. This system has some merits; it ensures strong concealment to biometric information, more robust against biometric feature variations.

Gomez-Barrero et al. [4], for protecting biometric templates, international standards regarding protection of biometric data has decided two main specification, unlinkability and irreversibility. This paper addressed a new method for determining the unlinkability of biometric template. To demonstrate the efficiency of the approach, the unlinkability is applied over four methods for protection of biometric template: Bloom filters, biometric salting, block re-mapping and Homomorphic Encryption. To show its advantages, the proposed framework is equated with other existing methods.

Jin et al. [5], a generalized point to string conversion mechanism for fingerprint information has been published on the basis of kernel learning techniques to produce specific strings of binary with fixed-length. They are: (a) Minutiae description extraction. (b) A method of kernel transformation that is based on kernelized hashing to produce a vector of fixed width. (c) Binarizarion (Segmentation). (d) To match. The good experimental outcomes using datasets from FVC2002 & FVC2004 describe the posibility of proposed framework in cases of template randomness, accuracy of matching & proficiency.

Lee et al. [6], has initiated a new methodology to yield cancellable fingerprint templates for which alignment becomes unnecessary. For each minutia, they calculate a constant value of rotation and translation. Based on the orientation information, the adjacent local areas surrounding the minutia are considered to get the value. This value is used as reference which composed of two major operations that include shift- that outputs two values the said minutia in the cancelable example, for the translation and rotational motions of these values respectively. Replacement by a new template is done that generates various functionalities. It provides the function to alter when a template consists of corrupted data values. The strategy they use is to retain, withhold the primary geometric interconnections (i.e., rotation & translation) between the enrolled models and query or request after transformation is performed. Thus, the templates which are transformed can be operated to verify on the person without having input fingerprint images. The two major criteria that help in analyzing our experiment results are success and changeability. Evaluations of results support us to analyze how the accuracy is met during transformation of models. The difference between original and new transformed template are been tested. After testing, the variations in both the templates collected from the same initial fingerprint given as input. The experimental outcomes show that these two major criteria has affected one another in same way and also can be controlled by considering diversified controlling parameters for changing the functionality.

Kelkboom et al. [7], proposed FRR improvement method consists of the use of Different biometric instances during the enrollment or the verification process. The noise is that, and sum of errors bit and HD also decreased. The number of samples is collected experimentally without taking the basic results entirely into account. They propose a Gaussian analytical method for evaluating the output of the biometric device in binary that specifies the amount of instances used during the process of measuring and checking. The machine efficiency is approximately assessed by the error-detection trade-off curve that incorporates FAR and FRR. The analytical expressions will be tested using the biometric databases Face Recognition Grand Challenge v2 and FVC2000. All the related work is also briefly mentioned in the Literature survey table below (Table 1).

3 Proposed Method

Wong's [8] multi-line code algorithm is referred here. Proposed system (Fig. 1) will try to simplify the complexity of Multi-line code generation step. Also proposed system will try different algorithms than Wong's system during feature extraction steps. The fingerprint verification system has following four stages: extraction of minutiae, generation of line code, permutations in line code, matching of fingerprint.

3.1 Extraction of Minutiae

The minutiae extraction includes following five stages:

Segmentation: In the fingerprint image some part is only the actual fingerprint which is known as the region of interest. Segmentation technique is used to select the actual fingerprint by excluding the background.

Contextual filtering: To improve the fingerprint image the contextual filtering technique is used. Noise is removed by conserving the fingerprint structures. Here, Gaussian filter is used for contextual filtering.

Binarization: This method is used to transform the gray scaled actual image to a binary image in black and white.

Thinning: The fingerprint ridges are reduced to one pixel long using image thinning method which makes next move easier. Here, zhang-suen thinning algorithm is used for thinning the image.

Minutiae detection: Here the bifurcation points and endpoints of the ridges are identified. These points are used as minutiae i.e. feature points.

3.2 Generation of Multi-line Code

After extraction of minutia, Fig. 2a will illustrate Wong's system where a straight-line is drawn with length l through the selected minutia. Then take sample points on

Table 1 Literature survey

S. No.	Paper title	Biometric data types	Database	Methodology	Observations	Cons
1	Harkeerat Kaur et al. "Random Distance Method for Generating Unimodal and Multimodal Cancelable Biometric Features", IEEE Transaction March 2019 [1]	Face Palmprint Palmvein Fingervein	CASIA- Face V5 IRIS CASIA Palmprint CASIA-MS V1 CASIA-MS V1 SDUMLA-HMT	Random Distance Method	Transformed feature vector is created by calculating the Euclidean distance between random point and feature points	For two different persons, distance between feature point and random point may be same which will cause ambiguous condition
2	Shun-Chi Wu et al. "Cancelable Biometric Recognition with ECGs: Subspace-Based Approaches", IEEE Transaction, May 2019 [2]	ECG	Physikalisch Technische Bundesanstalt (PTB)	Signal subspace collapsing and Multiple signal classification (MUSIC)	Signal subspace collapse methods used for template generation and Multiple signal classification method is used to determine person by ECG	Due the changed physiological conditions, inter-beat morphological differences are inevitable and this may affect the accuracy of identification
3	Zhe Jin et al. "Ranking Based Locality Sensitive Hashing Enabled Cancelable Biometrics: Index-of-Max Hashing", IEEE Transaction, 2018 [3]	Fingerprint	FVC2002 (DB1, DB2, DB3) FVC2004 (DB1, DB2, DB3)	Index-of-Max (IoM)	Using IoM the real valued feature vector is changed to max ranked hashed code	There may be probability of hash collision

(continued)

Table 1 (continued)

S. No.	Paper title	Biometric data types	Database	Methodology	Observations	Cons
4	Marta Gomez-Barrero et al. "General Framework to Evaluate Unlinkability in Biometric Template Protection System", IEEE Transaction, June 2018 [4]	IRIS, Signature, Face Fingervein	Multimodal Biosecure Database UTFVP database	Explored unlinkability	Unlinkability is applied over various template protection techniques	System works only by using Lebesgue functions
5	Zhe Jin, Meng-Hui Lim et al. "Generating Fixed-Length Representation From Minutiae Using Kernel Methods for Fingerprint Authentication", IEEE Transactions, 2016 [5]	Fingerprint	FVC2002 FVC2004	Kernal based framework	Point-to-string converting framework used to generate fixed length binary string	In training stage, the longest time will be taken for creation of kernel matrix
6	Chulhan Lee et al. "Alignment-Free Cancelable Fingerprint Templates Based on Local Minutiae Information", IEEE transactions [6]	Fingerprint	FVC2002 DB1	Alignment free cancellable technique	For two changing functions, a uniform value is used as input which will give two values as output as rotational and translational movement of actual minutia	The false accept rate can be increased due to misuse of user's PIN and stealing

(continued)

Table 1 (continued)

S. No.	Paper title	Biometric data types	Database	Methodology	Observations	Cons
7	Emile J. C. Kelkboom et al. "Binary Biometrics: An Analytic Framework to Estimate the Performance Curves Under Gaussian Assumption", IEEE transactions [7]	Face Fingerprint	FRGC V2 FVC2000 DB2	Gaussian analytical framework	The performance of biometric system is calculated with Gaussian analytical framework and error detection curve is used to calculate FAR and FRR	Performance may decrease due to capturing multiple fingerprints

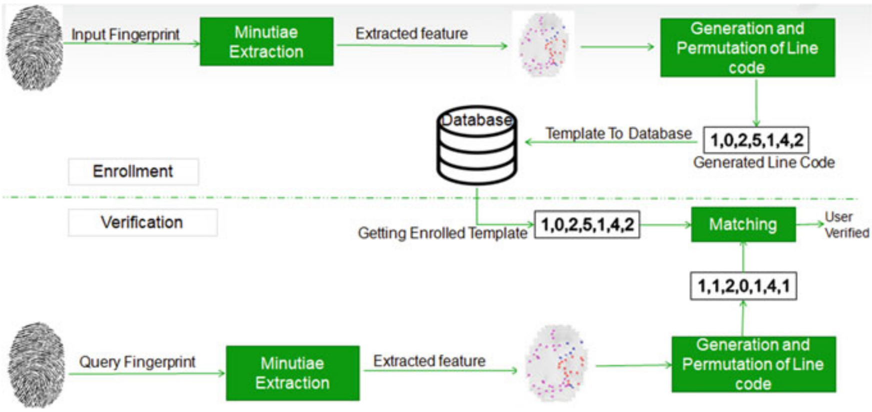


Fig. 1 System architecture

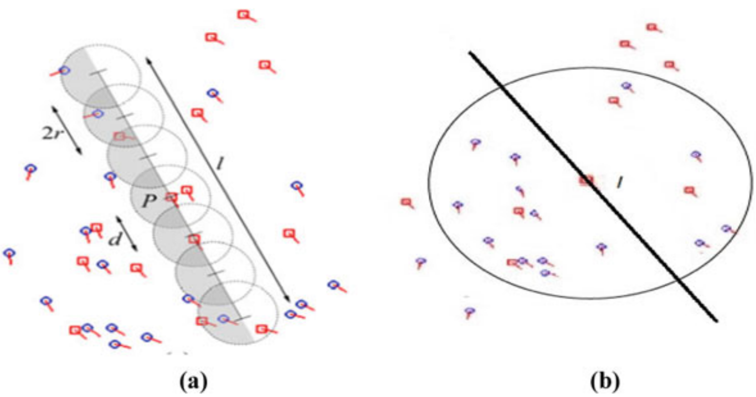


Fig. 2 **a** Plane view of Wong's system. **b** View of proposed system

the line. For each sample point, draw a circle through it for obtaining the total count of minutiae inside the area of circle. For generating a line code, the total minutiae in each circle are arranged sequentially that describes the reference minutia. Figure 2b will illustrate the proposed system where instead of drawing multiple circles on a line, simply draw a single circle. And then total minutiae in this single circle will be used to generate the line code.

3.3 Permutation in Multi-line Code

For achieving revocability, template diversity, add an external factor that will create wide variations into the created line code. Simple permutations are done in line code using a user precise secret key. Permutation order of the minutia code will be changed by the secret key. This secret key is produced using a pseudo-random number generator (PRNG). It's necessary to make sure that the same key should not be allocated to two persons or two applications of same person. Therefore, random numbers should be generated by the PRNG without duplicate values.

3.4 Fingerprint Matching

Consider TF as template fingerprint, QF as query fingerprint where T is total count of minutiae in TF and Q is total count of minutiae in QF. Lt is a line code for the template stored in database and Lq is a line code for query fingerprint. While matching fingerprint, each and every line code from QF will be matched with each line code from TF.

4 Conclusion

Technique like cancelable biometrics has the potential to boost the security and confidentiality of a traditional biometric system. We have seen different algorithms of cancelable biometrics applied on different biometric data.

The researchers are putting more efforts to implement new approaches to reduce the complexity of cancelable biometrics and more secured, along with the improved performance of the biometric system. Variations made in multi-line code algorithm will simplify the complexity along with revocability and diversity to the generated template. This will also reduce template size than the actual fingerprint.

References

1. Kaur H, Khanna P (2019) Random distance method for generating unimodal and multimodal cancelable biometric features. IEEE Trans Inf Forens Secur 14
2. Wu S-C, Chen P-T, Swindlehurst AL, Hung P-L (2019) Cancelable biometric recognition with ECGs: subspace-based approaches. IEEE Trans Inf Forens Secur 14
3. Jin Z, Lai Y-L, Hwang JY, Kim S, Teoh ABJ (2018) Ranking based locality sensitive hashing enabled cancelable biometrics: index-of-max hashing. IEEE Trans Inf Forens Secur 13
4. Gomez-Barrero M, Galbally J, Rathge C, Busc C (2018) General framework to evaluate unlinkability in biometric template protection system. IEEE Trans Inf Forens Secur 13
5. Jin Z, Lim M-H, Teoh ABJ, Goi B-M, Tay YH (2016) Generating fixed-length representation from minutiae using kernel methods for fingerprint authentication. IEEE Trans Syst Man Cybern Syst
6. Lee C, Choi J-Y, Toh K-A, Alignment-free cancelable fingerprint templates based on local minutiae information. IEEE Trans Syst Man Cybern B 37(4)

7. Kelkboom EJC, Molina GG, Breebaart J, Veldhuis RNJ, Kevenaar TAM, Jonker W, Binary biometrics: an analytic framework to estimate the performance curves under Gaussian assumption. IEEE Trans Syst Man Cybern A Syst Humans 40(3)
8. Wong WJ, Wong MLD, Kho YH, Lin Q, Wen S, Multi-line code: a low complexity revocable fingerprint template for cancelable biometrics. IEEE Trans Inf Forens Secur

Image Captioning Using Capsule Neural Network and LSTM

Bharat Sharma, Ashwini Sapkal, AShiva Krishna, Rahul Chauhan, and Pankaj Solanki

Abstract Machine translation and extracting meaningful language descriptions has always been a challenging task in the computer vision field. This paper presents a very modern approach which uses deep learning techniques to generate image description. The concept of Capsule Neural Network is applied in Image Captioning. A capsule denotes a nested layer inside the capsule. One major difference between traditional networks and Capsule Neural Network is that the Capsule Neural Network handles better segmentation and recognition. Evaluation of the model on Flickr 8 k data-set qualitatively and quantitatively shows the proper learning and language efficiency. For model evaluation a very popular metrics BLEU score has been considered which shows significant improvement from 28.9 to 37.8 for the latest Flickr 8 k data set.

Keywords Language modelling · Image description generation · Image captioning · CNN (capsule neural network) · LSTM (long short-term memory)

1 Introduction

One of the most challenging and continuously evolving problems in the Computer Vision community is image classification and object localization. Since the time when the image classification was introduced, communities have been inclined towards generating automatic descriptions from the images [1]. Automatic caption generation from just an image is very challenging but it can be a great benefit to the visually impaired people and can provide them with a better understanding of the surroundings around them. Captioning an image or tagging it with description not only deals with identifying or localization of objects and classifying them instead it deals with finding relations between the distinct objects. Semantic understanding i.e. language modelling is a very important aspect which needs to be considered while building an image captioning model (Fig. 1).

This paper introduces a novel approach for automatic image captioning. Traditional Image captioning process deals with the Convolution Networks to generate

B. Sharma (✉) · A. Sapkal · A. Krishna · R. Chauhan · P. Solanki
Army Institute of Technology, Pune, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_67

Fig. 1 Caption of an image generated by the proposed model [6]

a vectoring representation of the localized objects and Recurrent neural network or Long short-term memory (LSTM) for generating natural image descriptions [2, 3]. But the proposed approach focuses on the first fold of the process which is the use of Capsule neural networks instead of generic Convolution Networks. Convolutional Neural Networks have been the most state of the art for feature extraction methodologies since decades but the recent advancement and discovery of Capsule Networks has set the bar higher and not to mention the challenges along with it. Usually, the image-captioning process deals with the generation of depiction which must address every single object contained in the frame, likewise, it should communicate how these objects identify with one another just as their characteristics and the exercises they are associated with [4–6]. Also, the above semantic information must be communicated in a natural language like English, which implies that a language model is required notwithstanding visual comprehension. To address this issue of preserving the orientation and nature of the interaction of each localized object with each other we have used Capsule network which completely eradicates the mentioned issue. Preserving orientation here refers to the complete identification of the different orientation, shapes, sizes and regions. Use of Capsule Neural network which addresses the issue of the orientation of the images that ate fed to the network hence gives us a more robust model.

The paper presents a Capsule Neural Network that takes an input image as an input and uses LSTM to create its portrayal in the text. Giving semantic attention using LSTM decoder which takes input as an embedded feature vector and a start sequence which helps the decoder to generate the caption recursively till it finds an end sequence. The investigations show that the captions produced reasonable subjective forecasts. The model is trained on Flickr 8 k data-set gathered correspondences and assesses its presentation on region-level annotations. In Sect. 2, the literature survey is carried out which focuses on the current state of the art in this domain. In

Sects. 3 and 4, our proposed approach is discussed along with the results. The article is concluded at the end.

2 Literature Survey

This part shows the significance of Long short-term memory, recurrent neural network, and Convolutional neural network in image caption generation. As of now, many strategies have been tried for programming the image captioning. In 2014, researchers from Google launched a paper [7] that describes the image captioning. Previously this structure was once based on the MSCOCO data-set. It accepts picture as input and produces significant sentences as a yield with the assistance of Convolutional neural network CNN alongside Recurrent neural network (RNN) and Long short-term memory (LSTM). The dense feature vector is made by utilizing the Convolutional neural network CNN and this dense feature vector also referred to as an embedding feature, is used as an input into other functions, programs or networks. For an image caption generation model, this embedding becomes an important representation of the image and is used as an initial stage of the LSTM. A lot of work has been done on the image to sentence generation. The huge work in tackling the image to sentence generation was done by Ali Farhadi [8]. Here three spaces are characterized in a particular way: the picture space, the meaning space and the sentence space. The mapping is done to the meaning space via a particular picture and sentence space. With the assistance of mapping, the likeness between the pictures and the sentence is assessed, the implications are put away as triplets of (picture, activity, object) and a score is assessed by anticipating the picture and sentence triplets. In the event that an image and sentence have an elevated level of similarity as far as the predicted triplets then they will be exceptionally perfect and have a high score. Thus, appropriate sentences can be produced. This model has numerous drawbacks such as the necessity of the centre meaning space and the outcomes obtained from it are not in the slightest degree exceptionally accurate. Various different works were presented however later work utilized the technology of neural networks for understanding the journey. With the approach of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), a good execution was accomplished and discovered applications in different fields of study. In [9] O-Vinyals and group, in the work, presented a novel methodology using (CNN) and (RNN) for picture subtitling undertakings. Convolutional neural networks were utilized to remove highlights from the pictures. In this way, CNN goes about as an encoder, first for the arrangement of errands and the last layer as yield provides the contribution to (RNN). RNN goes about as a decoder that creates sentences using LSTM systems (Long Short Term Memory) which was the sort of RNN used. A Similar methodology was utilized by Donahue et al. [10] which utilized LSTM as on recordings rather than pictures. Works proposed a model on which detectors were prepared to extricate different highlights out of the pictures and a translation model was prepared on a lot of inscriptions to produce the fitting descriptions for the current picture. In true situations there is commotion or clutter present in the pictures so not at all like the conventional strategies, not all the features are taken care of into the (RNN) however just the significant and striking highlights are fed into the (RNN).

Geoffrey Hinton, a world-leading British-Canadian researcher specializing in artificial neural networks and a significant figure within the history of deep learning, introduced a totally new kind of neural network called a Capsule Neural Network [11]. This did not fail to arouse the curiosity and enthusiasm of the bogus intelligence research community. Capsule networks will definitely transform the abilities and capabilities of data science in the field of artificial intelligence or neural nets. After going through a long way finally in late 2019 comes the term CapsNet which stands for Capsule network. According to the recent papers based on CapsNet, it reduces the error by more than 45 in detecting some of the images. This will be very helpful because it gives a solution to the problem of orientation with which the Convolutional neural networks and machine learning always struggled with. The idea behind the CapsNet is one of the obvious directions where Deep Learning falls short [12]. Presently, with the introduction of the dynamic routing calculation [13] and the expectation–maximization routing algorithm [14], capsule nets have been executed with striking starting results. This has improved the accuracy of image captioning. The utilization of a capsule net for the image captioning generation task, the point of convergence of this paper, is a difficult issue. In any case, the comprehension between capsules is achieved by using the dynamic routing computation that could help in the generation of the captions of complex pictures.

3 Methodology

Language model used in this paper is similar to the most recent methods in machine translation. Initially, a Capsule net with three-layer is used through which the features are extracted. These features are taken as input into an LSTM encoder, where the LSTM network is a recursive network that learns a portrayal of the first sentence and produces partial captions until the complete caption is created in an iterative way. The benefits of utilizing LSTM is that the model design above is that it can handle a sequence of subjective length, whereas in a conventional neural network the length is fixed. Taking inspiration from these approaches, this paper proposes an extension where given an image encoding decoding is done. Word embedding has been used which maps the word generated in each iteration to a higher-order space which then is used for finding the similarity to other words.

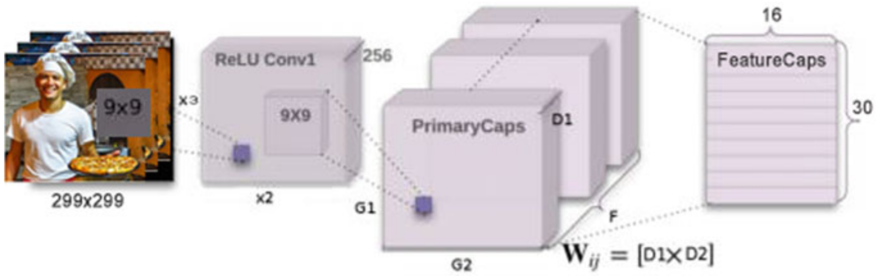


Fig. 2 The architecture of Proposed CapsNet

3.1 Overview of the Proposed Model

The model uses a pipeline of encoder-decoder where the Capsule neural network acts as an image encoder which takes an input image, converts it to a feature vector and then passes the output to the LSTM which acts as a sequence generator.

3.2 Proposed Capsule Neural Network

The model can be divided into two sections. The first is the encoder part and the section is the decoder part. In the encoder, the part is inspired by the architecture proposed in [11]. The Encoder section is divided into 3 layers. The first layer is the ReLu convolutional layer. The convolutional network accepts an image of size 299 $\times$ 299 as input and has 256 kernels with the size of 9 $\times$ 9 $\times$ 1 and stride one followed by ReLu activation. The main reason to choose the ReLu activation function was to eliminate the problem of vanishing gradient. The second layer is Primary caps layer and has dimensions G1, G2 as (6, 8) which depend on the input, and F = 32 primary capsules, and whose function is to detect features which were perceived by convolutional layers and then bring forth a combination of features. The layer 3 is called a feature layer which takes the input from the previous layer and assigns them a (D1 $\times$ D2) 8 $\times$ 16 weight matrix, that maps 8-dimensional input space to the 30-dimensional capsule output (Fig. 2).

3.3 Lstm

LSTM is a memory cell where information is stored in a time tamp manner for the data seen at each respective timestamp that follows the model used in [15]. LSTM precisely uses three gates which decide whether to consider or to forget the existing value of that cell (forget gate f), if the state of forget gate is off then it should use information from (input gate i) and whether to yield the value of the new cell (output gate o). Letting σ be the sigmoid nonlinearity used, where for each time step t LSTM refreshes, given inputs x_t, h_{t-1}, c_{t-1}, where the output and cell updates are as per the following:

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1}) \tag{1}$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1}) \tag{2}$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1}) \tag{3}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot h(W_{cx}x_t + w_{cm}m_{t-1}) \tag{4}$$

$$m_t = o_t \cdot\ c_t \tag{5}$$

$$p_{t+1} = \mathrm{Softmax}(m_t) \tag{6}$$

where W matrices are trained weighted parameters and where '·' means the product with a gate value. In conventional recurrent neural networks, there are problems which include vanishing gradients but in LSTMs because of these multiplicative gates it is comparatively easier to train the LSTM as these gates handle exploding and vanishing gradients [16]. The nonlinearities are involved in sigmoid σ and hyperbolic tangent h activation functions. The last Eq. 6 is used as a softmax function. Softmax is used for probability distribution as it gives the probability of occurring over all the words. The long term dependency at different timestamps are preserved because of the forget gates in LSTM. The order of the sentence is protected in sequence generation due to the long term dependency (Fig. 3).

3.4 Combining Encoder and Decoder

This output from capsule neural network is re-scaled to linear space and then feed- ed into an LSTM, which concatenates the input vector and the word embedding along with the position of the word embedding into a concatenated vector, and passed to LSTM. For the output of LSTM dropout = 0.25 and batch normalization is used to avoid overfitting and thus improve optimization. The output of the softmax layer is a vector which indicates the most probable word, which is mapped to a word dictionary and then generates the resulting caption. This caption is again feed-ed to the LSTM as word embedding along with position (POS) and the input features, and this continues until the LSTM gets the end of a sentence or a threshold is reached.

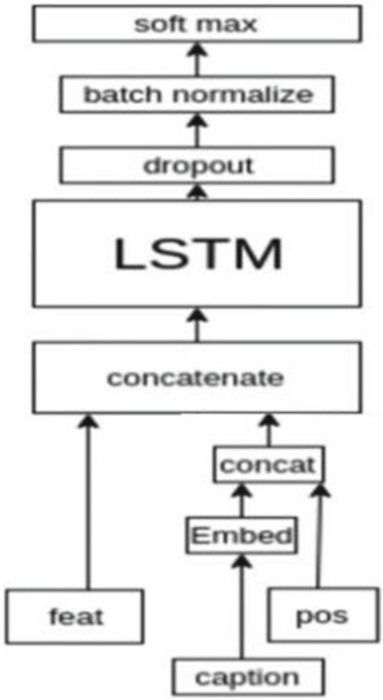


Fig. 3 The architecture of LSTM [15]

4 Results Experiments

4.1 Visual Results

The model was assessed on its result upon the training set of Flickr 8 k data-set [17]. This data-set has 8000 images, which is broken into 3 sets, a training set of 6000 picture samples, a validation set of 1000 picture samples and a test set of 1000 picture samples where every image is accompanied by 5 captions to train and test upon (Figs. 4 and 5).

From the above two generated captions it is evident that how the proposed model can understand the spatial differences in the given image and correctly identifies the in one image a man is standing “besides” the tree and in another image the object is in “front” of the tree and thus can learn from the images at training time to produced better results than traditional convolutional neural networks which totally ignores the spatial differences in an image [10].

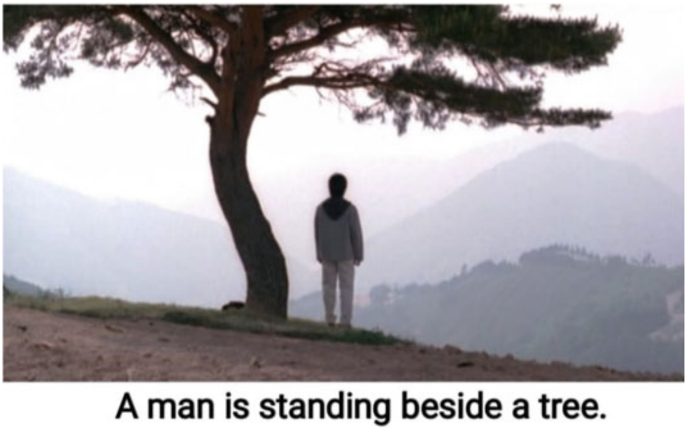


Fig. 4 Caption generated by the proposed model [17]

Fig. 5 Caption generated by the proposed model [17]

4.2 Statistical Results

The proposed model gave results of a 37.8 BLEU score on this data-set. The Bleu-1 score was 66.7, Blue-2 was 46.6, Bleu-3 was 33.7 and Bleu-4 was 19.5. An individual

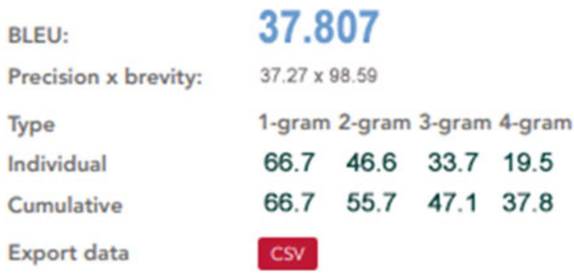


Fig. 6 37.8 BLEU-4 points of the proposed model

N-gram score is an assessment of simply matching grams of a particular order, for example, single words (1-gram) or word sets (2-gram or bi-gram). These cumulative BLEU scores are calculated taking into consideration individual n-gram scores at every order from 1 to n, weighting them by computing the weighted geometric mean [18]. Every stage of the proposed implementation was done utilizing the Tensor-Flow library for Python [19] (Fig. 6).

The table below shows a comparison between the proposed and existing architectures.

Metric	BLEU-1	BLEU-2	BLEU-3	BLEU-4
Mao et al. [20]	56.5	38.6	25.6	17.0
Jia et al. [21]	64.7	45.9	31.8	21.6
Proposed	**66.7**	**46.6**	**33.7**	**19.5**
Xu et al. [22]	67.0	45.7	31.4	21.3
Qi et al. [23]	74.0	54.0	38.0	27.0

To avoid the discrepancy between training and evaluation criteria in typical text generation models, BLEU metrics are used to calculate accuracy. By making use of BLEU which is Bilingual Evaluation Understudy Score we are avoiding the calculation complexity of traditional policy gradients. The matrix shows the various existing models with individual BLEU scores [20–23]. Our proposed architecture shows a Bleu-4 of 19.5 and overall cumulative BLEU of 37.80 which proves to be a good model for caption generation. We have not included the recent SOTA methods like OSCAR [24], CutMix [25], Unified VLP [26] are trained on MSCOCO and due to unavailability of resources the proposed model could not be trained on such a large data-set.

5 Future Work

The proposed neural network can be combined with current state of art language modelling techniques like GPT-3 to achieve better results [27] in addition to this a large data-set like MSCOCO [28] can be used to train the base model and achieve best results.

6 Conclusion

The paper proposes a novel architecture combining the power of Capsule network which was first proposed by Hinton [11] and the powerful architecture of LSTM. This has improved the accuracy of image captioning and achieved a Blue score of 37.807. The utilization of a capsule net for the image captioning generation task is the point of convergence of this paper. The paper acquires satisfactory results and presents a comparison with 4 other famous results. After the analysis of hyperparameters, the model presented in this paper was able to generate a sensible caption for the Flick8k data-set.

References

1. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2014) ImageNet large scale visual recognition challenge
2. Kulkarni G, Premraj V, Dhar S, Li S, Choi A, Berg AC, Berg TL (2011) Baby talk: understanding and generating simple image descriptions. In: CVPR
3. Farhadi A, Hejrati M, Sadeghi MA, Young P, Rashtchian C, Hockenmaier J, Forsyth D (2010) Every picture tells a story: generating sentences from images. In: ECCV
4. Hodosh M, Young P, Hockenmaier J (2013) Framing image description as a ranking task: data, models and evaluation metrics. J Artif Intell Res
5. Young P, Lai A, Hodosh M, Hockenmaier J (2014) From image descriptions to visual denotations: New similarity metrics for semantic inference over event descriptions. In: TACL
6. Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollar P, Zitnick CL (2014) Microsoft coco: common objects in context. arXiv preprint arXiv:1405.0312
7. Vinyals O, Toshev A, Bengio S, Erhan D, Show and tell: a neural image caption generator, arXiv:1411.4555
8. Farhadi A, Hejrati M, Sadeghi MA, Young P, Rashtchian C, Hockenmaier J, Forsyth D (2017) Every picture tells a story: generating sentences from images. In: Proceedings of the 11th European conference on computer vision: Part IV, ECCV'10. Berlin, Heidelberg, Springer-Verlag, pp 15–29
9. Vinyals O, Toshev A, Bengio S, Erhan D (2014) Show and tell: a neural image caption generator. CoRR, abs/1411.4555
10. Donahue J, Hendricks LA, Guadarrama S, Rohrbach M, Venugopalan S, Saenko K, Darrell T (2014) Long-term recurrent convolutional networks for visual recognition and description
11. Hinton GE, Krizhevsky A, Wan SD (2011) Transforming auto-encoders. In: ICANN. Springer (2011)
12. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. In: Advances in neural information processing systems, pp 3859–3869
13. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. In: NIPS
14. Hinton GE, Sabour S, Frosst N (2018) Matrix capsules with EM routing. In: ICLRW
15. Zaremba W, Sutskever I (2014) Learning to execute. arXiv:1410.4615
16. Hochreiter S, Schmidhuber J (1997) Long short-term memory. A Neural Comput.

17. Hodosh M, Young P, Hockenmaier J (2013) Framing image description as a ranking task: data, models and evaluation metrics. J Artif Intell Res 47:853–899
18. Papineni K, Roukos S, Ward T, Zhu W-J (2002) BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting of the association for computational linguistics (ACL), Philadelphia, July 2002, pp 311–318
19. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Kudlur M (2016) TensorFlow: a system for large-scale machine learning. In: OSDI, vol 16, pp 265–283
20. Mao J, Xu W, Yang Y, Wang J, Huang Z, Yuille A (2015) Deep captioning with multimodal recurrent neural networks (m-rnn). In: International conference on learning representations (ICLR)
21. Jia X, Gavves E, Fernando B, Tuytelaars T (2015) Guiding the long- short term memory model for image caption generation. In: Proceedings of the IEEE international conference on computer vision, pp 2407–2415
22. Xu K, Ba J, Kiros R, Cho K, Courville A, Salakhudinov R, Zemel R, Bengio Y (2015) Show, attend and tell Neural image caption generation with visual attention. In: International conference on machine learning, pp 2048–2057
23. Qi Wu, Shen C, Wang P, Dick A, van den Hengel A (2018) Image captioning and visual question answering based on attributes and external knowledge. IEEE Trans Pattern Anal Mach Intell 40(6):1367–1381
24. Li X, Yin X, Li C, Zhang P, Hu X, Zhang L, Wang L, Hu H, Dong L, Wei F, Choi Y, Gao J (2020) Oscar: object-semantics aligned pre-training for vision-language tasks. eprint 2004.06165, archivePrefix=arXiv
25. Yun S, Han D, Oh SJ, Chun S, Choe J, Yoo Y (2019) CutMix: regularization strategy to train strong classifiers with localizable features. eprint 1905.04899, archivePrefix=arXiv
26. Zhou L , Palangi H, Zhang L, Hu H, Corso JJ, Gao J (2019) Unified vision-language pre-training for image captioning and VQA. eprint 1909.11059, archivePrefix=arXiv
27. Brown TB, Mann B et al (2020) Language models are few-shot learners. eprint 2005.14165 archivePrefix=arXiv
28. Yi T, Michael L et al (2015) Microsoft COCO: common objects in context. arXiv:1405.0312v3 [cs.CV] 21 Feb 2015

Problem Solving Techniques Using Ant Colony Optimization in Computational Intelligence

Manjunath R. Kounte, E. Niveditha, Kalaigar Afrose, and A. Sai Sudeshna

Abstract Over the years, technology has played a crucial role in the advancement of the society. In achieving human like intelligence in machines, computational Intelligence plays an important role due to development of software algorithms imitating biological inspired intelligence behavior. In this paper we have briefed about that the evolution of computational intelligence over the years, starting from Turing Test to the present-day Swarm Intelligence. Also, the importance of CI and AI is highlighted by considering some major approaches like Fuzzy systems, Artificial Neural Networks, Evolutionary computation, Perception, Natural language processing, cognitive thinking. In this paper, the analysis of travelling salesman problem along with vehicle routing problem is presented using the approach of ant colony optimization. Finally, the applications of CI in the areas of space applications, driving cognition, and Bioinformatics are described.

Keywords Cognition · Pheromone · Swarm intelligence · Ant colony optimization · Evolutionary computation · Stigmergy

1 Introduction

Ever Since the Science Fiction term has evolved for "Robots" involving machines that possess human-like intelligence, there is a lot of curiosity about how to make machines more intelligent. Intelligence is concerned with performing mental tasks such as logical reasoning, solving problems, applying skills and decision making. The world is changing more rapidly that traditional Computational approaches are inefficient in solving complex real-world issues as they have been preprogrammed to provide solution to a specific problem. Hence, there is a need to develop machine intelligence that can solve complex real-world tasks very easily.

M. R. Kounte (✉) · E. Niveditha · K. Afrose · A. Sai Sudeshna
Dept of Electronics and Computer Engineering, School of ECE, REVA University, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_68

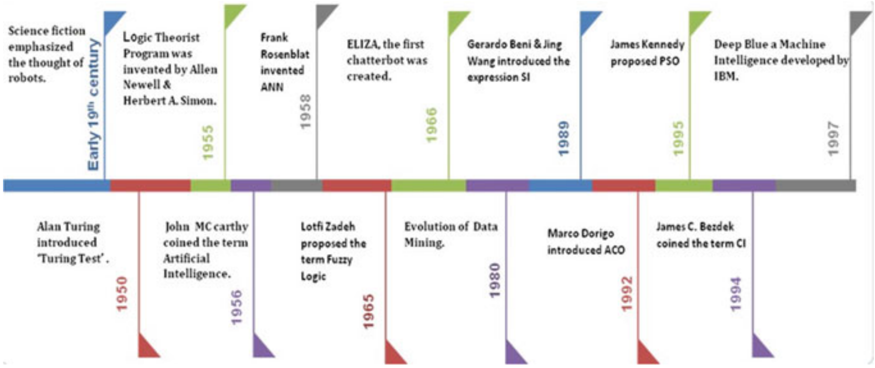


Fig. 1 Evolution of computational intelligence

For this we need to take inspiration from nature where different living organisms have different levels of intelligence to survive and adapt in hostile environments. One such technique where machine intelligence can be developed is through Computational Intelligence (CI) approach.

Computational Intelligence is a newly emerging technology formed by integrating many diverse models such as neural networks, fuzzy logic, machine learning etc., to achieve the goal of creating intelligent system to analyze, understand and model a solution (decision making) that results in adaptive mechanisms to sustain in a complex changing environment. CI technique utilizes cognitive thinking approach which is generally inspired by natural and biological models and develops their structure, continuously adjusting and modifying its knowledge representation. The term Computational knowledge was first utilized by the IEEE neural system committee in1990. Yet, the main clear meaning of computational knowledge was coined by James C. Bezdek in 1992 [1]. The History and Evolution of CI is described in Fig 1.

2 Artificial Intelligence and Computational Intelligence

Over the last few years, there has been research going on artificial intelligence and computational intelligence. Is there is any distinction between them? Both the approaches are aimed at simplifying complex tasks to make human life easier, however, when we see from a broader perspective there is a clear line separating them i.e. the methods and models implemented to achieve the goal are different. CI has decision making ability, even when inappropriate information is provided, whereas in AI it takes a decision only when large amount of information is provided as input [2]. CI and AI approach includes five branches as shown in Fig. 2 respectively. Artificial Neural Network is a special case of neural computation, formed by interconnecting large number of processing units in a layered fashion that resembles human brain consisting of neurons. There can be multiple layers for computation,

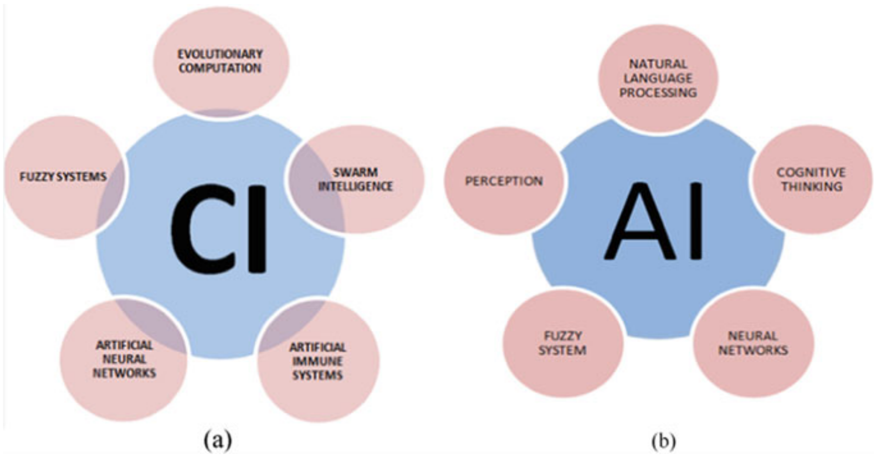


Fig. 2 **a** AI branches. **b** CI branches

the artificial neuron receives one or more inputs (representing dendrites) and sums them to produce an output [3].

SI is a branch of CI where models are mostly derived from observing natural behaviors of animals, birds that tend to live in flocks and groups, which have enabled them to learn, adapt and survive over generations running into millions of years [4]. In flocks and groups, decision making is based on collective opinion of every member's individual knowledge [5].

Fuzzy logic is the major paradigm commonly used in AI as well as CI to represent vague and imprecise information, as its procedure dependent on the degree of truth or false rather than perfectly true or false.

Evolutionary computation approach is inspired by biological evolution; it gives the best possible solution for optimized problem. This approach is based on inheritance property and the objective is "Survival of fittest".

Artificial immune system is nature inspired methodology and used to solve to solve singular and multi-objective problems. The representation and procedure that the brain uses to depict our general surroundings at algorithmic level is called perception [6]. It is basically a semantic process in the overall objective of implementing artificial intelligence. The perception can be for image data, auditory or video data.

Cognitive computing is a subbranch of AI that is focused on introducing human way of understanding and problem solving into computers so that complex situation can be easily and efficiently handled. Unstructured data is increasing in the digital world, thereby limiting our ability to analyze and take decisions.

The cognitive computing can leverage the capability of computer to infer, understand huge amounts of data and extract the knowledge contained in the data thereby improving the reasoning and decision making.

Natural language processing is a important technique that fill the communication gap between a machine and a human being. Humans speak and write wide variety of languages it is difficult for a computer to understand and compute the required task, which paved for the intervention of Natural language processing (NLP). NLP interprets human language by analyzing trends in grammar, syntax rules slang and there by converting unstructured data into structured that can be analyzed to extract meaningful data [7].

3 Ant Colony Optimization

Ant Colony Optimization (ACO), is one of swarm intelligence model which was introduced by Marco Dorigo [8, 9]. It is a population-based metaheuristic approach for solving classical traveling salesman problem, path planning, network routing, function optimization problems, combinatorial optimization problems and so on.

ACO approach is derived from an interpretation of real ants and their self organizing strategies used to locating food supplies, build nests, defending against predators without a centralized leader. Most categories of Ants don't have sensory organs to hear or to see; there communication is dependent on sharing of information through the chemical released by individual ants called as pheromone which is a form of stigmergy. The quality and the distance of the food source from the nest determine the type and amount of pheromone released. The trail of pheromone acts as road map for ants to reach the food source. The path having highest concentration is the shortest path and eventually all ants tends to follow the shortest path. Based on the above behavior of real ants, an ACO algorithm is designed, where artificial ants cooperate and communicate with each other through memory distribution to find a solution for a given real-world problem by depositing pheromone (weights) on each edge of the graph..

3.1 Travelling Salesman Problem (TSP)

TSP is the most successful application of ant colony optimization algorithm. This is applied in complex computation problems to find/optimize the best route in terms of distance as well as cost [10]. Problem statement: let's consider a sales man who has to deliver 5 parcels to 5 different locations (cities) $V = \{a, b, c, d, e\}$. Provided the distance between each city with every other city *(A=* $\{i, j\}$*: i,,j* ϵ *v,* d_{ij}*length of the edge),* we have to find the shortest route which connects all cites only once.

The number of cities and the distance between each city to the every other city are given as inputs, based on this information a virtual map is fabricated consisting of vertices and edges resembling cites $\{i \in V, i = (x_i,\ y\text{i})\}$ and the corresponding distances (d_{ij}) between the them respectively. Each edge has a weight which is intern dependent on the pheromone and heuristic value (inverse of the length of the edge).

Algorithm:
[1] Creation of weighted pheromone model for
V= {a, b, …e} with m ants randomly placed at n vertices
[2] Initialize equal pheromone values on each edge, set t=0
[3] While convergence condition is not met (start of ant based solution construction)
[4] While N(i) is not empty
i. Selection of next vertex using Probability State transition rule by n ants in the set N(i)
ii. determine N(i) end while
[5] Calculate length L_k of solution by each ant
[6] Update the pheromone trail
using the Eq. 2
[7] End while

Pheromone value depend on the quality of the solution built by the individual artificial ants and are updated after each iteration dynamically. The weight of the pheromone on each edge decreases uniformly with time [11].

A set of artificial ants{k} (software agents) start randomly at one vertex and move along the edges of the graph by selecting the subsequent vertices based on probabilistic state transition rule given by (1) which not visited till that point of time to complete one iteration.

$$p_{ij}^{k} = \begin{cases} \frac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}]^{\beta}}{\sum_{u \in N(i)}[\tau_{iu}]^{\alpha}[\eta_{iu}]^{\beta}} & if\ j \in N(i) \\ 0 & otherwise \end{cases} \tag{1}$$

where, τ_{ij} is the pheromone value of the corresponding path; η_{ij} is the heuristic visibility of the edge; α, β are adjustable parameters that relate pheromone trail and heuristic value; t is the iteration number; $N(i)$ are the set of cities are vertices that are not visited previously.

The probability of each eligible or feasible vertex is calculated by taking the ratio of product pheromone value and heuristic visibility of one vertex in the numerator and sum of the product of pheromone and heuristic value of all the vertices in the allowed set N(i). After all the ants complete one iteration the pheromone values are updated using the equation given in (2). The formula mathematically depicts both deposition and evaporation of pheromone trail on each edge travelled by m ants.

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t) \tag{2}$$

$$\Delta\tau_{ij}^{k} = \begin{cases} \frac{G}{L_K} & if\ (i, j) \in tour\ length\ by\ Kth\ ant \\ 0 & else \end{cases} \tag{3}$$

L_k is the total length of the solution by kth ant,
G is an optional arbitrary parameter.

TSP algorithm is proven to be one of the best methods to find the shortest path compared to other greedy algorithms, which make local optimal choices. The greedy technique terminates in less number of steps but it is not capable of finding a global best solution for complex problems.

3.2 Vehicle Routing Problem

VRP is one of the NP-hard problems in which a fleet of vehicles with limited capacity are used to find a set of routes that serve all customers at minimum cost [12].Each vehicle(from a specific depot) has to serve *n* number of customers with known demand level [11, 13].There are certain constraints for VRP:

i. Each customer can be served only once by a single-vehicle.
ii. Each vehicle must return to the depot after delivering the goods to the customers [14]. Same capacity is provided to every vehicle, total demand for any vehicle should not exceed the capacity.

```
Algorithm:
Input: Construct a structure for sample 'x' of VRP initialize pheromone values and no of
iterations (I)
While termination condition not met(I ≠ 0)
Do
for each ant
{
Construct a route solution
Apply local pheromone updation
}
End for
Apply global pheromone updation
End while
```

Vehicle routing problem generalizes the travelling salesman problem, mainly serves in goods delivery and fetching from n number of points to the central depot using m number of vehicles.

The case studies of both travelling salesman problem and Vehicle routing problem signifies the important role of biological inspired algorithm, such as Ant Colony Optimization in optimizing the path planning thereby reducing cost, time and improving the efficiency.

4 Applications

Computational intelligence is a flexible area for it has its applications in all most every domain say medical field, business & marketing, banking, agricultural field, home automation, space applications and many more. Here we have briefly discussed 3 applications as follows.

4.1 Bioinformatics

Bioinformatics is hot topic in recent times formed by integration of biological knowledge and computational techniques used to understand complex biological structures and to address biological questions by collecting and analyzing large amounts of biological information [15]. CI approaches have proven to be accurate and reliable in identification of genetic structures even when dealing with imprecise data which further processed to identify common occurring diseases such diabetes, cancers and heart diseases caused due to genetic mutations associated with single nucleotide polymorphism [16–18]. CI can also be used to find cures to those diseases by performing test on a simulated human being millions of times instead of doing on a rat eventually forming a theory to cure the diseases [19].

4.2 Driving Cognition

Automated cars were a science fiction 10 years ago but is it going to be our reality now, in order to do this, we need computer with perception, pattern recognition, image processing and ability to cognitively compute all these data. Computational intelligence is one major field contributing towards this. A research on detailed study and understanding of human driving cognition through EEG scanning is possible through CI which is more robust in removing noise and artifacts from the EEG signals comparatively through other techniques. These helps us building a new breed automated cars that drive better than humans [20–22].

4.3 Space Applications

Sparkling in the night sky we see hundreds and billons of stars (like fireflies) yet unexplored and mysterious. Are we alone? Is earth unique or is there any life out there in the infinite black space. CI is possibly providing a solution to this question through powerful telescopes such as Kepler and Hubble space telescope and ultra high definition cameras. These high precision telescopes can capture deep space images focusing on stars and galaxies millions and billions of light years away. Huge volumes of data is captured which is compressed with minimum loss in the image quality and sent to the ground station. Now this information is further processed and analyzed that scientists can tell how far the star from earth is, temperature of the planet and traces of water on the planet that support life all just by using the light emitted by the star reaching the telescope. ACO algorithms have been implemented with new self-adaptation strategies to find optimized solutions for Extravehicular missions [23]. CI approach are being implemented in rovers giving ability to make its own decisions adapting to the surrounding environment efficiently rather than instructed from earth causing delay and risking the loss of rover [24].

5 Conclusion

The paper provides a overall review about computational intelligence, its evolution is explained from the Turing test in late 1940s to its present day models and how it is different from artificial intelligence by considering various models and approaches used in each of the technologies such as ANN, evolutionary computation, fuzzy systems, swarm intelligence etc. A detailed case study on the travelling salesman problem and vehicle routing problem is presented inspired by ant colony optimization technique in swarm intelligence and some of the most important real- world applications are described highlighting the uniqueness of computational intelligence.

Acknowledgements The authors would like to thank Management and Chancellor of REVA University for the support to carry out research activities.

References

1. Bezdek C (2016) (Computational) Intelligence: what's in a name? IEEE Syst Man Cybern Mag 2
2. Wang K (2001) Computational intelligence in agile manufacturing engineering. Agile Manufacturing the 21st Century Competitive Strategy. Elsevier Science Ltd, Oxford, UK
3. Heudin JC (1995) Artificial life and evolutionary computing in machine perception. In: Proceedings of conference on computer architectures for machine perception. IEEE, pp 418–428
4. Siganos D, Stergiou C (1996) Neural networks, the human brain, and learning. Imperial College London: Surveys and Presentations in Information Systems Engineering
5. Beni G, Wang J (1993) Swarm intelligence in cellular robotic systems. In: Robots and biological systems: towards a new bionics? Springer, Berlin, Heidelberg
6. Zhang W et al (2017) Comprehensive overview on computational intelligence techniques for machinery condition monitoring and fault diagnosis. Chin J Mech Eng 30(4)
7. Pramod P, Tripathy PK, Bajpai H, Kounte MR (2019) Role of natural language processing and deep learning in intelligent machines. In: IEEE international conference on electrical, communication, electronics, instrumentation and computing (ICECEIC), Kanchipuram, India.
8. Colorni A, Dorigo M, Maniezzo V (1992) Distributed optimization by ant colonies. In: Proceedings of the first European conference on artificial life
9. Dorigo M, Gambardella LM (1997) Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans Evol Comput

10. Fuentes G, Martinez-Contreras U, Parada-Gonzalez M, Woocay-Prieto A (2018) A case study using ant colony optimization approach to provide a shortest route plan to evaluate the maintenance needs in elementary schools in Northern Chihuahua mexico. In: 2018 systems and information engineering design symposium (SIEDS). IEEE
11. Pan T, Pan H, Gao J (2015) An improved ant colony algorithm based on vehicle routing problem. In: 2015 34th Chinese control conference (CCC). IEEE
12. Stodola P, Mazal J, Podhorec M, Litvaj O (2014) Using the ant colony optimization algorithm for the capacitated vehicle routing problem. In: Proceedings of the 16th international conference on mechatronics-Mechatronika. IEEE
13. Wang M (2012) Hybrid behavior ant colony algorithm for Vehicle Routing Problem. In: Fourth international conference on computational and information sciences. IEEE
14. Carwalo T, Patil V (2015) Solving vehicle routing problem using ant colony optimization with nodal demand
15. Kelemen A, Vasilakos AV, Liang Y (2009) Computational intelligence in bioinformatics: SNP/haplotype data in genetic association study for common diseases. IEEE Trans Inf Technol Biomed
16. Simha CY, Harshini VM, Raghuvamsi LVS, Kounte MR (2018) Enabling technologies for internet of things & it's security issues. In: 2018 second international conference on intelligent computing and control systems (ICICCS). IEEE
17. Virat MS, Bindu SM, Aishwarya B, Dhanush BN, Kounte MR (2018) Security and privacy challenges in internet of things. In: 2018 2nd international conference on trends in electronics and informatics (ICOEI). IEEE
18. Bandyopadhyay S, Maulik U, Roy D (2007) Gene identification: classical and computational intelligence approaches. IEEE Trans Syst Man Cybern C (Appl Rev)
19. Naveen S, Kounte MR (2020) In Search of the future technologies: fusion of machine learning, fog and edge computing in the internet of things. In: Pandian A, Senjyu T, Islam S, Wang H (eds) Proceeding of the international conference on computer networks, big data and IoT (ICCBI-2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham
20. Kamble SJ, Kounte MR (2019) Routing and scheduling issues in vehicular ad-hoc networks. Int J Recent Technol Eng (IJRTE) 8(3):4272–4283. ISSN: 2277-3878
21. Kamble SJ, Kounte MR (2020) Enabling technologies for internet of vehicles. In: Pandian A, Senjyu T, Islam S, Wang H (eds) Proceeding of the international conference on computer networks, big data and IoT (ICCBI-2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham
22. Lin C-T, Ko L-W, Shen T-K (2009) Computational intelligent brain computer interaction and its applications on driving cognition. IEEE Comput Intell Mag
23. Zhu Y, Luo Y, Tan KC, Qiu X (2017) An intelligent packing programming for space station extravehicular missions. IEEE Comput Intell Mag
24. Yuen P, Gao Y, Griffiths A, Coates A, Muller J-P, Smith A, Walton D, Leff C, Hancock B, Shin D (2013) ExoMars Rover PanCam: autonomous & computational intelligence [application notes]. IEEE Comput Intell Mag

Covid-19 Face Mask Prediction Using Machine Learning Techniques

A. Rama Chetan, A. Arjuna Rao, and P. K. J. Mohapatra

Abstract Entire world including India is going through a pandemic that has arisen due to the outbreak of COVID-19. Medicines and Vaccine for Covid-19 are still under developmental stage. Wearing a Face Mask is the best viable option for humans to prevent the spread of infection due to Corona virus. As a result, controlling government agencies may want to know the percentage of people wearing masks during a period as well as which group of people are most likely to wear masks when they go outside. To help answer these questions, this paper introduces a model that can classify faces among masked faces and unmasked faces using Python 3.0 Language. In the present face detecting model, Vietnam based mask classifier dataset, CelebA dataset, WiderFace dataset and MAFA datasets are used for achieving better results. Single Stage Headless Face Detector (SSH) is successfully implemented to segregate human faces with or without mask. Experimental results with the Mask Classifier model show that it can achieve about 96.5% accuracy during testing stage. Selected on road going people video is tested successfully where the present model clearly segregated human faces with and without mask. The present model is useful to safeguard people from spread of Covid-19 virus in public places.

Keywords Face mask detector · Machine learning · ResNet architecture · Single stage headless face detector · KerasFramework · Mask classifier model

1 Introduction

In view of Covid-19 pandemic problems, it is aimed at developing a machine learning based intelligent model to predict human faces with and without mask using python 3 language. In fact, Anti-viral medicines and vaccine for Covid-19 are still under

A. Rama Chetan (✉) · P. K. J. Mohapatra
School of Electrical Sciences—CSE Department, Indian Institute of Technology Bhubaneswar, Bhubaneswar, India
e-mail: arc10@iitbbs.ac.in

A. Arjuna Rao
Miracle Educational Society Group of Institutions, Bhogapuram, India
e-mail: aarao@miracleeducationalsociety.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_69

developmental stage only to cure the corona virus. As a result, controlling government agencies may want to know the percentage of people wearing masks during a period as well as which group of people are most likely to wear masks when they go outside. To help answer these questions, this paper introduces a model that can classify faces among masked faces and unmasked faces. In addition to the Mask Classifier dataset consisting of about 11,000 face images of Vietnamese people, we also used CelebA [1] dataset consisting of about 202,599 celebrity images, WiderFace [2] dataset consisting 32,303 images and 393,703 faces and MAFA [3] data consisting of 30,000 images in which each image contains at least one masked face to improve accuracy percentage.

In the present research paper, we described two phases: data preparing phase and training testing model phase. For face detection purposes, we used Single Stage Headless Face Detector (SSH) [4]. In the area of crowd monitoring systems and crowd abnormality detection in densely populated crowd videos, the classification of crowd motion as an ordered flow or disordered flow has significant importance towards reaching the end goal of building powerful crowd monitoring systems. The main objective of this machine learning based project work is to classify the people who are not wearing masks in live videos and images. Face detection techniques are more relevant today because it is not only used on images but also in video applications like real time surveillance and face detection in videos. High accuracy image classification is possible now with the advancements of Convolutional neural network (CNN) based technologies. Pixel level information is often required from photographs after the face detection which most of the face detection methods fail to provide. Obtaining pixel level details has been a challenging task in semantic segmentation process. Figure 1 explains detailed flowchart implemented [5] in the present work to find out face masks using machine learning techniques.

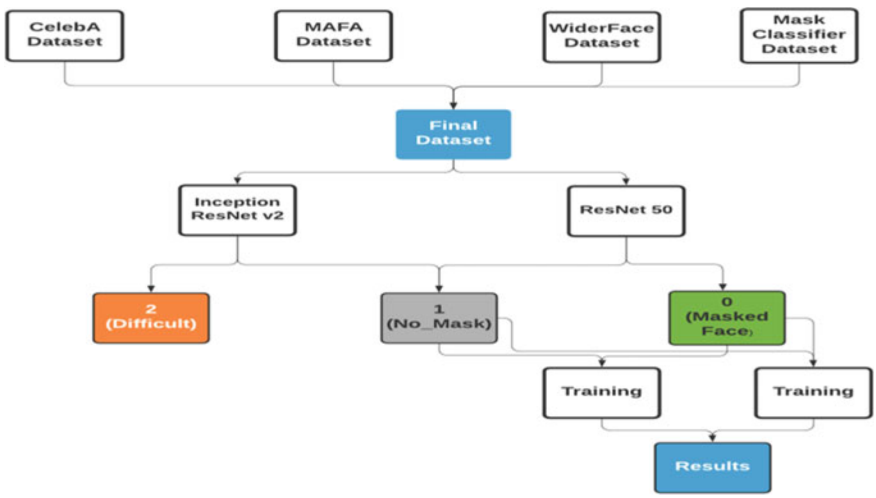


Fig. 1 Flow chart used in present work

2 Data Compilation

In this section, we describe four different face datasets including CelebA dataset, Wider Face dataset, MAFA dataset and our own dataset—Mask Classifier dataset [6]. After preparing all listed datasets, we use scripts \gen data.sh to combine them together into a single dataset for training and testing our machine learning model. In the combined dataset, Output "0" indicates masked face and "1" as an unmasked face.

2.1 CelebA Dataset

CelebA is a large-scale face attributes dataset with 202,599 celebrity images. We use about 8% of CelebA dataset for our combined dataset. After that, we use gen data CelebA function from gen data.py to generate a cropped CelebA dataset (to do this, we must set up SSH before hand). Figure 2 shows CelebA dataset used in the present model.

Fig. 2 CelebA dataset representation

Fig. 3 WiderFace dataset representation

2.2 WiderFace Dataset

WiderFace is a face detection benchmark dataset consisting of about 32,203 images and 393,703 faces with a wide variety of scales, poses, and occlusions. We only use about 12,000 normal, clear, and medium-occluded faces from this dataset. After downloading WiderFace's modified version, we use gen data WiderFace function from gen data.py to use a part of this dataset. We used about 9000–10,000 images for our training and testing purpose. Figure 3 shows WiderFace dataset used in the present model.

2.3 MAFA Dataset

MAFA is a face dataset consisting of about 30,000 images in total and each image contains at least one masked face. There are six annotation attributes, but we only care about Location of faces, Occlusion degree, and Mask type attributes. Figure 4 shows MAFA dataset used in the present model.

Fig. 4 MAFA dataset representation

2.4 *Mask Classifier Dataset*

Mask Classifier is a dataset consisting of about 11,000 face images of people on the road in Vietnam. Images from this dataset are generated from 9 videos with a total average length of about 3 min. To generate these images, we consider an interval of 10–20 frames from those videos and use SSH to crop faces from a video frame. After that, we put those cropped images to different folders for different videos. We label an image **0** for masked face, **1** for normal face, and **2** for difficult-to detect face or non-face. We used only faces with label 0 and 1 for training and testing our model. Figure 5 shows sample image of Mask Classifier dataset representation.

3 Mask Classifying Model

In the present work python 3.0 [7] language-based mask classifying model is developed. In this ResNet 50 architecture is implemented for achieving expected results. The above mentioned 4 different compiled datasets are used as input for the model.

Fig. 5 Mask Classifier dataset representation

3.1 ResNet 50 Network

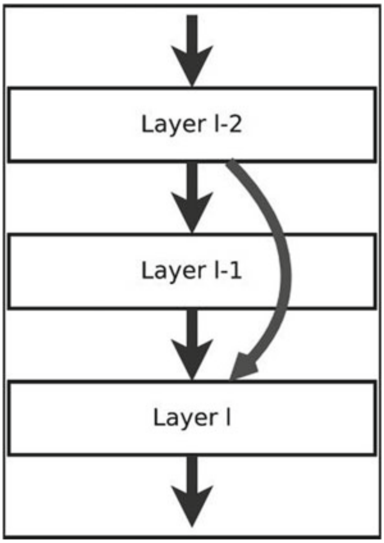
ResNet-50 [8] is a deep residual network. The “50” refers to the number of layers it has. It is a subclass of convolutional neural networks, with ResNet most popularly used for image classification. In general, in a deep convolutional neural network, several layers are stacked and are trained to the task at hand. The network learns several low/mid/high level features at the end of its layers. In residual learning, instead of trying to learn some features, we try to learn some residual. Residual can be simply understood as subtraction of features learned from input of that layer. ResNet does this using shortcut connections (directly connecting input of the nth layer to some (n + x)th layer. It has proved that training this form of networks is easier than training simple deep convolutional neural networks and the problem of degrading accuracy is resolved.

ResNet first introduced the concept of skip connection. Figure 6 illustrates skip connection stacking convolution layers together one after the other. On the right we still stack convolution layers as before, but we now also add the original input to the output of the convolution block. This is called skip connection.

3.2 ResNet 50 Architecture

ResNet 50 Architecture is proved model for computer vision tasks such as image classification, object localization and object detection. Hence this architecture is selected for present face mask recognition task. ResNet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool

Fig. 6 Skip connection



layer. It has 3.8×10^9 Floating points operations. It is a widely used ResNet model and we have explored **ResNet50 architecture** in depth.

As shown in Fig. 7, the ResNet 50 architecture is flexible in usage and contains the following elements:

A convolution with a kernel size of 7 * 7 and 64 different kernels all with a stride of size 2 giving us 1 layer. Next, we see max pooling with also a stride size of 2. In the next convolution there is a 1 * 1, 64 kernel following this a 3 * 3, 64 kernel and at last a 1 * 1, 256 kernel, these three layers are repeated in total 3 times so giving us 9 layers in this step. Next, we see kernel of 1 * 1, 128 after that a kernel of 3 * 3, 128 and at last a kernel of 1 * 1, 512 this step was repeated 4 times so giving us 12 layers in this step. After that there is a kernel of 1 * 1, 256 and two more kernels with

Architecture

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		[3×3, 64; 3×3, 64] ×2	[3×3, 64; 3×3, 64] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3
conv3_x	28×28	[3×3, 128; 3×3, 128] ×2	[3×3, 128; 3×3, 128] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×8
conv4_x	14×14	[3×3, 256; 3×3, 256] ×2	[3×3, 256; 3×3, 256] ×6	[1×1, 256; 3×3, 256; 1×1, 1024] ×6	[1×1, 256; 3×3, 256; 1×1, 1024] ×23	[1×1, 256; 3×3, 256; 1×1, 1024] ×36
conv5_x	7×7	[3×3, 512; 3×3, 512] ×2	[3×3, 512; 3×3, 512] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Fig. 7 Understanding the ResNet50 Architecture

3 * 3, 256 and 1 * 1, 1024 and this is repeated 6 times giving us a total of 18 layers. And then again, a 1 * 1, 512 kernel with two more of 3 * 3, 512 and 1 * 1, 2048 and this was repeated 3 times giving us a total of 9 layers. After that we do an average pool and end it with a fully connected layer containing 1000 nodes and at the end a Soft Max function, so this gives us 1 layer. We do not actually count the activation functions and the max/ average pooling layers. So, totaling this it gives us a $1 + 9 + 12 + 18 + 9 + 1 = 50$ layers Deep Convolutional network.

3.3 Choosing Appropriate Hidden Layers

ResNet-50 [9] is a convolutional neural network that is 50 layers deep. As the name itself says that ResNet 50 we used 50 layers by default in the present model. We can load a pre-trained version of the network trained on more than a million images from the ImageNet database. The pre-trained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224 as shown in Fig. 8.

4 Training of Mask Classifier Model

To train the Mask Classifier model, we deploy Keras framework. To prepare the dataset for training and testing on Keras, we use gen data function from gen data.py-A 316-line Python3 code developed for generating data. After that, we trained present Mask Classifier model with train.py—a 87-line python code developed for training. We deploy two different CNN architectures (resnet50 and inception ResNet v2 [10]).

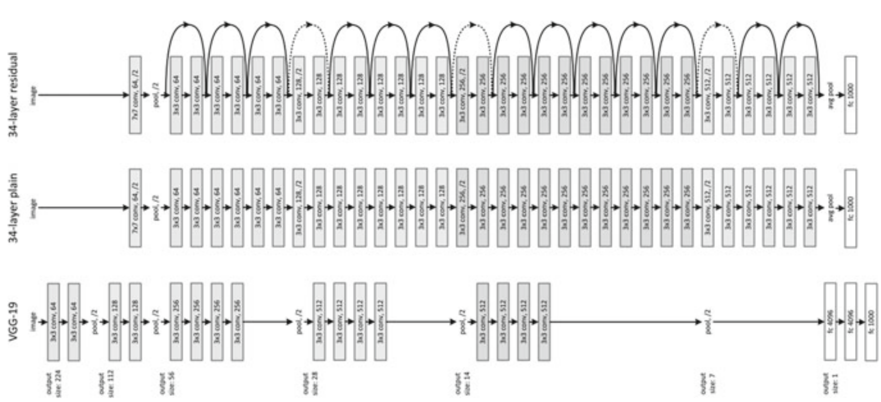


Fig. 8 ResNet 50 Architecture with hidden layers

The default network architecture is resnet50. To train the network, first we initialize the network with image net weights. After that, we replace the last SoftMax activation layer to sigmoid activation layer and change the loss function to Binary Cross Entropy loss. We train the whole network without freezing any layers. During the training phase, we use Stochastic Gradient Descent [11] with a base learning rate of 0.0001, weight decay and momentum. We also apply different types of data augmentation. The network begins to stabilize and achieve around 98–99% on training set and 95–96% on validation set after about 1 $\times$ epochs.

5 Testing of Mask Classifier Model

We use demo.py (a 159-line python3 code developed for testing) to demonstrate present model to implement on live videos with different datasets. During the experiment, we observe that our model performs quite well on medium to large size face images. The model can achieve about 96.5% accuracy when predicting. However, in some cases, because of the wide variety of textures on masks, as well as the large range of poses and occlusions, the model still predicts incorrectly in around 5% cases. To demo our model on videos or large images, first we use SSH to detect faces from an image or a video frame (we only use the SSH pyramid option when demoing on images to improve performance). After that, we deliver previous detected faces to our trained network to classify between masked faces and normal faces. We draw bounding boxes around detected faces with color yellow, green or red based on the classified results. We also display the probability of a normal face and the percentage of masked faces on a picture.

In this demo.py file we implemented code to extracted people faces form the input video but taking the coordinates of the faces and making necessary boxes around them and classified them into Masked and Unmasked by running our model. In func.py (a 101 line python code developed for this model) file we mentioned about plotting the training and validation accuracy values by getting the accuracy, precision and recall along with F1 score for each classes names Masked and Unmasked and also plotted the training and validation loss values. We made to dynamically grow the memory used on the GPU.

5.1 Confusion Matrix

As shown in Fig. 9, A confusion matrix is a table that is often used to describe the performance of a classification model (or “classifier”) on a set of test data for which the true values are known.

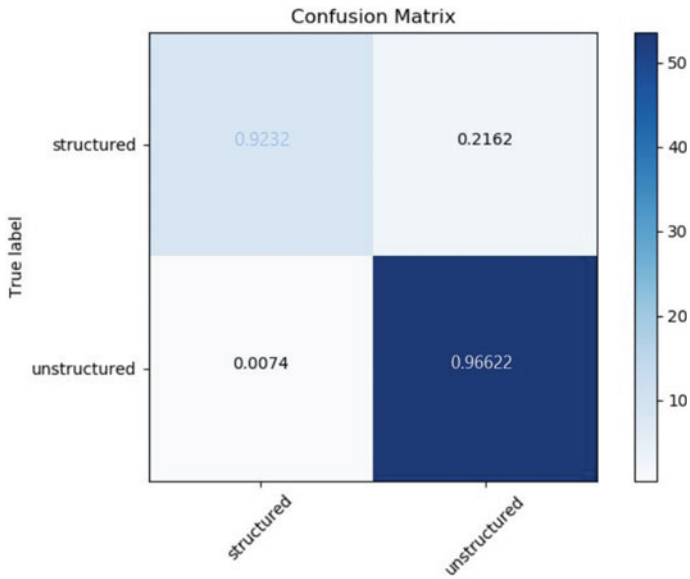


Fig. 9 Confusion matrix with data about structured and unstructured classes

5.2 Accuracy

As shown in Fig. 10, we achieved an accuracy of 96.2%. Although this work needs a larger dataset in order to get perfect results, we are confident that our work can show satisfactory results.

5.3 Precision, Recall and F1-Score Table

As obtained in Fig. 11, **Precision** is the ratio of correctly predicted positive observations to the total predicted positive observations (True Positive/ Total Predicted Positive). **Recall** is the ratio of correctly predicted positive observations to all observations in actual class (True Positive/ Total Actual Positive). **F1 Score** is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account:

$$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision})/(\text{Recall} + \text{Precision}) \quad (1)$$

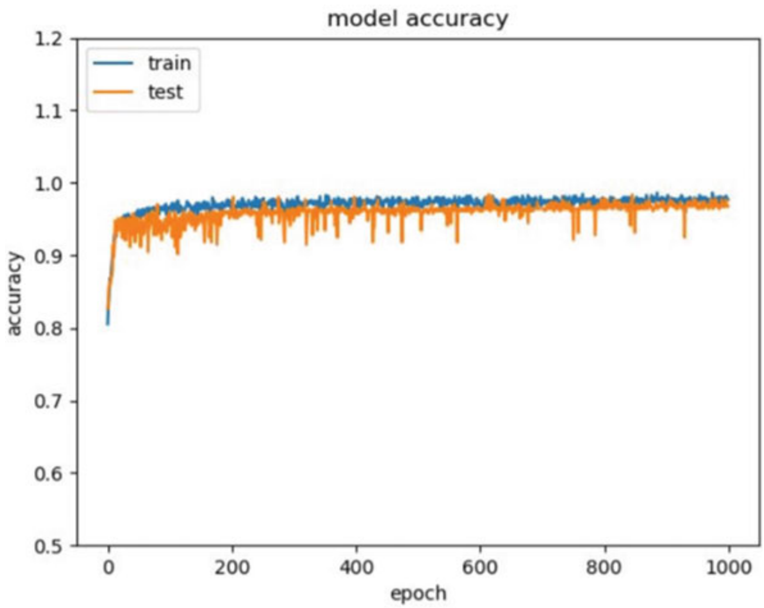


Fig. 10 Accuracy of the model

	precision	recall	f1-score
masked	0.91	0.89	0.92
unmasked	0.94	0.92	0.92

Fig. 11 Precision, recall and F1-score

5.4 Loss Function

Figure 12 shows loss function. It is a method of evaluating how well specific algorithm models the given data. If predicated value varies largely from actual value loss function would show exceptionally large number and vice versa.

5.5 Results

As Explained above, we implemented our workon different scenario as shown in Figs. 13 and 14.

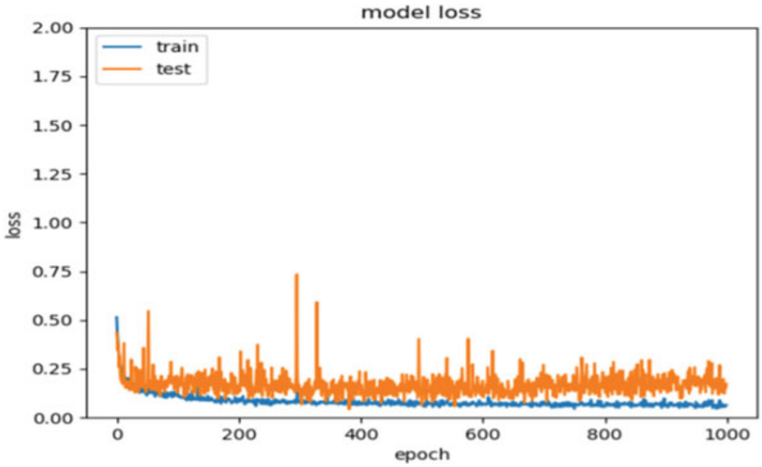


Fig. 12 Loss function

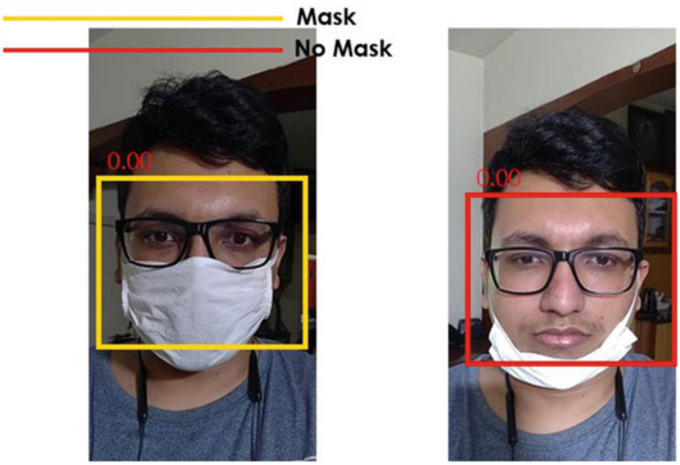


Fig. 13 Same face with mask and no mask

As shown in the YouTube video [12] the present model is implemented, and people are separated with yellow border on people with mask and red border on people without mask with good accuracy of prediction (Figs. 15 and 16).

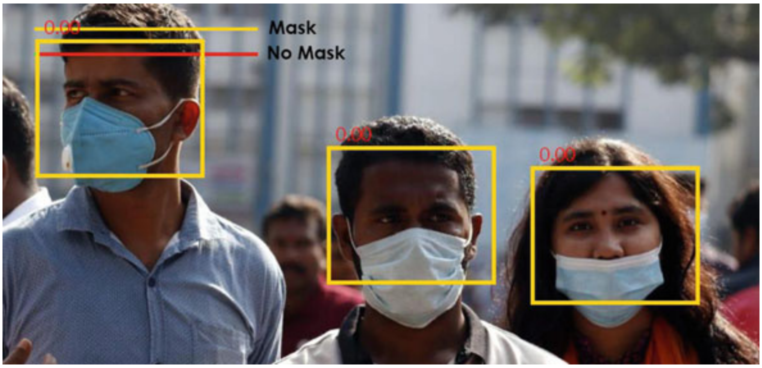


Fig. 14 Faces with mask and no mask

Fig. 15 YouTube video of on road faces with and without Masks

6 Conclusions

In view of Covid-19 pandemic problems, it is aimed at developing a machine learning based intelligent model to predict human faces with and without mask using python 3 language. Controlling government agencies may want to know the percentage of people wearing masks during a period as well as which group of people are most likely to wear masks when they go outside. To help answer these questions, this paper successfully introduced Mask Classifying Model that can classify faces among masked faces and unmasked faces. In addition to the Mask Classifier dataset, well known CelebA dataset with celebrity images, WiderFace dataset consisting wide range face images and MAFA data consisting of more number of images in

Fig. 16 QR code for testing video online

which each image contains at least one masked face to improve accuracy percentage was successfully introduced. Also in the present research paper, we could demonstrate data preparing phase and training testing model phase. For face detection purposes, Single Stage Headless Face Detector (SSH) was effectively utilized in present Machine Learning based model.

The Face Mask Detection System can be used at airports, hospitals, offices etc. This project can help in various fields to detect the persons who are not wearing a face mask in the present Covid 19 scenario. We achieved an accuracy of 96.2%. Although this project needs a larger dataset to get perfect results, we are confident that our project can show satisfactory results. Presently the world is facing a crisis, and no one knows exactly how to come out of this. We are in a situation that needs us to learn to live with the virus. Projects like this will help us to live with the present situations. We strongly believe that the automatic mask detecting software helps government agencies and private institutions to stop the spread of the corona virus disease.

Acknowledgements Authors would like to thank Professors of CSE, **Dr. M. Satpathy**, Dr.Joy Chandra Mukherjee, Dr.Sudipta Saha and HOS Dr.N.C Sahoo, School of Electrical Sciences, IIT Bhubaneswar for useful discussions and arranging required facilities during Machine Learning Project work and also encouraging for publication of present paper. We are thankful to Dr.R.V. Raja Kumar, Director, IIT Bhubaneswar for his kind support and inspiration for research work on Covid'19 innovations. We also thank opengenus.org which helped us to know more about the ResNet 50—Deep Residual Learning and helping in providing the architecture images.

References

1. CelebA dataset link: https://www.kaggle.com/jessicali9530/celeba-dataset
2. WiderFace dataset link: https://www.kaggle.com/mksaad/wider-face-a-face-detection-benchmark
3. MAFA dataset link: http://www.escience.cn/people/geshiming/mafa.html
4. Najib M, Samangouei P, Chellappa R, Davis L (2017) SSH: single stage headless face detector, 1708.03979
5. Lucidchart.com for present Flow Chart used
6. Rama Chetan A (2020) B.Tech-CSE Thesis, School of Electrical Sciences, IIT Bhubaneswar
7. www.python.org for Python 3 Language and Compiler
8. ResNet-50: https://www.mathworks.com/help/deeplearning/ref/resnet50.html
9. Zhong Y, Qiu S, Luo X, Meng Z, Liu J (2020) Facial expression recognition based on optimized ResNet, pp 84–91. https://doi.org/10.1109/WSAI49636.2020.9143287
10. https://keras.io/api/applications/inceptionresnetv2/
11. An overview of gradient descent optimization algorithms, Sebastian Ruder Insight Centre for Data Analytics, NUI Galway Aylien Ltd., Dublin
12. YouTubevideo link for live on road public video tested with present model where mask and mask less people are identified. https://www.youtube.com/watch?v=psgSeyW3xkI

A Review on Big Data Analytics in Internet of Things (IoT) and Its Roles, Applications and Challenges

Rajesh Mothe, S. Tharun Reddy, B. Vijay Kumar, A. Rajeshwar Rao, and Kanegonda Ravi Chythanya

Abstract Big Data Analytics is a popular term all around utilized for a huge arrangement of data and it's analytics, it manages a large volume of heterogeneous information as organized, semi-organized and unstructured. We are having some knowledge on how information is pulled back from various sources and get converged for information preparing. In this paper, our major concentration on basic importance of the Big Data Analytics and IoT (Internet of things) and the multi-faced motivation behind Big data Analytics in IoT like its roles, applications, and challenges.

Keywords Big data analytics · Internet of things · IoT · Big data

1 Introduction

1.1 Big Data Analytics

Big data is considered as a technical problem. Nowadays it has become business and future opportunities. A few years ago, in early 2000, big data was considered as a serious problem by the organizations. As the data volume was growing at a very fast rate, storage, and CPU technologies swamp by the countless terabytes of data, at this point the Information Technology field faced a crisis on the data capacity. As time passed the storage and CPU become powerful, intelligence and process the data at a very fast speed as data is increasing. Today the numbers of organizations are

R. Mothe (✉) · B. Vijay Kumar
Research Scholar, School of Computer Science and Artificial Intelligence, S R University, Warangal, India

S. Tharun Reddy
Department of CSE, SR Engineering College, Warangal, India

A. Rajeshwar Rao
Warangal, India

K. R. Chythanya
Associate Professor, School of Computer Science and Artificial Intelligence, S R University, Warangal, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_70

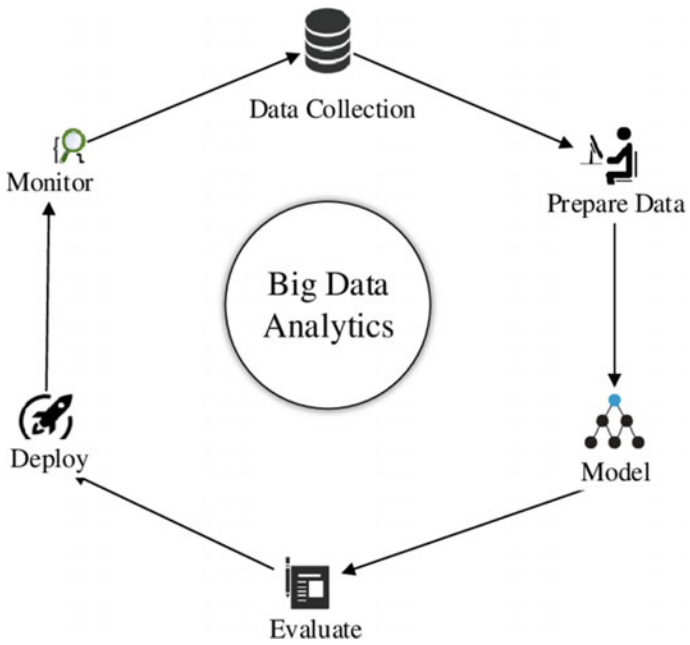


Fig. 1 Overview of big data analytics

exploring the big data to identify the valid information that was not possible before. Now all the business enterprises are producing data in a huge amount that is based on their mass consumers. By using the analytics techniques, the organizations can identify the current situation of running a business is producing valid big data and can observe the consumer requirement on processed data. Big data analytics is the approach of popular analytics technologies to verify big data sets. The term "Big data" and "Big data analytics" are two technical entities. First, big data is a huge amount of data with detailed information. Second, analytics is a mix of different types of analytical tools such as statistical analysis, data mining, NPL and so on. If put them all together will get the big data analytics, which is currently used in BI. Based on this, the numbers of organizations are maintaining big data in their organization data warehouse (ODW). The DW should be auditable, well maintained and clean that can be used to generate business reports demand. The basic overview of big data analytics and it's flow of phases depicted in the below Fig. 1.

2 Internet of Things

With the progressions of microelectronics, interchanges and data innovation, the Internet of Things (IoT) idea has been broadly applied in different fields, including

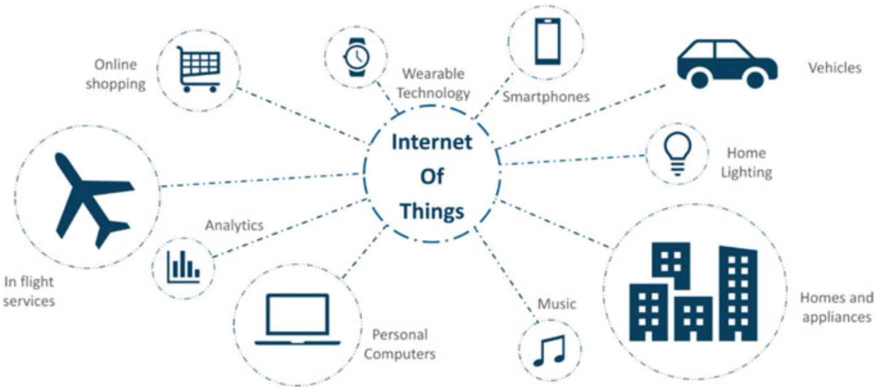


Fig. 2 Overview of internet of things

brilliant homes [1], medicinal services [2], and things discernibility [3]. These days, the scholastic and modern examinations principally center around the accompanying headings, for example, IoT framework engineering, correspondence conventions, sensor systems, security and security assurance and utilizations of the IoT. For the most part, the IoT alludes to the things in physical universes that could be interoperated with the help of correspondence conventions [4], and it very well may be viewed as a system of minimal effort electronic gadgets where detecting information and correspondence could be accomplished consequently. This sort of interoperation among gadgets is additionally called "Machine to Machine" (M2M) correspondence [5], which is significant for taking care of the vulnerability in IoT frameworks. In any case, M2M gadgets are hard to interface in certifiable executions on the grounds that there is an absence of normalized correspondence conventions. A passage is commonly utilized among M2M gadgets to fill in as a middle of the road gadget. In any case, these doors need smart control capacities. These capacities are significant in IoT frameworks on account of the constrained computational force and memory of M2M gadgets. Rather, a microcontroller could straightforwardly control the endpoint gadgets. Furthermore, it could speak with the Internet so insight administrations could be given to the end-clients [6]. The Fig. 2 shows the overview of IoT applications in different fields.

3 The Relation Between Big Data Analytics and Internet of Things

Data however gathered by the gadgets should be separated to make it significant and valuable. The repetition in the information being gathered is prevalent because of the sheer idea of the system of IoT. The information is ceaseless thus the extraction of important data isn't straightforward. This requires a decent component of

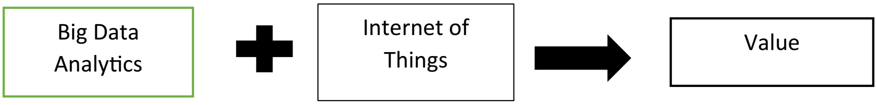


Fig. 3 Value creation using Big data analytics and internet of things

conventions and programming to guarantee that the information is made sure about and furthermore noteworthy [7]. Data is commonly gathered by the sensor gadgets in which these gadgets gather and transmit information to a unified server. So also, the information is circulated back to the gadgets too. These exercises require the presentation proficiency of the system to be ideal. IoT includes a few heterogeneous systems like remote Sensor Networks (WSN), Wireless work systems, Wireless LAN. These systems would help in the transmission of information and furthermore include different kinds of value issues going from execution to vitality effectiveness. Huge Data and IoT are corresponding to one another and are two elements of observation. Dealing with the information and extricating data from it is an extremely indispensable undertaking related with IoT [8]. A proper systematic stage is required to empower to get information from IoT information. IoT gadgets create nonstop surges of information in a versatile manner. It is fundamental to deal with the high volume of stream information and endeavor the information. In an ordinary situation Big Data, the information probably won't be stream information, yet the activities are. While in IoT information, it is a persistent stream. Applying ongoing investigation is the need in the IoT condition. The upsides of IoT can be seen just when ongoing investigation is applied to the information put away[9]. Real-time Big Data analytics and IoT equates to value creation which is depicted in Fig. 3.

4 Applications of Big Data Analytics and Internet of Things

Big Data is omnipresent. Each business, for example, wellbeing or general expectations for everyday comforts could apply large information investigation. Big data is a field that can be utilized in any zone at all given this huge amount of information can be an outfit to further one's potential benefit. The significant utilization of big data is recorded underneath [10].

The Third Eye-Data thought: Organizations all around are perceiving the significance of large information examination. Huge Data examination is a one-stop answer for pretty much every association, it predicts client buying personal conduct standard, identifying misrepresentation and misuses. It gives a chance to business specialists to address and comprehends information as per their business needs regardless of the trouble and volume of the information [11]. This can be cultivated by expertly picturing and reasonably introducing the information. Mammoths like Google, Facebook, Twitter, eBay, Wal-Mart, and so forth., likewise receive information perception to ease the trouble of taking care of information. Information representation has indicated tremendous positive results in such business associations. Executing information investigation and information representation; endeavors can, in conclusion, start to take advantage of the gigantic potential that huge information has and make a definite more noteworthy profit for ventures and business consistency.

In Banking: The utilize client information could likewise climb security issues. By spotting an unnoticeable relationship between apparently indisputable bits of data, Big Data Analytics might reveal delicate individual information. An investigation shows that 62% of lenders are cautious in their use of Big Data on account of detachment issues [8]. Further, redistributing data examination execution or sharing of customer data crossways workplaces for the creation of more joyful seeing furthermore opens up security perils.

In Agriculture: A biotechnology firm uses sensor information to improve obtaining efficiencies. Its plant's assessment gathers and runs re-institutions to figure how plants react to different changes in condition. Its data environmental factors, again and again, adjust to change like various data it accumulates, and temperature, water levels, soil game plan, advancement, yield, and quality sequencing of each plant in the demonstrating ground. These diversions endorse to find the perfect natural situation for the right quality sorts.

In Finance: Finance related associations are using pariah praise scoring while at the same time surveying new recognition applications. In any case, the banks are at present using their recognition scoring examination for open customers using a far-reaching extent of data, and checking, hold reserves, charge cards, home advances, and adventure data.

In Economy: Expected beginning from the most punctual stage to deal splendidly with product gear, Hadoop can help relationships for a transition to low spending servers.

IoT applications are required to prepare lots of regular articles with availability and insight. It is as of now being conveyed broadly, in different areas, some are given underneath.

Wearables: this innovation is a sign of IoT applications and most likely is probably the soonest business to have conveyed the IoT at its administration. Examples are Fit Bits, pulse screens and smartwatches wherever nowadays. One of the lesser-realized wearables incorporates the Guardian glucose observing gadget. The gadget is created to help individuals experiencing diabetes. It recognizes glucose levels in the body, utilizing a small terminal known as a glucose sensor put under the skin and transfers the data by means of Radio Frequency to an observing gadget.

Smart Home Applications: In IoT Applications, Smart Homes are most likely the main thing that we consider. The best model I can consider here is Jarvis, the AI home robotization utilized by Mark Zuckerberg. There is additionally Allen Pan's Home Automation System where works in the house are incited by utilization of a string of melodic notes. The accompanying video could give you a superior thought.

Health Care: IoT applications can transform responsive clinical based frameworks into proactive wellbeing based frameworks. The assets that momentum clinical research utilizes, need basic certifiable data. It for the most part utilizes extra information, controlled conditions, and volunteers for clinical assessment. IoT opens approaches to an ocean of important information through examination, constant field information, and testing. The IoT Things likewise improves the present gadgets in force, accuracy, and accessibility. IoT centers around making frameworks as opposed to simply hardware.

Smart Cities: By now I accept, the greater part of you more likely than not found out about the term Smart City. The speculation of the improved traffic framework I referenced before, is one of the numerous perspectives that establish a brilliant city. The thing about the brilliant city idea is that it's unmistakable to a city. The issues looked in Mumbai are altogether different than those in Delhi. The issues in Hong Kong are not quite the same as New York. Indeed, even worldwide issues, as limited clean drinking water, weakening air quality and expanding urban thickness, happen in various forces across urban communities. Thus, they influence every city in an unexpected way. The Government and specialists can utilize IoT to break down the regularly intricate components of town arranging explicit to every city. The utilization of IoT applications can help in regions like water the board, squander control, and crises.

Agriculture: Statistics gauge the ever-developing total populace to arrive at about 10 billion continuously 2050. To take care of such a huge populace one needs to wed farming to innovation and acquire the best outcomes. There are various prospects right now. One of them is the Smart Greenhouse. A nursery with installed gadgets makes it simpler to be observed as well as, empowers us to control the atmosphere inside it. Sensors measure various parameters as per the plant prerequisite and send it to the cloud. It, at that point, forms the information and applies a control activity.

5 Role of Big Data Analytics in IoT

We know that brilliant gadgets are significant parts in IoT, these gadgets create a gigantic measure of information that should be investigated and researched progressively. This is the place prescient and Big Data Analytics become an integral factor [12].

Big data analytics influence IoT for simple working, yet also gives a few difficulties, Big data is observable in IoT because of colossal organization of sensors and web pertinent things, information preparing in big data is confronting difficulties because of short computational, systems administration and capacity implies at IoT gadget end.

When the total IoT framework goes about as an information produced source, the job of big data in IoT gets basic, Big data analytics is a rising apparatus for dissecting information made by associated gadget in IoT which help to start to lead the pack to improve dynamic.

A lot of information on an ongoing premise and store that utilizing different capacity methods can be taken care of Big data process. Below are the steps that are considered for information preparing:

1. A monstrous measure of heterogeneous information is made by IoT associated gadgets that are put away in the big data framework for a huge scope. This IoT created big data unequivocally relies upon 3'V components or qualities of Big data that are volume, velocity, and verity.
2. A Big data framework is a common and disseminated database, consequently, the big amount of information is recorded in large information documents in the capacity framework.
3. Interpreting and analyzing the gathered IoT Big information utilizing progressed explanatory instruments like Hadoop, Spark, and so on.
4. Inspecting and creating the portrayals of analyzed information for exact and opportune dynamics.

6 Challenges in IoT with Big Data Analytics

Quick development in various applications in IoT additionally offers begining to different moves that should be tended to, below are someof the key challenges in IoT with Big data analytics.

1. **Data storage and management**: Data created from web prepared gadgets is expanding at an ever-extending rate, and the capacity limit of Big information framework is restricted, consequently it turns into an earlier test to store and oversee such a lot of information. It is important to plan a few systems and structures to accumulate, spare and handle this information.
2. **Data visualization**: We realize that produced information is heterogeneous. for example, organized, unstructured and semi-structured in various organizations, so it gets hard to picture this information straightforwardly. It is required to get ready information for better representation and comprehension for precise mechanical dynamics on schedule and improving the productivity of the business.
3. **Confidentiality and privacy**: Each shrewd item into an internationally associated organize comprises an IoT framework particularly utilized by people or machines, it adds more regard for protection and spillage of data, so this vital information should keep private and give protection as delivered information contains individual data of clients.
4. **Integrity**: Connected-gadgets are capable of detecting, conveying, data sharing, and leading examination for various applications. These gadgets guarantee clients not to share their information inconclusively, information amass strategies must send scale and states of honesty effectively by few methodologies.

5. **Power captivity**: IoT gadgets ought to be associated with the ceaseless force supply for the good working of IoT activities. These gadgets are restricted as far as memory, preparing force, and vitality, so gadgets must be conveyed with small components.

Aside from these significant difficulties, Big information investigation experienced other gigantic difficulties too, for instance, gadget security and reinforcement against assaults as these are the clearest devices for assaults and give a portal for evil exercises. Simple Availability of these gadgets another test, gadgets must be accessible without a doubt because of their basic application nature, for example, smart homes, smart urban communities, keen ventures, and so forth [13].

7 Conclusion

To create productive and constant data analytics of all around associated gadgets, different Big data applications are effectively accessible tools, we know the consolidated effect of Big data analytics and an IoT in analyzing large amounts of data accurately and efficiently with appropriate techniques. Data Analytics likewise differs with sorts of information taken from heterogeneous information sources and interpreted for results. For example big data framework is fit for performing great yet in addition faces a few issues while data processing.

References

1. Misra A, Sharma A, Gulia P, Bana A (2014) Big Data: challenges and opportunities. Int J Innov Technol Explor Eng (IJITEE) 4(2):41–42
2. Alkar AZ, Buhur U (2005) An Internet based wireless home automation system for multifunctional devices. IEEE Trans Consum Electron 51:1169–1174
3. Atzori L, Iera A, Morabito G (2010) The Internet of Things: a survey. Comput Netw 54:2787–2805
4. Bubbi JB, Buyya R, Marusic S, Palaiswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29:1645–1660
5. Emiliano S, Abusayeed S, Song H, Ulf J, Mikael G (2018) Industrial Internet of Things: challenges, opportunities, and directions. IEEE Trans Ind Inform 4:4724–4734
6. Gubbi J, Buyya R, Marusi R, Palaniswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29:1645–1660
7. 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 8(2 Special Issue 11):2634–2641
8. Praveen P, Jayanth Babu C (2019) Big Data clustering: applying conventional data mining techniques in big data environment
9. Sheshikala M, Mohmmad S, Shabana (2018) Survey on multi level security for IoT network in cloud and data centers. J Adv Res Dyn Control Syst 10(10 Special Issue):134–146
10. Naeem M, Ejaz W, Karim L, Ahmed SH, Anpalagan A, Jo M (2017) Distributed gateway selection for M2M communication in cognitive 5G networks. IEEE Netw 31:94–100

11. Mukherjee S, Shaw R (2016) Big Data—concepts, applications, challenges and future scope. Int J Adv Res Comput Commun Eng 5(2)
12. Swathi, Mothe R (2018) An overview of iot towards irrigation system. Indian J Public Health Res Dev 9(11):1184
13. Xu LD, He W, Li S (2014) Internet of Things in industries: a survey. IEEE Trans Ind Inform 10:2233–2243

Scientometrics and Publications: A Comparative Study of Ranking of Multi-source Databases

Priti Kumari and Rajeev Kumar

Abstract In the field of data mining and knowledge discovery, scientometric is an emerging research area for technical publications. Knowledge discovery from the publications and their profiles is of interest across academic and scientific world. Such knowledge is utilized for policy-making and various decisions about institutions and individuals both. Thus, various metrics as well as scientific databases have been designed for analyzing the publication data. In this paper, we examine such indices and data sources, and their benefits and limitations. We argue that ranking of a journal varies with respect to databases. Our study is limited for Computer Science (CS) sub-fields. Therefore, we explore ranking of journals of CS sub-field-wise with respect to such databases. Further, we carry-out a comparative study of ranking of journals from three CS sub-fields. We empirically found that the relation between the ranking of journals, e.g. Scopus versus Google Scholar, is valid between 36% to 81% in CS sub-domains.

Keywords Computer science research · Scientometric resources · Journal ranking

1 Introduction

Knowledge discovery is one of the key aspects of assessing scientific activity in the field of scientometric. Since the publication data is expanding day-to-day, dedicated Data-Mining (DM) methods are required to evaluate publication data. DM provides necessary tools and techniques to explore and exhibit the relevant patterns that can be benefited to the whole scientific community. Such extracted knowledge is utilized while taking decisions regarding promotion, hiring, tenure, funding, and performance assessment of researchers as well as institutions [1, 2]. Hence, evaluating research publication data is a prime concern across the global scientific and academic community. However, a variety of scientometric metrics and databases have been designed for analyzing and quantifying the publication venues as well as individual researchers. Specifically, numerous indicators are designed by researchers

P. Kumari (✉) · R. Kumar
School of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_71

as well as academic-institutions for quantifying journal publications, namely Impact Factor (IF), Cited half-life, CiteScore, Eigen-Factor, Immediacy Index, and others. Such indicators are so prevalent and used frequently that these are treated as an essential yardstick in academic and scientific communities.

Other than indicators, several databases are designed that categorize publication venues on the basis of several metrics. For example, Scopus database indexes journals based on CiteScore value, Google Scholar utilizes h5-index for ranking journals and conferences.

Besides the widespread use, such indicators and databases are limited to several factors. Most of these indices are misleading at times, and the values of such indicators may not be interpreted trivially. Moreover, these indicators are usually based on citation count and it has been analyzed that citation-based indices are misleading at times [3]. Most of these metrics are unsuitable for cross-disciplinary comparisons. Thus, several research efforts are dedicated to find-out alternative ways to quantify journal publication. Researchers have included different approaches and algorithms, namely the weighted page-rank, P-rank-based indicator, mutual reinforcement-based model to quantify journals [4–6]. Yet, there is no single metric across disciplines. This is a matter of current discussion and research.

We believe that ranking of these journals differs with respect to databases. There are a few studies present in the context of Computer Science (CS) journals ranking [1]. Therefore, in this study, we explore CS sub-fields-wise journal ranking based on Scopus and Google Scholar databases. We compare the ranking of journals with respect to these databases.

The rest of the paper is organized as follows: we discuss various scientometrics indicators, resources, and the related work in Sect. 2. Then in Sect. 3, we explore and compare ranking of a few journals taken from a few selected CS sub-domains. Finally, we conclude our findings in Sect. 4.

2 Background

Scientometrics mainly focuses on assessing the quality and quantity of scientific literature [7]. More importantly, scientometric methods are widely utilized by various academic and scientific processes. Thus, measuring and analyzing the high volume of publication data is one of the important tasks. For this purpose, various indicators have been defined, and several databases have been established as resources for analyzing the publication data.

2.1 Scientometric Indicators

Scientometric indicators are widely utilized for (i) quantifying the publication venues, and (ii) assessing performance or impact of individual researchers and their groups. The most basic indicators are publication and citation counts. Based on these indices, a variety of indicators have been designed, such as Impact Factor (IF), Cited half-life, Immediacy Index, and others. IF is mainly used for measuring the impact of journal publications. The cited half-life helps to evaluate the age of the majority of cited articles published in a journal [7, 8]. Additionally, Eigen-Factor is considered more robust than IF and treated as a better metric for measuring the journal importance. Apart from such indices, CiteScore, SCImago Journal Rank (SJR), Source Normalized Impact per Paper (SNIP), etc., are emerging metrics for evaluating the impact [6]. For example, the impact of Scopus-based journals is expressed by CiteScore, and SNIP is utilized for comparing the impacts of journals of different disciplines. Furthermore, Journal Citation Reports (JCR), Eigenfactor Score (ES), and Article Influence etc., metrics have been designed by academic institutions to quantify journals. However, most of these indices are biased at times, and influenced by several factors [3]. Therefore, there are several studies by researchers, who have analyzed these metrics and inferred that a single metrics *alone* might not be suitable for evaluating purposes [9].

2.2 Scientometric Resources

One of the most commonly used databases for scientometrics is Web of Sciences (WoS). The WoS is a multi-disciplinary resource that included journals worldwide across major disciplines, namely social sciences, sciences, humanities, and arts. The WoS indexes mostly journals articles. Recently emerged, Scopus and Google Scholar can be seen as major alternatives to the WoS. Scopus, designed by Elsevier publication house, covers a variety of sources, e.g., conference proceedings, journals, book series, and trade publications. The Scopus ranks all these sources based on the CiteScore metrics. Moreover, Google Scholar (GS) is a web-based open access datasource that includes almost all types of sources such as journal articles, conference papers, books, reports, dissertations, theses, and unpublished reports [2]. On the basis of the h5-index, the GS ranks the top 20 publications.

2.3 Emerging Approaches for Quantifying Journals

In addition to the above, several researchers have provided alternatives to quantify journals. For example, Bollen et al. have proposed Y-factor by combining impact factor and Weighted Page-Rank (WPR) algorithm for expressing the status of journals. On the basis of the citation patterns, they have explored that popular journals are cited frequently by little prestigious journals; in the process, such low prestigious journals acquire very high IF with significantly low WPR value. In contrast, the prestigious journals are cited much less as they receive citations from highly prestigious journals. Thus, such journals have poor IF and very high WPR [4].

In addition, a group of researchers presented mutual reinforcement-based model to rank journals. Their model included several links of networking, namely authors-paper network, paper-journal network, and others. They concluded that the prestige of a journal depends on the authors and their prestige [5].

Feng et al. have explored several evaluation indicators for the journals. With the help of a few examples, they have shown that the ranking of journals varies with respect to different indices. Based on experimental results, several researchers have concluded that CiteSore, h-index, and IF indicators are good for evaluating journals [10]. Thus, it can be inferred that there are several factors affecting prestige of a journal. This is an active research area in scientometrics.

3 Results and Discussion

In this study, we have utilized Scopus and Google Scholar (GS) as the sources for exploring and analyzing the various indicators and ranking of journals. We have randomly selected five journals from each of the three selected CS sub-fields. The selected CS sub-fields are: Artificial Intelligence (AI), Computer Networks (CN), and Computer Vision and Pattern Recognition (CVPR). These are sub-fields which are categorized in both Scopus and GS databases. We also included some of the indicators, namely CiteScore, h5, SJR, SNIP etc. Based on these sources, this study is focused on a comparative analysis of journal ranking of selected CS sub-fields.

3.1 Indicators and Ranking of Corresponding Journals

Tables 1, 2 and 3 present the ranking and indicators' value of journals of three CS sub-domains. Specially, ranking of journals is based on the value of Cite Score and h5. It can be seen that ranking of journals is varying with respect to the chosen metrics. The maximum difference between these ranking is 12, and minimum is 3 in the AI sub-field (Table 1), whereas the difference is between 1 and 24 in CN area (Table 2), while this is in the range of 1 to 19 for CVPR area (Table 3). Thus, this shows that the ranking of journals varies.

The CiteScore and h5 indices consider different parameters to categorize the journals, there are significant differences in ranking as seen with respect to these metrics. It may be noted that the values of SJR and SNIP also vary according to the rank.

It can be seen, for example, that the Journal Machine Learning Research (JMLR), having SJR = 2.219 from Table 1, scores higher than the other three journals in the table, yet JMLR gets lower rank with respect to other metrics. From tables, we can also infer that higher SJR and SNIP values do not necessarily mean higher rank.

3.2 Local Rank Based Statistical Exploration

The ranking of journals that we have shown in Tables 1, 2 and 3 has taken from Scopus and GS databases. Such data is raw data and limited to a few samples, therefore, we assigned local rank within the sub-field for knowing the correlation between two ranks. We assigned rank 1 to the highest value of CiteScore and h5-index, rank 2 to second-highest, and so on.

Based on Spearman's rank correlation, we find that the relationship between the rank of journals (Scopus vs. Google Scholar) is valid between 36% to 81%. From Fig. 1, it can be seen that rank correlation is significantly lower in the CS sub-field of AI, and better in CVPR.

Table 1 Journals of artificial intelligence (AI)

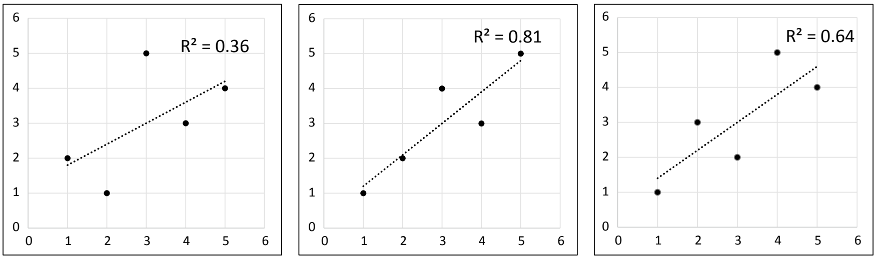
Journals (artificial intelligence)	CiteScore		h5		SJR	SNIP
	Value	Rank	Value	Rank	Value	Value
IEEE Trans. Neural Networks Learning Sys.	17.6	4	107	7	3.555	3.726
Expert Systems Applications	11.0	16	111	5	1.494	3.139
Neural Networks	10.0	18	64	15	1.718	2.246
Neurocomputing	9.5	20	100	8	1.178	2.044
Journal Machine Learning Research (JMLR)	9.3	21	82	13	2.219	3.147

Table 2 Journals of computer network (CN)

Journals (computer networks)	CiteScore		h5		SJR	SNIP
	Value	Rank	Value	Rank	Value	Value
IEEE Communications Magazine	23.4	1	130	2	4.025	4.403
IEEE Journal Selected Areas Communications	21.7	2	101	5	3.885	3.666
IEEE Networks	21.6	3	68	13	2.773	3.323
Journal Network & Computer Applications	13.8	9	83	7	1.389	3.154
IEEE/ACM Trans. Networking	9.0	26	59	17	1.498	2.112

Table 3 Journals of computer vision and pattern recognition (CVPR)

Journals (computer vision and pattern recognition)	CiteScore		h5		SJR	SNIP
	Value	Rank	Value	Rank	Value	Value
IEEE Trans. Pattern Analysis & Machine Intelligence	35.2	2	127	4	7.536	11.910
Int. Journal Computer Vision	15.3	5	70	8	4.121	4.559
Pattern Recognition	13.1	6	85	6	2.323	3.638
Computer Vision and Image Understanding	8.7	11	50	15	1.453	2.255
Pattern Recognition Letters	6.3	17	59	10	0.848	2.021



(a) AI (b) CN (c) CVPR

Fig. 1 Scatterplot of ranking of journals (Sopus vs. Google Scholar) of CS sub-fields

4 Conclusion

In this paper, we explore indicators of multi-sources as well as analytical discussion about ranking of the journals of CS sub-fields. This study asserts that the ranking of the journals varies widely with respect to the chosen databases. Moreover, a higher value of metrics does not necessarily indicates a higher rank. Generally, it is believed that the quality of journals decided by their ranking, whereas this study argues that ranking is diverse in the sub-fields of CS research. Therefore, we believe that such ranking should not be considered blindly as the sole criteria for assessment process.

References

1. Franceschet M (2010) A comparison of bibliometric indicators for computer science scholars and journals on Web of Science and Google Scholar. Scientometrics 83(1):243–258
2. Yang K, Meho LI (2006) Citation analysis: a comparison of Google Scholar, Scopus, and Web of Science. Proc Am Soc Inf Sci Technol 43(1):1–15
3. Ioannidis JPA, Baas J, Klavans R, Boyack KW (2019) A standardized citation metrics author database annotated for scientific field. PLoS Biol 17(8):e3000384
4. Bollen J, Rodriquez MA, Van de Sompel H (2006) Journal status. Scientometrics 69(3):669–687
5. Yu D, Wang W, Zhang S, Zhang W, Liu R (2017) A multiple-link, mutually reinforced journal-ranking model to measure the prestige of journals. Scientometrics 111(1):521–542
6. González-Pereira B, Guerrero-Bote VP, Moya-Anegón F (2010) A new approach to the metric of journals' scientific prestige: the SJR indicator. J Informetr 4(3):379–391
7. Mingers J, Leydesdorff L (2015) A review of theory and practice in scientometrics. Eur J Oper Res 246(1):1–19
8. Roldan-Valadez E, Salazar-Ruiz SY, Ibarra-Contreras R, Rios C (2019) Current concepts on bibliometrics: a brief review about Impact Factor, Eigenfactor score, CiteScore, SCImago journal rank, source-normalised impact per paper, H-index, and alternative metrics. Ir J Med Sci 188(3):939–951
9. Setti G (2013) Bibliometric indicators: why do we need more than one? IEEE Access 1:232–246
10. Feng L, Zhou J, Liu SL, Cai N, Yang J (2020) Analysis of journal evaluation indicators: an experimental study based on unsupervised Laplacian score. Scientometrics 124(1):233–254

Quantitative Analysis of Breast Thermograms Using BM3D Denoising Method and Features Extraction

N. Sriraam, N. Kavya, N. Usha, D. Sharath, B. Venkatraman, and M. Menaka

Abstract Breast thermography is an evolving imaging technique used for detection of breast cancer. The higher metabolic activity of cancer cells will increase the temperature of the affected and its surrounding regions which can be captured using infrared imaging. It is well known fact that the application of filtering plays a crucial role towards making effective quantitative metrics. This study makes use of BM3D image filtering followed by CLAHE enhancement process. Wavelet entropy and gray level run length matrix features were extracted and support vector machine pattern classifier is introduced to discriminate normal and cancerous breast thermograms. Two datasets, one local hospital and other open source visual lab database was considered for the study.

Keywords Breast cancer · Thermography · BM3D technique · Wavelet entropy features · GLRLM · Support vector machine

1 Introduction

Today, the breast cancer is one of the major causes of death in women worldwide. Breast cancer starts at the lobes of mammary glands and it is due to fast growth of cells in the breast tissues [1]. Breast cancer is the most commonly occurring cancer in women and the second most common cancer overall. There were over 2 million new cases in 2018 [2]. The reason for occurrence of breast cancer could be external or internal causes. Hence early detection of breast cancer is very important which saves life.

The sudden change in the skin temperature indicates illness and the functional changes undergoing inside the body. Generally, the metabolism of cancerous tissues are very high which increases the surface temperature. Since 1960s in medical field,

N. Sriraam (✉) · N. Kavya · N. Usha · D. Sharath
Center for Imaging Technologies, Ramaiah Institute of Technology, Bengaluru, India
e-mail: sriraam@msrit.edu

B. Venkatraman · M. Menaka
Safety, Quality and Resources Management Group, IGCAR, Kalpakkam, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_72

the infrared thermography has been used to diagnose various problems associated with the body parts. Infrared thermography is a non-invasive technique which is used to measure the body surface temperature based on IR radiation emitted from the skin surface [3]. Then the obtained temperature distribution is displayed as digital image using highly sensitive and high speed thermal camera.

Thermography is harmless, fast, non-contact, non-ionising and safe diagnostic procedure. To produce the correct thermographic results, some standards must be followed in the medical field. Some generalized settings for thermal camera, maintaining proper environmental conditions and patient preparation are needed.

Computer aided diagnostic tool has been useful in the regular clinical practice for analysis of various abnormalities [4]. In this work, various statistical and texture features were extracted and analysis was done based on the BM3D denoising technique. The block matching 3D filter is the denoising method which preserves edges and blurs the homogenous regions by finding similarities in the input image by using 3D transfer collaborative filtering. The paper is arranged as follows. A literature survey on the feature extraction and BM3D technique is presented in Sect. 2. Section 3 illustrates the methodology in detail. Section 4 describes the results executed and Sect. 5 includes discussion and Sect. 6 concludes the work.

2 Related Background

Several studies have been carried out on the analysis of thermal images for the detection of breast cancer. Analysis of breast thermograms is a very challenging task due to the constraints like low contrast and noisy images. Hence the preprocessing is very important step before performing any analysis.

In BM3D, the groups selects the blocks which are same as reference blocks. Matching technique is useful in considering the variances among two blocks compared in the block matching step [5]. In [6], it is shown that using the proposed wiener filter in BM3D capable of achieving high quality image de noising. In [7], three morphological components are considered for the denoising of the images. Contour, texture and smooth are the three components used. Image is partitioned and the smallest block size is grouped as contour components, medium blocks as texture components. Largest sized blocks are grouped under the smooth components. Adaptive block matching is used for preserving the edges of an image in [8].

Feature extraction techniques are helpful in recognition of images as features define the image behavior. Feature information should be insensitive to irrelevant input variability with high computational efficiency [9]. Asymmetry analysis of breast thermogram images are carried out by extracting GLCM features from both normal and abnormal images. The images are denoised before feature extraction and denoising is done by using BM3D technique. The asymmetry is calculated on both raw images and denoised images [10]. The statistical information about image characteristics are obtained from different entropy features and these parameters are extracted from the sub images produced by the procedure of wavelet decomposition

[11]. The feature extraction plays a vital role in getting accurate results and it depends on the type of the image selected [12].

In this work, BM3D method was used to denoise the breast thermal images. Texture features such as wavelet entropy and GLRLM were extracted and analyzed.

3 Methodology

The overview of the proposed technique is shown in Fig. 1. The thermal images are converted into grayscale images and preprocessing is done using denoising and enhancement methods. Then 7 features from GLRLM and 5 wavelet entropy features were extracted.

3.1 Data Acquisition

For the experiment, totally 26 thermal images were collected (13 normal cases and 13 abnormal cases) from M S Ramaiah Hospital, Bangalore. The FLIR 650sc thermal camera was used for image acquisition with spectral range of 7.5–14 μm and thermal resolution of 640 $\times$ 480. Before the data acquisition, the subjects were asked to disrobe and rest for at least 15 min. The screening was carried out in the controlled room temperature. Images were captured in frontal, oblique and lateral angles. In the present work, the frontal view is considered.

Visual lab is publically available database. Infrared thermograms were acquired at the Fluminense Federal University (UFF) using a FLIR Thermal SC620 infrared camera. In this work, 100 images (50 normal cases and 50 abnormal cases) were considered.

3.2 Image Denoising Using Block Matching 3D Filtering (BM3D) Technique

BM3D groups the similar 2D fragments of images into 3D data arrays. Collaborative filtering gives the finest details of fragments and presents important features of each block [5].

The steps involved in the BM3D techniques are (a) grouping (b) Collaborative filtering and (c) Aggregation.

- Grouping: The similar patches in the image together to form 3D arrays. In the block matching group technique, the groups are not necessarily disjoint. The similarity between the reference and candidate fragments are tested by considering the threshold.

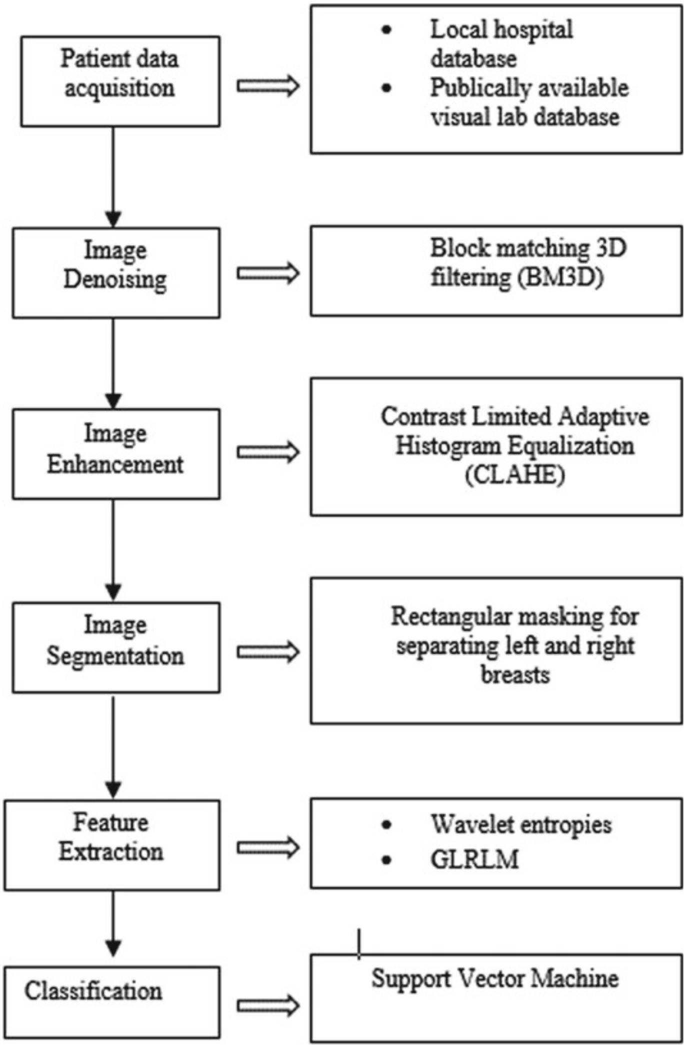


Fig. 1 Proposed block diagram

- Collaborative filtering: After grouping procedure, the linear transformation is applied in the 3D array of the grouped blocks. Then the noise is attenuated by shrinkage using weiner filtering on the transform coefficients of 3D array. Then the inverse of the 3D transform is applied to reproduce all filtered fragments.
- Aggregation: By adjusting the weights, all the block wise estimates are aggregated to obtain the final estimate.

3.3 Contrast Limited Adaptive Histogram Equalization (CLAHE)

In this technique, enhancement is applied over all neighbourhood pixels and transformation function is obtained. Exponential distribution parameter is used for histogram equalization graph shape and 'clip limit' function is used to prevent the oversaturation in homogeneous areas of the image.

3.4 Image Segmentation

The rectangular masking is used to segment the left and right breasts by detecting lower and upper breast boundaries and mathematical operations were used to separate the left and right breasts. Manual segmentation was used to remove the under breast of the image.

3.5 Feature Extraction

- Wavelet Entropy Features: DWT is employed to extract the entropy features. The important region is represented in multi resolution by DWT at various decomposition levels. In this work, four decomposition levels are used. A 2D DWT can be implemented by applying 1D DWTs in two dimensions separately [6]. The image is divided into four sub bands (LL, HL, LH and HH), then the spectrum of original image is splitted by the wavelet. Then the statistical information of normal and abnormal images were computed by extracting different entropy features such as shannon, log sure, threshold and norm.
- Grey Level Run Length Matrix (GLRLM) Features: In GLRLM, the texture is considered as grey intensity pixel pattern in a particular direction from the reference pixels [13]. The number of adjacent pixels having same grey intensity in a specific direction is considered as run length. The texture features extracted from the GLRLM are Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray Level Non-Uniformity (GLN), Run Percentage (RP), Run Length Non-Uniformity (RLN), Low Gray Level Run Emphasis (LGRE) and High Gray Level Run Emphasis (HGRE).

3.6 Classification Using Support Vector Machine

The support vector machine is a nonlinear classifier widely used due to its ability of learning supervised model to separate the binary classes. It estimates the best hyperplane which identifies the most significant margin between two classes to classify high dimensional data [14]. The kernel functions convert the non-separable data into separable data by transforming the low dimensional input space into higher dimensional space.

4 Results

In this study, 26 front breast thermograms were taken from MSRIT proprietary database and 100 front view images were used from visual lab database. Original image contains the temperature scale and labels which were removed using FLIR software. As in the Fig. 2, the pseudocolored thermal image converted into grayscale

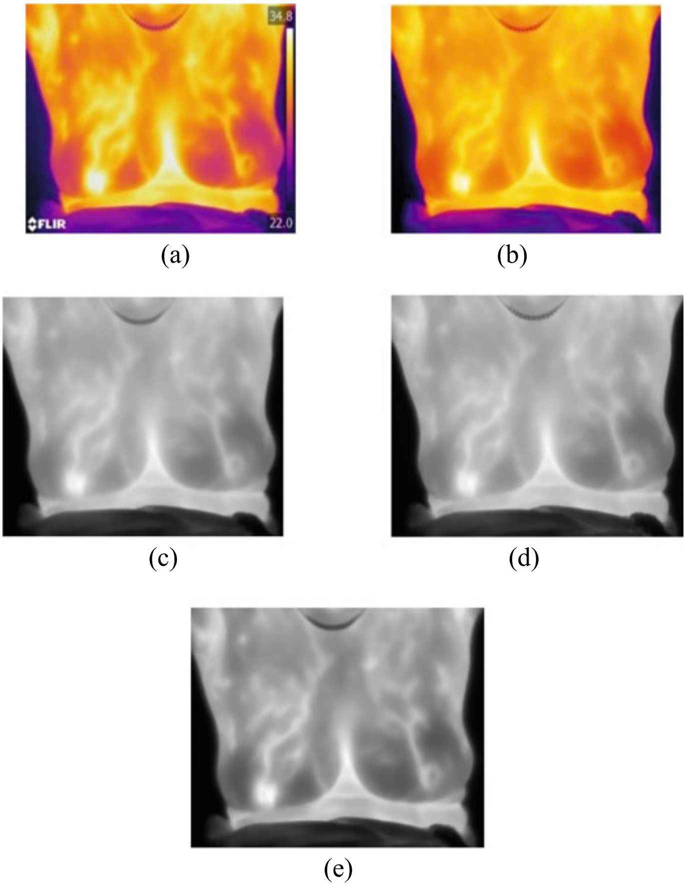
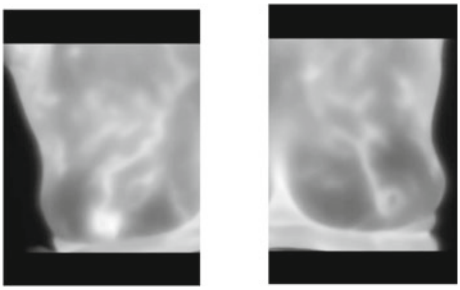


Fig. 2 **a** Original image, **b** label removed image, **c** gray scale image, **d** BM3D filtered image, **e** CLAHE enhanced image



(a) (b)

Fig. 3 **a** Right segmented breast, **b** left segmented breast

image to make computation easy. The filtered and enhanced results are shown in Fig. 2.

The Breast boundaries were detected using rectangular masking. The unwanted under breast and under arm part was removed manually as shown in Fig. 3.

The performance of the scheme is evaluated using Receiver Operating Characteristics (ROC) plot. ROC curve is the two-dimensional plot which shows the trade-off between hit rates and false alarm rates of classifier in which the area under the curve was measured [15].

The equations for specificity, sensitivity and accuracy are given below, TN is true negative, TP is true positive, FN is false negative and FP is false positive.

$$Specificity = \frac{TN}{TN + FP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

4.1 DWT Based Entropy Features

The performance parameters such as accuracy, sensitivity and specificity were measured for the entropy features before and after denoising technique. From Tables 1 and 2, it is shown that the values of performance parameters is better for denoised images in both proprietary and visual lab databases.

The AUC measures the complete two dimensional area underneath the entire ROC curve from (0, 0) to (1, 1). Figure 4 shows the ROC curve for wavelet entropy features before and after BM3D technique using local hospital database and Fig. 5 shows the

Table 1 Performance parameters for entropy features using local hospital database

Entropy features using proprietary database		
	Before BM3D (%)	After BM3D (%)
Accuracy	88.5	92
Sensitivity	92	92
Specificity	85	92

Table 2 Performance parameters for entropy features using visual lab database

Entropy features using visual lab database		
	Before BM3D (%)	After BM3D (%)
Accuracy	84	91
Sensitivity	84	92
Specificity	84	90

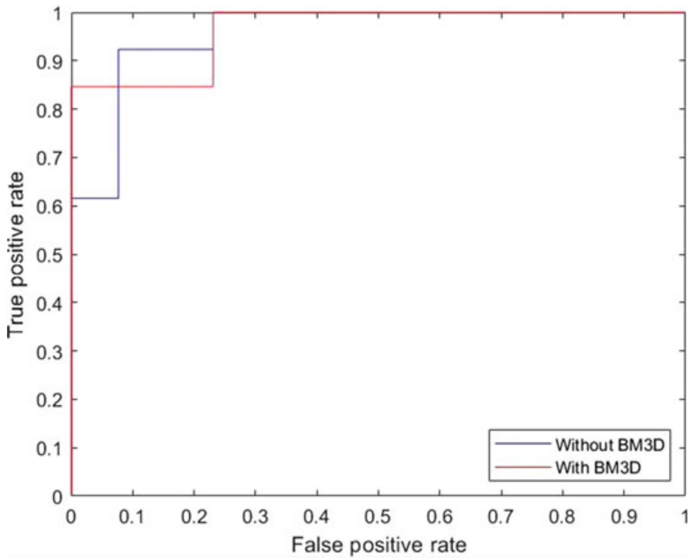


Fig. 4 ROC curve for wavelet entropy features with and without BM3D using local hospital database

ROC curve for entropy features before and after BM3D technique using visual lab database. The AUC is better for denoised images comparatively in both databases.

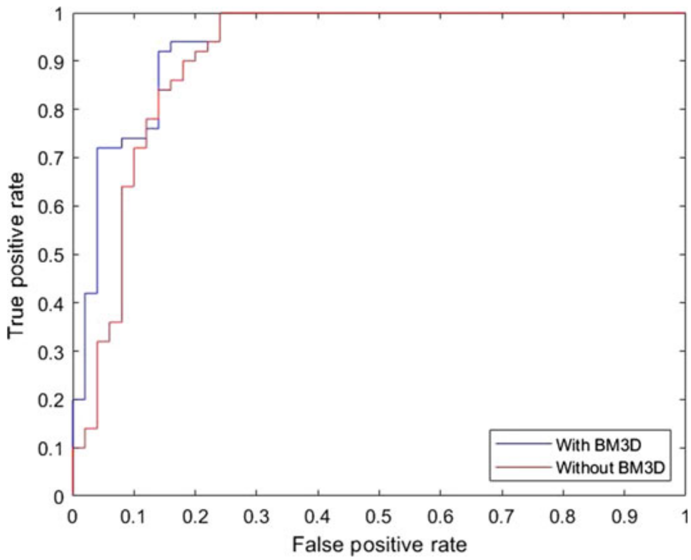


Fig. 5 ROC curve for wavelet entropy features with and without BM3D using visual lab database

4.2 GLRLM Features

Tables 3 and 4 shows the performance parameters for GLRLM features using both databases. In both cases, the performance parameter values are high after denoising technique.

Figures 6 and 7 are the ROC curves with and without BM3D technique using local hospital database and visual lab database respectively.

Table 3 Performance parameters for GLRLM features using local hospital database

GLRLM features using proprietary database		
	Before BM3D (%)	After BM3D (%)
Accuracy	88.5	92.5
Sensitivity	100	100
Specificity	77	85

Table 4 Performance parameters for GLRLM features using visual lab database

GLRLM features using visual lab database		
	Before BM3D (%)	After BM3D (%)
Accuracy	80	84
Sensitivity	80	84
Specificity	80	84

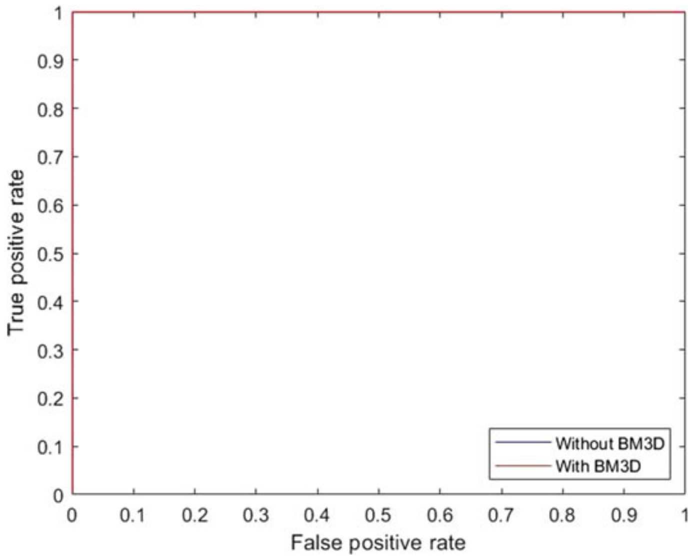


Fig. 6 ROC curve for GLRLM features with and without BM3D using local hospital database

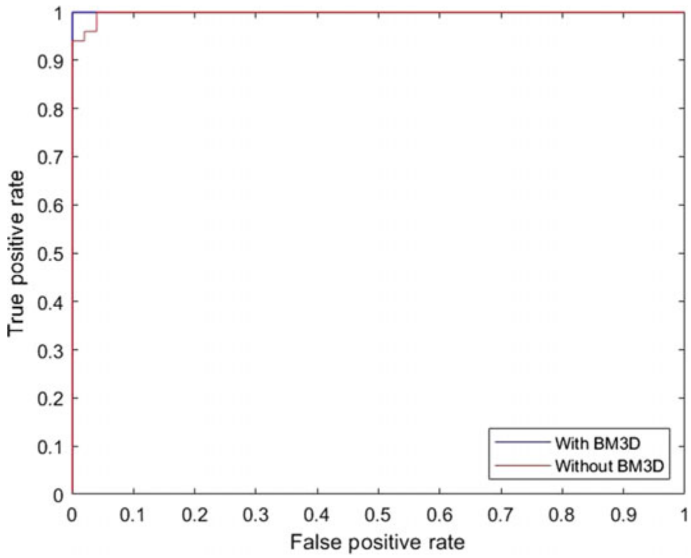


Fig. 7 ROC curve for GLRLM features with and without BM3D using visual lab database

From the above obtained results, the accuracy, sensitivity and specificity increases after denoising the images using BM3D technique. This shows that the BM3D technique is able to denoise the images efficiently along with preserving clear edges.

The analysis was done using different features. The highest accuracy of 92.5% is obtained for proprietary database using GLRLM features and highest accuracy of 91% is obtained for visual lab database using entropy features. This work seems to be helpful in accurate clinical analysis of breast thermograms.

5 Discussions

Thermography is a non-ionizing and non-contact technique which provides the significant information about breast cancer by analyzing thermal vascular patterns. Extraction of several texture and statistical features helps in quantizing the differences between normal and abnormal breasts and contribute to the analysis of breast thermograms. In the present work, the computer aided diagnosis tool is used for breast cancer detection. The accuracy of CAD technique depends on the image quality. Filtering plays very important role in suppressing unwanted noise and enhancing the image details. In the current work, the BM3D technique is used for filtering purpose. The study is performed with a limited data size. Upon considering the high volume of data, the result will be enhanced and yield us better and efficient values than any other filtering techniques. The wavelet entropy and GLRLM features were extracted and classified using support vector machine. The proposed system requires 7.18 s under MATLAB environment to diagnose the single image.

Table 5 shows the summary of various research works done using visual lab database.

6 Conclusion

Thermography is gaining popularity because of its noninvasive nature, low cost and portability properties. In this work, the breast thermal images were denoised using BM3D technique. Different features such as wavelet entropies and GLRLM features were extracted from the images before and after BM3D denoising process and classification was done using support vector machine. The classification accuracy was more for the denoised images compared to non-denoised images. It appears that this work could be used to perform the quantitative analysis of breast thermograms and hence this study seems to be clinically helpful for the detection of normal and abnormal images.

Table 5 Summary of various research works done using visual lab thermal database for breast cancer detection

Authors	Features	No of images	Classifiers	Performance
Maira Araujo de Santana. et al. [16]	Haralick and Zernike attributes	100 Patients (219 cyst, 371 benign, 235 malignant)	MLP	Accuracy:76.01%
Usha Rani Gogoi et al. [17]	Statistical and Texture	35 Abnormal 45 Normal	SVM	Accuracy:87.5%
Marcus C Araujo et al. [18]	Temperature values from morphological and thermal matrices	50 Patients	Linear Discriminant, Euclidean distance, Mahalanobis, Uniform Parzen Window	Misclassification Rate: 16%, Sensitivity: 85.7%, Specificity: 86.5%
Dayakshini Sathish et al. [19]	DWT Based Local Energy	100 (47 Abnormal 53 Normal)	SVM	Accuracy: 91% Sensitivity: 87.23% Specificity: 94.34%
Vijaya Madhavi et al. [20]	BEMD and URLBP	67 (24 Abnormal 43 Normal)	LSSVM	Accuracy: 86% Sensitivity: 92% Specificity: 73%
Proposed method	Wavelet entropy and gray level run length matrix features	100 (50 Abnormal 50 Normal)	SVM	Using GLRLM Accuracy: 84% Sensitivity: 84%, Specificity: 84% Using Entropy Accuracy: 91% Sensitivity: 92%, Specificity: 90%

Acknowledgements We thank the radiology department and surgery department of M.S. Ramaiah Memorial Hospital and staff for their valuable support. We would like to express our gratitude to the patients who participated in this study. We thank IGCAR, Kalpakkam for funding the project (IGC/HSEG/RSD/CP-01/2018).

References

1. Raghavendra U, Rajendra Acharya U, Ng EYK, Tan J-H, Gudigar A (2016) An integrated index for breast cancer identification using histogram of oriented gradient and kernel locality preserving projection features extracted from thermograms. Quant InfraRed Thermogr J 13(2):195–209
2. https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics.

3. Singh J, Arora AS (2019) Automated approaches for ROIs extraction in medical thermography: a review and future directions. Multimedia Tools Appl
4. Raghavendra U, Gudigar A, Rao TN, Ciaccio EJ, Ng EYK, Rajendra Acharya U (2019) Computer-aided diagnosis for the identification of breast cancer using thermogram images: a comprehensive review. Infrared Phys Technol 102
5. Makinen Y, Azzari L, Foi A (2019) Exact transform-domain noise variance for collaborative filtering of stationary correlated noise. In: 2019 IEEE International Conference on Image Processing (ICIP), 2019
6. Hasan M, El-Sakka MR (2018) Improved BM3D image denoising using SSIM-optimized Wiener filter. EURASIP J Image Video Process
7. Hou Y, Shen D (2018) Image denoising with morphology- and size-adaptive block-matching transform domain filtering. EURASIP J Image Video Process 2018(10)
8. Djurović I (2016) BM3D filter in salt-and-pepper noise removal. EURASIP J Image Video Process
9. Kumar G, Bhatia PK (2014) A detailed review of feature extraction in image processing systems. In: 2014 fourth international conference on advanced computing & communication technologies
10. Sathees P, Sujatha CM, Swaminathan R (2014) Asymmetry analysis of breast thermograms using BM3D technique and statistical texture features. In: 2014 international conference on informatics, electronics and vision (ICIEV 2014), pp 1–4
11. Devi M, Audithan S (2017) Analysis of different types of entropy measures for breast cancer diagnosis using ensemble classification. Biomed Res (India). 28:3182–3186
12. Kavya N, Sriraam N, Usha N, Hiremath B, Suresh A, Sharath D, Balasubramaniam V, Menaka M (2020) Breast cancer lesion detection from cranial-caudal view of mammogram images using statistical and texture features extraction. Int J Biomed Clin Eng 9:16–32
13. Öztürk Ş, Akdemir B (2018) Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA. Procedia Computer Sci 132:40–46
14. Shivarudhrappa R, Sriraam N (2018) Classification of focal and non-focal EEG signals using neighborhood component analysis and machine learning algorithms. Expert Syst Appl
15. Rouhi R, Jafari M, Kasaei S, Keshavarzian P (2015) Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Syst Appl 42:990–1002. https://doi.org/10.1016/j.eswa.2014.09.020
16. de Santana MA, Pereira JMS, da Silva FL, de Lima NM, de Sousa FN, de Arruda GMS et al (2018) Breast cancer diagnosis based on mammary thermography and extreme learning machines. Res Biomed Eng 34(1):45–53
17. Gogoi UR, Bhowmik MK, Ghosh AK, Bhattacharjee D, Majumdar G (2017) Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. In: 2017 international conference on innovations in electronics, signal processing and communication (IESC)
18. Araújo MC, Lima RCF, de Souza RMCR (2014) Interval symbolic feature extraction for thermography breast cancer detection. Expert Syst Appl 41(15):6728–6737
19. Sathish D, Kamath S, Prasad K, Kadavigere R (2017) Role of normalization of breast thermogram images and automatic classification of breast cancer. Visual Computer
20. Madhavi V, Christybobby T (2017) Assessment of dynamic infrared images for breast cancer screening using BEMD and URLBP. Int J Pure Appl Math 114(10):261–269

Evaluating Input Representation for Language Identification in Hindi-English Code Mixed Text

Ramchandra Joshi and Raviraj Joshi

Abstract Natural language processing (NLP) techniques have become mainstream in the recent decade. Most of these advances are attributed to the processing of a single language. More recently, with the extensive growth of social media platforms focus has shifted to code-mixed text. The code-mixed text comprises text written in more than one language. People naturally tend to combine local language with global languages like English. To process such texts, current NLP techniques are not sufficient. As a first step, the text is processed to identify the language of the words in the text. In this work, we focus on language identification in code-mixed sentences for Hindi-English mixed text. The task of language identification is formulated as a token classification task. In the supervised setting, each word in the sentence has an associated language label. We evaluate different deep learning models and input representation combinations for this task. Mainly, character, sub-word, and word embeddings are considered in combination with CNN and LSTM based models. We show that sub-word representation along with the LSTM model gives the best results. In general sub-word representations perform significantly better than other input representations. We report the best accuracy of 94.52% using a single layer LSTM model on the standard SAIL ICON 2017 test set.

Keywords Code mixing · Language identification · Sub word representation · Convolutional neural networks · Long short term memory

1 Introduction

The automatic processing of text to derive insights has been widely used in the industry. It can be used to process product or movie reviews in order to derive a general sentiment. Other applications include analysis of tweets to derive perception of a brand or thoughts of people on a specific topic. All of these applications can be boiled

R. Joshi
Pune Institute of Computer Technology, Pune, Maharashtra, India

R. Joshi (✉)
Indian Institute of Technology Madras, Chennai, Tamilnadu, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_73

down to classification or summarization tasks. The state of the art NLP techniques performs very well on these tasks for single language texts. However, they may not perform well on code-mixed text due to the unavailability of enough labeled data. The code-mixed text has become relevant these days because of different social media platforms [1, 4]. Most of these platforms prefer English as the preferred medium of communication. In a multi-lingual country like India people tend to mix local language with English while using social media platforms. This is because people are more comfortable in local languages. It is natural to describe local terms or entities in local languages which result in code-mixing. Code-mixing essentially allows us to borrow terms from different languages thus aiding ease of communication. A local touch can be given to the movie reviews, product reviews, and comments by adding some details in the local language. All of these factors have led to a rise in the popularity of code-mixed text [6]. In order to understand such code-mixed text, it is important to identify the language used in different parts of text followed by language-specific processing [3]. In this work, we present different approaches for language identification in code-mixed text. In the code-mixed text, languages can be interleaved in different forms. One form of code-mixing is represented in this example, "*this is not union budget, ye to aam admi ka budget hai*" with language tagged as "eng eng eng eng eng hin hin hin hin hin hin hin". Another form can be seen as "*maine aaj WhatsApp and Facebook uninstall kiya h*" tagged as "hin hin eng eng eng eng hin hin". It is challenging to determine the language of individual words as the same word can be used in both Hindi and Engish depending on the context. For example English words like "*are*" and "*maze*" can also be used in Hindi as "*are mai ghar jaa raha hu*", and "*appke to maze hai*". To make it more challenging the social media text is normally noisy where words are written in different ways just to emphasize them. For example the in word "*good*" the letter '*o*' can be repeated multiple times to get different variations like "*the movie was gooood*". This makes it important to consider different input representations. The sentence can be processed word by word or character by character. The more recent form of representation is sub-word where a word is split into logical sub-word units [8]. The character and sub-word based representations are more agnostic to noisy text variations as compared to word representation which will treat each variation as a separate word. The focus of our work is to evaluate the performance of these input representations. This is the first work to explore sub-word based representations for Hi-En language identification. The task is to determine the language of each word in the sentence. The task is formulated as a token classification task. The deep learning models based on convolutional neural networks (CNNs) and long short term memory (LSTM) networks are the most popular techniques used for token classification. We use these simple models in combination with different input representations to evaluate their effectiveness. These models are often used with a conditional random field (CRF) to improve the performance. However, the work is restricted to simple architectures with a focus on input representation. We show that sub-word based representation coupled with these simple models perform better than other complex architectures reported in the literature. Simple architectures are also favorable as it reduces runtime speed and complexity. The language identification module should be fast and efficient as it will often be followed by other NLP modules. With this perspective, we experiment with single-layer CNN and LSTM models. The experiments show that these architectures are sufficient to reach desired accuracy levels. The main contributions of this work are:

- The effectiveness of character, sub-word, and word-based representations are evaluated for the task of Hindi-English language identification.
- The combinations of popular model architectures and input representations are also compared.

2 Related Work

In this section, we review some of the deep learning-based approaches used for code mixed language identification. Simple feed-forward neural networks utilizing character n-gram and lexicon features have shown to produce a good performance for this task [15]. Although the task can be performed at the word level, it is a common practice to use neighboring words contextual information to aid the classification process [12]. The word vectors can be directly passed to a bi-directional LSTM to encode the contextual information. Alternatively, both word vectors and character-based word vectors can be provided to the LSTM. The character-based word representations can be generated using another CNN or LSTM [10]. Multi-channel CNNs are commonly used to capture such character-based word representations [7]. The primary motive behind using character-based representation is to avoid the out of vocabulary problem. It also helps in better classification of words that have very low representation in the training data. There have been few works related to code mixed Hindi-English languages applying similar concepts [14]. The Bengali-English code mixed text has also been equally explored in literature [2]. Other Indian languages like Telugu and the Assamese have also be analyzed from the code-mixing perspective [5, 13]. We have limited ourselves to Hindi-English text because of the lack of standardized and meaningful data sets in other languages. We study the usage of plain word embedding and its combination with character-based word embedding. Sub-word based representations are also evaluated to highlight its importance in identifying code-mixed languages.

3 Dataset

The ICON 2017 code mixed sentiment analysis data set is used for evaluation [11]. The data set also has each and every word tagged with its corresponding language. The sentiment tags are ignored and only language information is used for the experiments. The code mixed text in this data set was extracted from twitter and manually annotated for sentiment and language. There are a total of 12,936 training sentences and 5525 test sentences. The split is predefined and the results reported are using the same split. A small validation set comprising of 10% of train sentences is using for hyper-parameter tuning of the models. There are around 80k Hindi and English tokens individually in the train data. In the test data, there are around 30k tokens in each class.

4 Input Representations

The input representations evaluated in this work are characters, sub-words, and words. The structure of each representation is shown in Fig. 1.

4.1 Word Representation

The word embeddings are the default representations used for processing text. In this approach, each word is considered as a token and the sentence can be seen as a series of tokens. Each token is represented using a 300-dimensional vector also known as the word vector. This distributed representation allows us to capture the semantic relationship between individual words. The sentence can now be seen as a series of word vectors that are processed using CNN or LSTM to get their contextual representations. The output of CNN or LSTM is again a series of vectors but these vectors now encode the contextual information as opposed to word vectors. This contextual vector is passed through dense layers to get final predictions. So, each word corresponds to a time step, and for each time step its corresponding label in terms of language id is predicted.

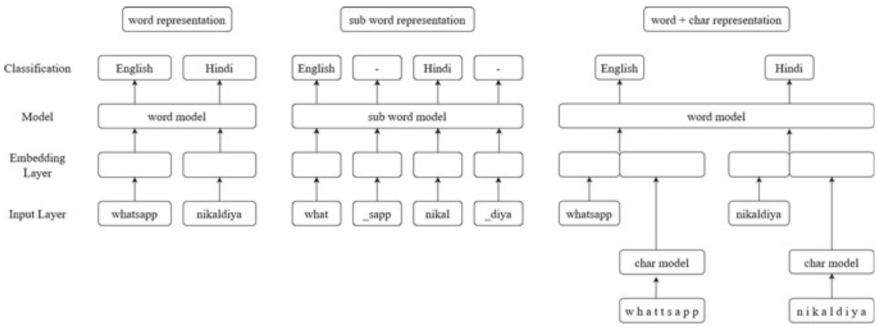


Fig. 1 Input representations

4.2 Character Representation

Another approach is to use character-based representations to augment the word vectors. The idea is to use another shallow network to process each word character by character. The series of characters can now be seen as a time series that is passed through CNN to get contextual character representation. The representations are max-pooled over time to get corresponding word representations. These character-based word representations are then concatenated with token-based word representations discussed earlier and then processed using neural networks. Simply using character-based word representations performs similar to the first word-based approach at a cost of increased complexity. So we only report results for model where character representations are used in conjunction with word embeddings. The usage of character-based word embeddings mitigates the out of vocabulary problem to some extent. The token-based word representations for unknown words will be mapped to a single unk token and the neural network will have to rely on contextual information to make predictions. Whereas the character-based word representations will always give us meaningful representations since the number of characters is fixed and known.

4.3 Sub-word Representation

The final representation strategy explored here is the sub-word embeddings [9]. This can be seen as an intermediate form with granularity somewhere between words and characters. Some of the possible sub words for word "hello" are "hell/o", "he/llo", and "he/l/l/o". This type of representation is very useful to mitigate the open vocabulary problem. The exact sub word split is determined by the statistical character n-gram properties of the training corpus. We train a uni-gram based subword model using Google sentence piece tokenizer [8]. The subword vocab size is set to 12k. This subword model is used to split each word into constituent sub words. The first subword of each word is assigned the parent language label. The subsequent subwords are assigned a dummy label. The problem is again formulated as a token classification problem where we are only concerned about the label of the first sub-word token of each word. However, during cross-entropy training, all the tokens contribute to the loss. The masking of loss from dummy labels is not explored in this work.

5 Model Architecture

- **CNN**: This is a basic CNN model based on 1D convolutions. The word or sub-word embeddings of 300 dimensions are passed through a single 1D convolution. The kernel size is 4 and the number of filters is 64. This is followed by two dense layers of size 100 and 1. The output of the individual time step is subjected to a dense layer so as to have a prediction for each word. The convolutional layer and dense layers are followed by relu and sigmoid activation functions respectively in all the models described here. Adam is used as an optimizer. The binary cross-entropy is used as the loss function as there are two output classes. The optimizer and loss function are common across all the models.

- **Multi-CNN**: In this model, three parallel 1D convolutions are applied on the word or sub-word embeddings. The filter sizes are 2, 3, and 4 with 64 filters each. The output of these convolutions is concatenated. This is followed by two dense layers of size 100 and 1.
- **LSTM**: This is a basic LSTM model. A single Bi-LSTM with 300 hidden units is used. The word or sub-word embeddings of 300 dimensions are passed through this layer. The output at each time step is subjected to two dense layers of size 100 and 1. A dropout of 0.4 is used in the recurrent layer.
- **CNN + LSTM**: This combines the basic CNN and LSTM models sequentially. The 1D CNN as described above is followed by the Bi-LSTM layer. The output is then subjected to two dense layers.
- **CharCNN + LSTM**: The three parallel convolutions as described in the Multi-CNN network is used to process characters. The output of each convolutional layer is max-pooled over time and concatenated to produce word embedding of size 192 dimension. The time axis corresponds to the number of characters in the word. The word embedding generated by this multi-CNN network is then concatenated with 300 dimension learnable word embeddings as used in previous models. The concatenated representations are passed through a single Bi-LSTM layer and dense layers. The post embedding setup is the same as the basic LSTM model described above.

6 Results and Discussion

A combination of models and input representations is evaluated on the ICON 2017 data set. The word and sub-word representations are used as input to CNN and LSTM based models. The models used are simple CNN, simple LSTM, Multi-CNN, and CNN + LSTM. The character representations are used with the CharCNN + LSTM model. In this model, characters are processed using Multi-CNN and passed to LSTM along with word representation. This model architecture was chosen as it performs better than other variations of character embedding based models. Their variations can have either CNN or LSTM for processing characters and the main network can again be based on either CNN or LSTM. The metrics used for evaluation are precision, recall, f1-score, and overall accuracy. Table 1 shows the results for model and input combinations. The LSTM model utilizing subword embeddings performs the best. Augmentation of character-based word embeddings with vanilla word embeddings boosts the performance of word-based models. Thus performance-wise word $<$ char + word $<$ sub-word relationship holds. From the model perspective CNN $<$ Multi-

Table 1 Classification metrics of different models and input combinations

Model	Input	Lang	Precision	Recall	F1-score	Acc
CNN	Word	en	90.42	92.15	91.27	91.49
		hi	90.52	90.87	91.69	
	Sub-word	en	95.21	94.97	95.09	93.86
		hi	91.62	92.01	91.81	
Multi-CNN	Word	en	91.68	91.00	91.34	91.66
		hi	91.64	92.28	91.96	
	Sub-word	en	95.60	94.99	95.30	94.13
		hi	91.71	92.69	92.20	
LSTM	Word	en	91.25	91.40	91.33	91.61
		hi	91.95	91.81	91.88	
	Sub-word	en	**95.15**	**96.13**	**95.64**	**94.52**
		hi	**93.41**	**91.81**	**92.60**	
CNN + LSTM	Word	en	91.36	91.61	91.48	91.76
		hi	92.13	91.90	92.02	
	Sub-word	en	95.82	95.17	95.50	94.39
		hi	92.01	93.06	92.54	
CharCNN + LSTM	Char + word	en	91.94	93.06	92.49	92.71
		hi	93.44	92.39	92.91	

CNN $<=$ LSTM when the only word or sub-word embeddings are considered. The CNN+LSTM gives the best score when word embeddings are given as input to the models. In general sub-word based models are significantly better as compared to models utilizing word and character representations. This is primarily because sub words not only handle unknown words but also go well with the noisy data.

7 Conclusion

In this work, we study different deep learning approaches for language identification in Hindi-English code mixed text. The problem can be seen as a token classification problem. A combination of different models and input representations are compared for their effectiveness. The models include simple CNN, simple LSTM, multi-channel CNN, CNN+LSTM, and charCNN + LSTM. The input representations used are characters, words, and subwords. We show that sub-word based models are superior as compared to word and character-based approaches. The sub word representation when coupled with simple LSTM performs the best. Just changing the input representation helps basic models achieve high performance. The sub word representation helps tackle the out of vocabulary problem and at the same time handles noisy data well. These attributes are very analogous with the data set explored in this work.

References

1. Barman U, Das A, Wagner J, Foster J (2014) Code mixing: a challenge for language identification in the language of social media. In: Proceedings of the first workshop on computational approaches to code switching, pp 13–23
2. Chanda A, Das D, Mazumdar C (2016) Unraveling the English-Bengali code-mixing phenomenon. In: Proceedings of the second workshop on computational approaches to code switching, pp 80–89
3. Das A, Gambäck B (2015) Code-mixing in social media text: the last language identification frontier?
4. Gambäck B, Das A (2016) Comparing the level of code-switching in corpora. In: Proceedings of the tenth international conference on language resources and evaluation (LREC'16), pp 1850–1855
5. Gundapu S, Mamidi R (2018) Word level language identification in English Telugu code mixed data. In: PACLIC
6. Khanuja S, Dandapat S, Srinivasan A, Sitaram S, Choudhury M (2020) GLUECoS: an evaluation benchmark for code-switched NLP. arXiv preprint arXiv:2004.12376
7. Kim Y (2014) Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882
8. Kudo T (2018) Subword regularization: improving neural network translation models with multiple subword candidates. arXiv preprint arXiv:1804.10959
9. Kudo T, Richardson J (2018) Sentencepiece: a simple and language independent subword tokenizer and detokenizer for neural text processing. arXiv preprint arXiv:1808.06226
10. Mandal S, Singh AK (2018) Language identification in code-mixed data using multichannel neural networks and context capture. arXiv preprint arXiv:1808.07118
11. Patra BG, Das D, Das A (2018) Sentiment analysis of code-mixed Indian languages: an overview of sail_code-mixed shared task@ icon-2017. arXiv preprint arXiv:1803.06745
12. Samih Y, Maharjan S, Attia M, Kallmeyer L, Solorio T (2016) Multilingual code-switching identification via LSTM recurrent neural networks. In: Proceedings of the second workshop on computational approaches to code switching, pp 50–59
13. Sarma N, Singh SR, Goswami D Identifying languages at the word level in Assamese-Bengali-Hindi-English code-mixed social media text
14. Veena P, Anand Kumar M, Soman K (2018) Character embedding for language identification in Hindi-English code-mixed social media text. Comput Sist 22(1):65–74
15. Zhang Y, Riesa J, Gillick D, Bakalov A, Baldridge J, Weiss D (2018) A fast, compact, accurate model for language identification of codemixed text. arXiv preprint arXiv:1810.04142

Handling Class Imbalance in Fraud Detection Using Machine Learning Techniques

Reshma George and Bidisha Roy

Abstract A disproportion between the fraudulent and the non-fraudulent class is generally termed as class imbalance. In class imbalance problems the instances of the positive class (minority) are much smaller than the number of instances of the negative class (majority). The non-fraudulent class being more dominant and the fraudulent class being relatively rare the latter would be considered as outliers thereby resulting into a misclassification of the minority class. Hence the results of imbalanced classes are not always accurate. The applications of class imbalance ranges from fraud detection, anomaly detection, oil spillage detection, network intrusion to medical diagnosis. In all these scenarios we have cases where we have two different kinds of labels and one of them is highly skewed. The proposed model aims at using the sampling techniques like Random under sampling and Random oversampling to sample the dataset after which the balanced dataset will be given to the classifier and the results will be then compared and evaluated using the performance evaluation metrics.

Keywords Fraud detection · Imbalanced dataset · Random under sampling · Random over sampling · Synthetic minority oversampling technique

1 Introduction

Imbalanced class distribution arises when there is a significant difference in the frequency of the outcomes when dealing with binary classification. Essentially, it implies that the number of cases of the class which are less than half of the total dataset is significantly smaller than the number of cases of the class which is adequately represented [1]. Class Imbalance problem arises when the ratio of fraudulent activities (minority class) is very less in proportion to non-fraudulent activities (majority class) [2]. If we are trying to identify the fraudulent transactions from a dataset, we won't have a sufficient number. We may barely have 5% or less than that in terms of percentage. In such cases, the model may try to fit the majority class leading to

R. George (✉) · B. Roy
St Francis Institute of Technology, Mumbai, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_74

biased prediction and at the same time also provide misleading accuracy. In such cases even if we classify all the transactions as non-fraudulent, we would get 99.2% accuracy, but that is not our objective as the model won't be helpful despite giving accurate results.

Hence, when dealing with such class imbalance problems it becomes challenging to detect fraud and extract fraud patterns [2]. An example of class imbalance is the diagnosis of cancer, where we have a dataset which has 100 patients and out of which only one patient has cancer. So, the dataset here is highly skewed as the number of patients having cancer is significantly less when compared to non- cancer patients. During these situations it is very important to identify such a rare condition among the normal population as any incorrect diagnosis will give rise to a lot of stress and also lead to a lot of complications to the patients. Any incorrect diagnosis could also lead to a change in the course of treatments and medications thereby putting the patient's life at risk. Thus, it is very important for a classification model to achieve higher identification rate on the rare occurrences (minority class) in datasets and thereby producing accurate results. The problems that occurs due to imbalance classes also cause hinderance to the performance of the machine learning techniques. In this paper we are going to design a system to handle class imbalance in fraud detection using sampling methods like Random Under-Sampling, Random oversampling [3] and SMOTE (Synthetic Minority oversampling Technique) [4] to overcome the rare events in the dataset and get more accurate results.

2 Problem Description

The class imbalance problem arises when there is a significant difference between the majority and the minority class i.e. where the fraudulent activities are very less compared to criminal ones and thus leading to misclassifications and hence leading to incorrect and inaccurate results. Most of the machine learning algorithms are biased towards the majority class and hence provide inaccurate results. The main aim should be to recognize the minority class which is of utmost concern so that the misclassifications can be prevented. Hence, in this study we are going to design a system to handle class imbalance in fraud detection using sampling techniques line Random under sampling, random oversampling, and SMOTE (Synthetic Minority Oversampling Technique) [4] to overcome the rare events in the dataset and get more accurate results.

3 Related Work

One of the most important tasks that must be done while developing a project is literature survey. This is essentially required because of the changing needs of the world so that our creation is abreast with all the new and compatible hardware and software.

This section gives a basic survey of the different approaches accessible to handle the class imbalance problem in machine learning.

Class Imbalance occurs when the larger part of the observations is in one class which makes it challenging for the classifier to foresee the minority class.

A. Imbalanced Classification

The Authors in paper [2] have explained the different approaches for handling classification with imbalanced datasets along with the new trends. The performances of various algorithms like KNN, Support Vector Machine and Naïve Bayes were used out of which KNN performed the best. However, study could not be conducted on large amount of data therefore a scalable environment using big data technology was proposed. In paper [1], a hybrid of under-sampling (the negative cases) and over-sampling (the positive cases) was carried out and the performances of the three classifiers Naïve Bayes, K Nearest Neighbor and Logistic Regression (LR) were examined using accuracy, sensitivity, specificity, precision, balanced classification rate and Matthews Correlation coefficient metrics. In Paper [3] several shortcomings were discovered of existing methods and found that the approaches designed specially to tackle the imbalance problem were not adequately effective. They compared the performance of eight machine learning methods applied to credit card fraud detection and their weaknesses were identified and checked as to how effective they were in the case of extreme imbalance. It was found out that the LR, C5.0 decision tree algorithm, Support Vector Machine (SVM) and Artificial Neural Network (ANN) are the best methods according to the considered performance measures (Accuracy, Sensitivity). Also, the study showed that considering just one performance measure for imbalanced learning was misleading. Hartono et.al in paper [4] has explained the occurrence of Class imbalance when instances in a class are much higher than in other classes and found that this affects the accuracy. As SVM was comparatively weak in terms of bias data, Biased Support Vector Machine (BSVM) was used to solve the problem. After using BSVM it was studied that it could perform better in terms of control sensitivity yet there was a lack in accuracy when compared to general SVM. An integration of BSVM and SMOTEBoost was proposed as a solution to handle class imbalance problem. The Authors in paper [5] have explained the challenges faced in customer churn prediction due to the class imbalance problem. They have implemented various sampling techniques to address the issue and found it to be very effective when used with the correct classifier in dealing these kinds of problems. In paper [6] have explained the challenges faced when using data mining algorithms when dealing with imbalanced data and the techniques used to solve them. An optimized version called DSMOTE which is a combination of DBSCAN (Density based spatial clustering of application along with Noise) and SMOTE were used to address the issue and the results were found to be satisfactory in terms of precision and recall. Leevy et.al in paper [7] have used two types of approach to handle class imbalance problem namely the data level approach which involved data sampling and the algorithm level approach which includes ensemble or hybrid approach. However, after a thorough study it was found that the data level approach performed much better when compared to the algorithm level techniques. Random over sampling performed better in terms of accuracy when compared to SMOTE but there were many gaps that were identified as the results were a bit conflicting and hence, they suggested the use of more advanced techniques such as Adaptive Synthetic Sampling Approach (ADASYN) and many others. Based on the review of literature for Handling Class Imbalance, we have identified the following challenges. All these constraints the authors have come across are encouraging us to work in this domain to overcome these problems and get the desired accuracy in Handling Class Imbalance problem in Fraud Detection. The shortcomings are:

(a) The approaches that were designed to handle the class imbalance problem were not so effective. These approaches led to an increase in the number of false alarms.
(b) Sampling techniques are not used with the appropriate classification algorithms and hence the accuracy is affected.

4 Proposed System

In this study we are going to use various sampling techniques like oversampling and under sampling thereby balancing the data and then giving it to the classifier after which the results will be compared to check which sampling techniques gives accurate results.

The figure above explains the proposed work. The imbalanced dataset is given to the system which is then pre-processed so that the data is in the useable format for analysis. Now we shall select any data level approach to balance the data, we are using a sampling approach as shown in the Fig. 1. The sampling approach namely oversampling and under sampling techniques are used in the proposed system. The classification algorithms will be then applied to each sampling technique. We shall use the training data and the test data to be given to the classifier, this shall balance out the data and give better prediction and better accuracy. 80% of the data is for training and 20% data is used for testing. Hence by following the above steps and approaches, we shall eliminate the problem of class imbalance in the data [3].

Data Preprocessing.

The imbalanced dataset on Credit Card Fraud Detection from Kaggle [8] consisting of 284,807 transactions was given to the system which was then pre-processed to remove all the outliers so that the data is in the useable format for analysis. A data level approach was then selected to balance the data to avoid any misclassifications. Here to balance the data we have used a sampling approach.

(a) **Sampling Techniques:**

As the dataset is highly imbalanced with the number of non-frauds exceeding the number of frauds it is necessary to resample the dataset by decreasing the majority class or increasing the minority class observations. The techniques that are used in

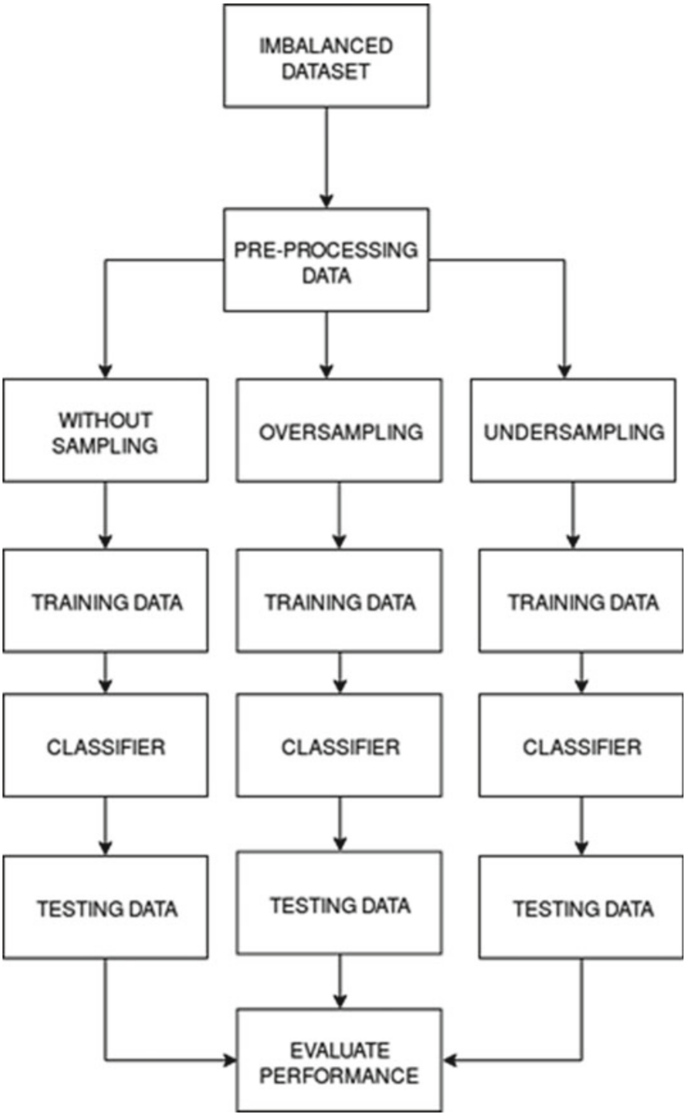


Fig. 1 Proposed System Block Diagram

this experiment are Random oversampling which involves (Synthetic Minority Oversampling Technique) SMOTE and (Adaptive Synthetic Oversampling Technique) ADASYN while Random under sampling technique involves Near Miss method.

1. SMOTE
   While dealing with classification problems, the percentage of the classes in the total sample plays an important role. Imbalance is nothing but the presence of

minority class in the dataset. SMOTE is an oversampling technique that aims at creating new synthetic observations or new data points. It involves identifying the feature vector and its nearest neighbors and finding the difference between the two. The difference is multiplied to random number between 0 and 1 and the new points are plotted.

2. Near Miss
   Near Miss is an under-sampling technique where the number of samples of majority class are decreased. Instead of sampling the Minority class, using a distance, this will make the majority class equal to minority class. There are three near miss methods [9]:

   (1) In the Near Miss -1 method the majority class samples are selected which are very close to the minority class observations and the majority class is under sampled.
   (2) The second method NearMiss-2 selects the majority class samples while their average distances to three farthest minority class samples are the smallest.
   (3) The third method NearMiss3 takes out a given number of the closest majority class samples for each minority class sample.

3. ADASYN
   ADASYN [10] is method of creating synthetic data samples for the minority class observations thereby helping to reduce the bias and misclassifications that affect the accuracy of the data. ADASYN helps to improve the learning performance by shifting the classifier decision boundary. It uses density distribution method that helps in deciding the number of samples that have to generated in the minority class to help in matching the number of observations present in the majority class.

(b) Classification Algorithms:

Classification algorithms segregates the data based on the characteristics into majority and minority classes. We have used two classifiers Random Forest Classifier and Logistic Regression and compared the performances of each when used along with the individual sampling techniques.

1. SVM (Support Vector Machine)**:** SVM technique is used for mapping the data to a high-dimensional feature space so that data points can be categorized, even when data is not linearly separable. Formally, a given training vector xi in Rn, i D 1;:::; l, n is the number of exploratory variables, and l is the number of observations in the train set. $y \in$ Rl taking the values of 1 and -1. The binary classification is done by solving the following optimization problem.
   The hyperplane equation is defined as:

$$w = T8xi + b \tag{1}$$

   where w is a vector of weights, 8(xi) maps xi into a higher dimensional space.

2. Naïve Baye's: These are known as probabilistic classifiers and have strong independence assumptions between the features. It also considers the text categorization, problem of judging the documents or reviews in any system. Naïve Baye's classifier is a kind of classifier which uses the Baye's theorem. Formula to calculate Baye's Theorem:

$$P(A/B) = \frac{P(B/A)\,x\,P(A)}{P(B)} \tag{2}$$

3. K-Nearest Neighbor Classifier: The K-nearest neighbor is a super and easy way to classify the data which carries out its classification based on a similarity measure, like Euclidean, Mahanttan or Minkowski distance functions. The first two distance measures work well with continuous variables whereas the third suits categorical variables [3].
4. Logistic Regression: Logistic Regression basically is used for problems related to binary classification where a functional approach is however used to check the probability of an answer that is binary based on one or more variables. It finds the best-fit parameters to a nonlinear called the sigmoid.
5. Random Forest: Random forest is used in majority when machine learning is concerned as the results achieved by random forest are quite overwhelming. A set of decision trees constitute a random forest and it uses the concept of averaging to prevent overfitting and thereby tend to provide better accuracy.

5 Experimental Study

In this study we have used the credit card dataset from Kaggle [10] where 80% of the data from the credit card dataset was used for training and 20% of dataset is for testing. A computer with minimum configuration of Windows 7 with 4 GB RAM and 10 GB HDD and Intel 1.66 GHz Processor Pentium 4 was required for the project work. Python 3.6.3 was used for project development and analysis. Overall, it is a low-cost research project as there is no use of any special or expensive resources.

6 Results and Discussion

In this section we are going to discuss the results of the machine learning algorithms before sampling and after sampling. Figure 2 shows the plotting of the ROC curve and the results show that Support Vector Machine and Logistic Regression performed well with an accuracy of 98% and 97% respectively. However, the results are not satisfactory as the major number of class observations being non fraud and only 5% being fraud, these results cannot be said as accurate results. Therefore, it is very necessary to balance the dataset in order to avoid any bias or misclassifications. We

Method	Accuracy	Precision	Recall	AUPRC
Near Miss LR	53%	79%	0.11	0.02
Near Miss RF	92%	78%	0.23	0.24
Random Under sampling LR	97.83	76%	0.45	0.75
Random Under Sampling RF	9.8	75%	0.80	0.78
SMOTE LR	98.3%	86%	0.11	0.78
SMOTE RF	99.93%	66%	0.98	0.81
ADASYN LR	96.83	89%	0.041	0.79
ADASYN RF	99.94	68%	0.98	0.81

Fig. 2 Table summarizing the performances of various machine learning algorithms with each sampling technique

have balanced the dataset using the resampling techniques and then used each one of them with the machine learning Algorithms and the results were compared.

We begin with the sampling technique Near Miss used along with Logistic Regression. The confusion matrices of Near Miss LR were not satisfactory and they are as follows:

Near Miss LR

Predicted	Actual	
	0	1
0	38,576	18,287
1	10	88

Near Miss RF

Predicted	Actual	
	0	1
0	14,465	42,398
1	0	98

In case of random under sampling using LR the results were better and the confusion matrices are as follows.

Random Under sampling LR

Predicted	Actual	
	0	1
0	55,661	1202
1	0	98

Random Under sampling RF

Predicted	Actual	
	0	1
0	54,661	2202
1	8	90

The results of SMOTE with LR and SMOTE with RF are shown below and SMOTE with RF has performed well with an accuracy of 99% and the confusion matrices are as follows:

SMOTE LR

Predicted	Actual	
	0	1
0	56,862	1
1	33	65

SMOTE RF

Predicted	Actual	
	0	1
0	56,233	630
1	13	85

The below confusion matrices are of Random Forest and Logistic Regression along with ADASYN sampling technique and the results of ADAYN using RF were quite promising with an accuracy of 99.5%

ADASYN LR

Predicted	Actual	
	0	1
0	54,816	2047
1	10	88

ADASYN RF

Predicted	Actual	
	0	1
0	56,862	1
1	31	67

We have now learned about all the approaches and classifiers and had a comparative study between all the machine learning algorithms. We have followed all the steps which were explained in the block diagram and then compared the performances of each machine learning algorithms when used with different sampling techniques. We have used the data-level approaches of oversampling and under sampling and compare the results of each of them to solve the class imbalance problem. For model evaluation, we used Area Under Curve (AUC), based on which we could decide whether our model is performing better after balancing the data. The results show that ADAYSN along with RF and SMOTE along with RF gave the best results with an accuracy of 99.93%. The accuracy before using sampling technique was 96% but our aim was to get a higher accuracy as the number of frauds are far less when compared to non- fraud cases. Therefore, the accuracy maybe higher but the precision and recall were far less when compared to after sampling technique. However, in terms of precision and recall they did have some shortcomings.

7 Conclusion

We have compared the performances of the classification algorithms with each sampling technique to identify the shortcomings with every pair of classification algorithms with the sampling technique. The aim of this study was to investigate if the classification model is able to classify the fraud and non-fraud transactions and to understand which sampling techniques improve the performance of the model. The performances of the classification model were better than the performance before using the sampling technique. Our test results show that ADASYN technique along with RF and SMOTE along with RF gave promising results of 99.93%. In future will also implement them with other machine learning algorithms.

References

1. Awoyemi JO, Adetunmbi AO, Oluwadare SA (2017) Credit card fraud detection using machine learning techniques: a comparative analysis. In: International conference on computing networking and informatics (ICCNI). https://doi.org/10.1109/iccni.2017.8123782
2. Ali A, Shamsuddin SM, Ralescu AL (2015) Classification with class imbalance problem: a review. Int J Adv Soft Comput Appl 7(3)

3. Makki S, Assaghir Z, Taher Y, Haque R, Hacid MS, Zeineddine H (2019) An experimental study with imbalanced classification approaches for credit card fraud detection. IEEE Access 1–1. https://doi.org/10.1109/access.2019.2927266
4. Hartono H, Sitompul O, Tulus T, Nababan E (2018) Biased support vector machine and weighted-SMOTE in handling class imbalance problem. Int J Adv Intell Inform 4(21). https://doi.org/10.26555/ijain.v4i1.146
5. Wei W, Li J, Cao L, Ou Y, Chen J (2013) Effective detection of sophisticated online banking fraud on extremely imbalanced data. World Wide Web 16(4):449–475
6. Khor KC, Ting CY, Phon-Amnuaisuk S (2014) The effectiveness of sampling methods for the imbalanced network intrusion detection data set. In: Recent advances on soft computing and data mining, Springer, Berlin, pp 613–622
7. Seo JH, Kim YH (2018) Machine-learning approach to optimize SMOTE ratio in class imbalance dataset for intrusion detection. Comput Intell Neurosci
8. Credit Card Fraud Detection Dataset: https://www.kaggle.com/dalpozz/creditcardfraud towards data science
9. Zhang J, Mani I (2003) KNN approach to unbalanced data distributions: a case study involving information extraction. In: Proceedings of the ICML'2003 workshop on learning from imbalanced datasets
10. He H, Bai Y, Garcia E, Li S (2008) ADASYN: adaptive synthetic sampling approach for imbalanced learning. In: Proceedings of the international joint conference on neural networks, pp 1322–1328. https://doi.org/10.1109/IJCNN.2008.4633969

Efficient Resistive Defect Detection Technique for Performance Enhancement of Static Random Access Memory

Sheetal Tak and Madan Mali

Abstract With growing technology, memories with very high density are getting used in a large number of applications and devices. As a result of advanced very deep submicron technology and scaling effects, the size of the devices has reduced tremendously according to Moore's law. Advanced manufacturing technology is using the devices with the single-digit channel length. But there is also the threat of process variations impact. To increase the reliability of the devices, memory present on system-on-chip should be fault-tolerant. The memory should be tolerant for variations in process, temperature, and supply voltage. Resistive open defects in memory are increasing with advanced technology. So it becomes necessary to detect these faults for reliable memory operation. In this paper, the detection of resistive open defects is performed with the help of the predischarged feeble cell detection method. The analysis of the proposed method is performed for faults considered at different locations with varied values. The proposed method gives the minimum area overhead of 7.8% and less time penalty of 40.96 μs for 2 KB of the memory.

Keywords Faults · Embedded memory · SRAM · Resistive defects

1 Introduction

Nowadays, the dominant part of the system is memory. Memory has a key role in almost every type of application in various domains. With the advanced technology of very deep submicron in the semiconductor field, the integration of many devices like many memory cores on a single system-on-chip is available. But with advanced technology, the devices are impacted more due to process variations like temperature, supply voltage, electromagnetic radiation, vibrations, etc. It has resulted in the generation of faults in the memory. Faults may also get generated during manufacturing because of changes in doping concentrations. Faults generated may lead to disturbances in the symmetry of the devices in cells which leads to instability in the cell. Resistive open defects are the defects between the connecting nodes which

S. Tak (✉) · M. Mali
Department of Electronics and Telecommunication, Sinhgad College of Engineering, Pune, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_75

normally should be connected and have zero resistance. A strong resistive defect can be detected easily since its behavior shows logical faults in the cell directly after a read or write operation performed on the cell. But weak resistive defects are harder to detect. Since they don't show any logical fault after accessing the cell. But as the number of operations on the cell increases, it converts into strong defects due to aging and shows the faulty behavior of the cell. In earlier methods, faults were detected by observing the change in voltage. But the change in voltage can't be predicted only because of defects. Change in supply currents is also used to detect the faults. But leakage currents have increased to the extent with advanced technology such that it is hard to predict the change in current is due to leakage or defect in the cell. Figure 1 shows the standard transistor architecture of Static Random Access Memory (SRAM) cell. Two inverters are present in the cell with their terminals connected to realize the latch structure which can store 1-bit data in the cell. The pmos transistor is responsible to store the data 1 in the cell. The nmos transistor takes care of the storage of logic 0 data. The diagonally opposite pair of transistors are on depending the data stored in the cell is 1 or 0. The access transistors on the two sides are used to connect the cell to bitlines for reading the data from the cell or write the data onto the cells. The faults may get generated at the source, drain or gate terminal of the device or any connecting wire between the terminals. There are 23 fault locations of resistive open defects present in the cell.

Detecting the faults by measuring the voltage at a pinpoint is not feasible in advanced technology as the number of pins is too large and placements are very critical in memory. Disturbing the cell by means by adding a current pulse into it and reading the data twice with the help sense amplifier [1]. Wordline signal extended [2] for the detection of faults in memory to get less period for precharge. Read equivalent stress method used to detect coupling faults [3]. The power dissipation in the SRAM cell can be reduced by using the gating technique [4, 5]. Linked faults are detected by using the March test method [6]. By applying near ground voltage on the bitlines for accessing the cell is used to detect the faults in the memory [7]. Constant read cycles on the same cells are avoided to increase the reliability of memory [8]. A detailed summary of March tests is shown [9]. Storing a combination of data in cells and performing read operation to generate voltage for the detection of faults is

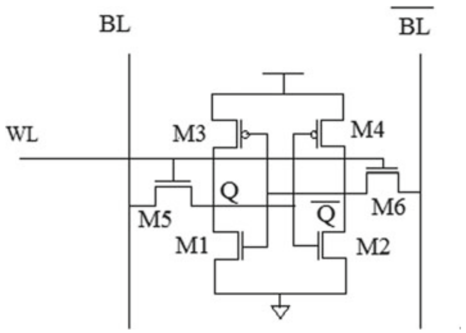


Fig. 1 Architecture of SRAM cell

discussed [10]. The resistive open defect faults can be detected by using the stress method [11]. Column redundancy can be used to repair the faults in memory [12]. The Built-In Self Repair scheme [13] is used for fault correction. Built-in Current Sensor [14] method is explored for the detection of faults in memory. Analysis of read and write stability is discussed [15]. The transient current detection [16] with the hardware approach is given in detail. The bridging defect can also be detected [17]. By using the aging effect, faults can be detected in memory [18]. The remainder of the paper is organized as follows. Section 2 gives the details of the resistive open defect in memory. Section 3 shows the detection of faults using the proposed method. Section 4 explores the results of fault detection and Sect. 5 concludes this paper.

2 Resistive Defects

The resistive defect can be caused by numerous operations performed on the cells. So due to the aging effect of layers, defects are getting generated. The defects can occur at any terminal of the transistors in the cell as shown in Fig. 2. The defects are caused as a result of scrawny contacts or connections. It shows all possible locations of the resistive defects in the cell. It may occur at word line, bit lines, or supply terminals. The connection between node terminals at any location in the cell may have defects generated. The presence of the resistive defect gives resistance to the charging or discharging path of the current. If the defect is present near supply voltage or source of any pmos transistor related logic 1 data stored, could not hold the data as logic 1 since the voltage drop across resistive defect reduces the voltage at nodes. Similarly, if the resistive defect present at the source or drain terminal of nmos, it

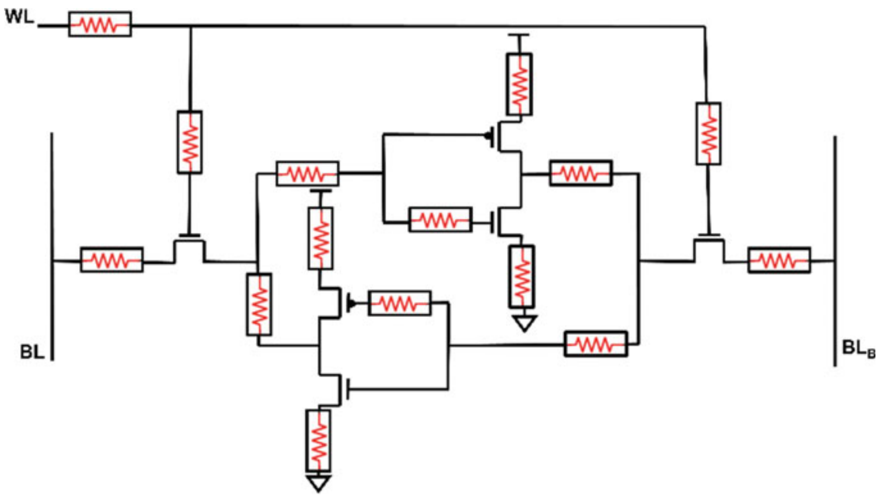


Fig. 2 Resistive defects in SRAM cell

resists the discharging current through nmos after a read operation. So improper voltage difference generated gives problems to the sense amplifier and wrong data may get forwarded to output.

As a result of the resistive defect, the cell couldn't hold the data properly depend on what location the defect is present. The defect results in the instability of the cell. The defect present at gate terminals gives a lesser impact on the cell data. A reduction in gate current does not directly impact the operation of the cell. The resistive defect present at the word line gives less voltage to activate the cell. But its impact is lesser compared to the defect present on bit lines of the cell.

3 Fault Detection by the Proposed Method

There are different methods used to detect the faults like inserting current spike, read equivalent stress, etc. Some of the methods require more hardware or complex timing circuitry. So there is a need to detect resistive faults with location information as well as less time and area overhead. In the proposed method analysis of the resistive defect done considering various locations as shown in the above diagram. For the optimum analysis of each fault, one fault is considered at a time as would be practically possible in memory. The resistive defect in the cell shows different logical behavior such as transition fault in which the cell could not transit data from 1 to 0 or 0 to 1 after write operation. In stuck at fault, the node gets permanently connected to supply or ground or nearby nodes and shows only one value after each operation.

In read destructive fault, the cell could not hold the value and data get upturned after a read operation and get incorrect read data output. In the dynamic read destructive fault, data stored in cells get upturned for a specific read cycle performed immediately after the write cycle. The read cycle at which it occurs depends on the value of the resistive defect. In the deceptive read destructive fault, data get upturned during read operation but it gives correct data at the output of read operation. In the incorrect read fault, read operation shows an inverted output even though data remain intact in the cell. Using these functional behaviors resistive defects can be detected in the cell using sequences like the March test. The proposed method in this paper based on the principle that voltage drop across the resistive defect lowers the voltage at the node. As a result, current drawn from the node will be lesser.

Figure 3 shows the block diagram of the proposed predischarged feeble cell detection method. By discharging the voltage across bit lines with the help of predischarge remove the additional charge on bit lines and after that access the cell. By giving the word line stress on the cell during normal operation, resistive faults are getting detected by the fault detection circuit connected. It is observed that strong value resistive defect gets detected earlier since their impact is more on logical behavior. Less value resistive defects show less effects on the normal operation of the cell. Initially, memory is designed using 45 nm technology.

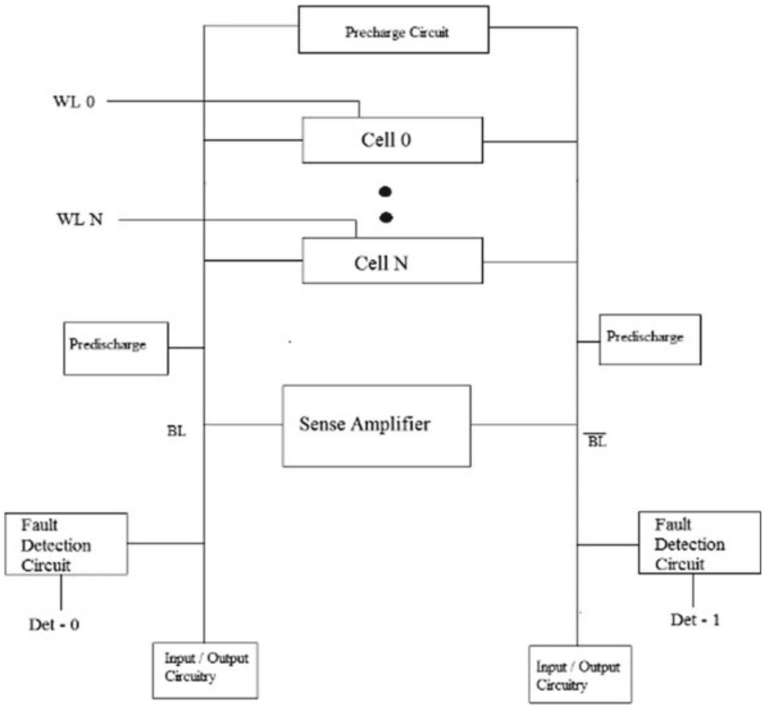


Fig. 3 Block diagram of proposed fault detection technique

4 Results

The memory is designed using 45 nm technology the Cadence virtuoso tool. The SRAM cell is designed for good read stability and write ability. The cell and peripheral circuits designed are connected and normal read and write operation of cells are verified. The number of cells is gradually increased and its effect on output waveforms is checked. The effect of the parasitic load is also taken into consideration. The effect of read or write operation on the neighboring cells have also observed. Figure 4 shows normal read and write operations performed on random locations in the memory. The access time observed was 400 ps. Later fault detection circuit is integrated with memory. The faults are generated at random locations. It is observed that resistive defect faults at source and drain terminals affect more on normal behavior than defect faults present at gate terminals. The resistive defect of high values can be sustained at the gate terminal of the device. The faults are introduced in the charging path of the pmos device. The fault is getting detected at an earlier stage. Figure 5 shows the waveform of the fault detected at the pmos device. Det-0 and Det-1 are the output waveform of the fault detection circuit connected on bitlines. Bitlines shows the waveform on bitlines BL and BLBAR. Q and QBAR are the node voltage in the cell as shown in Fig. 1. A fault is also introduced at the discharging path of the nmos

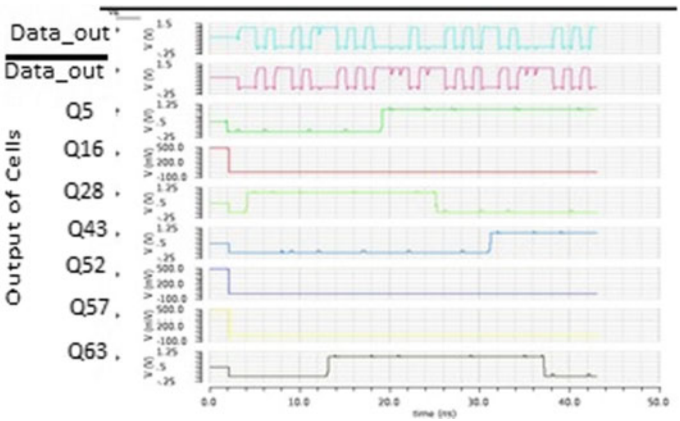


Fig. 4 SRAM read and write operation

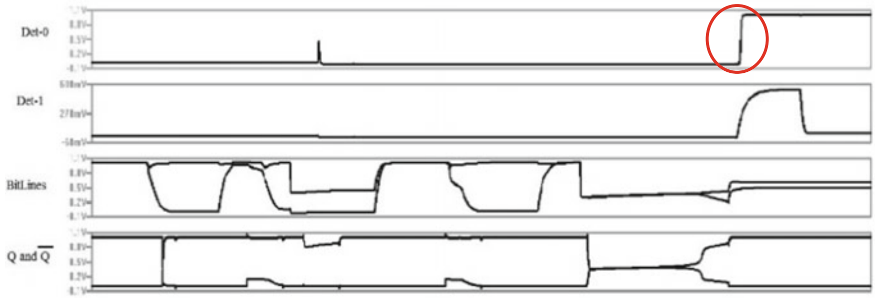


Fig. 5 Detection of resistive defects at nmos location in the cell

device. The weak resistive defect also gets detected on these paths by the proposed method. Figure 6 shows the waveform of the fault detected at nmos device in the cell. Similarly, faults are detected at various locations in the cell as shown in Fig. 2.

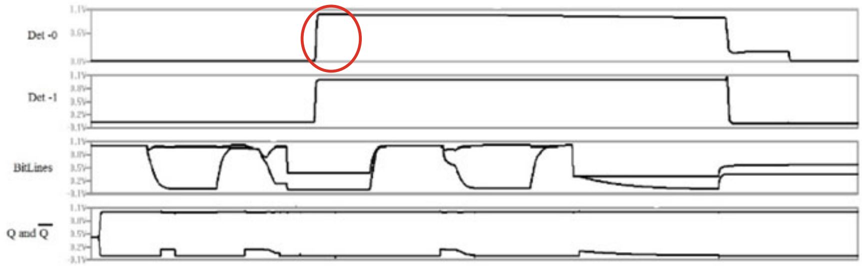


Fig. 6 Detection of resistive defects at pmos location in the cell

Table 1 Performance Parameters

Parameter	AgingTech [18]	Proposed method
Test time	0.25 ms	40.96 μs
Area overhead	–	7.8%

Similarly bridging defects are also detected by the proposed method. It is the defect of an unwanted connection between the two terminals. It may exist between the two cells as well (Table 1).

So with the proposed method resistive open defect at all possible get detected with less area overhead. The method also does not need complex timing circuitry or change in the supply voltages. With less time penalty resistive faults are getting detected with weak resistive defect coverage in terms of kilo ohms.

5 Conclusion

With the help of the proposed method, resistive open defect at all possible locations in the cell can be detected. Resistive defect of low values is also detected well by the proposed method. Analysis of the proposed method showed that resistive defects at various locations can be detected with less time complexity circuits and with less hardware requirements covering resistive values from tera ohms to few kilo ohms. The method detects the faults with less time penalty of 40.96 μs and less area overhead of 7.8% for 2 KB of SRAM memory.

References

1. Chen Q, Mahmoodi H, Bhunia S, Roy K (2005) Efficient testing of SRAM with optimized march sequences and a novel DFT technique for emerging failures due to process variations. IEEE Trans Very Large Scale Integr Syst 13(11):1286–1295. https://doi.org/10.1109/TVLSI.2005.859565
2. Ney A, Dilillo L, Girard P et al (2009) A new design-for-test technique for SRAM core-cell stability faults. In: Proceedings of design automation and test conference, Nice, France, 20–24 April 2009
3. Irobi S, Al-Ars Z, Hamdioui S (2010) Detecting memory faults in the presence of bit line coupling in SRAM devices. Proc Int Test Conf. https://doi.org/10.1109/TEST.2010.5699246
4. Mali M, Sutaone M, Tak S (2009) Gating transistor power saving technique for power optimized code book SRAM. In: Proceedings of international conference on advanced computing and communications control (ICAC3'09), pp 581–584. https://doi.org/10.1145/1523103.1523220
5. Mali M, Sutaone M, Tak S (2009) 5Gbits/s, 300 mV precharge, 256b, low power rhythmic SRAM. In: (ARTCom 2009) International Conference on Advances in Recent Technologies in Communication and Computing, pp 554–558. https://doi.org/10.1109/ARTCom.2009.223
6. Hamdioui S, Al-Ars Z, Van De Goor AJ, Rodgers M (2004) Linked faults in random access memories: Concept, fault models, test algorithms, and industrial results. IEEE Trans Comput Des Integr Circ Syst 23(5):737–757. https://doi.org/10.1109/TCAD.2004.826578

7. Yang J, Wang B, Wu Y, Ivanov A (2006) Fast detection of data retention faults and other SRAM cell open defects. IEEE Trans Comput Des Integr Circ Syst 25(1):167–180. https://doi.org/10.1109/TCAD.2005.852680
8. Li J, Tseng T, Hou C (2010) Reliability-enhancement and self-repair scheme for SRAMs With static and dynamic faults. IEEE Trans Very Large Scale Integr (VLSI) Syst 18(9):156–161
9. Linder M, Eder A, Schlichtmann U, Oberländer K (2014) An analysis of industrial SRAM test results - A comprehensive study on effectiveness and classification of march test algorithms. IEEE Des Test 31(3):42–53. https://doi.org/10.1109/MDAT.2013.2279752
10. Pavlov A, Sachdev M, De Gyvez JP (2006) Weak cell detection in deep-submicron SRAMs: A programmable detection technique. IEEE J Solid-State Circ 41(10):2334–2343. https://doi.org/10.1109/JSSC.2006.881554
11. Mali M, Barekar S (2020) Detection of resistive open defect fault in SRAM memory array structure for reliability against failures. Test Eng Manage 82(3–4)
12. Tak S, Mali M (2015) Fault tolerant SRAM by redundant array structure. In: 2015 International conference on pervasive computing (ICPC 2015). https://doi.org/10.1109/PERVASIVE.2015.7087128
13. Roberts D et al (2011) High repair-efficiency BISR scheme for RAMs by reusing bitmap for bit redundancy. IEEE Trans Very Large Scale Integr Syst 19(6):930–940. https://doi.org/10.1109/TVLSI.2013.2288637
14. Hsu CL, Ho MH, Lin CF (2009) Novel built-in current-sensor-based IDDQ testing scheme for CMOS integrated circuits. IEEE Trans Instrum Meas 58(7):2196–2208. https://doi.org/10.1109/TIM.2009.2013668
15. Grossar E, Stucchi M, Maex K, Dehaene W (2006) Read stability and write-ability analysis of SRAM cells for nanometer technologies. IEEE J Solid-State Circ 41(11):2577–2588. https://doi.org/10.1109/JSSC.2006.883344
16. Gyepes G, Stopjaková V, Arbet D, Majer L, Brenkuš J (2014) A new IDDT test approach and its efficiency in covering resistive opens in SRAM arrays. Microprocess Microsyst 38(5):359–367. https://doi.org/10.1016/j.micpro.2014.04.006
17. Mali M, Barekar S (2020) Bridging defect detection for fault diagnosis of on-chip cache memory. Test Eng Manag 83
18. Martins MT, Medeiros GC, Copetti T, Vargas FL, Bolzani Poehls LM (2017) Analysing NBTI impact on SRAMs with resistive defects. J Electron Test Theory Appl 33(5):637–655. https://doi.org/10.1007/s10836-017-5685-6

An Automated Safe Hybrid Energy Based On-Board Charging Station Using an Intelligent Controller

A. Jagadeesh, K. Sudarsana Reddy, and R. Mahalakshmi

Abstract Greener transportation is the motivation to shift from conventional vehicles to electrical vehicles (EV). EV's being the present technology, charging of the EV's battery with best charger in secured environment is the challenge. The EV's fast charging DC has advantages in reducing charge time, but maintaining the optimal temperature for the battery and availability of charging stations have benefited on-board charging (through AC supply). The constant current constant voltage (CCCV) battery charger has reduced the peaking currents during initial stages and promotes charging of cells with different capacities at the final stage. The environment to charge must be safe, so as the battery never damages. In onboard charging, due to fault currents inducing in the supply might damage the battery, so an intelligent controller is required to detect theses faults and can open the circuit breaker associated with it. The idea of this work is to charge the EV's battery with CCCV battery charger through on-board charging system where the supply is derived from grid and solar. The supply system has intelligent controllers operating to detect the faults and to work respectively. The backup generator has been provided during faults in grid. The intelligent controller used to generate the control signals for circuit breakers is Arduino Uno. The solar has been converted to AC by using a three-level neutral point clamped based State Vector Pulse Width Modulated (SVPWM) inverter. To generate the control signals from Arduino Uno, the Proteus Professional software has been utilized and to simulate the present work, MATLAB/Simulink software has been used respectively.

Keywords Electric vehicle (EV) · Solar array · On-board charging · Constant current constant voltage (CCCV) technique · State vector pulse width modulation (SVPWM) · Circuit breakers · Arduino Uno · Current sensor

A. Jagadeesh (✉) · K. Sudarsana Reddy · R. Mahalakshmi
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India

R. Mahalakshmi
e-mail: d_mahalakshmi@blr.amrita.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_76

1 Introduction

The drastic decline in the fossil fuels [1], brings out the term E for Electric in place of Engine. Electric vehicles have become a vital coinage in Auto mobile industry. The heart of this vehicles [2] are motor and the power supply to the motor (battery). Various charging paradigms have paved their way for fast and efficient charging. One among them is off-board charging (DC charging), which has a capability of charging 80% of Battery within 19 min with greater limitations. The less availability of DC stations and temperature management in the battery pave the way to shift again to on-board [3] charging stations (AC charging). There is a great demand for efficient way of charging and fault isolation. At present power grid system are solely following the conventional methods, where the integration [4] studies are highly required to adopt new generation units and consumers. In this context, the studies for integration of a solar to existing power grid is realized with the dynamic loading sequences at charging station. The charging station which is equipped with roof mounted photovoltaic panels is integrated with peak time solar power generation data [5]. The dynamic response of the grid is under the surveillance of micro controller to ensure the fault levels generated at the transmission lines and to simulate a response accordingly. This discourse, discusses the way to integrate power from the grid and the PV array [6] to charge a battery and also to isolate the system during the fault generation [7]. Battery controller [8] (CCCV) maintains the current level during the initial raise in voltage and maintains the voltage level when it reaches its desired value.

In Simulink, the voltage and current values of power grid for on board charging are observed. Then three phase faults like double line to ground fault are imposed at a specified section with a particular time period to identify the changes in the system parameters. Immediately the intelligent control system incorporates the circuit breaker [9] action preventing the system from any failure. The novelty in parallel scaling down and sensing out the fault current to isolate the system with the help of circuit breaker is achieved using the current sensor and Arduino micro controller [10] simulated in Proteus professional.

The paper has been organized into four sections. Section 2, describes the system and specifications of the proposed system. Section 3, discuss about methodology of proposed system, here the MATLAB simulated system and its results along with system simulated in Proteus have been discussed. Through Sect. 4 paper gets concluded.

2 System and Specifications

The proposed system's block diagram has been shown in Fig. 1. It can be seen that the power from grid has been integrated with solar power and together charging the battery of an EV with CCCV on-board charger. The integrated solar power has

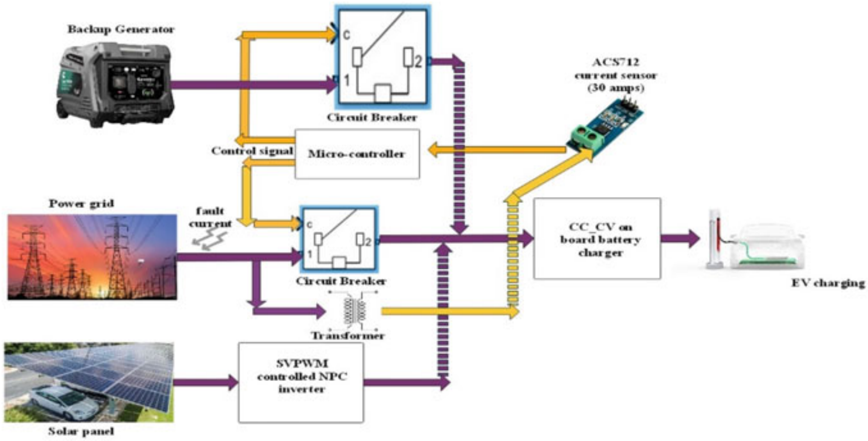


Fig. 1 Block diagram of the proposed system

been converted from its natural DC into AC with the help of a neutral point clamped three-level inverter respectively. This inverter [11] has been controlled through the PWM pulses generated from SVPWM. It can be observed that the input current to the charger is continuously monitored by the current sensor. The current sensor has been integrated to a microcontroller (Arduino Uno), which generate control signals to monitor the safe supply. The backup generator has been supplied for backup.

The VI-characteristics and maximum power point [12] (MPP) curve of the solar array used has been shown in Fig. 2. It can be observed that the plot has been plotted for irradiance of 1000 wm^{-2} and working temperature of 25 °C respectively. The maximum power (25 kW) can be tracked when solar operates at 415 V. The short-circuit current and open-circuit voltage are 72 A and 415 V respectively. Table 1 shows the parameter values of the solar array used in the simulation. The Fig. 3 shows the battery discharging characteristics. From the plot it can be inferred that the exponential voltage being 456 V and the pink area shaded in the plot corresponds to the exponential voltage drop. The discharge curve has been drawn at 0.4347 °C rating. It can be observed that the nominal voltage is 400 V and the battery drains completely at 2.3 h. with discharging current 43.47 A (0.4347*I_{rated}). The discharging curves have been even plotted for 0.065 °C, 0.13 °C and 0.325 °C ratings. It can be seen that battery drains completely at 15.8, 7.2, 3.1 h. respectively. The battery parameter values have been depicted in Table 1.

The battery charger input parameter curves along with the harmonics in input current have been shown in Fig. 4. The charger efficiency, power factor and THD of battery charger has been plotted with respect to the usage factor for the polynomial curve. The specified along with simulated curves have been shown. The specified being the darker. At 0.7 usage factor, the battery charger efficiency is observed as 82% and input power factor as 0.79 and the THD being 0.24 respectively. It can be seen that the dominating harmonic is 5th harmonic with frequency is 250 Hz

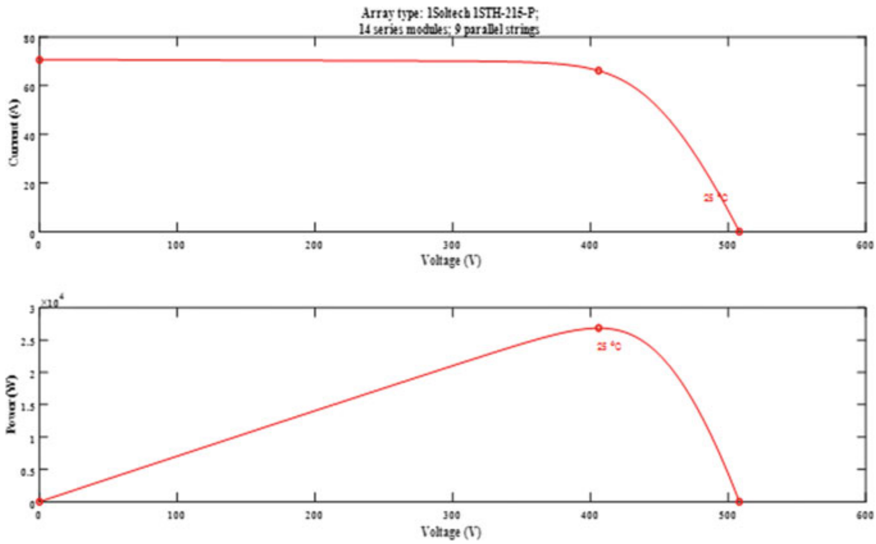


Fig. 2 V-I characteristics and MPP curve of solar array simulated

Table 1 Specifications and parameters of solar array and battery

Solar array		Battery	
Parameter	Value	Parameter	Value
Short-circuit current I_{sc} (A)	7.84	Maximum/rated capacity (Ah)	100
Open circuit voltage V_{oc}(V)	36.3	Nominal voltage (V)	400
Irradiance (Wm^{-2})	1000	Initial State -of –charge(%)	2
Maximum power (W)	213.15	cutoff voltage(V)	300
Series and parallel connected module per string	14.9	Fully charged voltage	465.5949
Current and Voltage at maximum power point I_{mp} (A), V_{mp} (V)	7.35,29	Normal discharge current and internal resistance (ohms)	43.4783 0.04

(5*50) with amplitude 0.9. The parameter values of the battery charger circuit used in simulation has been shown in Table 2.

The nominal voltage of current sensor ACS712ELC-30A is 5 V and it measures the between 30 and −30 A with the scaling factor of 66 mV per Amp.

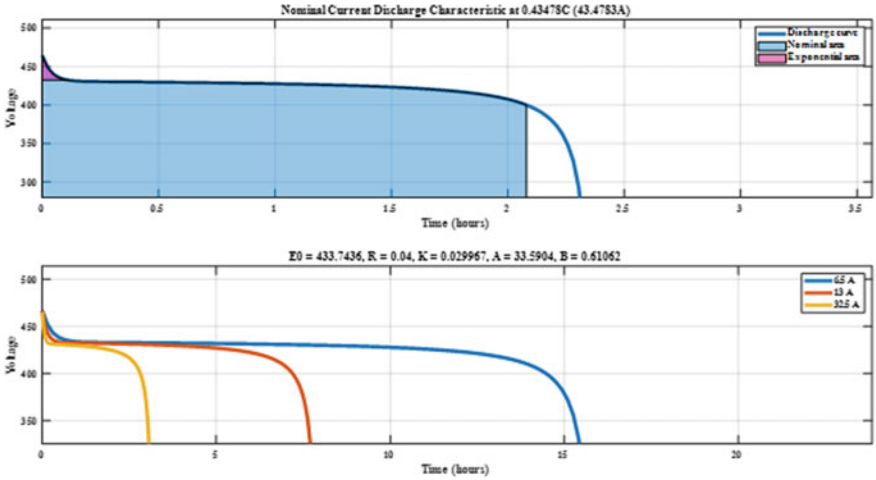


Fig. 3 Battery discharging characteristics

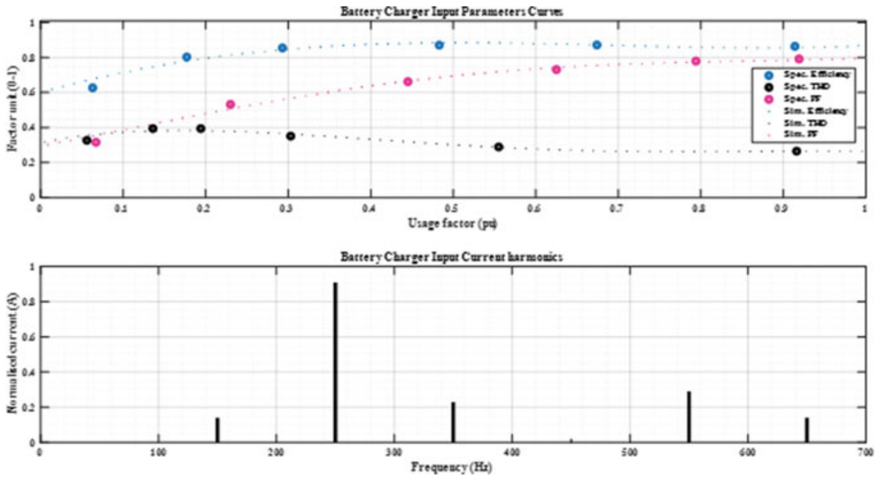


Fig. 4 Battery charger parameter curves and Input current harmonics

3 Methodology

3.1 Simulation and Results

The simulation model of the proposed system has been shown in Fig. 5. For simulating the specifications of Hyundai Kona vehicle has been selected. The lithium ion battery of the vehicle is 40 kW with terminal voltage of 320–400 V respectively. For powering

Table 2 Specifications and parameters of Battery charger

Parameter	Value
Nominal power (W)	40,000
Effective voltage (V)	400
Frequency (Hz)	50
Bulk current (A)	100
Float voltage (V)	400
Fully charged voltage (V)	465.5949
Overshoot (%)	1
Settling time (s)	5

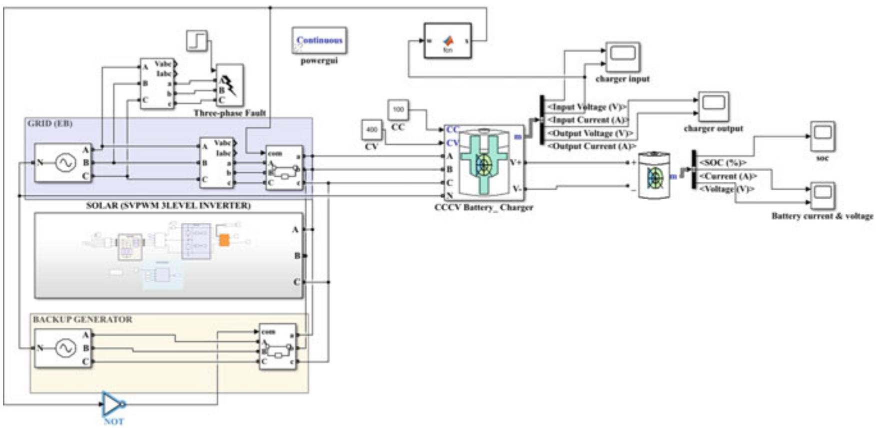


Fig. 5 Overall simulation figure of the proposed system

the battery of 40 kW, a ($\sqrt{3}$ * 40) 69.28 kW of AC power has been chosen. Among which the solar is providing a 25 kW of power and remaining being fed from grid respectively. The line rms voltage at which the system operating is 400 V respectively. The DC power of solar has been converted to AC for integrating it with grid power to provide on-board charging which has been shown in Fig. 6. The conversion has been done by utilizing a neutral point clamped (NPC) based three-level inverter. To achieve the desired output, the state vector pulse width modulation technique has been used for generating the gate pulses of the inverter. The state vector method converts the three-dimensional vectors into two-dimensional space. The inputs to SVPWM generator are the reference AC voltage waveforms, the positive and negative clamped DC link voltages along with the input dc current. The states for the respective block are generated by utilizing the XOR logic. The state variables being A, B, C and states being PNN, PPN, NPN, NPP, NNP, PNP respectively. This block generates the twelve pulses which are fed to the inverter.

The LC filter has been used after the output of the inverter to filter the harmonics. This integrated power is transferred to battery through CCCV battery charger which

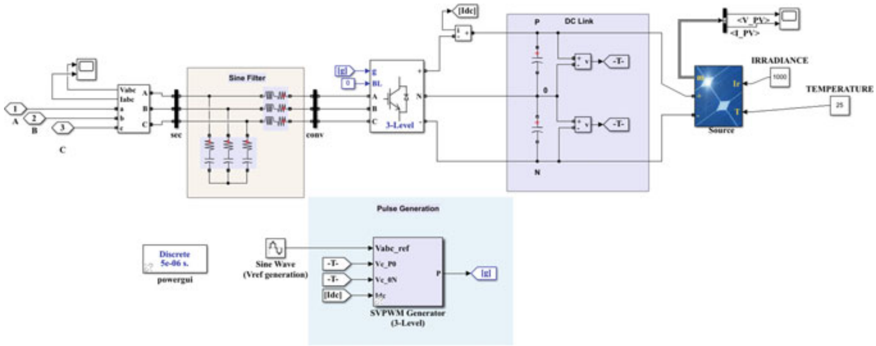


Fig. 6 Solar (SVPWM 3level Inverter) subsystem simulation figure

have inputs as constant voltage, constant current and three phases of AC input supply. The constant current and voltage are specified as 100 A and 400 V. This on-board charger converts the AC into DC and charges battery at constant current till the terminal voltage reached its specified value and once reached then battery charges at constant voltage respectively. The circuit breakers (CB) have been connected to the grid before integration to safeguard the equipment. These circuit breakers command signal will be generated according to the faults induced in the system. During fault conditions, the MATLAB function block generates low signal to open CB of grid and to close CB of the backup generators to supply the battery without any interruption. Due to limitations of MATLAB the logic utilized to generate control signals from Arduino Uno has been imported to the MATLAB function block and generated the command signals required for the CB's of the system. The generation of control signals during faults through Arduino Uno have been shown in the Proteus Professional software as the continuation of this model.

The simulation figure shown in Fig. 5 has been simulated and its respective results have been depicted from Fig. 7, 8, 9, 10, 11, 1, 13 and 14 respectively.

From Fig. 7 it can be inferred that the overall DC link voltage or solar voltage is maintained at 460 V and the solar current is observed to maintain at 42 A respectively. The Fig. 8 shows the gate pulses generated by SVPWM. It can be seen that the twelve pulses are multiplexed and fed to the NPC based three-level inverter. Figure 9 shows the output voltage and current waveforms of the inverter. It can be seen that 320 V peak voltage (i.e., 400 Vrms line) has been generated and current of 51 A peak (36Arms) respectively. It can be seen that the required ($\sqrt{3}$ * 400 * 36) 25 kW power has been generated from solar subsystem.

Figure 10 shows the voltage and current waveforms of input of the battery charger. It can be seen from the respective figure that the integrated voltage can be observed to be maintained at 320 V peak and current maintained at 140 A peak respectively. The fault has been introduced into the system at t = 5 s and the fault currents into the system have been introduced in phase A, B respectively with respect to ground and phase C is free of fault current. The fault currents in fault generator before and after

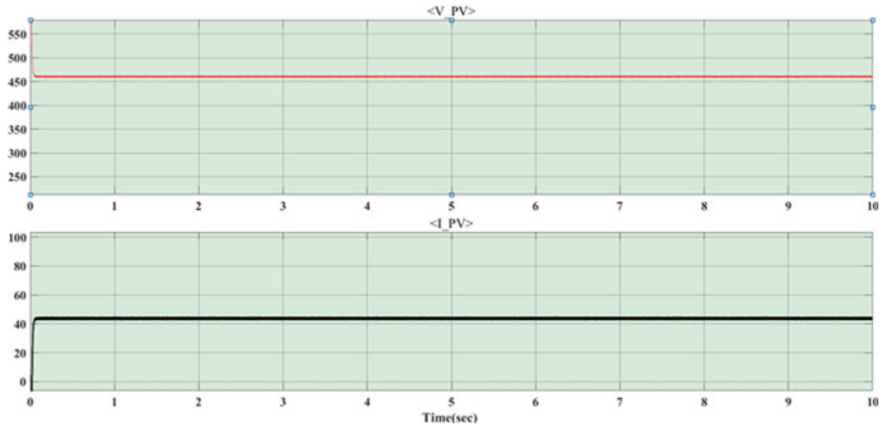


Fig. 7 Solar array voltage in volts and solar array current in amps

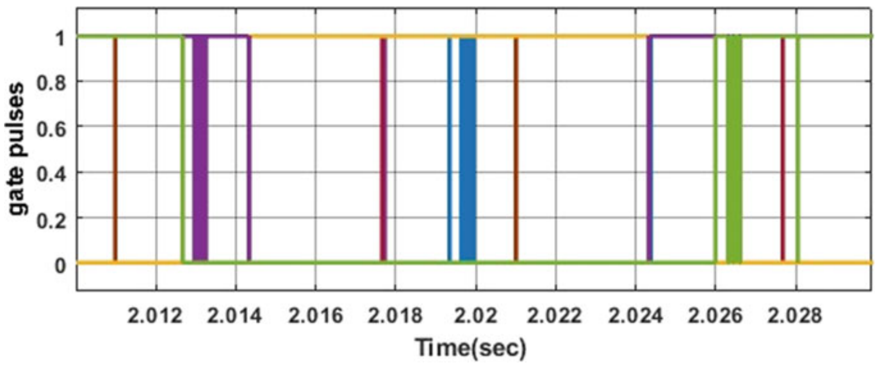


Fig. 8 Gate pulses generated from SVPWM

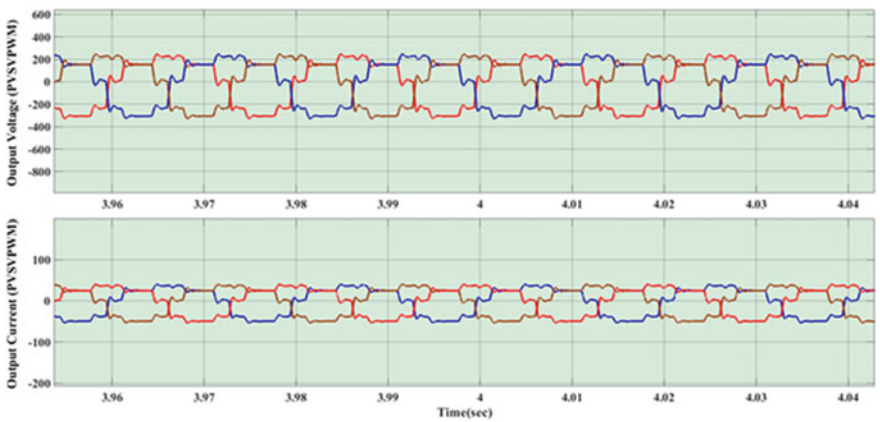


Fig. 9 Inverter output voltage in volts and inverter output current in amps

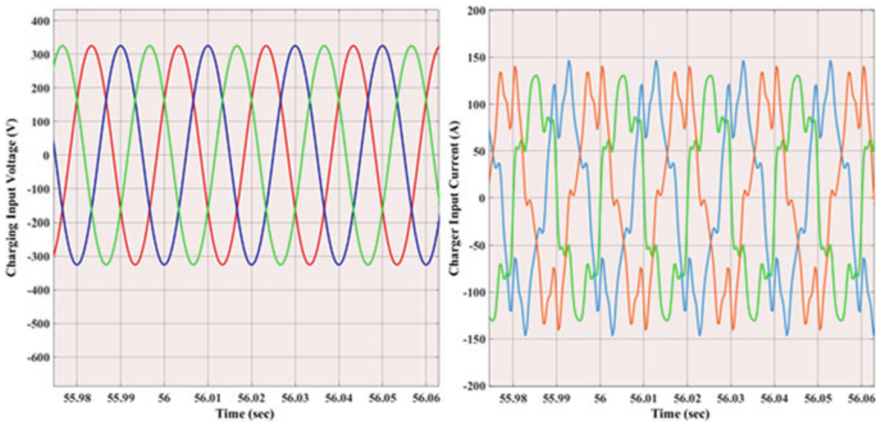


Fig. 10 Battery charger input voltage (volts) and current (amps)

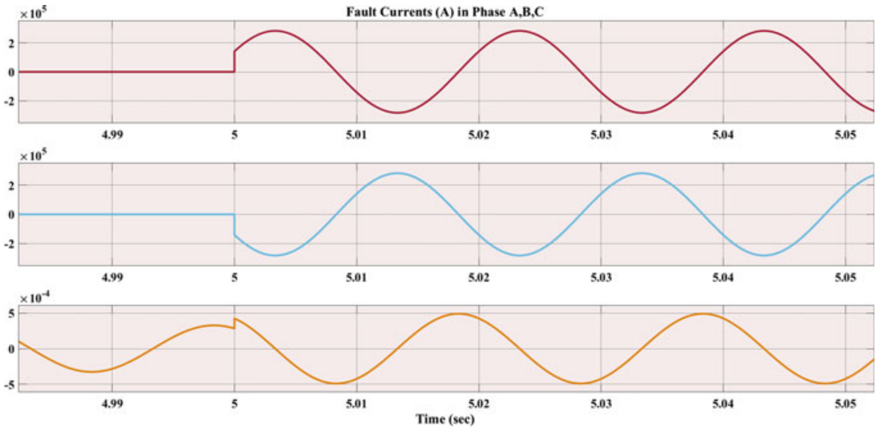


Fig. 11 Phase currents (amps) of Fault generator during the transition

transition have been shown in Fig. 11. It can infer that the currents in phase A and B after transition have been peaked to 30 kA respectively and phase C current in fault generator being zero before and after transition. The corresponding response of the output has been shown in Fig. 12. Figure 10 is the zoomed version of Fig. 12 after transition. It can be seen that the during fault currents at 5 s the CB of the main system have been opened and after 0.4 s the backup along with solar have been charging the battery. It can be inferred that the fault currents haven't been introduced into the battery.

Figure 13 shows the voltage and current waveforms of the battery. It can be inferred that battery is getting charged at constant current of 100 A till 70 s. After 70 s the battery reaches its nominal voltage 400 V, then mode of charging has been shifted to constant voltage. The Fig. 13 shows the state of charging of battery. It can be seen

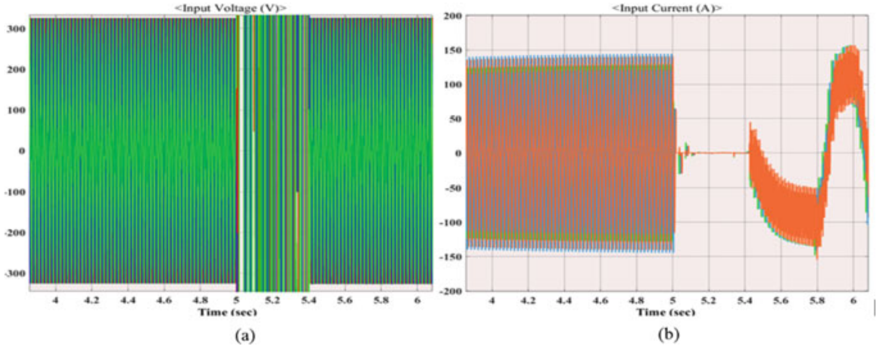


Fig. 12 Battery charger input voltage (volts) and current (amps) during the transition

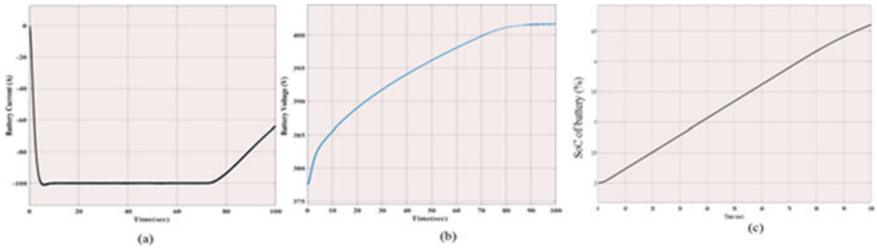


Fig. 13 Battery voltage (volts), current (amps) and SOC %

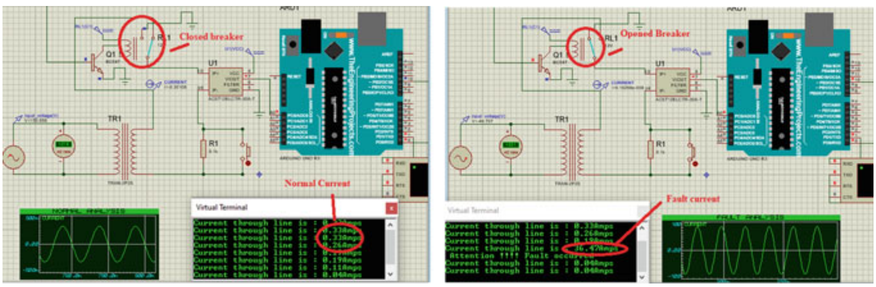


Fig. 14 Analysis of the system with closed breaker and open breaker

that the initial State of Charge (SOC) of 2% has been increased to 4.6% in 100 s of simulation.

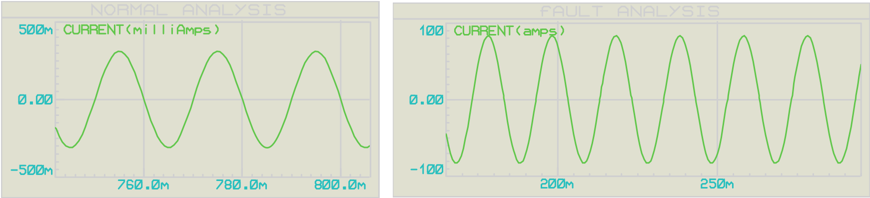


Fig. 15 Normal (CB closed) and fault current (CB open)

3.2 Proteus Simulation

MATLAB function used in Simulink is merely a micro controller (Arduino) used to sense the fault current is simulated in the Proteus Professional as shown in Fig. 14 due to the constraints in interfacing Arduino to the MATLAB. Figure 14 shows the system simulation with closed and opened circuit breaker. The scaled down current from transformer is sensed using the current sensor (ACS712 _30A), that set out the output voltage in the range 0–5 V. Using the analog pins, the voltage is read and is converted into current by using the resolution specifications of built in ADC used in micro-controller. Scaling factor of the sensor can be referred from the data sheet of the current sensor. Arrangement is made using a push button so as to mimic the generation of Line to ground fault. In Fig. 14 resistive load is taken and the current through the load is calculated as per the Eq. 1 and is read continuously ensuring that it is in a safe limit (in range of milliamps) and maintaining the breaker in closed condition. Current through the load is

$$IL = \frac{V}{R} = \frac{22}{0.1k} = 0.22\ \mathrm{A(rms)} \tag{1}$$

The push button is closed to induce fault, by providing short circuit to the load resulting in the high currents (70–80 A). Normal current under healthy condition and fault current are shown in Fig. 15. These currents are sensed down immediately by the Arduino and sends out the control signal to the transistor that connects the excitation coil of the breaker to the ground terminal. Coil gets energized and moves the contactor in the breaker to the ground terminal isolating the load from the high Fault currents. Virtual terminal (Fig. 15) displays the high current being sensed by the controller.

4 Conclusions

The sole aim of the proposed system is to have an efficient, safer and consistent way of on board charging with the power sharing from both grid and the solar is illustrated using MATLAB Simulink and Proteus Professional. The Solar array of maximum

25 kW has been selected to meet the demands of load along with the losses. The corresponding open-circuit and short-circuit test values have been considered to plot V-I characteristics and MPP curve respectively. DC power of solar is converted into Ac using NPC 3 level inverter with SVPWM technique to maintain the voltage and current at 400 V, 42 A. Using CC_CV battery controller, battery is allowed to charge efficiently with constant current (100 A) until it reaches the rated voltage (400 V) and thereafter with constant voltage. The double line to ground fault at the grid transmission lines are been detected and isolated the fault by activating circuit breaker using the MATLAB function which mimics the functionality of micro controller. It is also seen that when fault occurs, the voltage and current values deviates from their normalized values. Opening of Circuit breaker during fault instantly protected the on board battery blast out. Small prototype of fault analysis is simulated in Proteus. The command signal to energize the coil in relay has been generated from the Arduino Uno during the detection of faults. Current sensor plays a vital role in sensing the fault currents. The consistency can be obtained by providing backup generator during isolation. It can be inferred that proposed system has the best results and the whole control system is energy efficient as the input power has been parallely shared by solar and grid.

References

1. Dotzauer E (2010) Greenhouse gas emissions from power generation and consumption in a nordic perspective. Energy policy 38(2):701–704
2. Sreelakshmi S, Mohan Krishna S, Deepa K, Bidirectional converter using fuzzy for battery charging of electric vehicle. In: IEEE transportation electrification conference (ITEC-India), Bengaluru, India, pp 1–6
3. Kamble AS, Swami PS (2018) On-board integreted charger for electric vehicle based on split three phase insuction motor. In: 2018 international conference on emerging trends and innovations in engineering and technological research (ICETIETR), Ernakulam, pp 1–5
4. Varma RK, Salama M, Seethapathy R, Champion C (2009) Large-scale photovoltaic solar power integration in transmission and distribution networks. In: 2009 IEEE Power & Energy Society General Meeting, Calgary, AB, pp 1–4
5. Gautam M, Raviteja S, Mahalakshmi R (2019) Household energy management model to maximize solar power utilization using machine learning. Procedia Computer Sci 165:90–96
6. Mahalakshmi R et al (2015) Implementation of Grid Connected PV array using Quadratic DC-DC converter & 1ϕ multi level inverter. Indian J Sci Tech 8:1–5
7. Bagchi S, Goswami S, Ghosh B, Dutta M, Bhaduri R, "Symmetrical and Asymmetrical Fault Analysis of Transmission Line with Circuit Breaker Operation. In: 2019 1st international conference on advanced technologies in intelligent control, environment, computing & communication engineering (ICATIECE), Bangalore, India, pp 343–347
8. Sudarsana Reddy K, Mahalakshmi R, Deepa K (2020) Bio-diesel fed solar excited synchronous generator. J Green Eng 10:1–27s
9. H. Wang, B. Huang, X. Leng and P. Yang, "Research on Fault Isolation of Feeder Short Circuit in Parallel Battery System," 2018 2nd IEEE Advanced Information Management,Communicates,Electronic and Automation Control Conference (IMCEC), Xi'an, 2018, pp. 1100-1103, doi: https://doi.org/10.1109/IMCEC.2018.8469702

10. R. Majumder, S. Dolui, D. Agasti and S. Biswas, "Micro-controller based over current relay using Hall Effect current sensor," 2018 Emerging Trends in Electronic Devices and Computational Techniques (EDCT), Kolkata, 2018, pp. 1-4
11. R. Mahalakshmi, K. Deepa and K. C. Sindhu Thampatty, "Analysis of Multi level Current Source Inverter for low Torque Applications," Journal of Green Engineering, Vol.8, no.1, Dec 2018.
12. R. Mahalakshmi, Ashwin Kumar A. and A. Kumar. (2014) "Design of Fuzzy Logic based Maximum Power Point Tracking controller for solar array for cloudy weather conditions," *2014 Power and Energy Systems: Towards Sustainable Energy*, Bangalore, pp. 1–4.

Artificial Intelligence Based Learning Approach for Leaf Disease Identification and Detection

G. Karuna, K. Sahithi, B. Rupa, R. Amani, K. Swaraja, and K. Meenakshi

Abstract Plants and Crops get diseased due to many reasons. It might be because of diseases of stems, leaves, roots etc. This Paper mainly congregates on leaves. Leaf Disease identification and Detection has many applications for cultivators and farmers to know whether the plant is diseased or not. So that they can retort in dwarf time and decreasing the loss and then can obtain immense profits. This paper mainly focused at learning the disease of plant through leaves. Here, we scrutinize the leaf through Image Processing and extract features of particular leaf and then utilizing those features as a dataset and done preprocessing and then administering them in Artificial Intelligence based learning algorithms like Convolutional Neural Networks to find disease.

Keywords Leaf disease · Neural networks · Features

1 Introduction

In the Earth, India is the second powerful Country in terms of population and also Food is one of the essential one to a person in order to survive. We are acquiring food from plants and so we have to safeguard our plants from diseases. Diseases of plants are normally due to insects, pests, pathogens and reduce the fertility to an extreme extent if not managed with in time.

Now-A-Days technology plays important role in most of the fields yet till this day we are applying some past approaches in farming. Knowing the plant's condition plays a vital role for productive farming [1]. Detection is happened normally by the scientists but due to the so many changes in the environment, the detection is becoming tough. So we can practice processing of image techniques for recognition of disease of a plant. Normally we can get to know the symptoms of diseases on stems, leafs, flowers etc. so here we mainly utilize leaves for recognition of diseased plants.

G. Karuna (✉) · K. Sahithi · B. Rupa · R. Amani
Computer Science and Engineering, GRIET, Hyderabad, India

K. Swaraja · K. Meenakshi
Electronics and Communications Engineering, GRIET, Hyderabad, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_77

Leaf Disease Identification is affiliate to domain Data Science. Here, Convolutional Neural Networks is taken in the Data science [2].

If we check a picture of a man, there can be a lot many features varying two persons. We could not give each and every feature as input to some sort of ML Algorithm. Even though, we give most of the features there would be some of other features which we may be missing. So, Here Deep Learning arrives into existence [3]. If we consider an image of 20*20 pixel, No of inputs to the Normal Neural Network is 20 * 20 = 400. 400 is a minute value and it is satisfactory. If we examine a picture of 300 * 300 pixel. No of inputs is 300 * 300 = 90,000 which is large number and it is very complex kind of thing to a Normal Neural Network and so that's the reason we are administering Convolutional NNs.

The learning ability of the human brain, which consists of neurons linked by synapses, motivates artificial neural networks. For images, ANNs are not appropriate since these networks lead to over-fitting due to image size. The major discrimination between the Artificial Neural Networks (ANN) and Convolution Neural Networks (CNN) is that only the last layer of a CNN is completely connected, while in ANN, each neuron is connected to every other neuron. Convolution neural networks directly use images as an input. Convolutional neural networks are used instead of handcrafted features to automatically learn a hierarchy of features that can then be used for classification purposes. This is achieved by successively converging the input image to create a hierarchy of feature maps with learned philtres. CNNs are different layered oversaw networks which can know features directly from datasets. For the recent days, CNNs have obtained superb production in almost all vital determination tasks. It can do one classification and the other feature extraction below the same network architecture [4].

2 Related Work

There are many strategies of management like disease particular chemical applications, vector control through pesticide applications and fungicide applications which gives early data on health of plant and disease detection. This keeps us diseases in control and improve productivity.

For plant disease identification, Number of methods are presently in usage applying computer vision. One among them is detection of disease by taking color features [5] or by extracting texture features. Features of textures such as Homogenity, Inertia, Correlation obtained from co–occurence on image by calculating gray level [6]. Many of the works like using image processing disease recognition approach, microscopy, double-stranded RNA(ribonucleic acid) analysis and nucleic acid probes [7–10] and also others like Particle Swarrn Optimisation (PSO) [11, 12]. We can also apply Support Vector Machines [13] and also K-Means Algorithm as a Clustering Method Algorithm Proposed by authors of [14].

3 Convolution Neural Networks

In the most of neural networks, Convolutional Neural Network (ConvNets or CNNs) is that the one among the most categories to strive to images recognition, image classification. A picture as an input as set of pixels is watched by Computers and it relays on the picture resolution. On the basis of picture resolution, it'll see h * w * d (h = height w = width d = dimension). Figure 1 refers to RGB values.

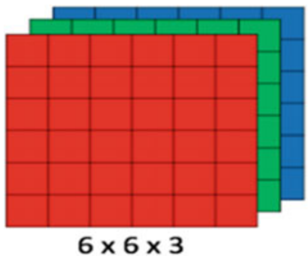
Technically Deep Learning Convolution Neural Network models check each and every input picture will permit it through various steps of layers of convolution with Kernels (filters) such as pooling, fully connected layers (FC) and utilize Softmax or any other methods to determine an object with probabilistic values between 0 and 1. Figure 2 is an overall flow of CNN to undertake a picture as an input and determines the objects supported prices.

Convolution Layer

Convolution is that the top most layer to take features from an image as an input. Convolution retains the interrelation among pixels by knowing picture features utilizing tiny squares of input file. It is a math kind of process that considers two inputs like image matrix and a filter or kernel.

Picture Convolution with different kernels can undergo operations like blur, edge detection and sharpen by using kernels. The following example gives different convolution pictures after using various kinds of kernels.

Fig. 1 RGB Layers



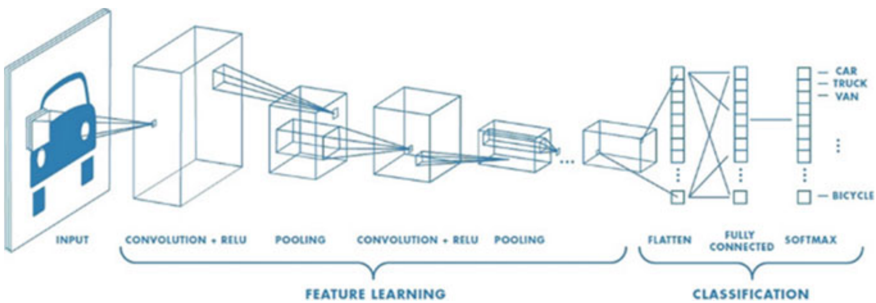


Fig. 2 Neural network with many convolutional layers

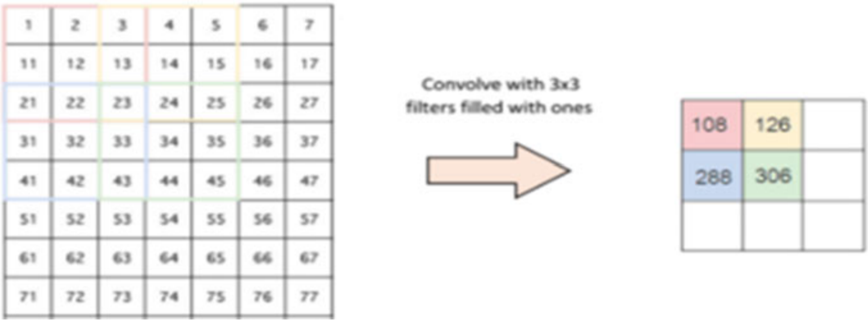


Fig. 3 Stride of two

4 Strides

Stride is that the no. of pixels moves through the input matrix. When the stride is 1 then we proceed the kernels to 1 pixel at a time. When the stride is 2 then we proceed the kernels to 2 pixels at a time then on. Figure 3 displays convolution can be done with a stride of two.

Padding
Kernel doesn't exactly fit the input picture sometimes. We have two options:

- Add the image with zeros (zero-padding) in order that it fits
- Drop the part of the picture where the kernel didn't fit. This is known as valid padding which provides only valid a fraction of the picture.

Non Linearity (ReLU)
ReLU full form is Rectified linear measure for a non-linear operation. The end result is $f(x) = \max(0,x)$. ReLU's purpose is to provide non-linearity in our ConvNet. Because, the important global info would want our ConvNet to find out would be positive linear values. Some kind of other non linear functions like sigmoid or tanh which will even be applied rather than ReLU. Most of the info researchers use ReLU because performance wise ReLU is best than the opposite two.

5 Pooling Layer

Pooling layers section would scale back the amount of parameters when the pictures are overlarge. Spatial pooling also known as to be downsampling or subsampling which trims the dimensionality of each and every map but retains essential information. Spatial pooling could be of variety of kinds:

- Max Pooling

- Sum Pooling
- Average Pooling.

Max pooling takes the most important item from the rectified feature map. Considering the most important item could also provide the typical pooling. Addition of all items within the feature map known as sum pooling (Fig. 4).

The layer which we say as FC layer, we flattened our matrix into vector and give it into a totally layer of connected sort of a neural network (Fig. 5).

In the above figure, matrix of the feature map are going to be converted into vector (× 1, × 2, × 3). With the fully connected layers, we combined all of these features together to produce a model. Ultimately, we have an activation method as sigmoid or softmax to determine the outputs as cat, dog, car, truck etc. (Fig. 6).

Related Background Work

There are many network architectures proposed like LeNet, AlexNet, GoogleLeNet etc. used for image recognition. The LeNet network architecture is the initial Convolutional Neural Network introduced by LeCun et al. to identify digits written by

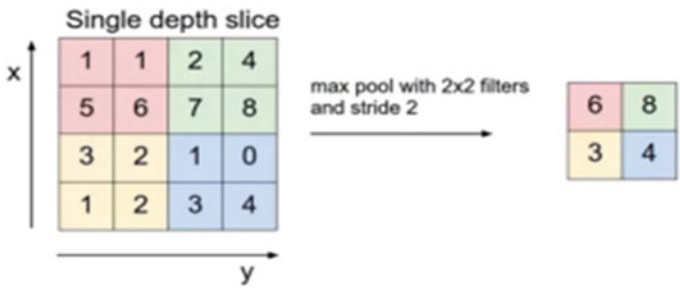


Fig. 4 Max Pooling with stride of two

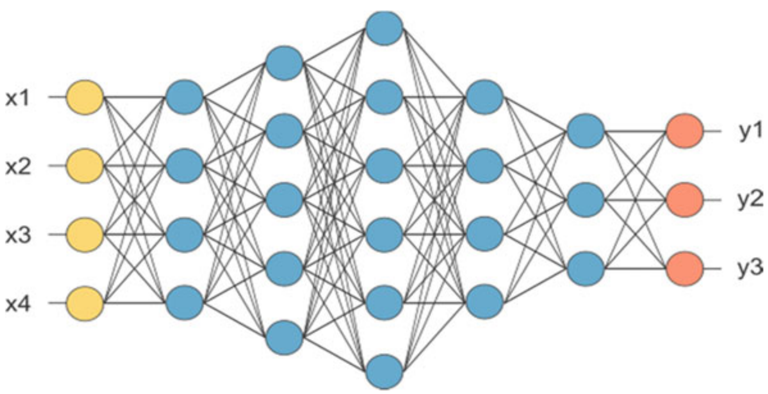


Fig. 5 Fully Connected Layers (FC)

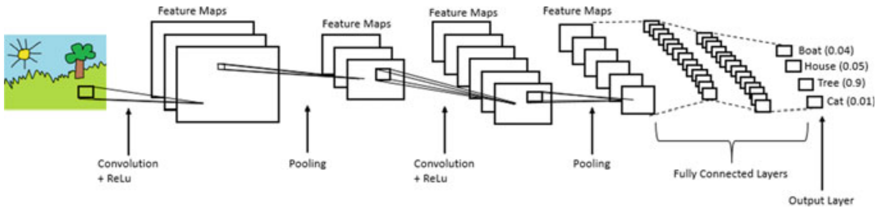


Fig. 6 CNN with multiple layers

hand [15]. It mainly contains two convolutional layers followed by two subsampling layers and a fully connected Multi-Layer Perceptron.

Few experimenters presented the use of CNN for recognition of leaf and disease classification of plant. Atabay [4] created a CNN architecture to detect plants based on images of a leaf. The one which is proposed importantly contains of five layers. After each convolutional layer a Rectified Linear Unit (ReLU) or Exponential Linear Unit (ELU) activation function is used and for each pooling layer, MaxPooling procedure is applied. The proposed one can be done on Flavia (Wu et al. 2007) and Swedish [16] datasets related to leaf consisting 32 species of a plants ¨ with 1907 samples and 15 kinds with 1125 samples respectively. The images of a single leaf in the dataset are pictures taken at uniform background. All the input images are 160 × 160 pixel grayscale images. The model achieved a 97.24% and 99.11% of classification accuracy for each dataset. The outcomes given that the architecture which is proposed for CNN-based leaf detection is almost competing with the new substantial approaches on devising leaf classifiers and features.

Reyes et al. and Camargo [17], suggested deep learning technique of hand-engineering. The designed system contains of 2 fully connected layers which follows 2 convolution layers. The CNN is trained using 1.8 million images from ILSVRC 2012 dataset and used a fine tuning policy to transfer known detection capabilities from normal areas to the particular challenge of Plant Recognition task. The dataset is combination of part of a plant or images of a plant taken both of them under a natural environment as well as in the controlled environment. They procured with an average accuracy of 0.486.

Mohanty et al. [18], applied the existing deep CNN architectures, i.e. AlexNet [19] and GoogLeNet [20] to systemize diseases of a plant. The public data set with 54,306 pictures of diseased and healthy plant leaves considered under controlled environments, the CNN was coached to detect 14 crop types and 26 diseases. The models grant accuracy of 99.35%. When tested on an images set taken at a different kinds of environment than the images used for the coaching, however, the accuracy of model dropped to 31.4%. Overall the results indicates the viability of deep CNN for plant disease classification.

The Proposed CNN Architecture Model

CNN architecture differs with the kind of the difficulty it has. The model which is proposed is a sequential model which contains of four layers of convolution,

Fig. 7 CNN with multiple layers with filter sizes

Layer type	Filter size	Output size
Conv2d	3*3	256 *256*32
Conv2d	3*3	254*254*32
Pool	8*8	31*31*32
Conv2d	3*3	31*31*32
Conv2d	3*3	29*29*32
Pool	8*8	3*3*32

maxpooling follows two of those layers and flatten function which is used to convert into a single column passed to fully connected layer from pooled feature map. Layer of Fully connected is appended by the dense layers of nearly two to the neural network.

The first convolutional layer filtrates the input picture with 32 kernels of size 3 × 3. After maxpooling is used, the end result is provided as an input for the second convolutional layer with 64 filters of length 4 × 4. Fully connected layer of 512 neurons follows the last layer of convolution has 128 filterss of length 1 × 1. The end result of this layer is provided to softmax function that makes a probability function of the four end result classes (Fig. 7).

6 Datasets

Here we considered grape disease dataset as the input. Grape disease dataset contains nearly 3000 training images and nearly 880 testing images. These images are colored. Images if they are taken from uncontrolled background and different lighting condition, then Background of training images may bias the neural network. So, these sort of pictures must be preprocessed and here the sort of dataset is preprocessed. So, There is uniform background to these images that is, only leaf figure is appeared (Tables 1 and 2). Classification of dataset is shown in Table 3.

Table 1 Training dataset

Disease name	No. of training images
Black-rot	966
Esca (Black Measles)	1154
Healthy	213
Leaf-blight	876

Table 2 Test dataset

Disease name	No. of testing images
Black-rot	210
Esca (black measles)	240
Healthy	220
Leaf-blight	210

Table 3 Disease description in grape leaves

Leaf disease name	Leaf image	Description	No. of training images	No. of testing images
Black Rot		Mainly due to fungus Guignardia bidwelli Leaves are reddish brown Circular to angular spots on upper surface of leaves	966	210
Esca (Black Measles)		Mainly due to fungi Phaeoacremonium aleophilum, Phaeomoniella chlamydospora, Fomitipora mediterranea It's like grapevine trunk disease	1154	240
Healthy		No disease. Healthy leaf	213	220
Leaf_blight (Isariopsis_Leaf_Spot)		It is a Bacterial Leaf spot	876	210

7 Proposed Work

The proposed work comprises of 6 steps given below to identify disease of grapes leaf.

- Resizing data and Split the dataset into inputs and targets
- Building the model
- Compiling the model
- Training the model
- Making predictions on new sample data.

Resizing data and splitting data inputs and targets

Resizing data means converting the image size into which we need whether it may be 256 * 256 or 128 * 128 as shown in Fig. 8.

Building Model
Sequential is pretty easy way in Keras to make a model. It drives you to construct layer by layer a model. Each and every layer has weights that correlate to the layer the follows it. 'add()' method is applied by us to append layers to our model.

Compiling Model
At this stage, we applied categorical cross entropy. As the loss function, we can also utilize MSE. Either Adam Function or rmsprop we can use as learning rate. Adam function is taking long time than rmsprop. The learning rate indicates how quick the best weights for the model are being calculated. A tinier learning rate may cause to high in precise weights (up to a particular point), but the time it considers to evaluate the weights will be more. If learning rate is high, learning done is quick. Whence, there might be prune in precision. If learning rate is less, Even though there may be increase in accuracy, takes lot of time. So, Learning rate must be accordingly.

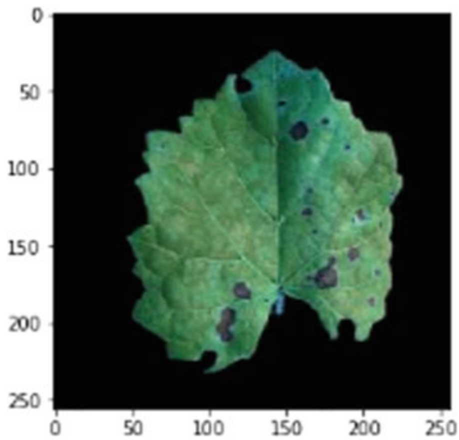


Fig. 8 Grape leaf image of 256*256 pixels

Training the Model

Right Now, Our model are going to be coached by us. To train, we'll use the 'fit()' function on our model with the next five parameters: training data (train_X), target data (train_y), validation split, the number of epochs and callbacks. The validation split will randomly split the information into use for training and testing. At the Duration of teaching, we'll be able to watch the validation loss, which offers the mean squared error of our model on the validation set. we put the validation split at 0.15, which suggests that 15% of the training data we provide within the model are visiting be forgot for testing model performance.

The number of cycles is that the quantity of times the model will cycle through the information. The high cycles we run, the high the model will get better, up to a specific point. Then point, the model will stop improving during each cycle. Additionally, the more cycles, the more the time the model will consider to execute. To watch this, we are going to use 'early stopping'.

Making Predictions on New Sample Data

In order to predict new data, we used predict() function.

predict_class = model.prdedict(img).

8 Results

Here we considered grape disease dataset as the input. Grape disease dataset contains nearly 3000 training images and nearly 880 testing images. These images are colored. Images if they are taken from uncontrolled background and different lighting condition, then background of training images may bias the neural network. These sort of pictures must be preprocessed. So, There is uniform background to these images that is only leaf figure is appeared and remaining part is black as shown in Fig. 9.

The dataset used in this paper is PlantVillage Dataset [21] and is available on Kaggle which is open source. It has approximately 55,000 well-labelled images of healthy leaves and infected leaves. This dataset contains leaf in broad level and here

Fig. 9 Grape leaf from dataset

taking two types of dataset where first one is segmented that comprises a leaf without background and other set comprises a leaf with a background.

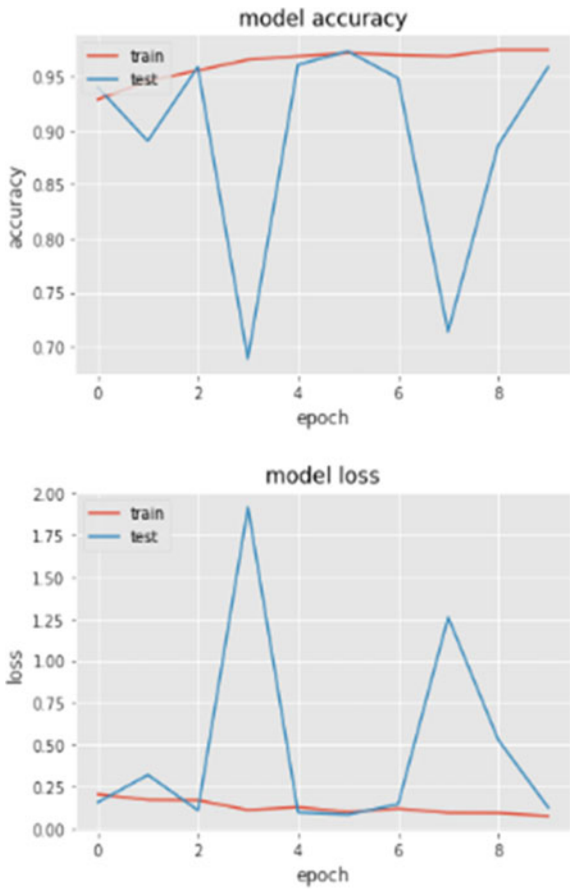
The total images in a given specific class is not same, and it varies from 423 to 1383 images. For our problem statement, we have used only Grape images which comprises of four classes i.e. black rot, Esca (Black Measles), Leaf Blight and healthy leaf images where the train-test-split data is shown in Tables 1 and 2.

Output is given as any one of four categories with the testing accuracy of 96.5.

If we consider grayscale images instead of colored images, Then There would be less classification accuracy compared to classification accuracy from colored images and from this we could understand that feature of color is one of the vital factor for classification. For various network architectures as shown in Table 2, different accuracies and different losses have been occurred which are shown in Figs. 10, 11, 12.

It can be watched from the above graphs that it is overfitting model. Overfitting happens when the model fits extremely well to the training collection. It is becoming

Fig. 10 [3 * 3, 3 * 3, 3 * 3, 3 * 3]



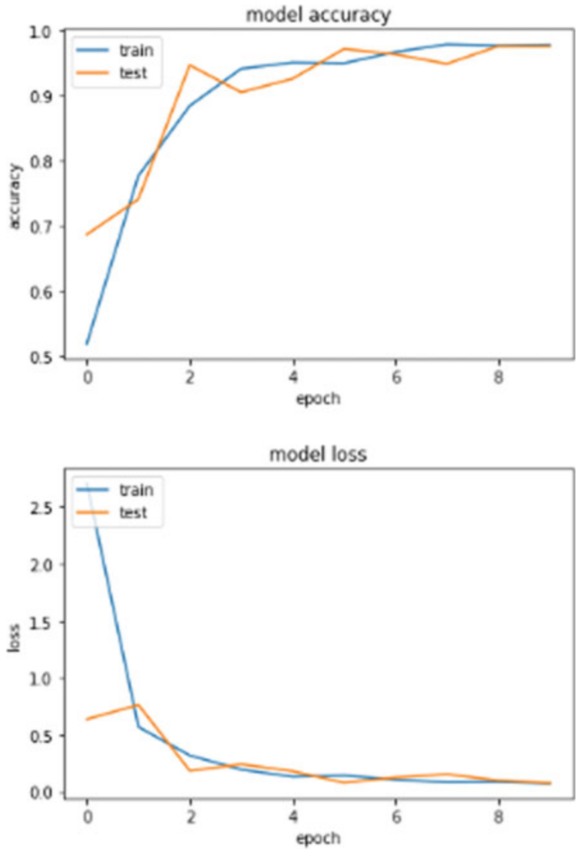


Fig. 11 [3*3, 3*3, 4*4, 4 * 4]

tough for the model to normalize to new examples which were not in the coaching set (Table 4).

9 Conclusions

The major objective of this paper is to identify and predict the Disease type of given grape leaf image. To identify and detect disease type, artificial intelligence based learning method convolutional neural network model is used. This learning method is unsupervised where target label is unknown. There are many kinds of methods in computer or automated vision plant disease identification and classification process but still this research field is lacking. Additionally, there are still no other commercial answers on the industry except those dealing with species identification of plants supported on the leaves images. So, this paper mainly identify the type of disease by using different kinds of CNN architectures, tested their accuracy and help cultivators

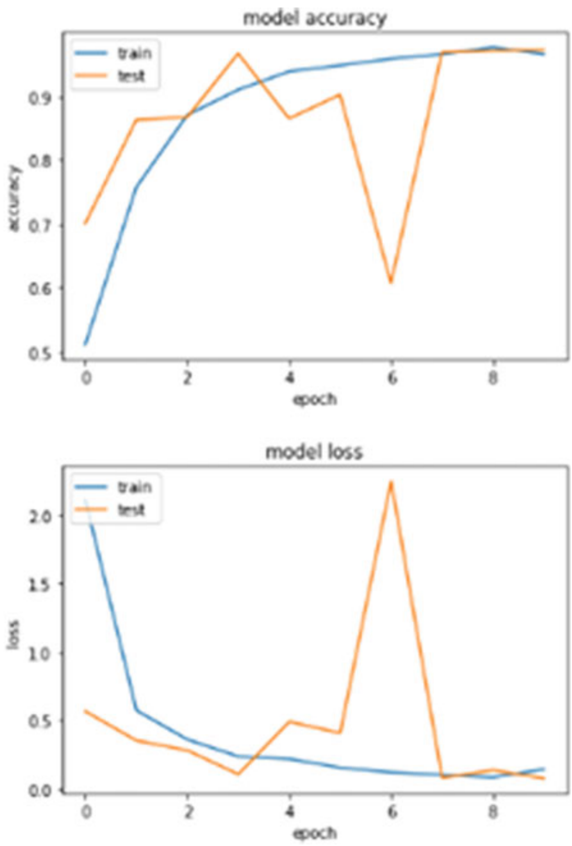


Fig. 12 [4 * 4, 4 * 4, 3 * 3, 3 * 3]

Table 4 Validation and Test accuracy of grape leaves using different CNN architecture models

Model	Network architecture	Validation accuracy	Testing accuracy
Model-I	[3 * 3, 3 * 3, 3 * 3, 3 * 3]	97.43	96.48
Model-II	[3 * 3, 3 * 3, 4 * 4, 4 * 4]	97.73	96.48
Model-III	[4 * 4, 4 * 4, 3 * 3, 3 * 3]	96.59	96.25

to detect the disease at an initial stage and may also provide solution to them what to use there to particular one so as to resolve the matter.

References

1. Amara J, Bouaziz B, Algergawy A et al (2017) A deep learning-based approach for banana leaf diseases classification. In: BTW (Workshops), pp 79–88
2. Jiang P, Chen Y, Liu B, He D, Liang C (2019) Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access 7:59069–59080
3. Geetharamani G, Pandian A (2019) Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Comput Electr Eng 76:323–338
4. Atabay HA (2016) A convolutional neural network with a new architecture applied on leaf classification. IIOAB J 7(5):226–331
5. Chaudhary P, Chaudhari AK, Cheeran AN, Godara S (2012) Color transform based approach for disease spot detection
6. Patil JK, Kumar R (2012) Feature extraction of diseased leaf images. J Signal Image Process 3(1):60
7. Reddy PR, Divya SN, Vijayalakshmi R (2015) Plant disease detection techniquetool—a theoretical approach. Int J Innov Technol Res 91–93
8. Mahlein A-K, Rumpf T, Welke P et al (2013) Development of spectral indices for detecting and identifying plant diseases. Remote Sens Environ 128:21–30
9. Xiuqing W, Haiyan W, Shifeng Y (2014) Plant disease detection based on near-field acoustic holography. Trans Chin Soc Agric Mach 2, article 43
10. Mahlein A-K, Oerke E-C, Steiner U, Dehne H-W (2012) Recent advances in sensing plant diseases for precision crop protection. Eur J Plant Pathol 133(1):197–209
11. Revathi P, Hemalatha M (2014) Identification of cotton diseases based on cross information gain deep forward neural network classifier with PSO feature selection. Int J Eng Technol 5(6):4637–4642
12. Zhou C, Gao HB, Gao L, Zhang WG (2003) Particle swarm optimization (PSO) algorithm. Appl Res Computers 12:7–11
13. Rumpf T, Mahlein A-K, Steiner U, Oerke E-C, Dehne H-W, Plümer L (2010) Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. Computers Electron Agric 74(1):91–99
14. Al-Hiary H, Bani-Ahmad S, Reyalat M, Braik M, ALRahamneh Z (2011) Fast and accurate detection and classification of plant diseases. Mach Learn 14:5
15. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
16. Soderkvist O (2001) Computer vision classification of leaves from Swedish trees
17. Camargo A, Smith J (2009a) Image pattern classification for the identification of disease causing agents in plants. Computers Electron Agric 66(2):121–125. Camargo A, Smith J (2009b) An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosyst Eng 102(1):9–21
18. Mohanty SP, Hughes DP, Salathe M (2016) Using deep learning for image-based plant disease detection. Front Plant Sci 7
19. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
20. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition
21. Ali A (2019) PlantVillage dataset, Version 1. Retrieved 22 Feb 2020. https://www.kaggle.com/xabdallahali/plantvillagedataset.

Robust and Imperceptible Region Based Watermarking on Medical Images

K. Swaraja, K. Meenakshi, Padmavathi Kora, and G. Karuna

Abstract Telemedicine is the remote delivery of health care services to evaluate, diagnose and treat patients using common technology, such as video conferencing and smart phones, without the need for an in-person visit. There is likelihood in altering the medical images purposely or unintentionally while transmitting over covert channel. The Physician confirms the diagnosis region obtained from the medical image as region of interest (ROI), prior to interpreting any report on evaluation. Watermarking scheme for medical images exploring DCT domain is conferred in this proposal. Fuzzy c algorithm is utilized in segmenting the assessment region (ROI) and non-interest region (RONI), further the watermark is inserted through modulation scheme termed as M-ary. The scheme efficacy is determined for MRI medical images through simulation by computation of quality metrics such as PSNR and NCC.

Keywords Medical image watermarking · Region of interest · Fuzzy c-means · M-ary modulation

1 Introduction

Internet is the most innovative improvement in the existing technology. In our existence, health care is the most significant application of internet to health concern providers as Electronic Patient Record (EPR) is transferred to dissimilar organizations. The watermarking schemes in the telemedicine area [1–3] entail extreme caution while inserting extra data inside the medical images as the added data need not distress the quality of the image. Therefore, to overcome the difficulty of memory exploitation as well as to defend the medical details against illicit handling, watermarking for medical images is employed. Thus in this work, a watermarking system exploiting the M-ary modulation scheme is proposed. The image with medical details

K. Swaraja (✉) · K. Meenakshi · P. Kora
Electronics and Communications Engineering, GRIET, Hyderabad, India

G. Karuna
Computer Science and Engineering, GRIET, Hyderabad, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_78

is separated into regions termed as assessment region (ROI) and non-interest region (RONI) after the confirmation given by Physician or Surgeon from the Diagnosis. ROI is extorted by Fuzzy C-means procedure, moreover the electronic record of the patient taken as watermark is inserted in non-interest region by means of M-ary modulation into the mid-band frequency coefficients of DCT. Thus, this modulation procedure is selected prior to the insertion of watermark to augment the security and covertness of the details in the medical image.

Rest of the work is organized consequently: Related effort concern to watermarking schemes is presented in Sect. 2 in the area of medical field. In Sect. 3 M-ary modulation is illustrated. The Proposed scheme is discussed with regard to ROI extraction along with the procedure of insertion and extraction of the watermark in Sect. 4. Simulation results are analyzed in terms of PSNR and Normalized Correlation (NC) in Sect. 5. Lastly, the given proposal is concluded in Sect. 6.

2 Related Work

There are several watermarking techniques of image [4, 5, 11] and video [12–15]. Inserting details of the patient into the modalities of medical image for diagnosis purpose is termed as Medical Image Watermarking scheme and it is ought to be robust in life. Swaraja et al. [3] projected data concealing of electronic record of the patient (EPR) into the test image to offer secrecy, reliability along with authenticity to electronic record of the patient details. These details might be related to diagnosis, ECG signals otherwise a signature preserved digitally with regard to physician, as a choice hospital emblem. The surgeons signature which is recorded digitally as well as the emblem of hospital can be utilized for authentication of electronic record of the patient. To attain secrecy, electronic details of the patient can be decoded. Lee et al. [16] projected watermarking scheme based on DWT to avert illicit misrepresentation. The watermark which includes the particulars of Patient is concealed into RONI to make the diagnosis perfectly. Li. et.al [17] projected medical image watermarking by means of IA-W watermarking to find the source of an invalid discharge of medical test images within the multicast atmosphere. Navas et al. [18] projected data concealing scheme for EPR which is invertible on CT brain images by utilizing integer wavelet transform (IWT). To obtain additional protection for medical images, EPR detail is concealed into RONI further ROI is selected randomly with a rectangular shape. Guo et al. [19] projected a ROI watermarking process for medical images which is lossless by exploiting Difference Expansion scheme in which the watermark is inserted into RONI whereas ROI as.

a polygon is selected physically. Dhavale et al. [20] projected safe transfer of medical images by exploiting Discrete Cosine Transform (DCT) as well as Least Significant Bit (LSB) replacement beside new (CDCS) for inserting details of the patient to augment concealing capability to present superior perceptual excellence of stago images. By applying this method if there is minor alteration of stago images either in inserting area or in ROI can be simply noticed at the receiver end.

3 Methodologies

M-ary Modulation

In digital communication the minimum entity used to represent data is termed with codes, however in case of binary signalling logic 0 or logic 1 are the two disparate codes.In case of M-ary (M > 2) signaling more than two dissimilar symbols are there. With regard to M-ary signaling method consider M feasible signals as $s1(t)$, $s2(t)$ … $s(t)$ moreover these signals can be passed on for each signaling that to for a scrupulous period of time *Ts*. Commonly, for all uses the probable count of codes can be signified as $M = 2^K$ where K is an integer. Spreading of data with regard to pass-band can be created through changing the amplitude and phase along with frequency of a sinusoidal carrier in M discrete steps, therefore ASK, PSK and FSK digital modulation methods are acquired. Dissimilar bandwidth ability next to the price of power efficacy could be attained with modulation techniques by exploiting M-ary.

M- PSK: With regard to modulation signals for M-PSK, it can be defined as given in Eq. 1.

$$R(t) = C\ \cos(2\pi f_c t + \theta_j + \theta'), \quad 0 \le t \le T \tag{1}$$

where, $\theta_j = \frac{2\pi}{M} j$.

for j = 0, 1, …, M − 1 where C is a constant and f_c is a carrier frequency. The initial phase angle is symbolized by θ′ and symbol period is signified by T. Equation 2 is acquired by expanding Eq. 1. Therefore,

$$R(t) = C\{\cos\theta j \cos(2\pi f ct + \theta') - \sin\theta j \sin(2\pi f ct + \theta')\} \tag{2}$$

Signal power can be symbolized as $P = \frac{C^2}{2}$ where $C = \sqrt{2P}$.
Then Eq. 2 can be written as

$$R(t) = \sqrt{PT \cos\theta j \sqrt{\frac{2}{T}} \cos(2\pi f ct + \theta')} - \sqrt{PT \sin\theta j \sqrt{\frac{2}{T}} \sin(2\pi f ct + \theta')} \tag{3}$$

$$R(t) = \sqrt{E \cos\theta j \sqrt{\frac{2}{T}} \cos(2\pi f ct + \theta')} - \sqrt{E \sin\theta j \sqrt{\frac{2}{T}} \sin(2\pi f ct + \theta')} \tag{4}$$

where $E = PT$ is the energy enclosed in a symbol interval. For ease, select angle θ′ to be zero and take $\varphi_1\, t = \sqrt{\frac{2}{T}} \cos 2\pi fc\, t$ and $\varphi_2\, t = \sqrt{\frac{2}{T}} \sin 2\pi fc\, t$ as the orthogonal basis functions from Eq. 3. The signal constellation figures for M-PSK modulation is depicted in Fig. 1 for M = 16.

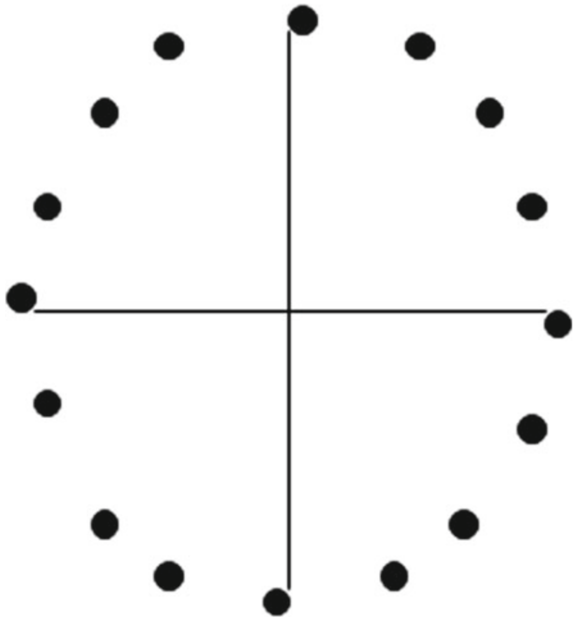


Fig. 1 Constellation diagram for M = 16

4 Proposed Scheme

Electronic record of the patient which is taken as a watermark in this work is inserted into the coefficients of mid- frequency band of DCT by implementing the M-ary modulation. During the process of inserting the watermark the scheme entails two stages, initially the assessment region (ROI) which is a portion to be diagnosed stands to be extorted by means of Fuzzy-C and furthermore, the electronic record of the patient (EPR) considered as a watermark is chosen to insert in the coefficients of mid frequency band of DCT by M-ary modulation within the non interest region (RONI).

4.1 Extraction of ROI

The portion of any medical image to be diagnosed which is termed as ROI includes useful information and so there is a need to be preserved without alteration. At present numerous segmentation schemes for image, tracked through morphological functions are there for extorting the portion of region to be diagnosed. The steps included for extorting the ROI from the host image are as follows.

- The image taken as input is a grey scale image.
- The algorithm which implements Fuzzy c-means is applied to get a binary image from grey scale image. Alignment of information starting with a big set to create a brief depiction is involved in Fuzzy clustering.

- To work out on automatic threshold value canny edge detection is applied to approximate the quantity of light as well as non-edge pixels of image.
- In this procedure, by applying Morphological functions the output image is attained based on evaluation of the consequent picture element in the input image with its neighbours. Binary mask can be produced by exploiting dilation and erosion function and acquired a white portion image which is called as ROI.
- RONI region is extracted by subtracting the ROI from the original image.

4.2 Watermark Embedding

The watermark considered in this work is an Electronic patient record which is inserted in the mid- frequency band of DCT coefficients. Primarily the RONI part is modulated with M-ary modulation prior to inserting the watermark. So as to develop the visual imperceptibility and robustness of the projected scheme, watermark is inserted in the middle chosen DCT coefficients [21]. The subsequent insertion process is discussed below:

Step 1: The grey scale medical image considered as cover image with dimension of 256 × 256 pixels is Interpreted.
Step 2: After segregating the test image into 8 × 8 block dimension, the DCT is applied to the cover image. Specific coefficients of DCT in the middle range of frequency are preferred while inserting the watermark.
Step 3: The watermark must not be inserted in the X and Y coordinates of region of interest (ROI) portion while selecting the precise coordinates.
Step 4: Towards improving the protection of the system, the scheme of M-ary modulation is executed on the watermark details, in addition these details are concealed in the mid- band frequency coefficients of DCT. The size of the watermark is 32 × 32 pels moreover modulation with 16-PSK is utilized.

4.3 Watermark Extraction

The detection procedure is a converse practice of concealing the watermark. The extortion steps include:

Step 1: The medical image which is watermarked, with dimension 256 × 256 pixel is to be taken.
Step 2: Watermarked medical image is converted into 8 × 8 pixels then DCT is applied on the RONI area.
Step 3: Extort the mid frequency coefficients from the tested position of the watermark.
Step 4: M-ary demodulation scheme is exploited.
Step 5: Watermark is retrieved.

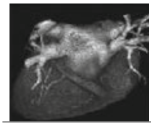
5 Experimental Results and Analysis

Simulations were performed on dissimilar MRI medical images. Medical original images are of size 256 × 256 pixels and the watermark size used is 32 × 32 pixel. In the proposed work watermark is inserted in the non interest region whereas the diagnosis portion termed as ROI must persist safe. Herein the watermark insertion besides extortion is interpreted with M-ary modulation towards dissimilar MRI medical images. The watermark which is considered as the electronic record of the patient (EPR), with unusual MRI images is demonstrated in Table 1. To assess alteration among original image as well as watermarked medical image specific preset additive white Gaussian noise (AWGN) is adapted, in addition outcome is examined through assessing some performance constraints like: Peak signal to noise ratio (PSNR) for evaluating Imperceptibility and Normalized correlation for evaluating Robustness. The simulation outcome for Watermarked image in terms of PSNR and NC values [22] along with the recovered watermark for dissimilar MRI medical image modalities is depicted in Table 2. The deformation among the original medical watermark image as well as recovered medical watermark image is estimated pertaining to PSNR and Normalized cross correlation (NCC) and is depicted in Table 2. The complete explanation with formulae for PSNR and MSE along with NCC constraints was elucidated in [23, 24]

Table 1 Dissimilar MRI Medical Images and Watermark taken for simulation are exposed

Image modality	Original MRI medical image	Watermark
Elbow		The Brain foundation: Secunderabad Patient_ref_no:019183 Doctor Name: Dr. K. Rao Patient Name: W. Rajesh Age: 65
Heart		The Brain foundation: Secunderabad Patient_ref_no:019176 Doctor Name: Dr. K. Rao Patient Name: S. Venkat Age: 65
Hips		The Brain foundation: Secunderabad Patient_ref_no:021517 Doctor Name: Dr. K. Rao Patient Name: S. Venkat Age: 65

Table 2 Watermarked MRI medical image Modalities with PSNR and NCC values along with recovered watermark for dissimilar MRI images

Watermarked MRI medical image modalities	Recovered watermark	PSNR (db) & NCC
Elbow 	The Brain foundation: Secunderabad Patient_ref_no:0191833 Doctor Name: Dr. K. Rao Patient Name: W. Rajesh Age: 65	65.6 0.98
Heart 	The Brain foundation: Secunderabad Patient_ref_no:0191762 Doctor Name: Dr. K. Rao Patient Name: S. Venkat Age: 65	64.3 1
Hips 	The Brain foundation: Secunderabad Patient_ref_no:0215179 Doctor Name: Dr. K. Rao Patient Name: S. Venkat Age: 65	63 1

6 Conclusion

Currently Tele diagnosis arrangement is admired in rural zones wherever medical check-up is tough to attain. In this work an image watermarking scheme is exploited by making use of DCT domain for dissimilar medical image modalities. The test image includes two regions ROI (diagnosis data) plus RONI. The diagnosis portion termed as ROI is extorted by fuzzy c- means technique and electronic record of the patient (EPR) as a watermark is to be inserted in mid-band frequency DCT coefficients utilizing M-ary modulation only on the RONI area. By exploiting M-ary modulation, the system robustness is enhanced through attaining improved values of PSNR with an average assessment of 64.5 dB and also recommends superior values towards the normalized cross correlation (NCC) of the extorted watermark.

References

1. Swaraja K, Meenakshi K, Kora P (2020) An optimized blind dual medical image watermarking framework for tamper localization and content authentication in secured telemedicine. Biomed Signal Process Control 55:101665
2. Swaraja K (2017) Protection of medical image watermarking. J Adv Res Dynam Control Syst (JARDCS) 11. ISSN: 1943-023X
3. Swaraja K (2018) Medical image region based watermarking for secured telemedicine. Multimedia Tools Appl 77(21):28249–28280
4. Meenakshi K, Rao CS, Prasad KS (2014) A robust watermarking scheme based Walsh Hadamard transform and SVD using ZIG ZAG scanning. In: 2014 international conference on information technology. IEEE, pp 167–172

5. Meenakshi K, Swaraja K, Kora P (2020) A hybrid matrix factorization technique to free the watermarking scheme from false positive and negative problems. Multimedia Tools Appl 79(39):29865–29900
6. Kulkarni BP, Krishna SS, Meenakshi K, Kora P, Swaraja K (2020) Performance analysis of optimization algorithms GA, PSO, and ABC based on DWT-SVD watermarking in OpenCV python environment. In: 2020 international conference for emerging technology (INCET). IEEE, pp 1–5
7. Meenakshi K, Swaraja K, Kora P, Karuna G (2021) A robust blind oblivious video watermarking scheme using undecimated discrete wavelet transform. In: Intelligent system design. Springer, Singapore, pp 169–177
8. Sravan V, Swaraja K, Meenakshi K, Kora P, Samson M (2020) Magnetic resonance images based brain tumor segmentation—a critical survey. In: 2020 4th international conference on trends in electronics and informatics (ICOEI)(48184), 15 June 2020. IEEE, pp 1063–1068
9. Meenakshi K; Swaraja K, Usha Kumari Ch, Kora P (2019) Grading of quality in tomatoes using multi-class SVM. In: 2019 3rd international conference on computing methodologies and communication (ICCMC), pp 104–107
10. Kora P, Meenakshi K, Swaraja K (2019) Detection of cardiac arrhythmia using convolutional neural network. In: Soft computing and signal processing. Springer, Singapore, pp 519–526
11. Kora P, Kumari CU, Swaraja K, Meenakshi K (2019) Atrial fibrillation detection using discrete wavelet transform. In: 2019 IEEE international conference on electrical, computer and communication technologies (ICECCT), pp 1–3
12. Meenakshi K, Srinivasa Rao C, Satya Prasad K (2014) A scene based video watermarking using slant transform. IETE J Res 60:276–287
13. Meenakshi K, Prasad KS, Rao CS (2017) Development of low-complexity video watermarking with conjugate symmetric sequency-complex hadamard transform. IEEE Commun Lett 21:1779–1782
14. Swaraja K, Meenakshi K, Kora P (2019) Robust optimized discrete wavelet transform singular value decomposition based video watermarking. Traitement du Signal 36(6):565–573. https://doi.org/10.18280/ts.360612
15. Swaraja K, Madhaveelatha Y, Reddy VS (2016) Robust video watermarking by amalgamation of image transforms and optimized firefly algorithm. Int J Appl Eng Res 11(1):216–225
16. Lee HK, Kim HJ, Kwon KR, Lee JK (2005) Digital watermarking of medical image using ROI information. In: Enterprise networking and computing in healthcare industry. In: Proceedings of 7th international workshop on HEALTHCOM 2005. IEEE, pp 404–407
17. Li M, Poovendran R, Narayanan S (2005) Protecting patient privacy against unauthorized release of medical images in a group communication environment. Comput Med Imaging Graph 29:367–383
18. Navas K, Thampy SA, Sasikumar M (2008) EPR hiding in medical images for telemedicine. Int J Biomed Sci 3:44–47
19. Guo X, Zhuang TG (2009) A region-based lossless watermarking scheme for enhancing security of medical data. J Digital Imaging 22:53–64
20. Dhavale SV, Mali SN (2010) high capacity robust medical image data hiding using cdcs with integrity checking. Int J Recent Trends Eng Technol 3
21. Podilchuk C, Zeng W (1997) Perceptual watermarking of still images. In: IEEE first workshop on multimedia signal processing. IEEE, pp 363–368
22. Swaraja K, Madhaveelatha Y, Reddy VSK (2011) Video watermarking based on motion vectors of H. 264 In: India conference (INDICON), 2011 Annual IEEE 2011, 16 Dec 16, pp 1–4
23. Swaraja K, Madhaveelatha Y, Reddy VSK (2014) A pristine digital video watermarking in H.264 compressed domain In: IEEE international conference on computational intelligence and computing research (ICCIC), Coimbatore, India, 18–20 Dec 2014, pp 1–4, ISBN: 978-1-4799-3974-9
24. Swaraja K., Madhaveelatha Y, Reddy VSK (2015) A secure method of optimized low complexity video watermarking. ARPN J Eng Appl Sci 10(4):1822–1827. ISSN 1819-6608

A Robust Watermarking Using RDWT and Slant Transform Using Hybrid Firefly and Differential Evolution Optimization Algorithm

K. Meenakshi, K. Swaraja, Padmavathi Kora, and G. Karuna

Abstract In this work, an optimized watermarking framework is proposed with the hybrid combination of two metaheuristic algorithms-firefly optimization and differential evolution, namely HFADE. The cover image is partitioned into 4×4 sub-blocks, and the watermark is concealed in the slant domain using Quantized Index Modulation (QIM). The optimized thresholds obtained with HFADE used in quantization to improve imperceptibility and robustness. Peak Signal to noise ratio(PSNR) and Normalized Cross Correlation (NCC) are used for evaluation of the proposed watermarking scheme. The fitness function for HFADE is taken as the reciprocal of mean square error between the watermarked and cover image. Simulation outcomes convey that the proposed scheme maintains improved imperceptibility, and the watermark extracted from a seriously distorted image.

Keywords Differential evolution · Firefly · Slant transform · Quantization index modulation

1 Introduction

The improvement of Internet and computer input-output devices has made the broadcast and alteration of digital content without difficulty. With the advanced editing technologies, an edited copy appears similar to the original. So, the security of ownership of digital content has become the utmost concern. Digital watermarking [1–16] is evolved to protect the interests of owners from the copyright infringement. The three trade-off parameters of watermarking scheme are transparency, robustness and capacity. Further, it must not suffer from the false positive problems. Optimization algorithms such as Fuzzy logic [17, 18], Genetic Algorithm [19], Differential Evolution (DE) [1] are used to optimize the mutually conflicting parameters of transparency and robustness.

K. Meenakshi (✉) · K. Swaraja · P. Kora · G. Karuna
Department of ECE, GRIET, Hyderabad, India
e-mail: mkollati@gmail.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_79

In [20], a novel blind image watermarking technique, is proposed based on the Redundant Discrete Wavelet Transform (RDWT) [21] and singular value decomposition (SVD) employing self-adaptive differential evolution (SADE) algorithm.The algorithm is found to be imperceptible, robust and has high capacity. The false positive problem in [3, 22] is avoided by sending 8 bit signature at the sender side along with the cover image. Reference [23] developed a multipurpose image watermarking scheme using Artificial Bee colony optimization(ABC) for copyright protection and tamper detection. The main drawback of ABC is that it suffered with slow convergence.

Reference [24] proposed image hashing based on slant transform (ST), which is widely used in content authentication, content based retrieval, and image watermarking. Reference [25] employed ST for video watermarking. The advantage of a slant transform is that it is recursive. This feature is explored in this scheme to produce robust and imperceptible watermarking.

The proposed algorithm explores hybrid combination of RDWT and ST, and identify multiple thresholds in slant transform by using a novel hybrid optimization algorithm of a combination of Firefly and Differential evolution HFADE. The fitness function used in the optimization is the reciprocal of mean square error between the watermarked and cover image.

The research contribution of the paper:

- RDWT is used for obtaining the low approximation coefficients of the image.
- Slant transform is applied to low approximation coefficients of RDWT.
- Later, using QIM, the slant transform is adjusted with a threshold to obtain modified slant transform coefficients.
- The thresholds which give the better PSNR and NCC are selected by HFADE optimization.

2 Preliminaries

The proposed watermarking used a hybrid combination of Redundant Discrete wavelet transform and slant transform, and the watermark is embedded in the selected slant coefficients using quantization index modulation. HFADE obtains the optimum thresholds. In this a brief introduction of Redundant Discrete wavelet transform, slant transform, and HFADE are described below.

2.1 Redundant Discrete Wavelets

In place of DWT, RDWT attracted researchers in watermarking field due to several advantages of it over DWT:

- RDWT is shift invariant.

- The size of RDWT transformed image is four times that of DWT. So there is an improvement of capacity in RDWT based watermarking scheme compared to DWT.
- Previous research reveals that if subband is having the same dimension of cover image, the watermarking scheme offers high imperceptibility. This feature in RDWT helps to boost imperceptibility in the proposed watermarking scheme.

2.2 Slant Transform

ST is a recursive orthogonal transform widely used in image processing due to its speed of processing. The transform comprises slant basis functions. The middle frequency coefficients of slant are used for watermark insertion in the proposed work.

2.3 Hybrid Firefly and PSO Algorithm (FFAPSO)

Yang devises firefly optimization in the year 2007 [26]. FA was conceived by mimicking the flashing (mating) activity of fireflies.

Fireflies employ light to attract another firefly or prey (mate). The light intensity attraction '*B*' of fireflies is inversely proportional to distance '*r*'. Hence, most fireflies are constraint only up to several hundreds of meters. To execute this algorithm, the fitness function is articulated based on the fluorescence light behavior of fireflies. For simplicity, it is imagined that the light intensity attractiveness of the firefly is determined by its brightness '*B*', which is connected with the fitness function.

Fireflies attractiveness μ_λ at light absorption factor λ is defined as

$$\mu_\lambda = \mu_0 \mathrm{e}^{-\lambda r_{k,l}^2} \tag{1}$$

where r is the distance between kth firefly and lth firefly and λ is the light absorption factor and μ_0 is the initial attraction at distance r_{kl}=0. The distance r_{kl} between two fireflies k and l at position x_k and x_l are computed as given in Eq.

$$||x_k - x_l|| = \sqrt{\sum_{k=1}^{d} x_{k,d} - x_{l,d})} \tag{2}$$

where $x_{l,k}$ is the kth component of the position x_l of lth firefly. The firefly k attracted to another more attractive (brighter) firefly l is determined by y_k

$$xk \rightarrow x_k + \mu_0 e^{-\lambda r_{k,l}} (x_l - x_k) + \eta \left(rand1 - \frac{1}{2} \right) \tag{3}$$

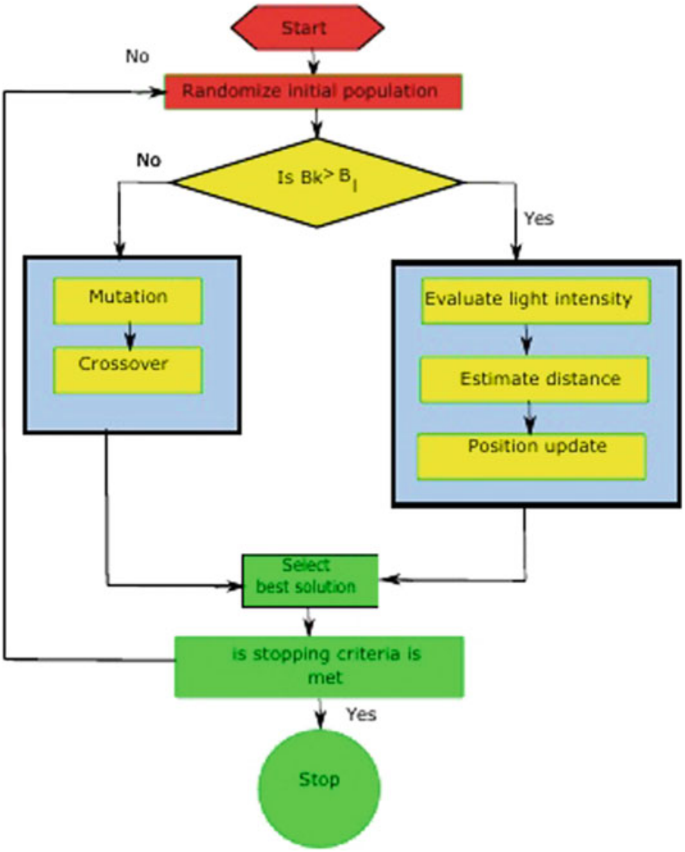


Fig. 1 Flowchart of HFADE

In Eq. 15, the x_k and x_l are the positions of kth and lth firefly and η is a randomization parameter. If the light absorption factor $\lambda \rightarrow 0$, the μ_λ=μ_0. Hence flashing fireflies can be seen within the search domain. On the other hand, if $\lambda = \infty$, the attractiveness is almost zero, and the firefly moves randomly. When fireflies move randomly, there is plenty of chance for the FA to limit local maxima, resulting in premature convergence. As FA has these limitations, the randomness in FA replaced by DE in the HFADE hybridization algorithm. FA and DE's hybrid algorithm (HFADE) has been proposed to avoid FA's premature convergence. The DE algorithm updates the position of firefly based on the mutation and crossover operators. Firefly settle in the local best when it is used alone. As shown in Fig. 1 crossover and mutation operator in DE are used to move the fireflies towards the global best.

3 Watermark Embedding and Extraction Algorithm

The watermark embedding steps are as follows:

Step 1 Perform 1-level RDWT onto the Host image to Partition it into its four sub-bands, which are I_{LL}, I_{LH}, I_{HL} and I_{HH}.
Step 2 Extract I_{LL} in RDWT band for watermark insertion.
Step 3 Segment I_{LL} band into 4 × 4 Blocks.
Step 4 Apply Slant Transform to each 4 × 4 Block.
Step 5 Two AC coefficients AC_x and AC_y in each block are identified based on HVS weightage matrix generated using Eq. 1–5 in [25].
Step 6 The slant coefficients AC_y are replaced with AC_x with T based on bits in the watermark.
Step 7 If watermark bit is zero

$$AC_y = AC_x - T \tag{4}$$

else

$$AC_y = AC_x + T \tag{5}$$

Step 8 The N/4 × N/4 times threshold values are generated using HFADE algorithm using random population.
Step 9 With modified coefficients with thresholds, the inverse slant is applied to obtain the lower approximation coefficients of RDWT LL_{new}.
Step 10 Merge ILL_{new}, I_{LH}, I_{HL} and I_{HH} bands.
Step 11 Apply inverse RDWT to obtain the watermarked image.

The steps of watermark extraction from the watermarked image are as follows.

Step 1 Perform 1-level RDWT onto the watermarked image to Partition it into its four sub-bands, which are Iw_{LL}, Iw_{LH}, Iw_{HL} and Iw_{HH}.
Step 2 Apply block-based ST on Iw_{LL} of the watermarked image.
Step 3 AC_x are identified and if $AC_y > AC_x$, watermark bit zero is extracted; else watermark bit 1 is extracted.

The steps of HFADE is formulated as follows:

Step 1 Define the fitness function, numbers of variables, and the values for population size, crossover rate, mutation rate, initial attraction parameter, and light absorption factor of firefly and number of generations (or any other terminating criteria).
Step 2 Initialize the population stochastically.
Step 3 Generate a watermarked image by applying random solutions in the population utilizing the concealing process. The solutions are the threshold values in each block.
Step 4 Calculate the NC values between the host image and each watermarked images.

Step 5 Apply the attack functions upon the watermarked images one by one.
Step 6 Extract out the watermarks from the attacked images using the extraction procedure.
Step 7 Calculate the NC values between the watermark and the extracted ones.
Step 8 Evaluate the fitness value for each corresponding solution.

The fitness function of HFADE is the 1/MSE, where MSE is the mean square error between the watermarked and original image.

4 Simulation Outcomes

This section validates the efficiency of the proposed watermarking scheme by taking the Lena image as the host image and cameraman as the watermark. The physical distinction between the host and watermarked image is negligible, and a PSNR of 65–75 dB is achieved with the proposed RDWT and slant based watermarking employing HFADE optimization. The attacked watermarked images with Gaussian noise, median filtering, symmetric cropping, and rotation by 30° is shown in Fig. 2. Further, the extracted watermarks are shown for the attacked images in Fig. 3.

The simulations we conducted with the MATLAB 2018, pentium processor with 8 GB RAM. We compared the proposed scheme against the related works using svd based GA and DCT and SVD based DE. The PSNR of the proposed algorithm is

Fig. 2 Attacked watermarked images **a** Gaussian noise, **b** median filtering, **c** symmetric cropping, **d** rotation

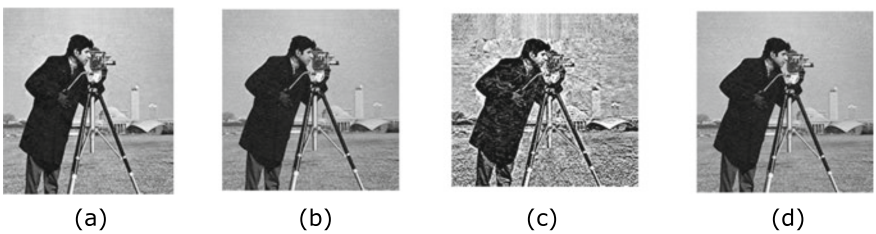


Fig. 3 Extracted watermarks from **a** Gaussian noise, **b** median filtering, **c** symmetric cropping, **d** rotation

Table 1 NCC applied for different attacks

Parameters	SVD+GA	DCT+SVD+DE	Proposed scheme
Gaussian noise	0.9662	0.9912	1
Median filtering	0.9677	0.998	0.998
Symmetric cropping	0.9788	0.9822	0.9122
Rotation	0.9956	0.9956	0.933

65.55 db, where as the PSNR of SVD based GA and DCT, SVD based DE is 37.88 and 44.33 dB. The NCC of the proposed algorithm compared to SVD based GA and DCT,SVD based DE are tabulated in Table 1. The results demonstrate the efficiency of the algorithm in terms of imperceptibility and robustness.

5 Conclusion

The proposed algorithm is false positive free, imperceptible and robust. The computational complexity of the HFADE is less because both FA and DE requires less parameters and simple in implementation.

References

1. Meenakshi K, Swaraja K, Kora P (2020) A hybrid matrix factorization technique to free the watermarking scheme from false positive and negative problems. Multimedia Tools Appl 79(39):29865–29900
2. Meenakshi K, Prasad KS, Rao CS (2017) Development of low-complexity video watermarking with conjugate symmetric sequency-complex Hadamard transform. IEEE Commun Lett 21(8):1779–1782
3. Meenakshi K, Rao CS, Prasad KS (2014) a fast and robust hybrid watermarking scheme based on Schur and SVD transform. Int J Res Eng Technol 3(4):7–11
4. Meenakshi K, Rao CS, Prasad KS (2014) A robust watermarking scheme based Walsh-Hadamard transform and SVD using zig zag scanning. In: 2014 international conference on information technology, pp 167–172. IEEE
5. Swaraja K, Madhaveelatha Y, Reddy VSK (2015) A secure method of optimized low complexity video watermarking. ARPN J Eng Appl Sci 10(4):1822–1827. ISSN 1819-6608
6. Swaraja K, Madhaveelatha Y, Reddy VSK (2014) A Pristine digital video watermarking in H.264 compressed domain. In: IEEE international conference on computational intelligence and computing research (ICCIC), Dec 18–20, Coimbatore, India, pp 1–4. ISBN: 978-1-4799-3974-9
7. Swaraja K, Madhaveelatha Y, Reddy VSK (2011) Video watermarking based on motion vectors of H. 264 In: India conference (INDICON), 2011 Annual IEEE 2011 Dec 16, pp 1–4
8. Meenakshi K, Kora P, Kishore D (2019) Video watermarking with curvelet transform. Int J Innov Technol Explor Eng (IJITEE) 8:602–607

9. Kulkarni BP, Krishna SS, Meenakshi K, Kora P, Swaraja K (2020) Performance analysis of optimization algorithms GA, PSO, and ABC based on DWT-SVD watermarking in OpenCV python environment. In: 2020 international conference for emerging technology (INCET). IEEE, pp 1–5
10. Meenakshi K, Swaraja K, Kora P, Karuna G, A robust blind oblivious video watermarking scheme using undecimated discrete wavelet transform. In: Intelligent system design. Springer, Singapore, pp 169–177
11. Kuraparthi S, Kollati M, Kora P, Robust optimized discrete wavelet transform-singular value decomposition based video watermarking robust optimized discrete wavelet transform-singular value decomposition based video watermarking
12. Swaraja K, Madhaveelatha Y, Reddy VS (2016) Robust video watermarking by amalgamation of image transforms and optimized firefly algorithm. Int J Appl Eng Res 11(1):216–25
13. Swaraja K, Meenakshi K, Kora P, An optimized blind dual medical image watermarking framework for tamper localization and content authentication in secured telemedicine. Biomed Sig Process Control 55:101665
14. Swaraja K (2017) Protection of medical image watermarking. J Adv Res Dyn Control Syst (JARDCS) Special issue 11. ISSN: 1943-023X
15. Swaraja K, Medical image region based watermarking for secured telemedicine. Multimedia Tools Appl 77(21):28249–28280
16. Sravan V, Swaraja K, Meenakshi K, Kora P, Samson M (2020) Magnetic resonance images based brain tumor segmentation-a critical survey. In: 2020 4th international conference on trends in electronics and informatics (ICOEI)(48184) 2020 Jun 15. IEEE, pp 1063–1068
17. Meenakshi K, Swaraja K, Kora P, Ch UK (2019) Texture feature based oblivious watermarking with slant transform using fuzzy logic. In: 2019 IEEE 5th international conference for convergence in technology (I2CT). IEEE, pp 1–5
18. Meenakshi K, Bethel GB (2014) Design and simulation of constant bit rate compressor using fuzzy logic. In: 2014 first international conference on networks & soft computing (ICNSC2014). IEEE, pp 309–313
19. Meenakshi K, Rao S, Prasad KS (2014) A hybridized robust watermarking scheme based on fast Walsh-Hadamard transform and singular value decomposition using genetic algorithm. Int J Comput Appl 108(11)
20. Ernawan F, Kabir MN (2018) A blind watermarking technique using redundant wavelet transform for copyright protection. In: 2018 IEEE 14th international colloquium on signal processing its applications (CSPA). IEEE, pp 221–226
21. Yasasvy T, Sushil KV, Meenakshi K, Swaraja K, Kora P (2019) A hybrid blind watermarking with redundant discrete wavelet and Hadamard transform. Int J Innov Technol Explor Eng 8(11):2216–2220
22. Meenakshi K, Swaraja K, Kora P (2019) A robust DCT-SVD based video watermarking using zigzag scanning. In: Soft computing and signal processing. Springer, Singapore, pp 477–485
23. Ansari IA, Pant M (2017) Multipurpose image watermarking in the domain of DWT based on SVD and ABC. Pattern Recogn Lett 94:228–236
24. Riaz U, Razzaq FA, Khan A, Gul MT (2017) Sparsity of magnetic resonance imaging using slant transform. In: 2017 international conference on frontiers of information technology (FIT). IEEE, pp 368–372
25. Meenakshi K, Srinivasa Rao C, Satya Prasad K (2014) A scene based video watermarking using slant transform. IETE J Res 60(4):276–287
26. Wang WC, Xu L, Chau KW, Xu DM (2020) Yin-Yang firefly algorithm based on dimensionally Cauchy mutation. Expert Syst Appl 150

Statistical Analysis of Text Corpus to Determine Appropriate Syllable Length for TTS

K. V. N. Sunitha and P. Sunitha Devi

Abstract Speech synthesis is the computer generated simulation of human speech for any text input. The objective of a text to speech system is to convert an arbitrary text into its corresponding spoken waveform. Generating speech that is close to natural human speech is always a challenging issue for synthesis systems. It is a factor that depends on the kind of available database, and the algorithms that choose the appropriate speech units from the database. Since decades researchers have been exploring building the database for speech systems. The issue lies on what should be the unit of speech to be stored in the database. Since Telugu language is syllabic by nature, a statistical analysis is carried out on the text corpus of 5 million words size to identify the required vowel and biphone syllable units that can cover the major vocabulary in Telugu language. This paper presents the analysis carried out at vowel and biphone level to extract the vowel and biphone syllable speech units and estimate the coverage. The experimental analysis carried out proved that top 50% biphone units covered 25% of the total corpus.

Keywords Biphone · Coverage · Speech synthesis · Syllable

1 Introduction

Speech synthesis is a process that can generate human-like speech for any text input to imitate human speakers. It is used to translate written information into audio information where it is more convenient for mobile applications, to assist vision impaired etc. Conversion of input text into linguistic representation is the first step in the TTS system, which is known as grapheme to phoneme or text to phonetic conversion [1]. The speech generation component uses either synthesis from parameters or by unit selection from a huge speech corpus to realize the orthographic form. The text processing component must produce an appropriate

K. V. N. Sunitha (✉)
BVRIT Hyderabad College of Engineering for Women, Bachupally, Hyderabad, Telangana, India

P. Sunitha Devi
G. Narayanamma Institute of Technology & Science, Shaikpet, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_80

sequence of orthographic units for the corresponding input text to generate natural sound [2, 3]. The difficulty of conversion is highly language dependent and includes many problems. For English and most of the other languages the conversion is much more complicated. A very large set of different rules and their exceptions is needed to produce correct pronunciation and prosody for synthesized speech. Most of the Indian languages are phonetic in nature [4]; the conversion is quite simple because written text almost corresponds to its pronunciation. Conversion can be divided in three main phases, text preprocessing, creation of linguistic data for correct pronunciation, and the analysis of prosodic features for correct intonation, stress, and duration. State-of-art TTS system for English and other well-researched languages use rich set of linguistic resources like word-sense disambiguation, morphological analyzer, Part-of-Speech tagging, letter-to-sound rules, syllabification, stress-patterns in some form or the other to build a text processing component of a TTS system. For minority languages which do not have enough linguistic resources requires several complexities to be resolved, starting from collection of text corpora in digital and processable format. Most of the languages of the world do not have linguistic components available in such rich fashion. In the real world, minority languages including few of the Indian languages do not have the luxury of assuming some or any of the components.

1.1 Features of Telugu Language

South central dravidian language Telugu is predominantly spoken in south India and official language in the states of Telangana and Andhra Pradesh [5]. Amongst the four classical languages of India, by the number of native speakers in India (74 million), Telugu ranks third and thirteenth in the list of most-spoken languages worldwide. Officially there are 23 languages of India, and most of them except Urdu and English share a common phonetic base, i.e., they use a common set of speech sounds. Almost all of these languages use a common phonetic base, but some of the languages like Hindi, Marathi and Nepali share a common script known as Devanagari. The phonotactics of the languages contribute majorly to separate from other languages rather than speech sounds and scripts. The permissible combinations of phones that can co-occur in a language is called phonotactics. The phonetic nature of Telugu language, i.e. one to one correspondence between what we speak and what we write. The letter to sound rules required to map Telugu letters to sound is straightforward.

2 GNITS Text Corpora

Creating a speech database for a text to speech system requires identifying an optimal set of textual sentences to be recorded from native speakers of the language.

These sentences should be minimum in number to save recording effort and should have enough number of occurrences of each type of sound unit to cover all the coarticulation effects. GNITS Text Corpora is a set of phonetically rich sentences which consists of DoE-CIIL corpus and newspaper articles which is nearly 5 Million word corpus for Telugu language. The number of phonemes, both vowels and consonants in Telugu, is a controversial issue. There are slightly different phonemic systems for Telugu distinguishing the social dialects into standard and non-standard, educated and uneducated, formal and informal, native and non-native. Different notations like WX, IT3 and ROMAN are in use for representing the Telugu text in digital format. The drawbacks in transliteration schemes include.

1. More than one notation for representing the same sound. For example, we may use either ‘aa’ or ‘A’ for representing long form of ‘a’
2. Not easily predictable notation like w for ‘t’ and ‘x’ for ‘d’

To avoid the drawbacks mentioned in the above notations, different codes are used wherever there is a possibility of such confusion. The text transcribed using these notations may not be straight forward to read, to overcome this problem KNS notation is used.

2.1 KNS Notation

Since Indian languages have more than 26 letters [6], there is a need to use capitals and two character codes at times. Basic lower case/upper case alphabets plus carrot (\^), tilde (\ ~), and back-quote(\‘). Maximum length of a symbol: 3 bytes. Finally 45 phonemes: 10 vowels, 2 diphthongs and 33 consonants broadly represent the standard Telugu language in KNS notation is given in Table 1a, b.

3 Syllabification Rules

In Indian languages there is one to one correspondence between what is written and what is spoken. Each character in the script has a corresponding sound unit of that language. Consonants are inherently bound with/a/vowel sound, and are pronounced with vowels always [7]. Sometimes this vowel is not pronounced, and is known as Inherent Vowel Suppression (IVS), which mostly occurs in word middle and final positions. In some occurrences this vowel is not pronounced, and this is referred to as Inherent Vowel Suppression (IVS). Letter to sound rules are almost straight forward in Indian languages, but the syllabification rules are not trivial. There is a necessity to break the words into syllables using some rules. Syllables can be broadly classified into four classes based on the number of phones they contain [8] i.e. monophone (only vowels—V), biphone (CV or VC), triphone (CVC or CCV), fourphone (CCVC) syllables. Simplistic rules were derived for syllabification based on heuristic analysis

Table 1 KNS Phoneme Set

Vowels					
a	A	i	I	u	U
అ	ఆ	ఇ	ఈ	ఉ	ఊ
e	E	ai	o	O	au
ఎ	ఏ	ఐ	ఒ	ఓ	ఔ

Consonants									
k	kh	g	Gh	-	c	ch	j	jh	-
క	ఖ	గ	ఘ	ఙ	చ	ఛ	జ	ఝ	ఞ
T	Th	D	Dh	N	t	th	d	dh	n
ట	ఠ	డ	ఢ	ణ	త	థ	ద	ధ	న
p	ph	b	Bh	m	y	r	l	W	s`
ప	ఫ	బ	భ	మ	య	ర	ల	వ	శ
S	s	H	L	R`					
ష	స	హ	ళ	ఱ					

of several words of Telugu language for forming groups of C*VC*. The rules used for syllabification of GNITS text corpus are listed as follows.

- If the first character is a ***C*** then associate it to the nearest Vowel on the right.
- If the last character is a ***C*** then associate it to the nearest Vowel on the left.
- If sequence is of the form ***V V*** then split it as ***V - V***. Ex అని అతడు , (**aniatadu**) **becomes as (a – ni – a – ta – du)**
- If sequence is of the form ***V C V*** then split it as ***V - C V***. Ex : అర (ara) becomes as (a – ra)
- If sequence is of the form ***V C C V*** then split it as ***V C – C V***. Ex : రాజ్యము (rAjyamu) becomes as (rAj – ya – mu)
- If sequence is of the form ***V C C C V*** then split it as ***V C – C C V***. Ex :ఇస్త్రీ (istrI) becomes as (is – trI)
- The strings separated by – are identified as syllable units.

4 Aligning of Speech Units in Text

It is observed that longer the speech unit the generated speech is close to the naturalness, but the size of the database increases. The increase in the size of the database requires more space and time to select units from the database. To address this problem text analysis is performed to identify the minimum possible units that can be stored to produce natural speech. Telugu language is orthogonal in nature, that is the written form and phonetic units correspond to each other. The minimum unit that can exist independently is a single vowel. For example the words in Table 2 include various sizes of speech units that range from single phone units to tri/four phone units that are uttered as a single unit. These examples indicates that possible units that are occurring as single vowel, a biphone/triphone or a morpheme unit that can exist as a single unit.

When a statement is given we need to identify the independent units that can be concatenated to generate natural speech. When two units are concatenated there is a chance of losing naturalness. It is essential to minimize the number of concatenations by selecting the possible units available in the database. Consider the following example statement which can be represented as basic independent units, that includes single phone, biphone, triphone or morph units.

Table 2 Words with possible syllable units

Word	Vowel Unit	Biphone Unit	Triphone Unit	Fourphone Unit
అన్ని	అ		న్ని	
ఆలయ	ఆ	ల, య		
ఈక	ఈ	క		
ఉడుత	ఉ	డు, త		
ఎలుక	ఎ	లు, క		
సినిమా		సి, ని, మా		
సంస్కృతి		తి	సం	స్కృ
చక్రం		చ		క్రం
పంజరం		జ	పం, రం	

Example Statement 1

Statement: చదవడానికి ఈ సంస్కృత పుస్తకం తీసుకో.

Read as : చ-ద-వ-డా-ని-కి-ఈ- సం-స్కృ-త- పు-స్త-కం- తీ-సు-కొ

Transliteration of this sentence in KNS notation
c-a-d-a-w-a-D-A-n-i-k-i I s-a-m~-s-k-r-u-t-a p-u-s-t-a-k-a-m~ t-I-s-u-k-o

Valid units that exist as single phones: ఈ

Valid units that exist as biphones: చ, ద, వ, డా, ని, కి, త, పు, తీ, సు, కొ

Possible Utterance with biphone units : చదవడానికి సరుత పుతక తీసుకొ

Valid units that exist as triphones: సం, స్త, కం

Valid units that exist as fourphones: స్కృ

Valid units that exist as morph units : చదవ, డానికి, సంస్కృత, పుస్తకం, తీసుకొ.

Example Statement 2

Statement: అక్రమార్జనను వ్యవస్థలోకి తేవడానికి పేదల ఖాతాల దుర్వినియోగం.

Read as :అ-క్ర-మా -ర్జ-న-ను వ్య-స్థ-లో-కి తే-వ-డా-ని-కి పే-ద-ల ఖా-తా-ల దు-ర్వి-ని-యో-గం.

Transliteration of this sentence in KNS notation
a-k-r-a-m-A-r-j-a-n-a-n-u w-y-a-w-a-s-th-a-l-O-k-i t-I-w-a-D-A-n-i-k-i p-I-d-a-l-a kh-A-t-A-l-a d-u-r-w-i-n-i-y-O-g-a-m~

Valid units that exist as single phones: అ

Valid units that exist as biphones: మా, న, ను, వ, లో, కి, తే, డా, ని, పే, ద, ల, ఖా, తా, దు, యో.

Possible Utterance with biphone units : మానను వలోకి తేవడానికి పేదల ఖాతాల దుయో

Valid units that exist as triphones: క్ర, ర్జ, వ్య, స్థ, ర్వి.

Valid units that exist as fourphones: NIL

Valid units that exist as morph units : అక్రమార్జనను వ్యవస్థలోకి తేవడానికి పేదల ఖాతాల దుర్వినియోగం.

For example 2 some of speech units cannot be created as isolated speech units; they have to be segmented from continuous speech only. From the examples it is clear that, selection of speech units of different size generate speech that varies in terms of pronunciation and naturalness. Table 3 explains the number of concatenations required for different units and the behavior of the system.

To minimize the number of concatenations it is preferred to have different speech units in the database and proper selection of them would result in good quality of speech. Hence it is desired to have speech units that are at morph level, fourphone,

Table 3 Different types of syllable units

Speech Unit	Number of concatenations		Behaviour of the system
	EX 1	EX 2	
Only single phone	31	53	• Database size is very less, but recording and segmentation at phone level is a complex task • As the number of concatenations increases naturalness of the speech produced decreases
Only biphone	16	18	• Database size is more when compared to phone level and requires to record and segment all biphones which is a tedious job • Number of concatenations decreases • Speech produced may be close to actual but not appropriate
Only triphone	No sentence is formed only with triphones		
Only fourphone	No sentence is formed only with fourphones		
Combination of different syllable units	15	21	• Database size increases and we need to record and segment all syllable units • Number of concatenations is very less and the speech produced is very close to required utterance
Morph units	5	6	• Very close to natural speech but the size of the speech database is very large

triphone, biphone and monophones that are frequently used. To identify these units the text corpus is analyzed at different levels. This paper focuses on analyses of biphone syllable units and the observations regarding the coverage of corpus is discussed in the latter section.

5 Experimental Analysis and Observations

After phonetization each word is a sequence of Telugu sound units. The GNITS text corpus consists of 12 vowels and 36 consonants (which includes 3 variations of anuswaras). Analysis of the text corpus is made to find the co-occurrence of these 48 phones against each of the 48 phones. In Telugu language all the 2304(48×48) biphones do not occur, and few occur very rarely. Out of 2304 biphone sequences only 737 biphones are valid biphone syllable units that occur in Telugu language. Experimental analysis was done on GNITS Text Corpora and top biphone units are listed in Table 4, based on which we can draw some observations.

Table 4 Comparison of a unit as biphone and biphone syllable unit

Biphone	Frequency	Biphone syllable units	Frequency	% of biphones as biphone syllable unit
a-n	146,905	an	16,003	10.89
a-m	136,888	am	15,228	11.12
l-a	**111,416**	**la**	**81,281**	**72.95**
n-a	**105,409**	**na**	**71,929**	**68.23**
i-n	99,361	in	4374	4.4
r-a	**95,726**	**ra**	**52,355**	**54.69**
a-r	91,124	ar	7492	8.22
a-l	84,407	al	8211	9.72
n-i	**80,553**	**ni**	**61,306**	**76.10**
k-a	**75,862**	**ka**	**50,291**	**66.29**

Vowel followed by Consonant as biphone sequence occurs with high frequency, but as a biphone syllable has very less frequency. For example the ***a-n*** (V-C) pair is occurring 146,905 but only 11% of it occurs as a biphone syllable unit. In words like ***atani, manaku, weNTanE, kanuka*** when statistical analysis is carried out ***a-n*** would be observed as a valid sequence as it is not considering the phonetic validations. But as per the language rules this combination is valid only for a few cases.

Example 1: ***anna***

when the syllabification rules are applied then we get the units as ***an-na*** which are the actual units that are uttered.

Example 2: ***nannaya***

when the syllabification rules are applied then we get the units as ***nan-na-ya*** which are the actual units that are uttered. In this case ***an*** is part of triphone. If we consider the 'C-V' biphones like ***la, na, ra, ni, ka*** exist with high frequency as syllables and 'V-C' biphones like ***an, am, in, ar, al*** occur with less frequency as syllables despite of high frequency as a biphone. From this we can conclude that a C followed by V will have a higher percentage of usage as a syllable unit than V followed by C. It is also observed that VCs are valid when they exist as CVC.

6 Biphone Syllable Coverage Analysis

Telugu text corpus of 5 million words which contains sentences that are phonetically rich and various articles related to sociology, history, poetry, and many other areas. All these words are converted into KNS notation and are syllabified according to the syllabification rules. For coverage analysis vowels and biphone syllable units are taken and the number of words in the total corpus covered is given in Table 5 and its corresponding graph is given in Fig. 1.

Table 5 Percentage of words covered

% of speech units	% of words covered
10	8.36
20	17.77
30	22.26
40	24.25
50	**25.27**
60	25.68
70	25.84
80	25.85
90	25.87
100	25.93

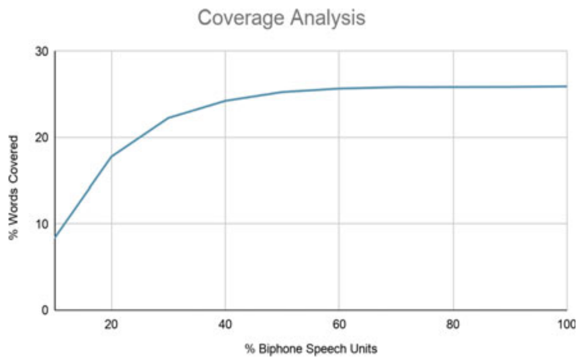


Fig. 1 Graph showing the percentage of words covered

Table 5 supports the observations made regarding the coverage of speech corpus using the minimum biphone syllable units. The observations are listed below.

- The set of vowels and biphone syllables covers 25.93% of the total corpus.
- Considering only the top 50% of the biphone units along with vowels is covering 25.27% of the total corpus which is kept in bold in Table 5.
- The biphone which are below the top 50% shows very less improvement in the coverage of corpus. So it is proved that the top 50% of the biphones are contributing more for coverage of words.

7 Conclusions

Speech databases required for a TTS cannot be constructed directly by storing the units from the text. As some of the units cannot be uttered or spoken as independent units they have to be segmented from continuous speech. Corpus analysis is carried out for extracting the required biphone syllable units. We have observed that some biphone sequences do not occur at all and few occur with very high frequency. Out of 2304 biphone sequences only 737 biphones can exist independently as biphone syllable units that occur in Telugu language. The size of speech corpus is minimized to a large extent by discarding the biphone syllables which do not exist as syllable unit. By the experimental results we can say that only 50% of the high frequency biphone syllables are covering 25% of the corpus.

References

1. Huang X, Acero A, Hon HW (2001) Spoken languages processing a guide to theory, algorithm and system development. Prentice Hall, New Jersy
2. O'Shaughnessy D, Speech communications human & machine
3. Jurafsky D, Martin JM, Speech and language processing
4. Kishore SP, Sangal R, Srinivas M (2002) Building Hindi and Telugu voices using Festvox. In: Sangal R, Bendre SM (eds) Proceedings of ICON 2002. LTRC, IIIT, Hyderabad, pp 303–309
5. Kishore SP, Kumar R, Sangal R (2002) A data driven speech synthesis approach for Indian languages using syllables as basic unit. In: Proceedings of ICON 2002. Mumbai, India, pages 311–316
6. Sunitha KVN, Sharada A (2013) KNS phoneme set—a new Telugu phoneme set for Telugu speech processing technology. ICRTC 2013, 4–5 Oct 2013, SRM University, NCR Campus. ISBN No.:978-93-80965-65-9
7. Sunitha KVN, Sunitha Devi P (2013) Text normalization for Telugu text to speech synthesis. Proc Int J Comput Technol, 2241–2249
8. Sunitha KVN, Kalyani N, Sreekanth N (2013) Minimum data set based on syllable position for Telugu speech systems. Proc Int J Adv Res Comput Sci Softw Eng IJARCSSE 3(10)

A Cognification Approach to Measure Contamination in Lakes

N. J. Anasuya, B. Ajith, B. Kiran Kumar, B. S. Medha, and S. Meghana

Abstract The paper mainly focuses on determining the amount of pollutants on a periodic basis. This helps to monitor and control the contamination. Any individual who needs to learn about the water purity is benefitted by this application. The method which exists in the industry for measuring the pollutants are laboratories. The existing methods are static. Hence people have to travel with the test sample to the laboratories. But the advantage of the proposed system is that it is portable. The device is portable to any place and the results will be available in the mobile application. This reduces the time taken to obtain results as well as the time taken for travelling. The proposed system is easier to use. It is cheaper and smaller in size. Hence portable and cost effective.

Keywords Machine learning algorithm · Water conditions · Classification · Analyzing

1 Introduction

For more than a decade, online water quality monitoring has been a serious topic since it is a serious issue related to environmental pollution. In the 21st century, there were lots of inventions, but at the same time were pollutions, global warming and so on are being formed. All life is dependent on water and exists in nature in many forms like ocean, river, lake, clouds, rain, snow and fog etc. In real time, water quality monitoring faces challenges and some of the reasons include global warming, scarcity of water etc. Therefore, better methodologies have to be implemented for measuring parameters in water quality monitoring. One of the main parameter is pH which measures the concentration of hydrogen ions. It shows the water is acidic or alkaline.

Water purity can be measured using characteristics like pH and turbidity. The pH value for pure water is 7pH. Less than 7pH indicates that water is acidic and more than 7pH indicates water is alkaline. So the range lies between 0 and 14 ph. Drinking

N. J. Anasuya (✉) · B. Ajith · B. Kiran Kumar · B. S. Medha · S. Meghana
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_81

water should have a pH value of 6.5–8.5. The measure of suspended particles also determines the quality of water. This is known as measuring turbidity.

After determining turbidity, the value can be used to identify disease causing factors. If the turbidity value of water is high, then the chances of being affected by water borne diseases are high. Different sensors are used to measure different criterial. To measure if the water is hot or cold, we use temperatures sensors. To measure the flow of water, we use flow sensors.

The water contamination is measured dynamically using IoT. The different lake parameters like pH, Temperature and Turbidity are measured using sensors. The collected sensor data is then moved onto the Arduino uno board for processing and then moved onto the cloud platform, through a iot module. Finally, the lake parameters are presented and displayed to the end-user in a graphical representation along with the current location of the module through a mobile application. The user can view the current data instantly by refreshing the application and the history of the previous data is also available in the application. When a water body is tested, it should be maintained on a regular basis to reduce the contamination level. The benefit of monitoring the water body qualities involves controlling the nature's pollution and finding solution for pollution. Water is the elixir of life. Out of 0.3% of usage water on the earth, more than half has already being contaminated. Water is contaminated by discharging sewage, garbage and liquid wastes of household, agricultural lands and factories into lakes and rivers. The usage of contaminated water in the daily life results in health hazardous. Water must be made available and contamination free for any household usage or for personal purposes. Water pollution control not only helps individual but also contributes for country's economy. Here the big need is to increase the quality of water and alert immediately as soon as the pollution is detected, in order to prevent further damage. So this system shall help in preventing the large quantum of damage of water body. The paper aims at developing a control system, in order to provide a technical solution for the above problem.

2 Related Work

"Remote Water Quality Monitoring System Using Wireless Sensors"

This paper proposes an architecture of water quality monitoring system. This system has a GSM module which helps in monitoring the water quality remotely. The traditional methods used for testing involves the staff to visit the water bodies, collect the samples and then return to laboratory. Hence, this method is time consuming. This GSM system also has features to send alerts or messages, when the quality of water decreases [1].

"SmartCoast Project—Smart Water Quality Monitoring System"

The project of the SmartCoast Project (a partnership between Dublin City University (DCU), the Marine Institute, Tyndall National Institute (TNI).This system was built to meet the requirements of WFD because it didn't have such facility before [2].

"Water Quality Monitoring System Using Zigbee Based Wireless Sensor Network"

This project uses wireless sensor network (WSN). Hence the system uses several sensor nodes with wireless networking for continuous purposes. Basic parameters of water quality like temperature, pH and turbidity are sensed from the respective sensors and are sent to the storage device or system. The system uses Zigbee based technology with IEEE 802.15.4 compatible transceiver [3].

"The Water Quality Monitoring System Based on WSN"

The particular project uses Wireless Sensor Networks (WSN) and consists of a remote data center.. The model collects the data and sends it to internet using GPRS DTU [4].

"Jayti Bhatt, Jignesh Patoliya."

To make sure that the water being supplied is safe, we need to monitor the water on a daily basis [5].

3 Algorithms

1 **Logistic Regression**

Logistic regression is one of the supervised classification algorithms. Generally, in a classification problem, the output value 'y', also known as target variable can take only discrete values for given input 'X', which is also known as set of features. Opposite to general opinion, Logistic Regression is a regression model. The model builds a regression model to predict the probability that a given data entry belongs to the particular category. Hypothesis function for Logistic Regression:

$$\frac{p}{1-p} = \exp(b_0 + b_1 x) \tag{1}$$

While training the model we are given:

x: input training data.
p: probability.
b0: constant.
b1: coefficient of x and defines the steepness of the curve

Logistic regression plots graph to predict the probability value p for given input x. The best fit for the curve is obtained by finding bo and b1 values. When the model is finally built and used for prediction, it predicts output value(y) for a given input(x).

2. **K-Nearest Neighbor**

One of the very common supervised Machine Learning algorithm is K-Nearest Neighbor (KNN) algorithm. This algorithm can be used for both regression prediction and classification problems.

The following steps indicate KNN's working:

Step 1 Algorithms works on data sets. Hence we must first load the data sets. These data sets are divided into training set and test set.

Step 2 In the next step, we choose the number of neighbors (K). This can be any integer.

Step 3 For each point in the test data do the following

- 3.1 We need to calculate the distance between each data point and its neighbors. For calculating distance, we use any one of the following methods namely: Euclidean, Manhattan or Hamming. But most commonly we use Euclidean distance.
- 3.2 Now, based on the distance value, sort them in ascending order.
- 3.3 In the ascending ordered list, it will choose the top K rows.
- 3.4 It identifies the most frequent class and assigns the data points to the respective class

Step 4 End.

3. **Support Vector Machine**

Another supervised learning methods include Support Vector Machine (SVM). The support vector machines can be used for outliers detection, classification as well as regression.

The advantages of support vector machines are:

- SVMs are highly effective in high dimension spaces.
- But it is also effective in situations where in the number of samples is less
- than the number of dimensions.
- From the training points, a subset is created and is used in the decision function.
- These are called support vectors. Hence it also becomes memory efficient.
- For the decision functions different kernels can be specified. Hence, it is versatile.
- By default, common kernels are provided. But user can use custom kernels.

The disadvantages of support vector machines include:

- If the number of samples are very much lesser than number of features, we need to avoid over-fitting in choosing Kernel functions and regularization term is also crucial.
- The probability estimates are not directly provided by SVMs instead these are
- calculated. They are calculated using a five-fold cross- validation.

However, to use an SVM to make predictions for sparse data, it must have been fit on such data.

4. **Gaussian Naïve Bayes**

The Naïve Bayes can be applied on to real-valued data attributes, but before applying we must assume a Gaussian distribution. Extending Naïve Bayes in this format is known as Gaussian Naive Bayes. Estimation of data can also be done using other functions, is the simplest way to work on because when using this distribution only standard deviation and mean of the training data has to be estimated. It is similar in calculating the standard deviation and mean values of all the input variable i.e., x for all the class values.

$$\text{mean(x)} = 1/\text{n} * \text{sum(x)} \tag{2}$$

where n denotes the instance number and x are the input values from training data.

Standard Deviation is calculated using the equation:

$$\text{standard deviation}(x) = \text{sqrt}(1/\text{n} * \text{sum}(\text{xi} - \text{mean}(\text{x})^2)) \tag{3}$$

The above equation the square root value of average squared difference. Here n denotes the number of instances, sqrt() returns the square root, sum() return the sum of the values, xi is the value of x at a specific instance i, mean(x) returns average value of all x instances and ˆ2 is nothing but the square.

5. **Decision Tree Classifier**

Decision Trees (DTs) are a non-parametric supervised learning method used for classification and regression. The main aim of decision tree is to build a model which predicts values based on simple rules.

Some advantages of decision trees are:

- Decision tree is visualized. Hence provides a clear understanding.
- Before using data on decision tree the data has to be prepared. If there are missing values, this method does not support it.
- The cost of using the tree (i.e.., predicting data) is logarithmic in the number of data points used to train the tree.
- Decision tree supports all types of data. It can use both categorical and numerical data unlike other algorithms. It can also handle problems that produce multiple-outputs.
- The methods and results in the decision tree are observable. It uses Boolean logic to explain the conditions of the algorithm. Where as in Artificial neural networks, it is difficult to analyze the condition.
- Possible to validate a model using statistical tests. That makes it possible to account for the reliability of the model.
- Performs well even if its assumptions are somewhat violated by the true model from which the data were generated.

6. **Random Forest Classifier**

Random Forest Classifier is formed by randomly selecting data from training set and building decision trees to it. To obtain the final result, it combines the output of all the decision trees.

Types of Random Forest models:

1. Random Forest Prediction for a classification problem:
   f(x) = majority vote of all predicted classes over B trees.

2. Random Forest Prediction for a regression problem:
   f(x) = sum of all sub-tree predictions divided over B trees

Advantages of Random Forest:

1. Random forest is one of the most accurate algorithms. It provides very efficient and precise results for models with large data set.
2. It provides accurate results for models with thousands of inputs without deleting any of the variable
3. It estimates the importance of variables and based on the estimation uses only the important variables for prediction.
4. As the building of the forest progresses, it generates an unbiased result.
5. If the data in the database is missing, then it has effective measures to estimate this data and also maintains accuracy of such large data.
6. **K-means Clustering**:

7. **K-means clustering,** is one of the simplest and popular unsupervised machine learning algorithms. The unsupervised learning always uses only input vectors and does not refer to outcomes or results.
   K-Means method is based on clusters

When the dataset is loaded into the model, the algorithm first randomly selects a group of centroids. To these centroids it assigns the initial data point. Then the algorithm performs iterative (repeated) steps to optimize the centroids.

The iterative step stops when either of these occur:

- When the centroid is stabilized.
- When the defined number of iteration are finished (Fig. 1).

8. **Hierarchical Clustering**

Hierarchical clustering, also known as *hierarchical cluster analysis,* is an algorithm that groups similar objects in to groups called *clusters.* All the objects in a cluster have similar properties whereas objects in different clusters have variable characteristics. To distinguish between two clusters, we look the cluster of endpoints.

In Hierarchical clustering, initially all the data is considered as separate clusters. Then the following two steps are executed continuously:

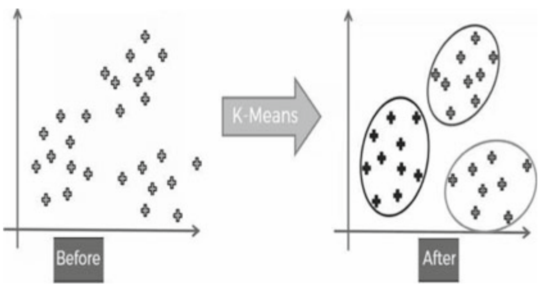


Fig. 1 K-means clustering

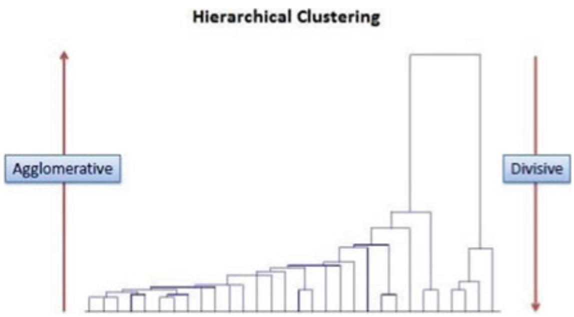


Fig. 2 Hierarchical clustering

1. Determines the closest cluster and check for the similarity
2. If two similar clusters are found, then merge the two clusters.

This process ends when all the clusters are merged together.

The main output of Hierarchical Clustering is a *dendrogram*, which shows the hierarchical relationship between the clusters (Fig. 2).

4 System Architecture

Data flow diagram is a representation of the data is used in the system. It is used as a basic step to view the working of the built model. The DFD's can be used for visualizing data processing since it shows what data is given as input and what is obtained as output (Fig. 3).

5 Results

1. **Logistic Regression result**
   Inputs: -all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output:-red dot indicates contaminated, green dot indicates not contaminated (Fig. 4).
   Result: According to trained model the given sample is determined as not contaminated with overall data.
2. **K-Nearest Neighbor result**
   Inputs:-all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output:-red dot indicates contaminated, green dot indicates not contaminated (Fig. 5).
   Result:-According to trained model the given sample is determined as not contaminated with overall data.
3. **Support Vector Machine result**

Fig. 3 Data flow diagram

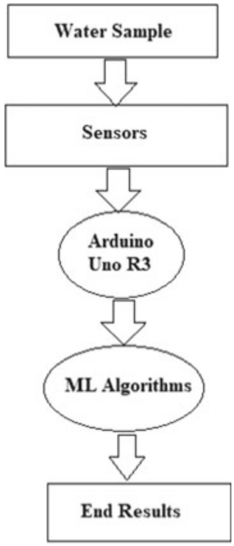
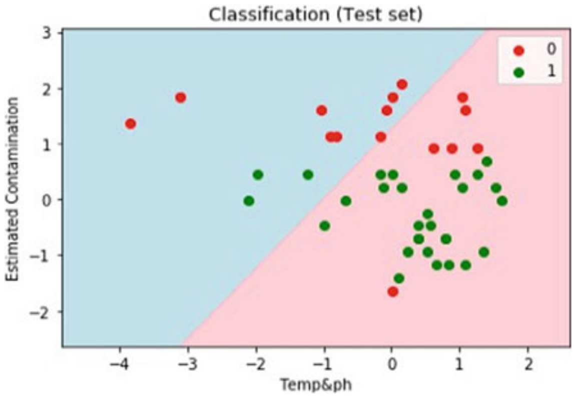


Fig. 4 Logistic regression result



Inputs:-all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output:-red dot indicates contaminated, green dot indicates not contaminated (Fig. 6).

Result:-According to trained model the given sample is determined as not contaminated with overall data.

4. **Gaussian Naïve Bayes Result**

   Inputs: all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output:-red dot indicates contaminated, green dot indicates not contaminated

   Result:-According to trained model the given sample is determined as not contaminated with overall data (Fig. 7).

Fig. 5 K-Nearest Neighbor Result

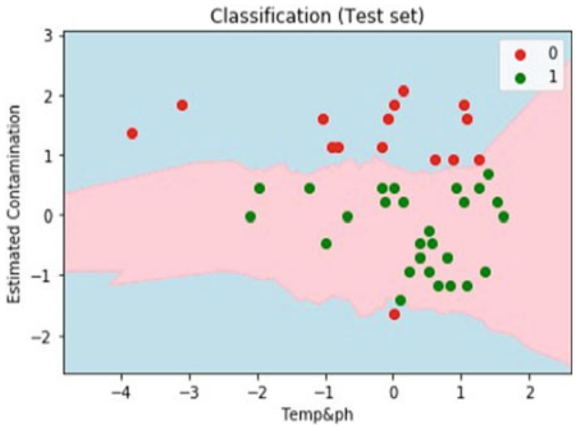


Fig. 6 Support vector machine result

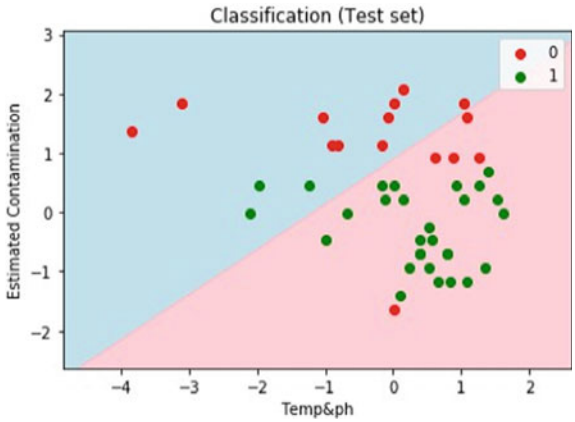
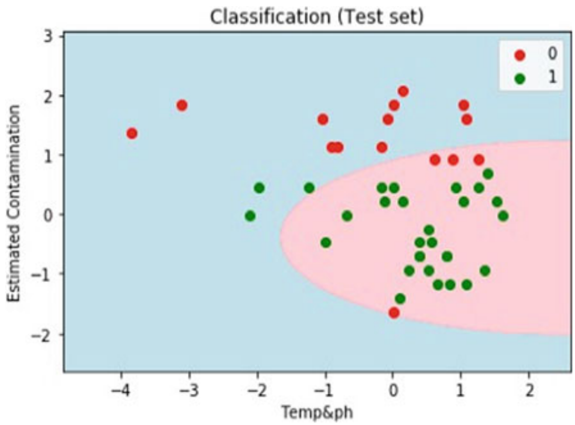


Fig. 7 Gaussian Naïve Bayes result



5. **Decision Tree Result**
   Inputs: all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output:-red dot indicates contaminated, green dot indicates not contaminated, Result: According to trained model the given sample is determined as not contaminated with overall data (Fig. 8).
6. **Random Forest result**
   Inputs: all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output: red dot indicates contaminated, green dot indicates not contaminated. Result: According to trained model the given sample is determined as not contaminated with overall data (Fig. 9).
7. **K-Means Clustering result**
   Inputs:-all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output: We obtain required cluster of similar sample Result: cluster 1: good, cluster 2: better, cluster 3: best, cluster 4: caution, cluster 5: Alert (contaminated) (Fig. 10).

Fig. 8 Decision tree result

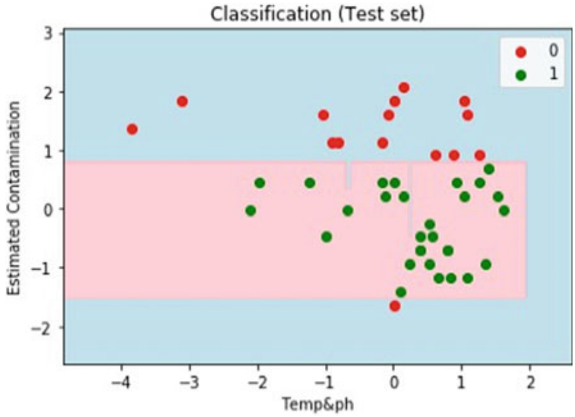
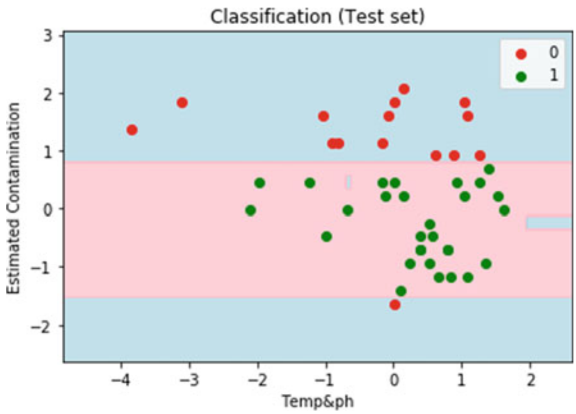


Fig. 9 Random forest result



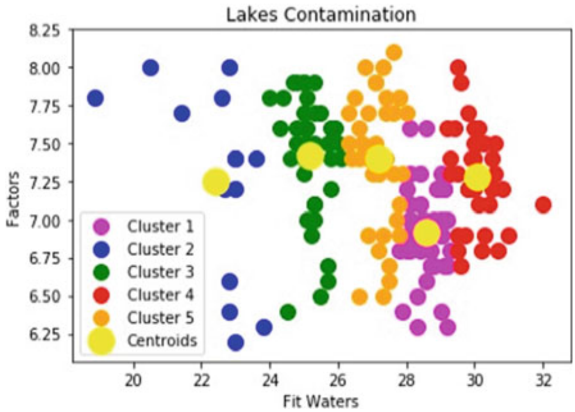


Fig. 10 K-means clustering result

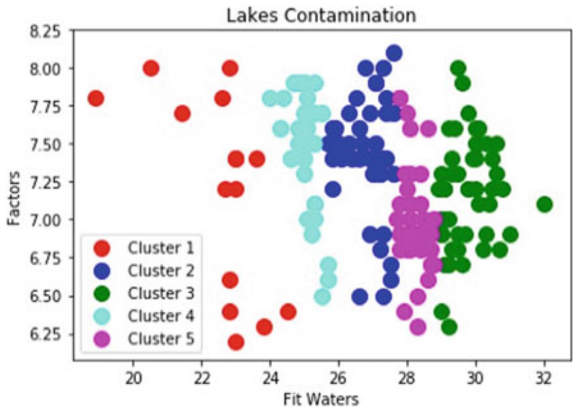


Fig. 11 Hierarchical clustering result

8. **Hierarchical Clustering result**
   Inputs: all X-axis is independent variables of dataset, Y -axis is dependent variable of dataset. Output: We obtain required cluster of similar sample. Result: cluster 1: good, cluster 2: better, cluster 3: best. Cluster 4: caution, cluster 5: Alert (contaminated) (Fig. 11).

6 Conclusion

Real time water quality monitoring system has been developed in this project. The model is very versatile and economical. The model finds different basic parameters of water and sends it to the central storing device. From the device the data is analyzed, processed and the result is sent to the user. The system is economical and low cost. It is portable hence has more advantage than the existing methods.

References

1. Haron NS, Mahamad MKB, Aziz IA, Mehat M (2008) A system architecture for water quality monitoring system using wired sensors. Computer and Information Science Department, University Technology PETRONAS Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan
2. Regan F, Lawlor A, McCarthy A (2009) SmartCoast project–smart water quality monitoring system. Environmental Protection, Agency, Synthesis Report
3. Rasin Z, Abdullah MR (2009) Water quality monitoring system using Zigbee based wireless sensor network. IJET-IJENS 9(10):14–18
4. He D, Zhang L-X (2012) The water quality monitoring system based on WSN. Institute of Mechanical and electronic information. China University of Geosciences (WuHan), WuHan, China, pp 3661–3664
5. Bhatt J, Patoliya J (2016) Iot based water quality. Monitoring system, IRFIC

Design and Implementation of Hybrid VM Migration Algorithm in Intra-cluster Environment

Abhay Deshapande, B. Sahana, K. R. Nataraj, and K. R. Rekha

Abstract Scenarios in which the number of users overflows in cloud data or management center, in order to trim down thrifty cost and guarantee better service quality the hardware management and virtual machine management becomes very challenging. Migration of VM is a preparation procedure for most of cloud management undertaking and thus achieves fault tolerance. VM migration releases the inherent hardware and manages the system from performance degradation. VM Migration plans are arranged from three points of view: Mode; Space; and Granularity. The investigations on live Migration are extensively overviewed dependent on the three fundamental difficulties it faces: cost of migration; number of migration check and down time. This paper proposes a hybrid algorithm which helps in finding shortest path based on calculating the cost/weight of the path. A working implementation of an algorithm calculates a sensibly best sequence of Virtual machine migrations from a given first position to a given final position. The hybrid algorithm developed enhances the result approximately by 26% with respect to Dijkstra algorithm. The cost function calculated by the hybrid algorithm is always lesser than that of traditional Dijkstra algorithm.

Keywords Cloud computing · Fault tolerant · Hybrid-algorithm · Virtual machine · Vm migration

A. Deshapande · B. Sahana (✉)
Department of Electronics and Communication Engineering, RVCE, Bangalore, India
e-mail: sahanab@rvce.edu.in

A. Deshapande
e-mail: abhaydeshapande@rvce.edu.in

K. R. Nataraj
Department of Electronics and Communication, Don Bosco Institute of Technology, Bengaluru, India

K. R. Rekha
Department of Electronics and Communication, SJBIT, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_82

1 Introduction

Migrating a virtual machine implies moving a virtual machine starting with one host or data store then onto the next host or data store. Virtualization innovation partitions a physical server into a few disengaged implementation situations by conveying a layer on hardware materials or Operating System [1]. Every implementation condition, i.e., Virtual Machine, autonomously runs with an OS and applications without shared obstruction on each other. Towards the beginning, virtualization advancement was not commonly used as a result of a variety of reasons. For example, it will have a fragment of equipment assets [2]. Virtualization is a key approving model of disseminated figuring that enables synchronous execution of resource over an hardware platform. Virtualization makes virtual condition on a solitary host by extracting information about the hardware, which licenses customers to utilize various occurrences of host to manage distinctive virtual machine on it [2]. The Virtual machines give versatility and increments efficiency by sharing com- putative assets.

In the current scenario, there are different types of VM migration techniques and are as shown in Fig. 1. It shows arrangement of migration plans. The colored box in (a) is to signify that VM is either shut down or suspended on the source host. The more sizeable voluminous line width for migration of memory data in (b) is to designate more immense system data transmission of LAN in correlation with WANS. The VMs in a numerous movement can be either independent or correlated.

(d) shows the migrated VMs which are correlated [1].

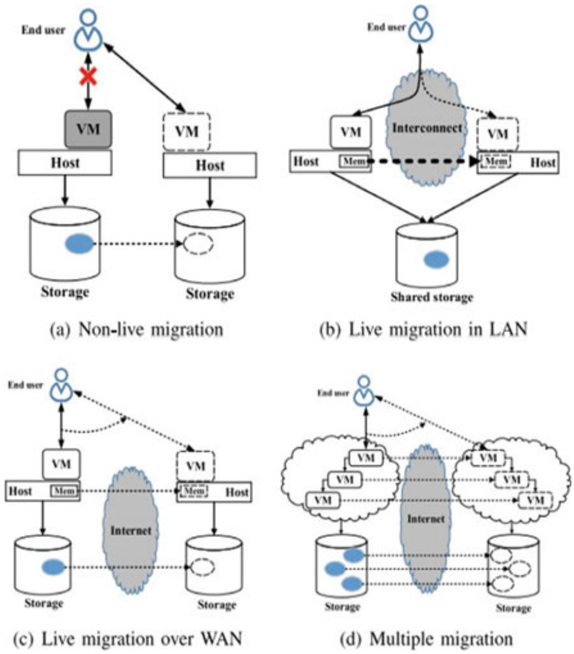
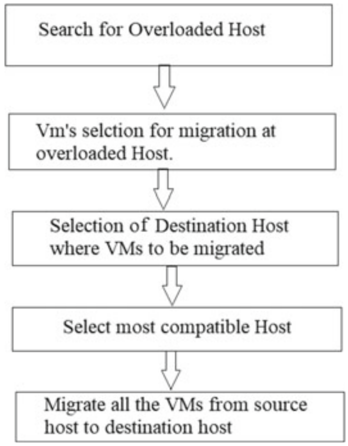


Fig. 1 Types of VM migration [1]

Fig. 2 VM consolidation framework [3]



VM Migration, a significant angle in consumerization of IT-with cloud innovation builds up arrangement of versatile services. It can achieve a couple of destinations, for instance, adaptation to real time server maintenance, green computing, load balancing, fault tolerance, energy efficiency etc.VM Migration empowers adaptable and productive allocation of resources. Diverse VM Migration Algorithms have improved utilization of assets in data centers. In Fig. 2 VM consolidation framework is shown.

The process of VM migration starts from identifying hosts which are already operating at full scale capacity or overloaded with new incoming tasks. Once the overloaded host is found next is to select the VMs at the source host to migrated from source to destination host. Now in order carry out VM migration find a most compatible i.e. host which has free resources such as RAM and ROM. Selection of under loaded host is carried out. Once this whole process is completed then migration VMs is taken place without any difficulties [3].

2 Related Work

Current developments in virtualization technology have transformed the way computing networks are designed and operated. Nonetheless, virtualization systems have a detrimental impact on the device output predictability, which poses many problems in managing efficiency and resource utilization [1]. In the initial stages virtualization was used in limited manner because of its hardware occupying characteristics [1]. Defines the required parameters to be considered during migration. The brief introduction to the hybrid algorithm concept is provided in [2]. Advantages and disadvantages of pre-copy and post-copy algorithm gives the basic structure for the migration of virtual machine from host to destination. This algorithm is used to prioritize the tasks by assigning weights to each virtual machine and choosing the virtual machine accordingly. With the help of selection policy of VM and selection strategy scheduling algorithm is presented [3]. More precisely, paper focuses on the reduction in the energy consumption and number of virtual machine migrations.

Various virtual machine scheduling methods are discussed [4]. It also discusses about the new set of evaluation metrics which is used for the evaluation of various scheduling methods which are built to predict the time distribution and prioritize the VMs for the allocation and reprogramming of resources. The performance of heterogeneous cloud computing tools has been enhanced through testing methods [4].

Kaur and Rani [5] Insights on the live virtual machine migration which takes more down time and migration time. Also paper focuses on researching and applying resource management and scheduling strategies in the cloud world, whereby big algorithms are implemented to improve the cloud environment's overall efficiency and presents the several VM migration techniques. Jabbehdari and Adabi [6] suggests a multi-objective genetic algorithm (GA) for the complex estimation of cloud data center resource use and energy usage. It also devises a multi-target resource management optimization issue that takes into account of the CPU and memory usage of VM, Physical Machine and data center energy intake. Similar to the historical data from previous time slots the proposed GA predicts the next time slot resource requirement [6]. Suggests a next-time VM positioning algorithm based on the GA prediction performance. The GA proposed is the ideal solution for resource estimation in simulation-based analysis under a secure, unpredictable utilization propensity is also discussed. A greedy algorithm is proposed to move a virtual machine from one location to another location by reducing total migration time is proposed. Shahapure and Jayarekha [7] in this paper based on distance and traffic an algorithm is proposed to reduce average Round Trip Tome (RTT). Addresses the utilization of soft computing with the better efficiency on cloud to automate and plan the tools. In [8] to minimize energy consumption and network cost a load balancing utilized VM migration algorithm is proposed. Liu [8] suggests that fewer number of migration required compared to other. In terms of reducing the average waiting period, total system flips and average turnaround time the experimental outcome offers improved results. Zakarya [10] an integrated prediction based and energy efficient virtual machine placement is proposed.it decreases degradation of performance which is caused because overload while decreasing the number of VMs [10]. Cloud Analysis is used to model the proposed process. Simulation Findings in [10] demonstrate that the suggested approach fits well for a real world image while occurring. Farahnakian et al. [11] a novel dynamic VM consolidation approach called Ant Colony Optimization (ACO) is used to reduce the energy consumption. Ant colony system (ACS) is used to obtain the optimal solution. The performance is evaluated by conducting experiments with number of workload in real time using ASC-VM [11] ACS-VMC is used as the extension for ACS-VM within the Open-Stack cloud platform. By reducing skew cost of each VM, we can effectively merge various multiple resources and boost resource usage of server. The proposed algorithms in [13] avoid overload by efficient load balancing and potential demand estimation and achieve maximize efficiency in terms of server resource usage with minimal energy consumption [14]. Introduction of SDF is made which decides the top-quality time to control the pre-copy stage to the post- copy stage. Use of Markov model- memory access pattern forecasting which decreases the number of invalid transfers. Compared with the original hybrid-copy algorithm, we can effectively reduce the page faults while achieving the same level of total migration time. By reducing the number of page faults this algorithm enhances the user experience [16]. A common behavior between the two policies is observed in the experiment, which allows the weights to be allocated according to the routing likelihood calculated by theatrical heuristic queue and literature optimization algorithms [16]. The interaction between probabilistic routing and weighted round robin load balancing policies will first be tested experimentally. Tsao et al. [17] is primarily centered on algorithms based on multi-class workload closed queueing networks that are appropriate for representing applications in diverse user-to-user service-standard agreements.

3 Proposed Work

The problem is modelled as a graph where each node represents an arrangement of VMs within the cloud (i.e. a mapping of VMs to VM hosts) and each edge represents a migration of a single VM from one host to another, and then use developed Hybrid Algorithm to find the shortest sequence of migrations which transforms the starting arrangement to the desired final arrangement.

The hybrid algorithm involves backtracking and depth first search. An optimal algorithm was introduced with the following aspects:

1. Make a list of each VM which isn't already in its final destination.
2. For each VM in the list, try to migrate it to its final destination.
3. If its final destination is already "full", e.g. doesn't have sufficient free RAM to accommodate the incoming VM, then displace one of the VMs already there in order to make space for the incoming VM.

Recursive path finding algorithm based around the concept of a TODO list containing which VMs have not yet reached their final destination host. The TODO list is iterated over until it is empty. If a VM cannot be immediately migrated to its final destination host, other VMs on that host are displaced in order to accommodate it. Displacement favors migrations already in the TODO list, to avoid accumulating unnecessary "migration debt".

Displacement happens recursively in a manner which travels down the dependency tree (i.e. if migration B is required before migration is A possible, then A is B's parent in the tree), until we reach a leaf node migration which can be performed immediately. Then its ancestor migrations can be performed in reverse order by backtracking up the tree again. Each migration enacted will result in an update of the TODO list.

Also need to lock VMs in place when exploring potential displacement paths, to exclude them from the list of candidates for displacement. This helps break mutual dependency cycles (e.g. swapping a pair of VMs between different hosts when they

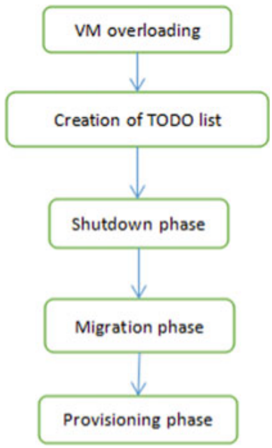


Fig. 3 Flow chart of hybrid algorithm implementation

cannot co-exist on the same host), which would otherwise cause infinitely deep recursion. Figure 3 depicts the flowchart of our hybrid algorithm.

It has three phases namely shutdown phase, migration phase and provisioning phase. In shutdown phase VMs which are not need to be migrated are shut off. Later in migration phase VMs which need to be migrated are migrated to its final destination freeing up certain VM hosts. Finally, in provisioning phase involves the process of preparing and equipping a network to allow it to provide new services to its users.

4 Results and Discussions

Totally 7 VM hosts and 12 VMs are created using random auto generative program in Python. The cost is mentioned in bracket which is directly proportional to the VM's RAM footprint. Since Dijkstra's algorithm immediately turned out to be completely useless due to the algorithmic complexity of exploring this graph since it runs in a infinite loop allowing discovery of multiple shortest paths in parallel as shown in Fig. 4. Each VM host is assigned some particular VMs into it such as VM01, VM02 and so on. Here cost is proportional to RAM of VMs as well. The Hybrid algorithm migration consists of 3 phases namely shutdown phase, migration phase and provisioning phase. In Fig. 4 consider VM01 in the above output, it is currently on host5. The Dijkstra algorithm of the above program only gives the shortest path. Dijkstra Algorithm checks with all the neighboring hosts and migrates to host4 which yields in cost function of 3576. This cost function is quite large.

In the shutdown phase VMs which need not needed for migration are shut downed. Later we enter into migration phase where VM migration. Recursive path finding algorithm based around the concept of a TODO list containing which VMs have not yet reached their final destination host. The TODO list is iterated over until it is

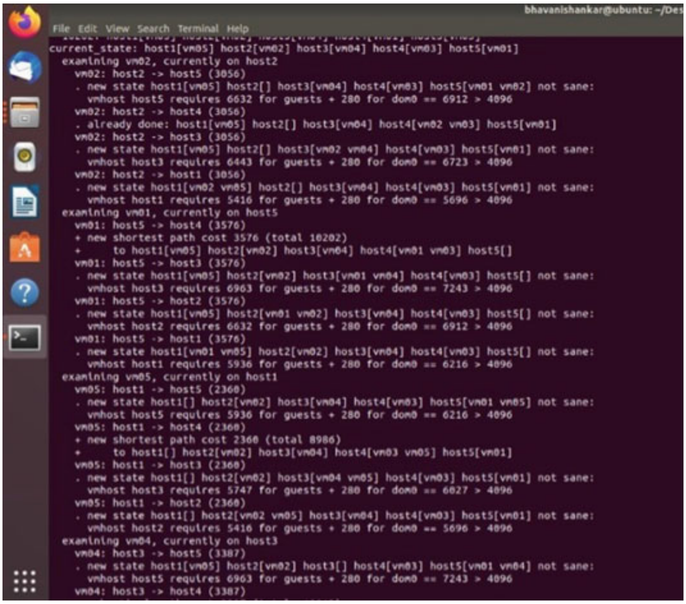


Fig. 4 Dijkstra's shortest path algorithm

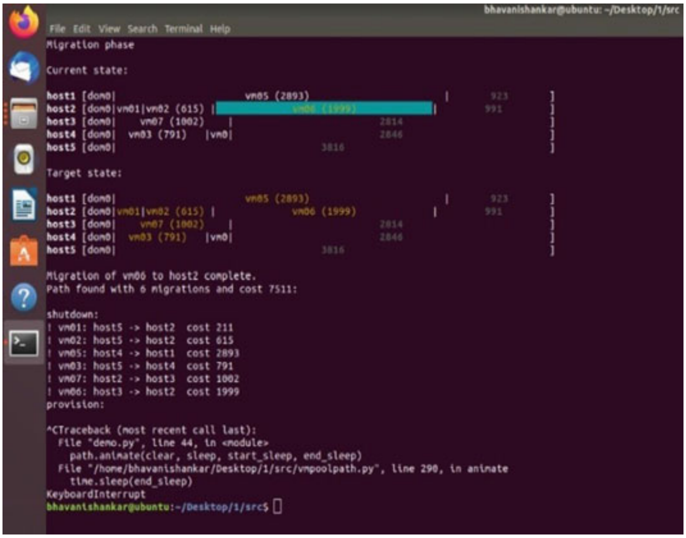


Fig. 5 Current and final destination host using hybrid migration algorithm

Fig. 6 Migration sequence from host to destination

```
Path found with 13 migrations and cost 4360:

shutdown: big6, small1
! big2: host2 -> host7  cost 510
! small5: host5 -> host2  cost 390
! small2: host2 -> host5  cost 360
! big2: host7 -> host2  cost 510
! big3: host3 -> host7  cost 520
! small4: host4 -> host3  cost 380
! small3: host3 -> host4  cost 370
! big3: host7 -> host3  cost 520
! small6: host6 -> host1  cost 400
! tiny1: host7 -> host1  cost 100
! tiny2: host7 -> host2  cost 100
! tiny3: host7 -> host3  cost 100
! tiny4: host7 -> host4  cost 100
provision: tiny5
```

empty. If a VM cannot be immediately migrated to its final destination host, other VMs on that host are displaced in order to accommodate it. In Fig. 5, we observe that VM01 currently on host5 migrates to host2 with a cost function of 211 which is very efficient against Dijkstra.

The hybrid algorithm showed 13 migration and total cost of 4360 as shown in Fig. 6, which is efficient when compared to Dijkstra's shortest path algorithm.

The hybrid algorithm for different architecture i.e. combination of different number of VM host and VMs is modelled. Figure 6 shows the current and final destination host where VMs need to be replaced. Each VM host is assigned some particular VMs into it such as big, small, tiny. Big refers to VM with high RAM usage and respectively small and tiny.

5 Conclusion and Future Scope

Dijkstra's algorithm could only be useful in finding shortest path for the VM migration. With proposed/Hybrid algorithm VM migration is performed with shortest path possible. The hybrid algorithm developed enhances the result approximately by 26%. The cost function calculated by the hybrid algorithm is always lesser than that of traditional Dijkstra algorithm. The migration path cost is reduced. The down time during the migration is reduced with use of hybrid algorithm.

Further algorithm can also be improved which gives less cost and minimum migration. The Rearrangement algorithm is much more simplified and efficient when compared to Dijkstra's. In future simple consolidation and re-balancing heuristics, and testing need to be designed such that it shows more effective and efficient migration and new debug level can be added which only shows partial path and fix slow test cases.

References

1. Zhang F, A survey on virtual migration: challenges, techniques and open issues. IEEE, Survey and communication tutorial. https://doi.org/10.1109/COMST.2018.2794881,2018
2. Naik PCM, Shakya DGA, A research paper on live VM migration approach in cloud computing. In: International conference on trends in electronics and information. IEEE. https://doi.org/10.1109/ICOEI.2018.8553741,2019
3. Wu X, Zenf Y, Lin G (2017) An energy efficient VM migration algorithm in data center. In: 16th international symposium on distributed computing and application to business, engineering and science
4. Li D, Wang W, Li Q, Cheng J (2017) A comprehensive evaluation of scheduling methods of virtual machine migration for energy conservation. IEEE, IEEE Syst J 11(2). https://doi.org/10.1109/JSYST.2015.2436930
5. Kaur PD, Rani A (2015) Virtual machine migration in cloud computing. Int J Grid Distrib Comput 8(5):337–342
6. Jabbehdari S, Adabi S (2014) Virtual machines migration based on greedy algorithm in cloud computing. 2014 IEEE Int J Comput Appl 96(12):0975–8887. https://doi.org/10.5120/16849-6709
7. Shahapure NH, Jayarekha P (2018) Distance and traffic based virtual machine migration for scalability in cloud computing. In: 2018, international conference on computational intelligence and data science, ICCIDS
8. Liu C (2015) A load balancing aware virtual machine live migration algorithm. In: 2015 4th international conference on sensors, measurement and intelligent materials, ICSMIM
9. Dabbagh M, Hamdaoui B (2018) An energy-efficient VM prediction and migration framework for overcommitted clouds. IEEE Trans Cloud Comput 6(4). https://doi.org/10.1109/TCC.2016.2564403
10. Zakarya M (2019) An extended energy-aware cost recovery approach for virtual machine migration. IEEE Syst J 13(2). https://doi.org/10.1109/JSYST.2018.2829890,2019
11. Farahnakian F, Ashraf A, Pahikkala T, Liljeberg P, Plosila J, Porres I, Tenhunen H (2018) Using ant colony system to consolidate VMs for green cloud computing. IEEE Trans Serv Comput 1. https://doi.org/10.1109/TSC.2014.2382555
12. Srinivas BV, Patil KS, Shruthi B, Mandal I (2018) A framework for efficient virtual machine migration. In: 2018 IEEE, international conference on current trends toward converging technologies
13. Wen W-T, Wang C-D, Wu D-S, Xie Y-Y (2015) An ACO- based scheduling strategy on load balancing in cloud computing environment. In: IEEE 2015 ninth international conference on frontier of computer science and technology. https://doi.org/10.1109/FCST.2015.41
14. Lei Z, Sun E, Chen S, Wu J, Shen W (2017) A novel hybrid-copy algorithm for live migration of virtual machine. School of Computer Engineering and Science, Shanghai University, Shanghai 200444
15. Xiaoqing Z (2017) Efficient and balanced virtualized resource allocation based on genetic algorithm in cloud. In: 2017 10th international symposium on computational intelligence and design (ISCID). IEEE https://doi.org/10.1109/iscid.2017.187,2017
16. Aldhalaan A, Menascé DA (2013) Analytic performance modeling and optimization of live VM migration. In: Proceedings of European workshop performance engineering Venice, Italy, pp 28–42
17. Tsao P-J, Sun Y-f, Chen L-H, Cho C.-y (2016) Efficient virtualization—based fault tolerance. In: International computer symposium

Towards Many to Many Communication Among Blind, Deaf and Dumb Users

A. S. Chaithra, Umme Athiya, R. Aishwarya, and Aswathi Rajesh

Abstract Humans are favorably quick-witted humanoids that have become the presiding classification on the Earth. The mortal's divulgence is grounded in communal and shared objectives. While people with hearing or visual impairments alone can find a way to share their thoughts with others and understand them, deaf blind people face a much more difficult communication task. Thus the project accords and executes the blueprint design, paradigm and testing of a lightweight software and speaker device with a display for the divulgence among two humanoid or also between visually impaired people.

Keywords gTTS · OCR (Optical character recognition) · Pytesseract · pyttsx3 · Googletrans · Contour extraction · Thresholding · Background elimination · Feature extraction · LSTM · Enchant · Convex hulk · Acoustic modeling · Language modeling

1 Introduction

A present-day exemplar is the rise of divulgence automation, which has lessened barriers to the mortal's interaction and as a result has helped many. We live in a digital era with advancement in information and communication technology where everything is available in a fraction of seconds. Humans are social creatures. While people with hearing or visual impairments alone can find a way to share their thoughts with others and understand them, deaf blind people face a much more difficult communication task. Thus solving the problems of people with Visual, Hearing and Vocal Impairment through a single supporting system is a tedious venture.

A. S. Chaithra (✉) · U. Athiya · R. Aishwarya · A. Rajesh
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_83

2 Problem Statement

2.1 Existing System

In the earlier days, blind people could only read only with the help of Braille script. Braille is a type of scripting language where blind people read through the 6 raised dots sensed by their fingers. It is traditionally written with embossed paper. Now a days Braille user can read computer screens and other electronics support using refreshable Braille displays. On a conventional basis, the gesture sign recognition method was aggregated into two types namely vision based and sensor based method. In vision based method, the input aided for this is capturing the image from the computer camera to analyze the position of the fingers. In sensor based systems, it requires having an aided device helps in gaining the required correct finger positions and associated angles between them. The disadvantage is wearing it consistently is not feasible.

2.2 Proposed System

The work focusses on developing "intelligent sign language recognition using image processing" which deals with the computer system in which sign language is captured and processed and translated to speech. The proposed system makes a clear-cut view on providing a solution for the people with deafened and dumb-sighted person to envision/ study which is in audio format by speech to text conversion process and we also provides a way for the dim-sighted person to represent their input/ conversation by the aid of text to voice process.

3 System Design

The idea for implementing this system design is to bridge the gap between the visually impaired person and a normal person for ease of communication between them. The proposed project is built with the most trending programming language also used in major industries ~ python. The system is provided with 4 unique and individual modules.

- **Text to speech (TTS) Text-to-speech** technology reads aloud digital text. It can take words on computers, smartphones and convert them into audio. gTTS is the commonly used API and accurate results are obtained. It aids in the conversion of a text entered or provided as an input by user into audio which can be saved as an mp3 file and reads it aloud for them (Fig. 1).

Fig. 1 Text to speech

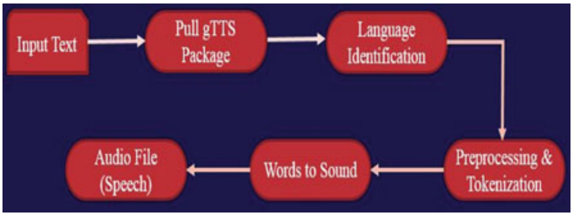
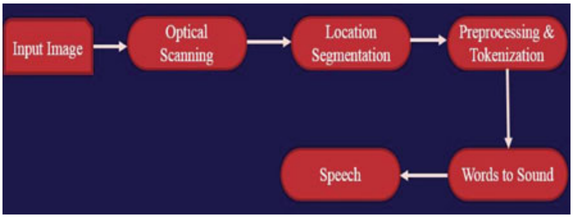
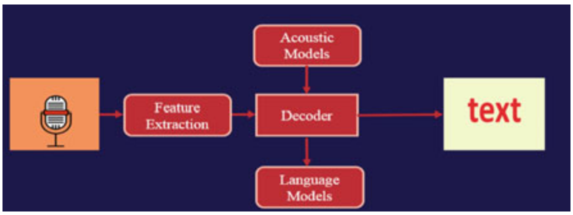


Fig. 2 Text to speech using camera



- **Text to speech using camera (OCR) Text to Speech using Camera** is performed by a source called Optical Character Recognition. OCR is used to scan the image and extract the available text from the image and perform background color and disturbances causing in the image. Thus after the final and successful recognition of the text from image it produces a sound speech using the gTTS module (Fig. 2).
- **Speech to text (STT)** Speech Recognition is based on the algorithm of acoustic and language modeling. Acoustic modeling represents the relationship between speech and audio signals. Language modeling suits resonates with term series. It captures the voices either from the audio file or the microphone or upon successful completion of the recognition module, produces the voice into its corresponding text (Fig. 3).
- Hand Gesture Recognition Existing methodology of finger tracking and outline perception for gesture sign recognition using OpenCV. Gestures play a vital role for exchange information among the human beings. Hand gesture and sign recognition executes ample enough with the help of Machine learning methods such as neural networks, support vector machine, and Adaptive Boosting (AdaBoost) (Fig. 4).

Fig. 3 Speech to text



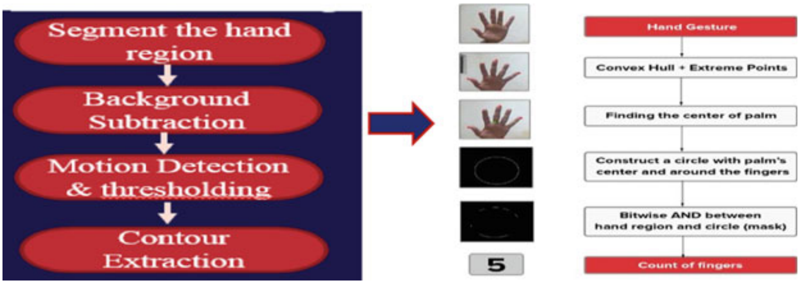
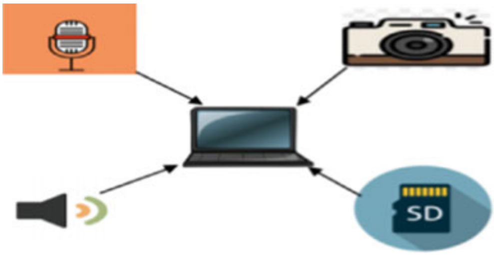


Fig. 4 Hand gesture recognition

Fig. 5 Block diagram of the project



4 High Level Design

Figure 5, represents the block diagram of the project which takes input from the devices like Microphone and Camera, produces output through console, speaker and SD card used to store all the user's activities.

5 Results

The proposed system also hands the role of transformation of Text to Speech and Image to Speech via camera for unsighted, Speech to Text for deafened users, and Hand Gesture Recognition for deaf and mute Person. All these modules are available for single system general purpose as well as for complete visually impaired person. Below are the comparison of algorithms results and their accuracy applied for Hand Gesture Recognition model (Figs. 6, 7, 8, 9, 10, 11 and 12).

Fig. 6 Results of comparison of algorithms

Algorithms	Train Results	Test Results
1. Decision Tree	95.43%	93.03%
2. Random Forest	100%	99.892%
3. SGD Classifier	72.00%	73.14%
4. Long Short Term Memory	99.887%	98.208%
5. Tensor flow Application Image Extraction.	98.34%	98.04%

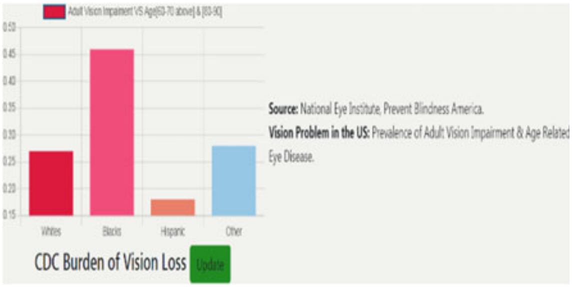


Fig. 7 CDC burden of vision loss

Fig. 8 Age specific prevalence

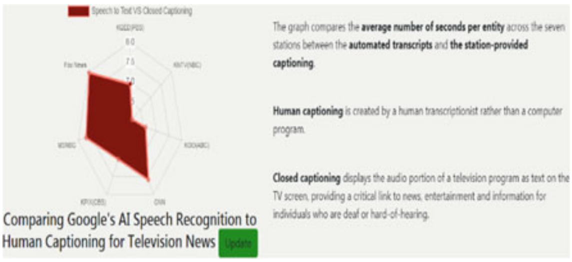


Fig. 9 Google's AI SR versus human captioning

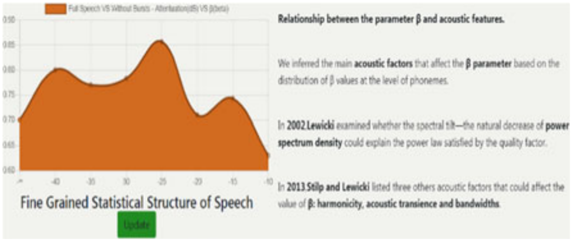


Fig. 10 Fine grained statistical structure of speech

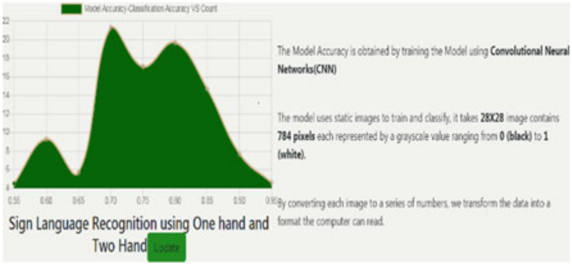


Fig. 11 SL recognition—one versus two hand

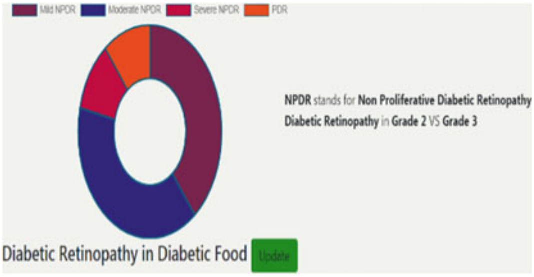


Fig. 12 Diabetic retinopathy in food

6 Conclusion

This project aims to lower the communication gap between blind, deaf or mute community and normal world, help them to lead a standard lifestyle. The design prototype is used to convert text/image to voice for blind, speech to text conversion for deaf and conversion of hand gestures to text for dumb people. We have presented the sketch and archetype model for unsighted, deafened and mute people (or integrated even as for visually impaired). The advantage of this is that it can be easily carried due to less weight. The functionality developed can be used as a smart assistant for differently abled people to communicate with others and it is a language independent system.

Acknowledgements The development of the proposed system was extremely supported by Don Bosco Institute of Technology. We present gratitude to our guide Mrs. Chaithra A S, Assistant Professor, Department of Information Science & Engineering, DBIT co-heartedly provided extra support to make this project successful with team members—Aishwarya R, Aswathi Rajesh & Umme Athiya. We thank the department and management to provide the idea in coming up with this system and extremity in resources.

Bibliography

1. Isewon I (2014) Design and implementation of text to speech conversion for visually impaired people. Int J Appl Inf Syst (IJAIS)
2. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.670.7344&rep=rep1&type=pdf-Feasible camera-based functional & supporting text and product tag studying from held things for unsighted users
3. http://www.ijirset.com/upload/2014/icets/131_EC343.pdf - a paper
4. A novel perspective for divulgence among unsighted, deafened and mute users

The Effective Approach to Content-Based Image Retrieval Based on ANN Classifier

D. K. Yashaswini and K. Karibasappa

Abstract Due to the existence of vast quantities of digital data, a program must be established which can retrieve similar data from the application. In this proposed system content-based retrieval method to get similar data based on the query image data is added. The proposed method is split into three stages, preprocessing, extraction of the element and classification. The preprocessed images are considered for the characteristics of extraction, form, color, and texture dependent features. CCM and BPF are removed and they train the machine. ANN is used as a classification to classify the system. Unit efficiency is measured using the efficiency parameters and tests indicate the improved result.

Keywords Content-based image retrieval (CBIR) · Color co-occurrence matrix (CCM) · Bit pattern features (BPF) · Artificial neural networks (ANN) · Recall · Precession · Accuracy

1 Introduction

The imaging equipment has been growing very rapidly day by day which are considered digital cameras, smartphones and imaging devices, due to the growth of technology in quality, availability and images become important. It is necessary to store the image in an efficient way. Due to a large number of image data available in all fields, the Creation of an efficient image recovery service system is critical. Content-Based Image Retrieval (CBIR) technique is developed to efficiently address an image's rendering and storage operation. CBIR is useful in many applications of the modern world, such as medicine, education etc. [1].

In the CBIR, the main features considered to retrieve the image content are based on the color descriptor, texture descriptor and shape descriptor to develop a standard that can be fallowed to CBIR. In the research area of CBIR, where many important aspects are fair easiness to conduct, for example sharing image functions for the benchmark database, comparative analyzes between several CBIR tasks,

D. K. Yashaswini (✉) · K. Karibasappa
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_84

(a) (b) (c)

Fig. 1 Content based on image retrieval

etc. The figure demonstrates the image retrieval dependent on contents. The norm also provides a major advantage in the distributed environment, where the user can remotely change the image content descriptor. In this case the original picture is not automatically moved to various places; for adjustment and recalculation the image descriptor is only necessary (Fig. 1).

2 Literature Survey

Xia [2] has presented a CBIR over encrypted techniques. The encrypted data is highly secured it does not allow the content to share in the cloud. The considered input image features are considered and extracted as a feature vector. By making use of the pre-filter table, search efficiency is increased. The extracted feature vector is secured using KNN. The image pixels are encrypted using the cipher. A watermarked image secured system is also introduced to provide the content to the only illegal or authorized person. In the watermark a unique mark as inserted into encrypted data by the cloud server before it is passed to the querying user. Therefore, when an unauthorized copy of the image is identified, the unlawful query user who transmitted the image may be tracked by the extraction of the watermark. The safety study and tests show the safety and effectiveness of the proposed scheme.

Guo [3] proposed a method for CBIR by using the advantage of low-complex ordered-dither block truncation coding (ODBTC) to construct a descriptor image output. First, the image is encoded, the image is compressed into blocks using the ODBTC form, and the image is quantized. Using the BTC encoded image, two features extraction is done, color co-occurrence feature (CCF) and bit pattern features (BPF). The features of the image are extracted without encoding.

Xu [4] has presented a graph-based model; the graph-based model is widely applied in the field of retrieving the image data. This method is efficiently applied to CBIR. This has an exceptional capacity to detect the underlying image database geometric structure. The new, efficient Manifold Ranking (EMR) model, which is graphically based, is presented for a large data set.

ElAlami [5] presented a CBIR-based, accurate and fast model. The method consists of four stages, extraction of the function, reduction of dimensionality, classification of ANN and a matching strategy. Functions depending on color and texture are removed in this framework. The techniques of extracted features are the color co-occurrence matrix (CCM) and the discrepancy between scan pattern pixels (DBPSP). Extraction can resolve the single frame constraint by using the two features, but due to the high dimensionality of the screen space, the time may consume more. The ANN is used as a category for the selected functions of the query image to be one of the most high SIMI multi-groups.

Gode [6] presented a CBIR focused on the features of color, texture, and form. The picture is broken down into equivalent block amounts. In RGB color space, the features derived from the local colour, shape, and texture descriptors act as conditional histograms between the image tiles and the corresponding additional tiles. The CCM is used to extract texture characteristics using the edge detection technique, canny method and invariant shape moments that are applied to get the shape-based function. The experimental results show how effective the methods are.

3 Methodology

Followed steps for this proposed method is as follow:

- The database is created initially from different available sources for the image to be retrieved from. Pre-processing is done for available images in the database and query image. Pre-processing includes resizing and color conversion. Resizing fixes, the size of all the images in the database and query image to one fixed size which helps for the further process of the model. Color conversion is needed to convert three plane images into two plane images to reduce complexity. It includes gray conversion, binary conversion. Overall pre-processing deteriorates the distortion and enhances the features of the image
- After the pre-processing Next step is to extract images foreground for further processing. Foreground extractions include the detection of an object in an image that is of interest.
- Once the foreground extraction is done feature extraction of an image is performed. In this proposed method color, texture, and shape feature extraction are done. This extracted feature helps in further steps for image matching.
- Some classifiers are used to equate the query image feature with image features in the database to fit the query image in the database. These classifiers are trained with an available set of features to check the best match for query images in the database.
- Once the classifiers are trained well, similarity matching is performed using these classifiers and the conclusion is done with the best matching for a query image from the database.

Different sets of algorithms are available to make use of, for the accomplishment of this method. Algorithm efficiency can be enhanced by accumulating more than one technique in the process shown in Fig. 2.

Color co-occurrence matrix (CCM): In the RGB image (Color image) most of the information is consists of the color distribution of the pixel. The features patterns of the color image are extracted from the distributed color pixel of an image. The extracted features matrix is used to calculate the pixel probability with its neighbor's neighboring pixel. By considering the feature matrix and its neighbor's pixel information specific color information is constructed. This information matrix also gives us information about the image in the spatial domain too. Figure 3 shows the Computations of CCM Features Block Diagram

Co-occurrence color function (CCF): CCF derives from the CCM. The minimum and maximum amounts of color shall be indexed first using a particular color codebook. In general the color indexing process on RGB is specified as a triple RGB

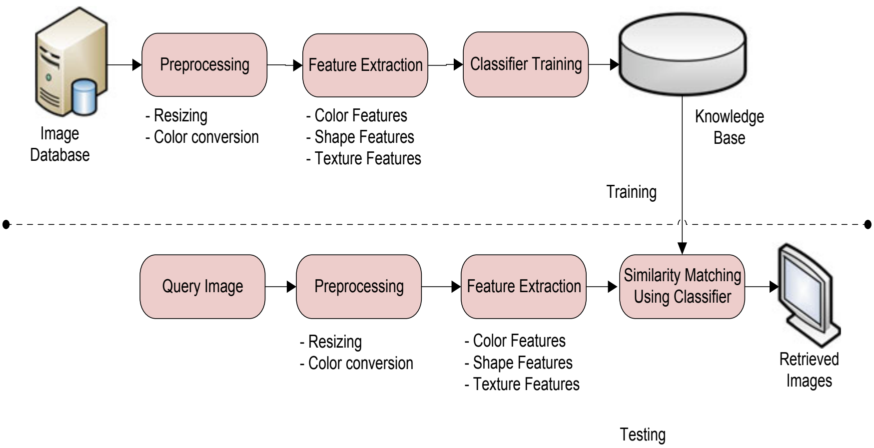


Fig. 2 Depicts the block diagram for CBIR

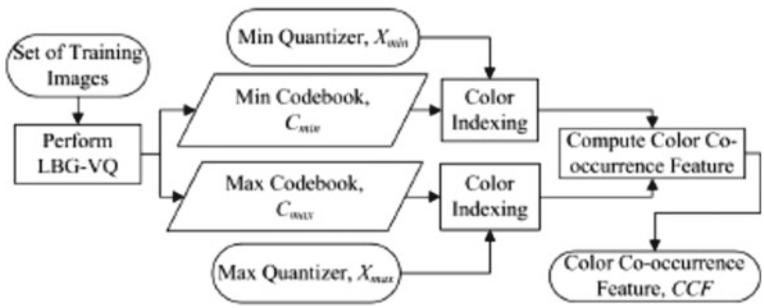


Fig. 3 Computations of CCM features block diagram

pixel mapping in an index codebook (the most representative codeword) sub-set (a single tuple).

LBGVQ produces a succinct codebook from multiple training vectors. Let

$$\mathrm{C}_{\min} = \left\{\mathrm{C}_1^{\min}, \mathrm{C}_2^{\min}, \ldots, \mathrm{C}_{\mathrm{N_c}}^{\min}\right\} \quad (1)$$

$$C_{max} = \left\{C_1^{max}, C_2^{max}, \ldots, C_{N_c}^{max}\right\} \quad (2)$$

Xmin and Xmax, respectively, are created from minimum quantizer and maximum quantizer codebooks. In $\mathrm{C}_i^{\min}$, $\mathrm{C}_i^{\max}$ and N c, respectively, denote the codeword from the minimum quantizer, maximum quantizer, and color codebook scale. The method of color indexing can be called the closest match between the minimum quantizer values of each block of images

$\mathrm{x}_{\min}(\mathrm{i}, \mathrm{j})$ and the codebook $\mathrm{C}_{\min}$ which meets the following condition.

$$\mathrm{i}_{\min}(\mathrm{i}, \mathrm{j}) = \underset{\mathrm{q}=1,2,\ldots,\mathrm{N_c}}{\arg\min} \left\| \mathrm{x}_{\min}(\mathrm{i}, \mathrm{j}), \mathrm{c}_{\mathrm{q}}^{\min} \right\|_2^2 \quad (3)$$

For all $\mathrm{i} = 1, 2, \ldots, \frac{\mathrm{M}}{\mathrm{m}}$ and $\mathrm{j} = 1, 2, \ldots, \frac{\mathrm{N}}{\mathrm{n}}$. Similarly, the indexing for the maximum quantizer of each image block $\mathrm{x}_{\min}(\mathrm{i}, \mathrm{j})$ with codebook $\mathrm{C}_{\max}$ is formally defined as

$$\mathrm{i}_{\max}(\mathrm{i}, \mathrm{j}) = \underset{\mathrm{q}=1,2,\ldots,\mathrm{N_c}}{\arg\min} \left\| \mathrm{x}_{\max}(\mathrm{i}, \mathrm{j}), \mathrm{c}_{\mathrm{q}}^{\max} \right\|_2^2 \quad (4)$$

For all $\mathrm{i} = 1, 2, \ldots, \frac{\mathrm{M}}{\mathrm{m}}$ and $\mathrm{j} = 1, 2, \ldots, \frac{\mathrm{N}}{\mathrm{n}}$. The color co-occurrence matrix (i.e., CCF) can be determined directly for a given picture when colors are indicated for minimum and maximum quantizers computed as.

$$CCF(t_1, t_2) = Pr\left\{ i_{min}(i, j) = t_1, i_{min}(i, j) = t_2 | i = 1, 2, \ldots, \frac{M}{m}; j = 1, 2, \ldots, \frac{N}{n} \right\}, \quad (5)$$

For $\mathrm{t}_1, \mathrm{t}_2 = 1, 2, \ldots, \mathrm{N}_{\mathrm{c}.}$.

The CCM is a sparse matrix where its entries are dominated by the zeros. In its columns and lines the CCM can be binned into a 1D image descriptor that enhances the dimension of the CCF and speed up the process of recovery of images. Thus, the dimensionality function of the CCF is Nc, that is, exactly similar to the color codebook scale. Figure 4 gives an example of CCF computation as shown, the calculation is quick and makes it simpler to the CBIR task.

Bit Pattern Feature (BPF)

In the above section, texture feature is extracted using CCM, here to get the features based on the edges, shape and image data BBF is applied. The quantization of

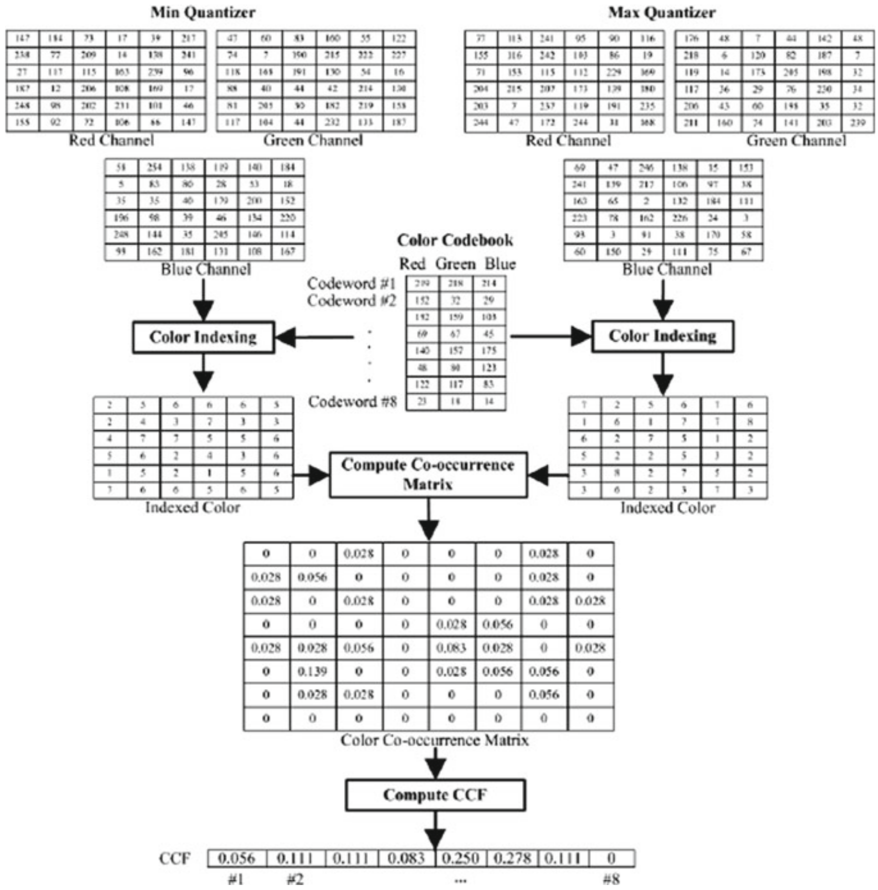
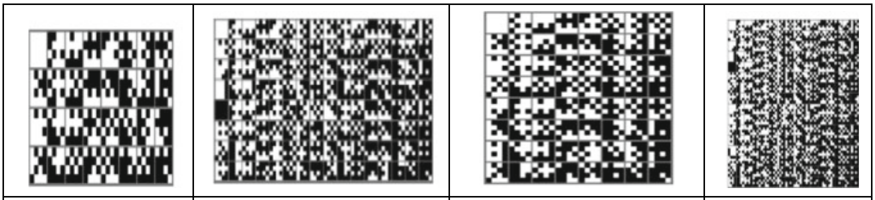


Fig. 4 Example of CCF computation

binary vectors generates a representative codebook of bit patterns from a collection of training bitmap images.

Let $Q = \{Q_1, Q_2, \ldots, QNb\}$ be the codebook for the bit pattern consisting of the Nb codewords. In the codebook, all code vector consists of intermediate real values as an opposite binary value between zero and one. The hard thresholding at the end of the training stage Binarizes all vectors of code in order to generate the final result. The bitmap of each block bm (i, j), is simply indexed by the similarity between the bitmap and the Qq codeword.

$$(i, j) = \underset{q = 1, 2, \ldots, N_b}{\arg min} \delta_h \{ bm(i, j), Q_q \}, \tag{6}$$



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

Fig. 5 An example for BBF code book trained

For all $i = 1, 2, \ldots, \frac{M}{m}$ and $j = 1, 2, \ldots, \frac{N}{n}$. The symbolic representation $\delta_H\{., .\}$. describe the hamming distance between the two binary patters (vector), i.e, bitmap image $bm(i, j)$ and bit pattern codeword. Q_q Subsequently, The BPF is extracted from the probability of occurrence of the bitmap image mapped in the specific bit pattern codeword Qq. Thus, BPF is formally considered as

$$BPF(t) = Pr\left\{b(I, j) = t | i(I, j) = t = 1, 2, \ldots, \frac{M}{m}; j = 1, 2, \ldots, \frac{N}{n}\right\} \quad (7)$$

For all $t = 1, 2, \ldots, Nb$.The BPF's dimensionality function is Nb, i.e., similarity to the codebook size for the bit pattern. The total dimensionality of the proposed Descriptor function is $Nc + Nb$. Figure 5 illustrates the BPF computation.

Gabor Filter

For an example given image I(x, y) with size $P \times Q$, its discrete Gaborwavelet transform is given by a convolution:

$$G_{mn}(x, y) = \sum_s \sum_t I(x - s, yT - t)\varphi^*_{mn}(s, t) \quad (8)$$

$$\varphi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right].\exp(j2\pi Wx) \quad (9)$$

where W is defined as frequency of modulation. Feature generation creates the self-similar working waves:

$$\varphi_{mn}(x, y) = a^{-m}\varphi(\tilde{x}, \tilde{y}) \quad (10)$$

where m and n represents the scale and orientation of the wavelet respectively, with $m = 0, 1, \ldots M - 1$, $n = 0, 1, \ldots, N - 1$, and

$$\tilde{x} = a^{-m}(x \cos\theta + y \sin\theta) \quad (11)$$

$$\tilde{y} = a^{-m}(x \sin\theta + y \cos\theta) \quad (12)$$

where a > 1 and $\theta = \frac{n\pi}{N}$ The variables listed are described below in the above equations

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{M-1}} \tag{13}$$

$$W_{m,n} = a^m U_l \tag{14}$$

$$\sigma_{x,m,n} = \frac{(a+1)\sqrt{2 \text{ in } 2}}{2\pi a^m (a-1) U_l} \tag{15}$$

$$\sigma_{y,m,n} = \frac{1}{2\pi \tan\left(\frac{\pi}{2N}\right)\sqrt{\frac{U_h^2}{2 \text{ in } 2} - \left(\frac{1}{2\pi\sigma_{x,m,n}}\right)^2}} \tag{16}$$

In the implementation, the constants are utilised as $U_l = 0.05,\ U_h = 0.4,$ s and t ranges from 0 to 60, i, e filter mask size are 60 * 60. The Gabor-based texture representation transforms, measures similarity in texture and normalizes rotation.

Texture representation: When Gabor filters are placed in the picture with different directions on different scales, we get a variety of dimensions:

$$E(m, n) = \sum_x \sum_y |G_{mn}(x, y)| \tag{17}$$

$M = 0, 1, \ldots, M - 1;\ n = 0, 1, \ldots, N - 1$. These quantities reflect a different degree and orientation for the energy content of the image.

Classifier

Artificial Neural Network (ANN): ANN is well known in the pattern classification field as the best tool for being strong. Specific forms of neural networks for the image classification have been proposed. The method continues to repeat in the feedforward and backpropagation until the output gives good accuracy and less error rate has been achieved or a given number of training cycles. Two separate samples called research and the training samples are considered by the networks. The testing samples are used for network preparation, and the network is configured according to its mistake. After the instruction, test samples are then used to test the device for an independent measure of network output.

Suppose that vectors in the extracted properties will express the input of the training sample and their corresponding vector class $s(X_i^p, Y_i^o)$ where the input number "p" is the number of input units in an input layer that indicate the features of an input variable; In the testing stage we extract its characteristics for a query image "q" in order to construct a query function vector, which then transforms the qualified neural network into an input vector, and it's allocated to one or more related groups by the network. Interior and output neuron are sigmoid activation characteristics. The efficiency of the model classification is accurately calculated [7]. This term refers to

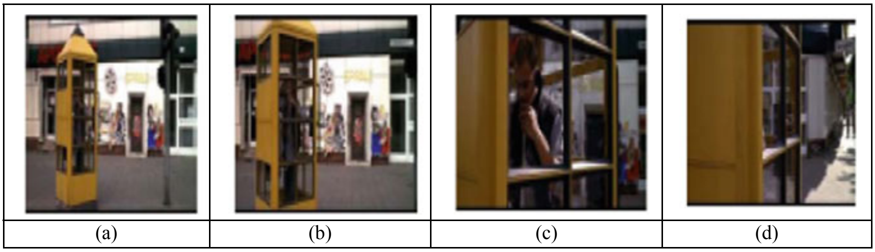


Fig. 6 **a** is the input image; **b**, **c** and **d** are the retrieved similar images

the model's ability to predict the shape of new unknown data correctly. The accuracy of the classification is calculated by calculating the proportion of cases where the test sets have been correctly categorized. The high p result is always a successful classification check:

$$\text{Accuracy} = \frac{\text{True positive} + \text{true negative}}{\text{Total number of images}}$$

If the right assumptions are positive, the real negative are the right ones if an instance is positive.

4 Experimental Results

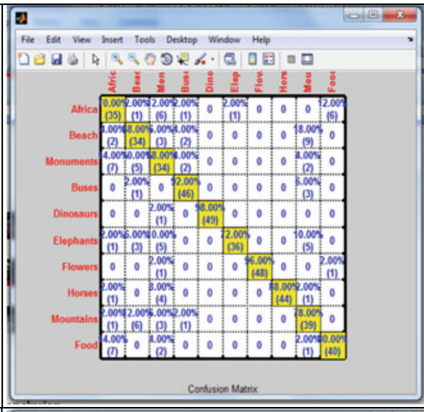
It is necessary today, with a huge number of digital images, to follow an appropriate model for access to the photos you want. To evaluate the effectiveness of the proposed system, we conducted extensive tests. The machine output is checked via the qualified dataset. Figure 6a displays the image of the input and (b–d) displays related images confusion matrix in diagram Fig. 7.

5 Conclusion

It is critical today to follow an effective model for accessing the desired images with a huge number of digital images. The model proposed involves four main stages, namely: Query Image pre-processing, Extraction functionality and ANN classification and match strategies. As for the removal stage, it extracts color and texture characteristics, called CCM and BPF respectively. Integrating multiple functions, however, can solve single-function problems. The unit output is computed by using the parameters known as precession, and recall. The Parameter Assessment indicates improved efficiency accuracy = 79.80%

Confusion Matrix:

31	3	6	0	0	4	0	0	0	6
1	32	4	0	0	4	1	2	5	1
3	1	33	2	0	2	1	1	6	1
1	1	3	43	0	0	0	0	2	0
0	0	0	0	48	0	0	0	0	2
0	3	0	0	1	40	0	0	5	1
0	0	0	0	0	0	49	0	0	1
0	0	0	0	0	1	0	48	1	0
1	7	5	1	0	2	0	0	33	1
3	0	0	1	0	0	0	0	4	42



accuracy = 82.00%

Confusion Matrix:

36	0	5	1	1	1	1	0	2	3
4	34	2	1	0	1	0	1	7	0
5	7	30	2	0	3	0	1	2	0
1	2	0	47	0	0	0	0	0	0
0	0	0	0	50	0	0	0	0	0
3	3	5	0	0	35	0	0	4	0
0	0	0	0	0	0	50	0	0	0
0	0	0	0	1	0	0	49	0	0
0	9	3	0	0	0	0	0	38	0
6	0	0	1	0	0	0	1	1	41

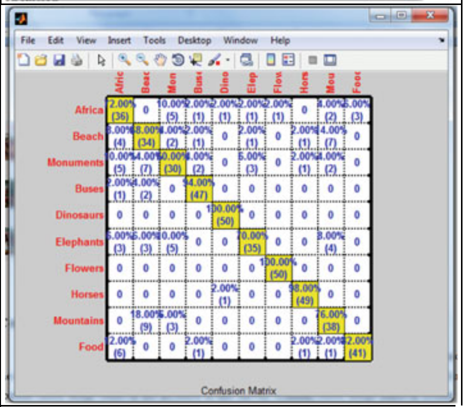


Fig. 7 Confusion matrix

References

1. Dubey SR, Singh SK, Singh RK (2016) multichannel decoded local binary patterns for content-based image retrieval. IEEE Trans Image Process 25(9):4018–4032
2. Xia Z, Wang X, Zhang L, Qin Z, Sun X, Ren K (2016) A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing. IEEE Trans Inf Forens Secur 1(11):2594–2608
3. Guo JM, Prasetyo H (2014) Content-based image retrieval using features extracted from Halftoning-based block truncation coding. IEEE Trans Image Process 24(3):1010–1024
4. Xu B, Bu J, Chen C, Wang C, Cai D, He X (2013) EMR: a scalable graph-based ranking model for content-based image retrieval. IEEE Trans Knowl Data Eng 27(1):102–114
5. ElAlami ME (2014) A new matching strategy for content-based image retrieval system. Appl Soft Comput 14:407–418
6. Gode CS, Ganar AN (2014) Image retrieval by using colour, texture and shape features. Int J Adv Res Electr Electron Instrum Eng 3:8637–8644
7. Dash JK, Mukhopadhyay S, Gupta RD (2015) Content-based image retrieval using fuzzy class membership and rules-based on classifier confidence. IET Image Proc 9(s9):836–848
8. Demir B, Bruzzone L (2014) A novel active learning method in relevance feedback for content-based remote sensing image retrieval. IEEE Trans Geosci Remote Sens 53(5):2323–2334

9. Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147:195–197. https://doi.org/10.1016/0022-2836(81)90087-5

Speech Recognition and Text Summarization Using Textrank Algorithm

K. H. Asha, N. S. Pushpahasa, Sahana S. Mathad, and B. Varalakshmi

Abstract Human beings use speech has basic form of communication, extending this concept to the world of computers will create a milestone in the field of technology. Speech recognition is the process of converting the speech into text format. This paper presents uses speech recognition in education sector along with text summarization module, Text summarization refers to process of taking long pieces of text and shortening it, by outlining the main points thereby creating the coherent summary of the document or text. This paper describes the method of using speech recognition system and text summarization model for a professor or teacher by recording the lecture delivered during the class and pass the recorded lecture to the text summarization model.

Keywords Speech recognition · Nltk · Textrank · Similarity matrix · Feature extraction · Vectors · Cosine similarity · Document-term matrix

1 Introduction

In general terms speech recognition is talking to a computer, It is a process of having the computer recognize what a person is saying [1]. Due to advancement in signal processing, algorithms, hardware, the technology of speech recognition has evolved for more than 4 decades. In recent years, the advances in deep learning and big data has widely effected the field of speech recognition, there are hundreds of academic papers published in this field and there are wide ranges of methods available for designing and deploying speech recognition system [1]. This Project provides a technique where the lecture in the class are recorded and the speech recorded is converted into text which is again taken as input for summarizing the lecture.

K. H. Asha (✉) · N. S. Pushpahasa · S. S. Mathad · B. Varalakshmi
Department of Information Science and Engineering, Don Bosco Institute of Technology, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_85

2 Problem Statement

The automatic speech recognition system takes speech waveform from unknown source and classify it as one of a set of spoken words, phrases, or sentences. To address the large amount of data that is produced every second and is growing exponentially in amount and to help discover the required information more quickly the automatic text summarization method are very helpful. This Project provides a technique where the lecture in the class are recorded and the speech recorded is converted into text which is again taken as input for summarizing the lecture.

3 Implementation

The speech recognition system primarily uses microphone to convert the speech sound into electric signal. There are certain building blocks for general speech recognition system [2]. These building blocks forms the most fundamental element of the speech recognition system:

1. *Signal Pre-Processing*: Speech signals are captured by a microphone or a telephone and these signals are analog in nature, so there is a need to digitize the speech waveform to make the signal able to process for the model.
2. *Feature Extraction*: Feature extraction is parameterization of speech signal which is used to find the stable and correlated set of properties of the speech signal.
3. *Language modeling*: Language modeling predicts the nth words using $n - 1$ preceding words, thereby helps to find correct word sequence for the model to predict the speech.
4. *Decoder*: For the sequence that is observed, this phase helps to find the word sequence which is more probable or likely to occur.

Normally, a reader will refer different source of blogs, websites, documents to get an overview or the summary of the data, which consumes time and needs lot of rigorous effort from the reader. By automating this task it would help the user to get the summary or the context of the different data sources without much effort and in less duration of time. Here we take the recorded data from the lecture and pass it to the recognizer to convert it into speech and the converted speech is feed into summarizing model. The application of text tank algorithm fallows different steps to accomplish the task. If there are multiple data sources, the first step of the process would be to concatenate all the text into one articles and then split them as individual sentences to perform pre-processing. The processed sentences are then used to find vector representation for every sentences. A matrix is created by calculating the similarities between the sentences. This matrix (similarity matrix) is converted into a graph. The graph contains sentences as vertices and edges are represented as scores of similarity, for calculating rank of the sentences.

4 Methodology

The Project is divided into 2 different modules:

1. Speech-to-Text (STT)
2. Text Summarization (TS)

4.1 Speech-To-Text (STT algorithm)

Step 1 Click on "Add file".

Step 2 Select the audio file in the directory.

Step 3 Click on "Proceed" button.

Step 4 Open the file to check the converted text.

4.2 Text Summarization Algorithm

Step 1 Concatenate all the text contained in the articles.

Step 2 Split entire text into individual sentences.

Step 3 Find representation of vectors (word embedding's) for every sentence.

Step 4 Matrix is created by calculating the similarities between the sentences. The basic working of the model is shown in Fig. 1

5 Result

The model takes in the audio file as input and provides the summary of the speech provided for the model.

1. *Feed the speech file to get the text file as output*: The model accepts the user to provide the speech file (preferably, mp3 file) by browsing the local directory in the system and the output text is stored as a text file. The design is shown in Fig. 2

2. *Output of the speech recognition*: The speech recognition converts the audio passed to the model to text format. Figure 3 shows the content of the text file obtained from the speech recognition model,
3. *Feed the text file to summarize the document*: The text file obtained in the previous step is given as input to the summarization model. Textrank algorithm

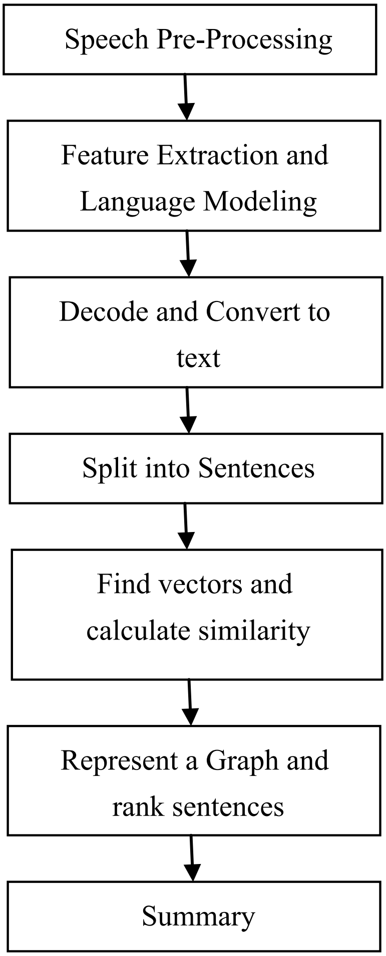


Fig. 1 Working of model

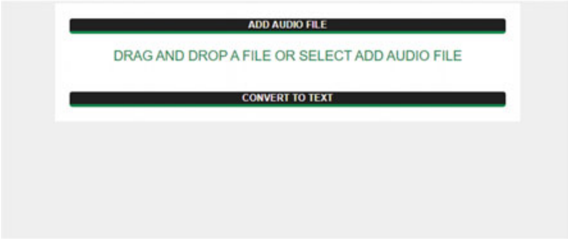


Fig. 2 Converting speech to text


After graduating from the Madras Institute of Technology in 1960, Kalam joined the Aeronautical Development Establishment o
f the Defence Research and Development Organisation (by Press Information Bureau, Government of India) as a scientist after
becoming a member of the Defence Research & Development Service (DRDS).
[13] In 1969, Kalam was transferred to the Indian Space Research Organisation (ISRO) where he was the project director of I
ndia's first Satellite Launch Vehicle (SLV-III) which successfully deployed the Rohini satellite in near-earth orbit in Jul
y 1980; Kalam had first started work on an expandable rocket project independently at DRDO in 1965.
[11][27] Between the 1970s and 1990s, Kalam made an effort to develop the Polar Satellite Launch Vehicle (PSLV) and SLV-III
projects, both of which proved to be successful.Kalam was invited by Raja Ramanna to witness the country's first nuclear te
st Smiling Buddha as the representative of TBRL, even though he had not participated in its development.
In the 1970s, Kalam also directed two projects, Project Devil and Project Valiant, which sought to develop ballistic missil
es from the technology of the successful SLV programme.
[29] R Venkatraman was instrumental in getting the cabinet approval for allocating â,¹ 3.88 billion for the mission, named
Integrated Guided Missile Development Programme (IGMDP) and appointed Kalam as the chief executive.
[29] Kalam played a major part in developing many missiles under the mission including Agni, an intermediate range ballisti
c missile and Prithvi, the tactical surface-to-surface missile, although the projects have been criticised for mismanagemen
t and cost and time overruns.
[29][30]Kalam served as the Chief Scientific Adviser to the Prime Minister and Secretary of the Defence Research and Develo
pment Organisation from July 1992 to December 1999.
Kalam served as the Chief Project Coordinator, along with Rajagopala Chidambaram, during the testing phase.

Fig. 3 Converted text file

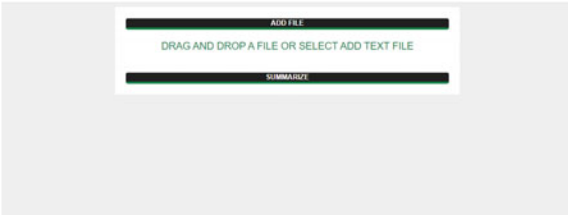
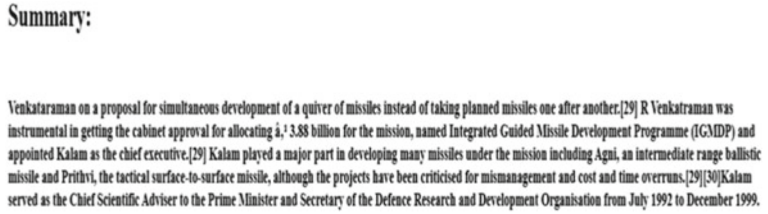


Fig. 4 Generating summary of the model

tokenizes the sentences, performs vectorization, calculate similarity between the text and provides the concise explanation of the document to the user. The design of the same is shown in Fig. 4

4. *Output of the summarization model*: When a text file is passed to the summarization model, the entire document is processed and summarized based on the parameters passed. Figure 5 shows the output of the summarization model.
5. *Analysing the text in the document*: The summarization model converts the entire corpus into matrix (similarity matrix) format, network graph can be used to visualise how the words in the document are connected. The below network


Summary:

Venkataraman on a proposal for simultaneous development of a quiver of missiles instead of taking planned missiles one after another.[29] R Venkatraman was instrumental in getting the cabinet approval for allocating â,¹ 3.88 billion for the mission, named Integrated Guided Missile Development Programme (IGMDP) and appointed Kalam as the chief executive.[29] Kalam played a major part in developing many missiles under the mission including Agni, an intermediate range ballistic missile and Prithvi, the tactical surface-to-surface missile, although the projects have been criticised for mismanagement and cost and time overruns.[29][30]Kalam served as the Chief Scientific Adviser to the Prime Minister and Secretary of the Defence Research and Development Organisation from July 1992 to December 1999.

Fig. 5 Output of the summarization model

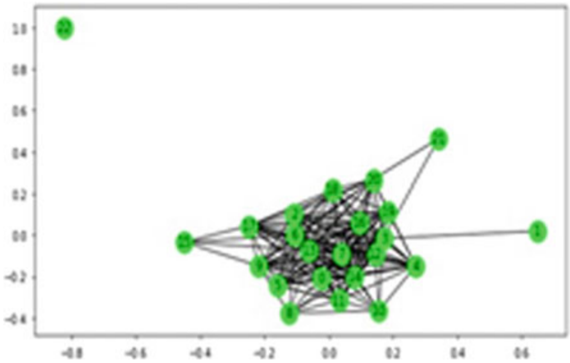


Fig. 6 Graph showing corpus of text

```
In [51]: filename_2 = 'summarized_doc.txt'
         filename_1 = 'recognized.txt'
         documentSimilarity(filename_1,filename_2)

File recognized.txt :
2994 lines,
474 words,
251 distinct words
File summarized_doc.txt :
807 lines,
125 words,
85 distinct words
The distance between the documents is:  0.563380 (radians)
```

Fig. 7 Similarity between the output text file of recognized speech and resultant summary

graph Fig. 6 is plotted on the similarity matrix of the document, the axis of graph has similarity score of the word.

6. *Comparing summarized document with it's original text file*: In text analysis to measure the similarities between two documents cosine similarity can be used which calculates the cosine angle between the two vectors. Figure 7 represents the code for the similarity comparison made between the output text file of recognized speech and resultant summary of it.

6 Conclusion

Although speech recognition field has gained a wide approval to automate the service and applications but there are several parameters which affect the accuracy and efficiency of speech recognition system. Many factors effect the variability of the speech such as environment, rate of the speech utterance, channel, accent of the speaker etc. Some of the features of the speech signal effects the robustness of the speech system. This Project work introduces the technologies and its application in different sectors of IT industry sectors and covers the concepts of text summarization and combine these two standard problems to implement the project. The audio is recognized and then summarized using Textrank algorithm. The cosine similarity of original and summarized documents is calculated to measure similarity between documents.

References

1. Patel B, Shah V, Shirsagar Rk (2011) Microphone based speech recognition system
2. Anusuya MA (2009) Speech recognition by machine. Int J Comput Sci Inf Secur 6(3)
3. Waibel A, Hanazawa T, Hinton G, Shikano K, Lang KJ (1989) Phoneme recognition using time-delay neural networks

Assistive Aid for Visually Impaired People

Nikith M. Jarali, P. S. Ashok Kumar, M. Likith Gowda, C. Pavan Gowda, and Niket Kumar Bhaskar

Abstract Addressing the problems of visually impaired persons through one aiding system is not an easy job. Now a days' modern researches concentrate over addressing the problems on visuals, hearing and vocal impaired persons but not on the minute issues they face in their daily life. Hence our work emphases on discover a exclusive method, that helps the blind persons by allowing them know the obstacle, by captures the image through a camera that is fixed in front of the stick, which captures the image and converts it to a voice signals. The paper provides various methods to overcome to visual impairment. It provides them to visualize or read the texts which are captured through the camera and converted into audio form using the text to speech method. So this three methods i.e. object detection, text to speech and image to speech are the three main methods which are used to help the visually impaired persons to overcome with the minute issues the face in their day to day life and to be completely independent.

Keywords Assistive aid · Rasberry Pi · Espeak · Dvyangan · Walking stick

1 Introduction

Around the worldwide there are about 285 thousand peoples, who are adjudicated to be visually impaired, out of which some and around 39 million are blind peoples and about 246 thousand are said to have low vision. And almost 90% of this people is from a dispirited income people. The number of individuals who are visually impaired are bought down within the past 2 decades according to the global estimation. And within which 80% of all the visually impaired can cured.

India is considered as the world's largest act of blind people. There are 15 million blind people in India out of 30 million blind people all around the globe. Though there are large number of researches getting along in this world to overcome with this major issue, Thus, in order to empower these people, we have proposed the assistive

N. M. Jarali · P. S. Ashok Kumar (✉) · M. Likith Gowda · C. Pavan Gowda · N. K. Bhaskar
Department of Computer Science and Engineering, Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_86

device to assist blind people to engage their jobs individually without the assistance of other peoples.

Signing or Insignia language is a linguistic process, it empowers deaf and dumb peoples to communicate outside world, like normal people or other Divangan peoples. Generally deaf or vocally impaired peoples are communicating with outside world by using sign language, it is a very big issue for such group of peoples, it is a blind and very crucial problem in the society, but its magnitude is very high.

To overcoming these difficulties with the impaired individuals, we are planned a model, but it used Raspberry Pi, it solves all critical problems to all impaired personnel's. Which they can read and understand in an easy manner. And for the dumb persons, their messages are conveyed through texts rather than signing.

2 Literature Survey

For deaf and blind persons, they face lot of problems when they were planning to do their job individually, in this regard we created an application it holds many devices like, IR sensor, Ultrasound sensor and water sensor to detect the obstacle and it provides an alert signal to needy peoples [1].

For visually impaired persons, a Pothole detection is a very big issues, in that aspect, concept of image processing used in this model, it encompasses a camera, it captures all images 15 frame per second [2].

Smart Walking Stick is a god gift to Blind peoples, it uses Raspberry Pi, with many more sensory devices. Based on these sensory devices an impaired people can do their job efficiently without the interference of normal peoples. The system is working based on image processing, where the objects are detected and verified based on stored image dataset. Generally, this system is inaccurate while finding generic obstacles to be detect, because system is going to find the objects using stored dataset (large set of images) [3].

Smart Belt embedded with ultrasound sensor, it detects the obstacle for blind peoples, more over the belt has a buzzer, it vibrates whenever obstacle is found, i.e. whenever the blind person receives a buzzer sound automatically, he will get alert and save his life [4].

3 System Design

3600 positional results have been given by the smart walking stick till now which is sufficient to provide support for the blind society. The ultrasonic sensor consists of a broad search angle which helps in a larger obstacle detection which is the main aim of the blind stick which is to provide the safety for the blind and to make them aware of their surroundings.

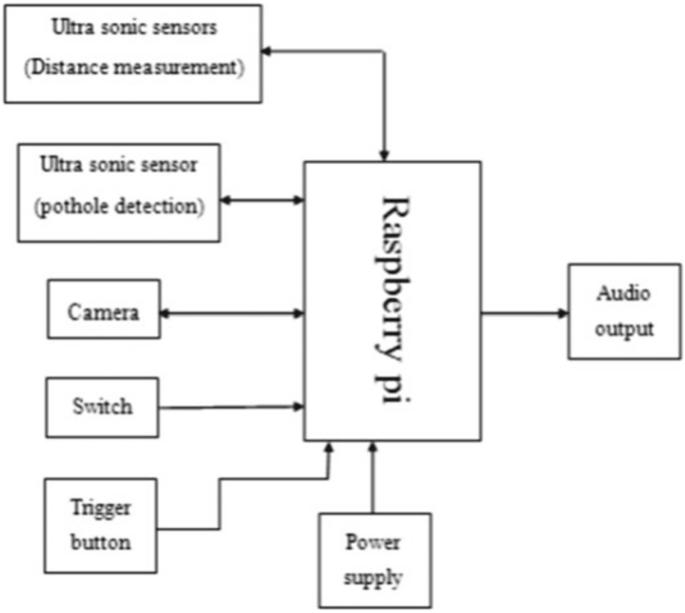


Fig. 1 Overall architecture of the system

Figure 1, shows an overall architecture of the Smart stick application, in tis system we plan add a GPS system which helps in identifying the precise location of the blind person which allows their guardians to identify their exact location. It is necessary that visually impaired people get access to an efficient and cozy object so as to measure their way of life comfortably. During a developing country like India, there's a requirement for a cost-effective solution in order that most of the people can have an effective product as proposed during this paper.

3.1 Introduction to RASPBERRY PI

The Raspberry pi is a master card sized computer, it consumes low power, operating with Linux OS and it has an ability to run different programs shown in Fig. 2.

For this advantages system is implemented with Raspberry Pi and implementing with python program language.

The RasberryPi board embedded with ultrasonic sensor HC—SR04, ultra-sonic transmitter, receiver and control circuit. It provides a distance measurement ranged from 2 to 400 cm,

$$Distance = \frac{(Timeathighlevel \times Velocityof\,Sound)}{2} \tag{1}$$

Fig. 2 RasberryPi board

If we need to find the obstacles within a range of 1 m, then it will be calculated by using Eq. 1, specified above.

To detect the object within the range of 1 m, then we must set the ultrasonic sensor, with an range of 1 m only, then raspberry pi with transmit and receives the signal very effectively, meanwhile pi board is attached with pi cam or wired camera. The importance of the camera is to capture the specific images, i.e. it might be any type of the obstacles, it harmful to the physically impaired person. Once the images captured by the camera, it directly comparing with captured images with stored dataset using raspberry pi. Here image processing technique is used for detection of obstacles effectively. Moreover, a head phone or ear phone is directly connected to the raspberry pi for the usage of buzzer.

Raspberry pi is embedded with Walking stick at, which Ultrasonic sensor is interfaced with GPIO pins of Raspberry pi, at the same time the camera should be connected with raspberry pi where both should face on the same direction on the road. It should be connected with Network so that captured image can be processed for object detection to intimate which object is an obstacle at this time. RF transmitter is also connected with RPi. Each command is sent by pressing the button on the RF transmitter which is held by them.

3.2 GPS and GSM Module

GPS module will provide the current location of physically impaired person is roaming with his current status is also be known to their care takers or guardians.

GSM module will act has a transponder, where an impaired person is contacted by using his mobile, so that guardians can easily track him through mobile while emergency.

3.3 Raspbian OS

The Raspbian OS is almost same as the windows operating system thus navigating through Raspbian OS is not that difficult it consists of a menu bar, a browser, a file manager and a lot of preinstalled apps thus allowing us for easy working environment.

3.4 Tesseract OCR

Optical Character recognition (OCR) engine is used in python Tesseract for various OS. Tesseract OCR extracts text from images and later we can use those texts for various activity such as an output through a speaker or document editing etc.

3.5 Open CV

Open CV is a open source software, it has wide collection of libraries, it is mainly used towards the operations of real time computer vision. The libraries are free to use by the users, because Open CV is under the open source license. Open source computer vision allows deep learning such as TensorFlow, Torch/PyTroch and café. The interfaces supported are Python, Java and MATLAB and also supports operating system such as Windows, Linux, Android and Mac OS.

3.6 Espeak

Espeak is a speech synthesizer for English and 11 languages it is supported in both windows and Linux platform. It converts text to voice it also supports many languages of small size. The programming of Espeak software is composed using rule files with feedback it also supports SSML. The text files consist of different characteristics such as pitch range and effects like echo, whisper and croaky voice or it will make adjustments to vary the sound of voices.

Espeak employs format synthesis method that allows many languages to run in very small size thus allowing for more memory space. The main drawback is the output voice is not as smooth as a natural human voice but the output is clear and audible.

4 Implementation

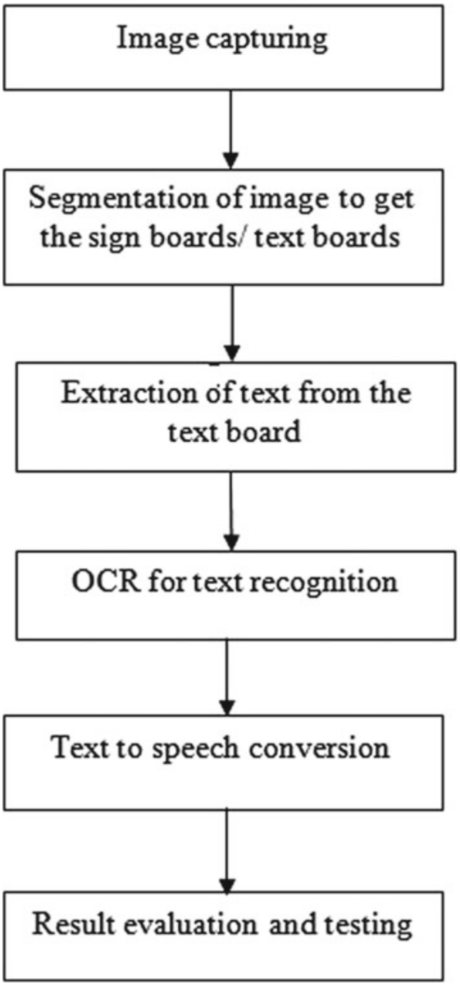
While implementing smart stick application, we need to follow the image capturing and validation process, it is specified in Fig. 3. Different steps are followed to reading System for Blind person Using Python and Raspberry Pi are,

- Image Capture
- Convert to Text
- Convert to voice

Generic operations of Smart stick are,

- Helping blind person to walk with Electronic Stick

Fig. 3 Low level design



- Inform the blind person to location of the obstacle
- Informing blind person about the location of the obstacle in Voice
- Detection of fall of blind person using Accelerometer
- Rescue blind person immediately from any exceptional condition

Algorithmic steps used in the assistive model are

1. : start the Application
2. : read IO pins of Pi Board
3. : if pin 1 is high go to point 4,
   if pin 2 is high go to point 7,
   if pin 3 is high go to point 10,
   else go to point 2.
4. : open webcam, takes picture and saves the image as "app1.jpeg" and stored it into PC.
5. : Convert image to text using python program and save the outcome as "out1.txt".
6. : for audio conversion—move "out1.txt" to "audio.txt" file
7. : for audio generation move outcome into point 2
8. : for distance measurement, use ultrasonic sensor to measure distance using python Program and Save the outcome as "out2.txt".
9. : for audio conversion, move "out2.txt" into "audio.txt"
10. : for audio generation, move out2.txt into point 2
11. : Image generation, open webcam and get picture, save image as "obj1.jpeg" and store it into PC,
12. : recognize object by using python program and save the outcome as "out3.txt".
13. : for audio conversion, move the text in "out3.txt" to "audio.txt"
14. : generate audio outcome and move to point 2

The above algorithm explains the various process that took place, when any visually impaired person is using this smart walking stick.

Figure 4 represents the working model of smart stick, meanwhile corresponding Figs. 5 and 6 represents the outcome of the smart stick project.

Figure 5 describes to finding the distance of obstacles, where person is going to reach.

Figure 6, shows the actual latitude and longitude of the person moving around is to be find by using google map.

5 Conclusion

This paper is designed in such a way that to form a working system using ultrasonic sensors, GPS module and voice commands through a output speaker for the

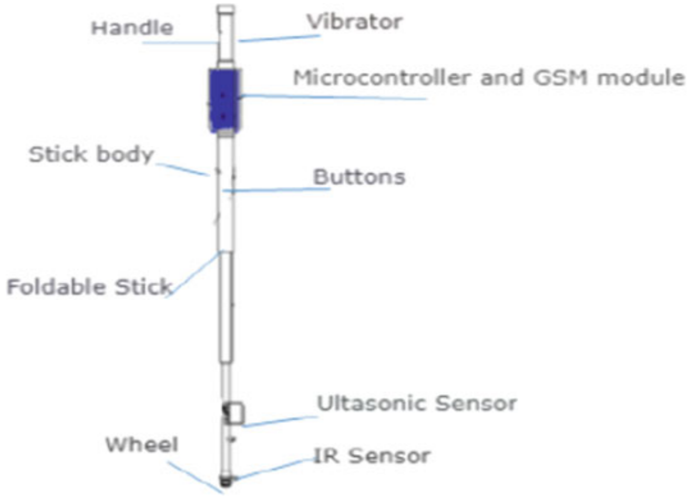


Fig. 4 Working model of smart stick

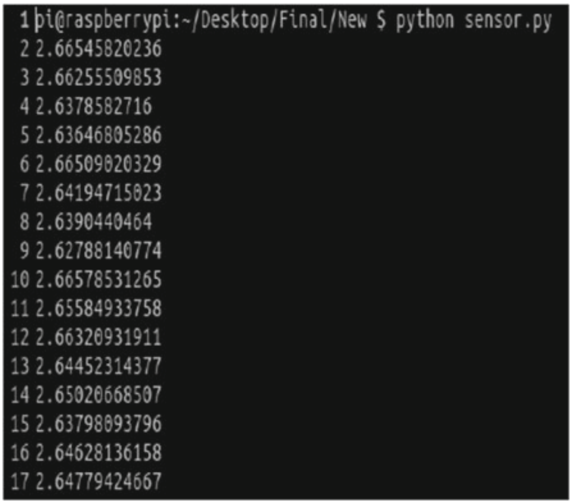


Fig. 5 Finding obstacle and collision detection

blind user this allows the visually impaired person to navigate through public places independently.

We try to overcome the errors caused by the previous versions. The device also main priority is the safety of the blind person. The blind stick uses ultrasonic sensors and GPS with an output speaker which is a great benefit to the blind person so that they can live independently. The other main important advantage of the device is that it is very affordable which is very helpful for visually handicapped person across the

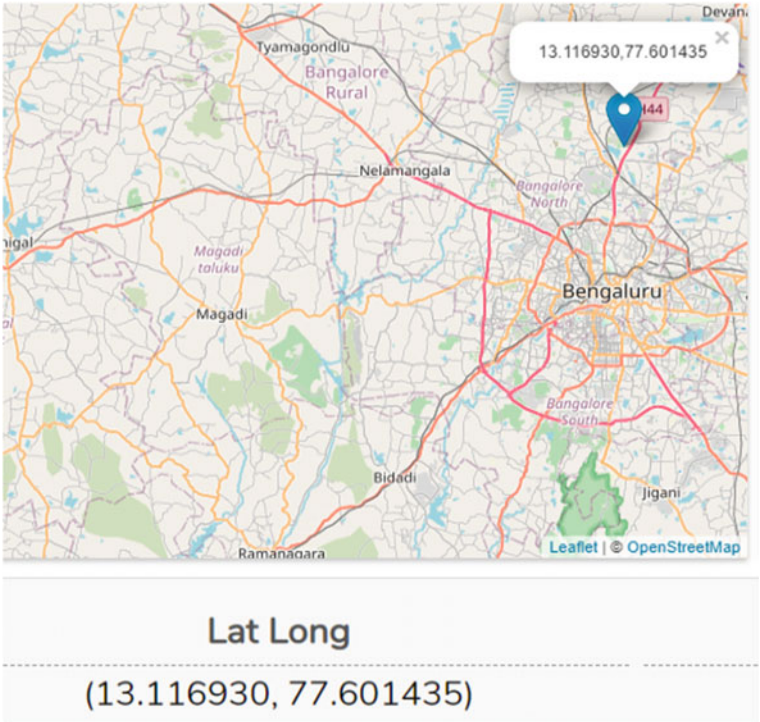


Fig. 6 Finding latitude and longitude

globe. The combination of both ultrasonic sensor and GPS allows for a system to track the person real time which is secure and gives a peace of mind too loved ones to monitor the easily.

We use Espeak algorithm to convert text to speech which is outputted through an on-board speaker which will allow the blind person to hear certain text which is not visible to him. The blind stick can track obstacles ahead of the user and can warn the user of the objects ahead of the users thus guiding the visually impaired.

References

1. Johnson J, Nikhil Rajan P, Thomas NM, Rakendh CS, Varghese SJC (2017) Smart stick for blind. Int J Eng Sci Invent Res Dev III(IX). Department of Computer Science, Jyothi Engineering College, Kerala, India
2. Nordrum A (2016) Title: pothole detection for blind. IEEE, 30 May 2016 https://spectrum.ieee.org/the-human-os/biomedical/devices/pothole-detection-for-the-visually-impaired
3. Prashanthi G, Tejaswitha P (2016) Sensor assisted stick for blind people. Trans Eng Sci 3
4. Mahesh K, Ashok Kumar PS, Naveen Raj HN, Naik PP, Apoorv MN (2019) IoT based smart surveillance security system using Raspberry Pi. IJARIIT 1356–1358

Crop Yield Prediction with Efficient Use of Fertilizers

S. M. Bharath, S. Manoj, Praveen Adhappa, Punit Laxman Patagar, and R. Bhaskar

Abstract Agriculture in India is the backbone of the economy, and contributes to the GDP of the country. Yet the farmer's won't get enough money. It is mostly due to insufficient irrigation or crop quality, or crop yield is less than expected at times. Use temperature, soil, moisture, ph sensor, we can evaluate the land and we can estimate how much yield we get. This forecast will help farmers select appropriate crops by soil type, temperature, humidity, soil pH, season, fertilizer and months. This prediction can be accomplished by means of the support-vector machine learning algorithm. World population is growing, and adequate production of crops is required. Monitoring crop growth and predicting yields is highly important to the economic development of a country. The estimation of crop yields has a direct impact on national and international economies, and plays a significant role in food health and food safety. Deep learning is gaining popularity in crop control, crop type recognition, and crop yield estimation applications with recent advances in image recognition using deep Convolutionary Neural Networks. Current crop yield prediction approaches focused on remote sensing are composed of classical machine learning methods such as supporting vector machines and decision trees. Because India is an agricultural country and is primarily dependent on agriculture and agro-industrial products development in its economy. New research into the study of crop yields is Data Mining. Renders are a very critical concern in agriculture forecasting. Each farmer needs to see how much he can produce. Analyze the various associated characteristics, such as the position, from which pH of the soil alkalinity is measured. Create a web-based system to display the farmers' outcomes on the basis of their vote.

Keywords Crop yield · Farmers machine learning · Deep learning

S. M. Bharath · S. Manoj · P. Adhappa · P. L. Patagar · R. Bhaskar (✉)
Department of CSE, DBIT, Bengaluru, India
e-mail: bhaskar@dbit.co.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_87

1 Introduction

Agriculture is an important Indian Economic segment. The Indian agriculture industry accounts for 18% of India's GDP and provides jobs for 50% of the country's labour force. Yet recent surveys have shown a gradual decline in the contribution of agriculture to the Indian economy while it is the broadest economic field demographically and plays a significant role in the overall socioeconomic structure of India. Other climate conditions affect the production of agricultural crops.As metrological parameters (humidity, wind speed, temperature, and humidity), precipitation parameters (runoff, wise runoff area, irrigation area, etc.) and soil parameters (PH, organic carbon, phosphorus, fibre, etc.). Yet because of the ongoing change in the atmosphere situation everything is messed up [1].

Farmers in India tend to use the same methods their predecessors followed. Nonetheless, the question is that everything that happened on time was absolutely great at the earliest moment in the world. Yet today a lot of the situation has changed because of global warming and many other factors. The major problem with agriculture in India is a lack of natural rainfalls. Moisture is often required for crops, but it also turns out to be an annoyance [2].

The winter season has affected Rabi cultivations to a large degree. The winter rainfall was as high as expected for a couple of years.

To order to address these challenges, we need to develop a system which can identify the secret data or results, patterns and insights. The farmer forecasts what he/she is going to seed to get more income. In the proposed plan, we use data mining tools on farm production databases to gain lessons so that farmers can improve their decision-making process [3].

Because farmers in India don't understand modern technology, we should be developing a system that farmers can understand.

The data processing is a tool to clean and model data and extract relevant details and conclusions. Specifications are used. Information analysis In order to derive any sequence, the related information from massive statistics should be evaluated, derived and projected. Companies use this approach to convert raw data from their clients into usable data. Throughout the agriculture industry, this work can also be used. Throughout the coming harvest season, the majority of farmers depended on their long-term experience with specific field crops to expect a better return but still don't get precious crop prices [4].

It is sometimes due to insufficient irrigation or crop selection or the crop yield is often below expected. The needs for an effective plant growth prediction and optimization system are the key focus of agricultural science, and biological processes for crop growth identification and crop yield improvements are at the heart of agriculture research. The resulting crop yields mainly depend on parameters like crop size, seed quality, and environmental parameters such as sunlight (temperature), soil (ph.), water (ph.), precipitation and moisture [5–7].

2 Literature Survey

In "Prediction of Crop Yield using Machine Learning", Rushikaghadge, juileekulkarni, pooja more, sachee Nene, priya R L—IEEE-2018. By analyzing location-wise weather constraints and the nutrients present in the soil, the system aims to help farmers grow proper crops for better yield production. This also recommends fertilizer depending on the soil condition, if appropriate. This also helps the user to pick a crop and display details thereon. The method does not take land area that is being cultivated and the date of sowing. The market price following harvesting of the cultivated crops is not regarded.

In "Prediction of major crop yields of Tamilnadu using K-means and Modified KNN", Mr A Suresh, Dr. P. Ganesh Kumar, Dr.M.Ramalatha. Agriculture is the primary source of livelihood for over 40% of this state's population. According to researchers at the Food and Agricultural Organization (FAO), the world population will rise by one third between 2010 and 2050. Demand for crop production will grow above current production by 60%. Prediction therefore plays a major role in finding out what crop production demands to maximize yield. Live data isn't taken into account.

In "Rice crop yield prediction in india using support vector machines", Niketagandhi, leisa J. Armstrong, owaizpetkar, amiya kumar tripathy. Manufacturing cereals in India, including rice, wheat and various pulses, are primarily dependent on milk. Rice areas rely on the right weather conditions for their sustainability and productivity. Drought occurrences may have detrimental effects in seasonal climatic settings with productivity declining. The development of innovative strategies that can allow farmers and other stakeholders to forecast production under different climatic environments can help to make smarter agricultural and crop choices. In order to increase crop production under different climatic conditions, machine learning approaches can be used. This essay provides a analysis of the machine learning methodology used in Indian rice growing regions.

3 System Architecture

These system uses SVM algorithm (Shown in Fig. 3) to predict yield, district code, state code, humidity, heat, PH, moisture values are the inputs we are providing [8]. Sensor values are obtained and uploaded to the Thingspeak cloud by using a Wi-Fi card in the hardware. The values are stored in .csv format and ML algorithms are applied to the values as shown in Fig. 1 SVM algorithm uses these values to predict the yield, datasets are used to compare the results with previously predicted values [9, 10]. Processing is done in real-time. Website is developed to show the results as shown in Fig. 5. Implementation is shown in Fig. 2. Hardware is shown in Fig. 4.

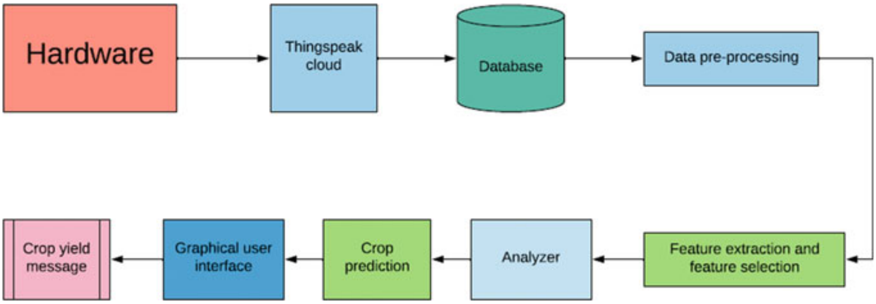


Fig. 1 System architecture

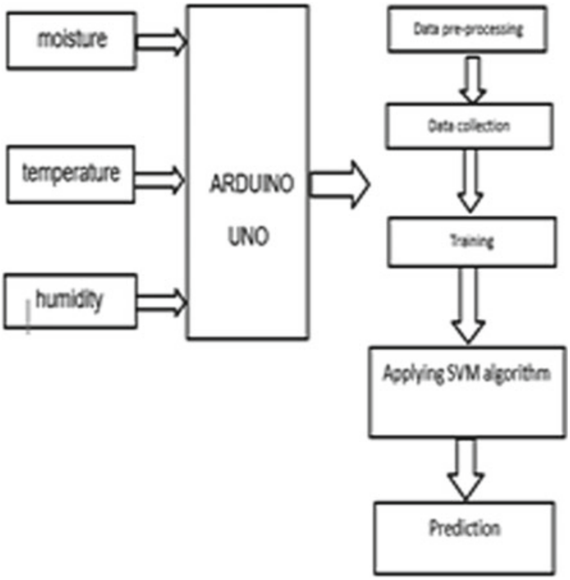


Fig. 2 Activities carried out in SVM

Throughout this paper, attempts are made to know the inspection of the production yield and are operated by applying the calculations of the SVM. In addition, information on manure was also obtained by measuring SVM and estimating how much nitrogen and phosphorus are needed in the field. We inspected the error rate gained while focusing on the calculation and back propagation of SVM where we obtained a lesser rate of misconception for SVM than back propagation while determining the yield for both models.

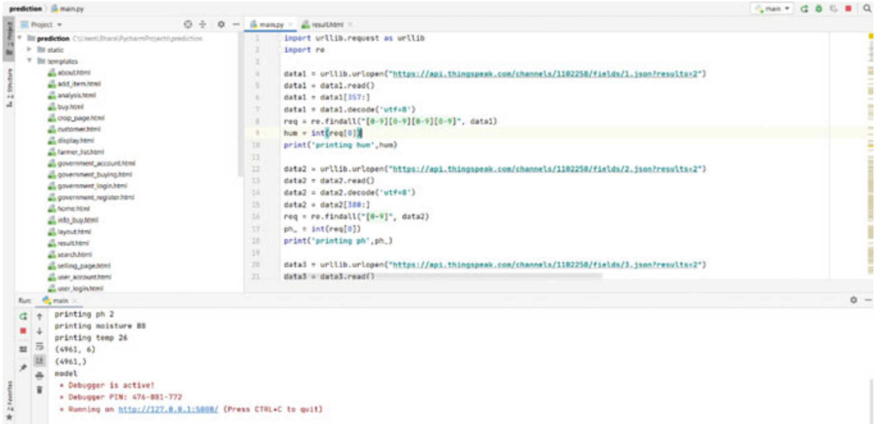


Fig. 3 SVM algorithm

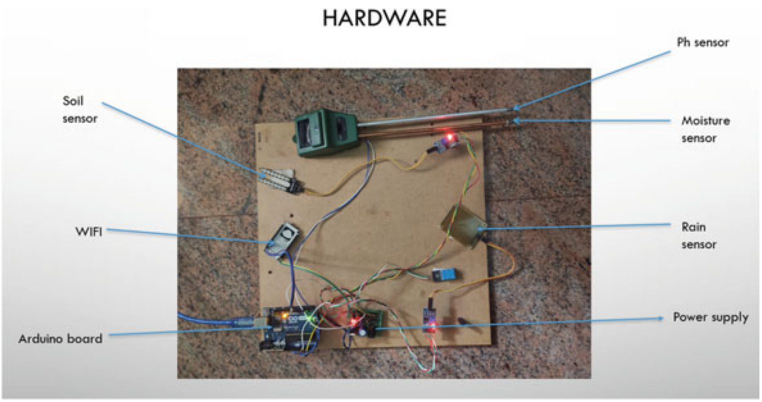


Fig. 4 Hardware components

4 Results and Discussions

The system comes with a model to be precise and accurate in predicting crop yield and deliver the end user with proper recommendations about required fertilizer ratio based on atmospheric and soil parameters of the land which enhance to increase the crop yield and increase farmer revenue [11, 12]. The yield will be determined after processing by the ml algorithms by giving the following inputs-state code, district code, season and land area, temperature, humidity, rain, ph as shown in Fig. 5. Finally after processing the results are obtained as shown in Figs. 6 and 7. This paper proposed a system which will help farmers to have an idea of yield estimates based on weather parameters and area under cultivation. Using this framework farmer can

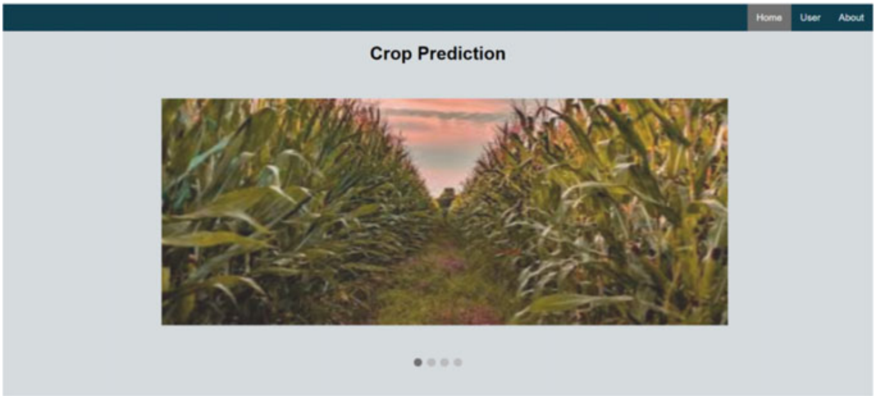
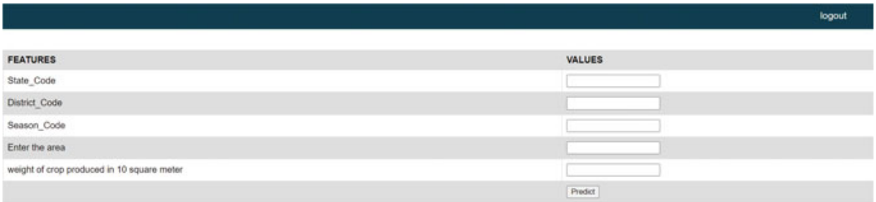


Fig. 5 Home page

State	District	Season	Temprature	Humidity	Moisture	Prediction	yield
28	1	2	27 celcius	1023 rh	94 g/m3	wheat , rice , paddy	400.0 Kg

Fig. 6 Results obtained using SVM



FEATURES	VALUES
State_Code	
District_Code	
Season_Code	
Enter the area	
weight of crop produced in 10 square meter	
	Predict

Fig. 7 Input parameters

make decisions on whether to grow that particular crop or go for alternate crop, in case yield predictions are unfavourable.

5 Conclusion

This paper suggested a framework for farmers to understanding of yield evaluations based on factors including ph, temperature, rain and crop-land. In countries like India, agricultural is a very important segment. However, the use of technology for agriculture should be given secondary importance to agricultural preclusion. The farmer will determine whether or not to plant this particular crop. Web-based interface is built to display the results, enabling farmers to understand easily. Sending out results directly to farmers mobile and can further improve this system. More parameters can be applied to better predict. Since we only use one algorithm to predict multiple algorithms can be used in better system.

References

1. Eswari KE, Eswari, Vinitha L, In Tamil Nadu, crop yield prediction using Bayesian network, intellectual development and research. Int J Field Eng Calcul 6(2). ISSN: 2348-2079
2. Culinary agriculture prevention and nitrogen status forecasts anna Chlingaryana, Salah Sukkarieha, Brettwhelanb—seed prediction machine learning approaches: a study, machines in agriculture, and electronics 151: 61–69, Elsevier (2018)
3. Menaka K, Yuvaraj N (2016) Indian innovation and development Journal 5(12), October, 2016 ISSN (online): 2277-5390 ISSN (print): 22775382 [8] Tom M machine learning. ISBN: 0070428077 Mitchell
4. Medar RA (2014) Int Comput Sci Manage Stud Adv J Data Min Techniq Crop Yield Predict 2(9). ISSN: 2231-7782
5. Pooja mcetal, implementation of data mining crop yield forecasting. Int Eng Technol Res J (irjet) p-issn: 2395-0072 (2018)
6. Hanjeong J et al, Random forests for global and regional crop yield predictions. PLOS-ONE (2016). https://doi.org/10.1371/journal.Pone.0156571
7. Afrin S, Khan AT, Mahia M, Ahsan R, Mishal MR, Ahmed W, Rahman RM (2018) Analysis of soil properties and climatic data to predict crop yields and cluster different agricultural regions of Bangladesh. In: 2018 IEEE/ACIS 17th international conference on computer and information science (ICIS)
8. Garg A, Garg B (2017) A robust and novel regression based fuzzy time series algorithm for prediction of rice yield. In: 2017 international conference on intelligent communication and computational techniques (ICCT)
9. Sekhar CC, Sekhar C (2017) Productivity improvement in agriculture sector usingbig data tools. In: 2017 international conference on big data analytics and computational intelligence (ICBDAC)
10. Sahu S, Chawla M, Khare N (2017) An ancient analysis of crop yield prediction using Hadoop framework based on random forest approach. In: 2017 international conference on computing, communication and automation (ICCCA)
11. Raja SKS, Rishi R, Sundaresan E, Srijit V (2017) Demand based crop recommender systemfor-farmers. In: 2017 IEEE technological innovations in ICT for agriculture and rural development (TIAR)
12. Dey UK, Masud AH, Uddin MN (2017) Rice yield prediction model using data

Autotron-Automated Vending Machine

J. K. Karthik Kumar, L. Anusha, R. Bhavya, R. Ganavi, and G. R. Thippeswamy

Abstract India is one of the most densely populated countries and has the second-highest population in the world. Many public sectors are operating to provide services for the citizens. Public Distribution System is one of the major public sectors that help to provide necessary monthly ration and other essential commodities to the people below the poverty line. It is a difficult task to maintain the PDS scheme in India and the system faces several problems such as the inaccurate supply of products, low processing speeds, waiting for time, corruption, ration hijacking, and so on. To overcome these problems, in this paper we propose the design and result of a completely automated system that makes use of RFID technology to replace traditional ration cards with smart cards. This proposed system aims to provide access to only authorized officials and supply the commodities to the people in need.

Keywords RFID smart card · PDS · CNN-image processing · Regression-random forest · Vending machine · Ration

1 Introduction

In a country like India, it's a really tedious process to maintain the Public Distribution System. The Public Distribution System (PDS) in the country aids the supply and distribution of essential commodities and food products to a large number of poor people through a web of Fair Price Shops at a stake price on a recurring basis. Food Corporation of India persuades and maintains the public distribution system. The food security bill of the Indian government is mainly dependent on PDS. With that in mind, the Indian government still does face a lot of problems such as food hijacking, corruption, etc. Also, the current PDS system is facing multiple challenges such as System Accountability, quality and quantity management, long queues, PDS wastages.

The PDS system was introduced during WW-ll, with time the system has seen a lot of changes naturally in order to meet the safety measures and needs of the people

J. K. K. Kumar · L. Anusha · R. Bhavya · R. Ganavi · G. R. Thippeswamy (✉)
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_88

in the country. The central government obtains and distributes the special essential commodities at a standardized price with the supervision of the PDS. At the moment, India has more than 5.5 lakh shops, forming the largest distribution network in the world. The Fair Price Shops (FPS) provides ration cards to the people depending on the economic conditions. There are two types of ration cards supplied namely:

1. Below Poverty Line (BPL) cards.
2. Above Poverty Line (APL) cards.

This system faces several problems such as corruption, faking of ration cards, renewal of cards, etc. Considering these problems, there is a need to improve the existing system in order to efficiently supply commodities to the right people. Thus, our system is built to provide solutions to these problems by making use of RFID cards instead of the existing ration cards.

Some of the challenges faced by the PDS system are:

- Commodities are not reaching the targeted recipients.
- Leakage of ration commodities during transportation.
- Inclusion/exclusion errors—that is, misclassification of non-poor as poor and vice versa.
- Supply of poor quality rations products.
- Corruption and malpractice being followed.

The aims and objectives of the system are as mentioned:

- Evaluation and validation of smart cards of the people.
- Eliminating the irregularity of supplying products.
- Avoiding fraudulent in the PDS system.
- Enhanced security in PDS.
- Automated quality checking mechanism.
- Auto-generated ration quantity.
- Providing security updates for quality control checks.

2 Related Work

Some of the existing works related to the system are discussed in this section.

Padmavathi in [1] has made of a prototype model that is based on the ATM model. Initially, the customer has to scan the RFID card and after the cards are authenticated he has to manually input the quantity of product, thus making the system semi-automated.

Pedwal and Borkar in [2] make use of RFID cards for the purpose of authentication, and after the verification of data from the database is done it allows users to enter manually the item and quantity b required using push buttons.

Automated Ration Material Distribution System [3, 4] model just allows only authorized users to access ration shops and to obtains materials from it. This system also implements fingerprint biometric for security purposes.

Maheshwari and Tiwari in [5, 6] used RFID tags and passwords for identification and verification purpose. Also the system uses a time password (OTP) for security purposes. This system allows only t specified persons to take the ration.

3 Design

3.1 Proposed System

To overcome the disadvantages in the existing system, we propose a completely automated system over a partially automated system. There is no middlemen intervention in our proposed system. Advanced technologies such as Artificial Intelligence are used for quality prediction by image processing and Machine learning is used for quantity prediction by the regression model. The automatic latch system is implemented once the quality is predicted. This ensures the safety of the products inside the system and protects it from middlemen's intervention.

The proposed system consists of the following modules:

1. Customer registration in the database through a web-based form.
2. Providing an RFID card with a unique barcode to customers.
3. Quality prediction of the product in the system.
4. Feature extraction.
5. Features classification.
6. Customer input by scanning the RFID card.
7. Quantity prediction of the product required by the customer.

Web App

Front End Design- HTML, CSS, JavaScript
Back End Design- Python, NodeJS
Framework—Flask and Django framework.
Database—Firebase authenticator database and function services.

3.2 System Architecture

Figures 1, 2 and 3.

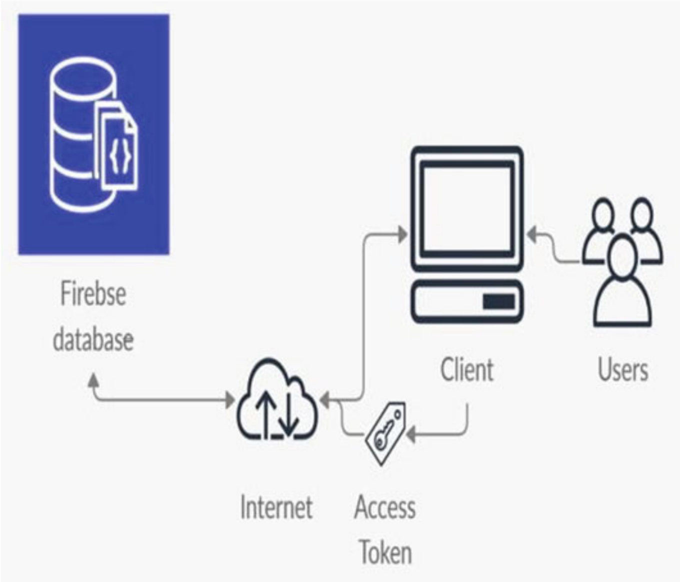


Fig. 1 User data storage flow diagram

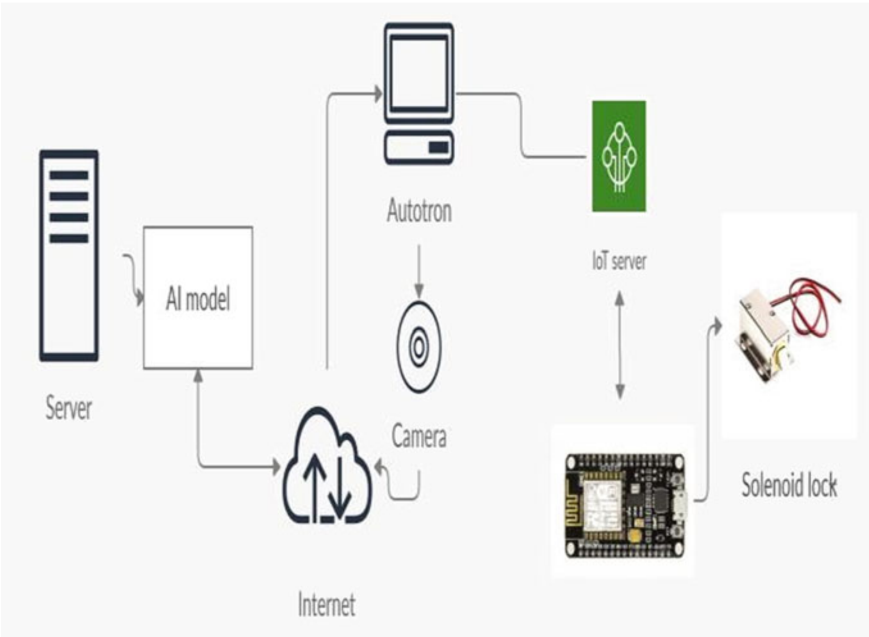


Fig. 2 Quality prediction flow diagram

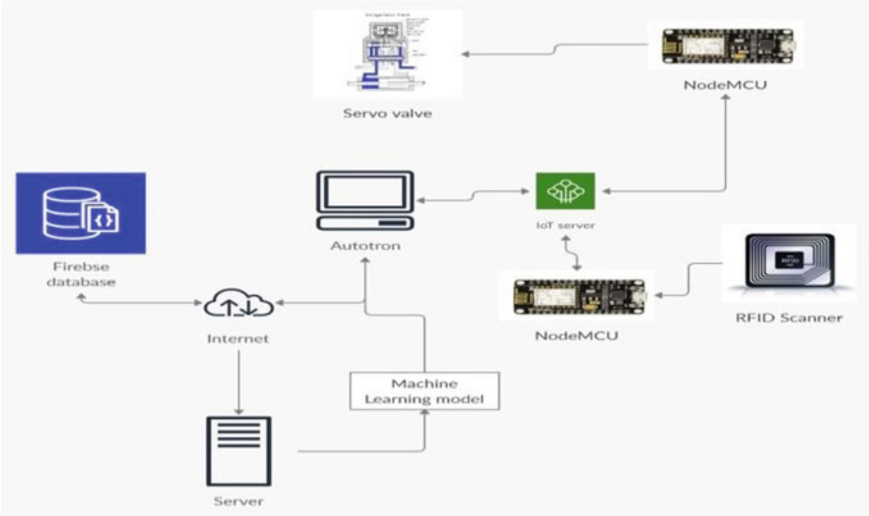


Fig. 3 Quantity prediction flow diagram

4 Methodology

4.1 *Create Database*

Figure 4 represents customer database is created by using Firebase authenticator database and function services which take the input from the front end form.

The authenticator initially enters the customer's phone number and the customer receives an onetime password which is entered in the form for authentication shown in Fig. 5.

Later customer and his family's details are filled in the form and previewed for any corrections. Once the details are correct, it is submitted and stored in the backend database through the internet. An RFID card is given to the customer with his details and a unique barcode.

The barcode is later scanned to access the details of that particular customer when he visits the ration shop to collect ration. The advantage of using the Firebase database is its real-time hot reloading. That is, when changes appear in the backend, all applications connected to the database is automatically updated without even refreshing the page.

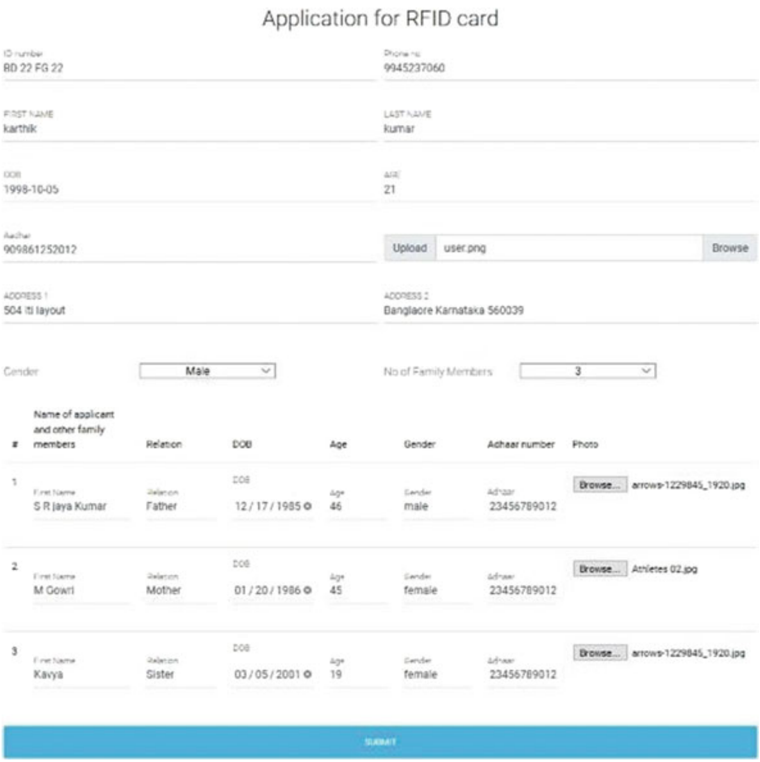


Fig. 4 Front-end forms to fill the customer's details

4.2 Image Acquisition

Image acquisition is the process of creating a digitally encoded representation of the visual characteristics of an object like a physical display or the inside structure of the object. A web camera is used inside the system to capture the images of the product. Initially it is positioned in a dimension that is appropriate to capture the images clearly. Once the image of the product inside the system is captured, it is sent to the AI server through the internet for image processing.

4.3 Quality Prediction

A CNN model is used in the server for classification of the images into good or bad quality where the model is pre-trained with the training images. If the quality of the product is identified to be bad, an alert message is sent to the respective authority. But if the product is identified to be good, the latch in the system is automatically locked,

DETAILS

ID NUMBER BD 22 FG 22

NAME karthik

DOB 1998-10-05

AGE 21

GENDER Male

PHONE 9945237060

ADDRESS 504 iti layout Banglaore Karnataka 560039

Total quantity in KG 15.0

Name of other family members	Relation	DOB	Age	Gender	Adhaar number
S R jaya Kumar	Father	1985-12-17	46	male	123456789012
M Gowri	Mother	1986-01-20	45	female	123456789012
Kavya	Sister	2001-03-05	19	female	123456789012

Fig. 5 Preview page that show complete details before submitting

this is done using NodeMCU. The latch cannot be manually opened by anyone until it is automatically unlocked once the product inside the system is completely emptied (Fig. 6).

CNN models are mainly used to train and test input images which pass through a series of convolution layers. These convolution layers consist of filters, pooling, fully connected layers, and later apply softmax function to classify an object with probabilistic values between 0 and 1.

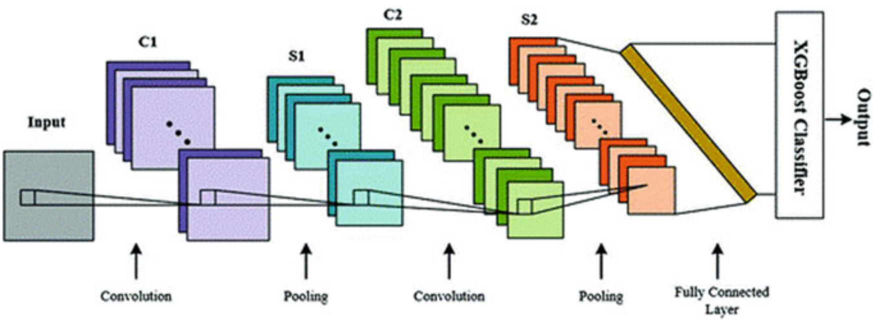


Fig. 6 Image processing using the CNN model

- Convolution layer: It is the first layer to extract features from an input image. It retains and protects the relationship between pixels by studying the image features by making use of small squares of input data. It is a mathematical function that takes two inputs such as image matrix and a filter and a ReLU function is the output.
- Pooling layer: This layer mainly functions to reduce the total number of samples required. There are three types of poling available namely: max, average, sum pooling.
- Fully connected layer: In this layer, we flatten our matrix into a vector and feed it into a fully connected layer like a neural network. We combine these features to create a model.
- Softmax layer: Finally we have an activation function such as Softmax to classify the outputs.

4.4 Customer Input

Once the customer arrives at the fair price shops, his RFID card is scanned for the unique barcode. This unique code is sent to the IoT server through NodeMCU, which in turn sends the barcode to our system is shown in Fig. 7.

The system accesses corresponding customer details from Firebase through the internet. This data is sent to the Machine Learning server through the internet where the quantity of the product is calculated.

NodeMCU is the main IoT module used in our system for data transmission between hardware and the software. It is an open-source firmware and development kit that helps us to build IoT products (Fig. 7).

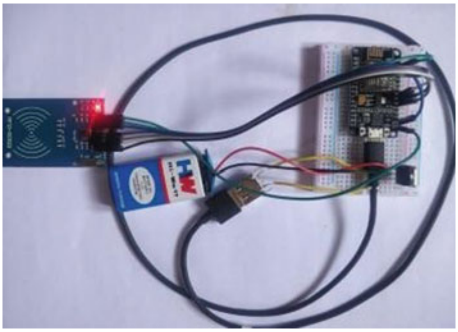


Fig. 7 RFID scanner connection NodeMCU

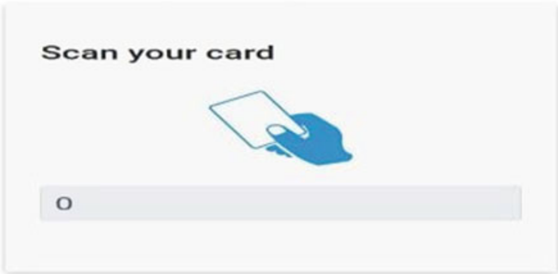


Fig. 8 Page that allows user to scan cards and retrieve ID number from the card

Fig. 9 Page that users with requests to enter the sent to the registered mobile number for security

4.5 Experimental Result

Once the details are accessed from the database and sent to the ML server, the quantity of the product required by the customer is calculated based on the number of members present in his family using the Random Forest algorithm.

The Random Forest model is pre-trained using datasets collected which contains two fields, Number of family members and quantity of product required by corresponding members.

Once user scan cards it retrieve ID number from the card requests sent to the registered mobile number for security, this representation shown in Figs. 8 and 9.

The output obtained from the ML model is then sent to the IoT server which contains the NodeMCU module, where timestamp against the event is generated based on the quantity outputted from the ML model. Once the timestamp is sent to the timer in the Servo valve of the AUTOTRON system, valve get open and timer is get started automatically, similarly the valve get closes once the timer expires (Figs. 10 and 11).

The outputted product is then weighed using load cell for double check and then handed over to the customer. The final prototype model shown in Fig. 12.

5 Conclusion

AUTOTRON-Automated Vending Machine is a standalone system and it is robust in nature, it consists of various analysis tools, so it works very efficiently, so it does not require any external tools for verification as well as validation for system process.

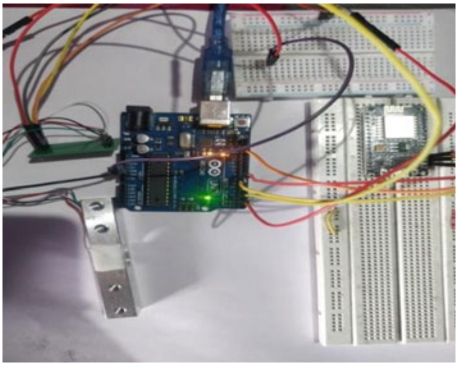


Fig. 10 Arduino and HX711 load cell connection

Fig. 11 Servo valve connection for product output

AUTOTRON totally minimizes the manual efforts in the Fair price shops, so it reduces the human intervention in the supply of food commodities. The system allows users to scan their RFID cards that retrieve the data stored in the database and automatically calculates the quantity of food that has to be supplied.

Based on the experimental result, system is ease of use, efficient, cost-effective and provides good performance.

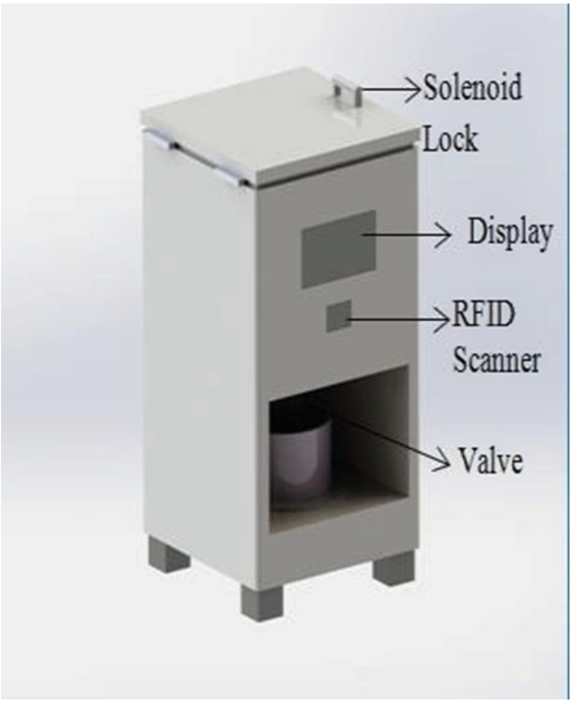


Fig. 12 Prototype model of the system

References

1. https://ieeexplore.ieee.org/document/8256670—Digitalized Aadhaar enabled ration distribution system using smart cards
2. http://ijcsmc.com/docs/papers/March2016/V5I3201699a46 (2016) IJCSMC 5(3):734–739
3. Hebbar S, Malya S, Bommegowda KB, Kumar S, Shrinivasan S (2018) Automated ration distribution system. Dept. of ECE, NMAIT, Nitte Mysuru, Karnataka, India
4. Santhosh Raikar M, Kamlapura GD (2016) Automated ration wending machine for ration shop using RFID card. Dept. of Electricals and Electronics, SDMCET, Dharwad, Karnataka, India
5. Agarwal M, Sharma M, Singh B, Santhanu, IEEE, "Smart ration card using RFID and GSM technique"
6. Maheshwari S, Tiwari M (2016) A smart public ration distribution system. Int J Innov Res Comput Commun Eng 4(3)

Breast Cancer Prediction and Trail Using Machine Learning and Image Processing

Y. Venugeetha, B. M. Harshitha, K. P. Charitha, K. Shwetha, and V. Keerthana

Abstract As per WHO, there are more than 2.09 million cases furthermore 627,000 deaths globally women are diagnosed with breast cancer worldwide annually. It is one of the common cancer in women in India, occurring at any age however in India rise in the early thirties and peak at 50−64 years of age. Only small number of accurate prognostic and predictive factors is used clinically for managing patients with breast cancer. Early detection of this fatal disease is very important which helps in decreasing the morality rate and increasing the survival period of breast cancer patients. This work exploits Mammography for screening and premature identification, analysis as well as processing are solution to improving breast cancer prognosis. To detect breast cancer with mammogram, segmentation of image is performed with the help of Fuzzy C-Means (FCM) technique. As the regions are segmented the required features are extracted, and trained. Finally trained images are classified by the efficient classifier of different classes in mammogram. Texture features are extracted using a feature extraction technique like Gray-Level Co-occurrence Matrix (GLCM), Multi-level Discrete Wavelet Transform and Principal Component Analysis (PCA). Morphological operators are used to make a distinction of micro-calcifications and masses from background tissue also for classification KNN algorithm are exercised. The boundaries of tumor affected region in mammogram are marked and displayed to the doctor, along with area of tumor.

Keywords KNN algorithm · Gray-level co-occurrence matrix · Multi-level discrete wavelet transform · Principal component analysis

1 Introduction

Machine learning is one twig of Artificial Intelligence that provides systems the ability to self-learn by using mathematical models and improve from experience to make decisions through trained data without being unambiguously programmed. The

Y. Venugeetha (✉) · B. M. Harshitha · K. P. Charitha · K. Shwetha · V. Keerthana
Department of Computer Science and Engineering, Don Bosco Institute of Technology, Bengaluru, Karnataka, India
e-mail: vgydbit@dbit.co.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_89

basic assertion of machine learning is to implement algorithms that can accept dataset and apply certain statistical analysis to forecast an output [1, 2]. Image processing is a method to translate an image representation into digital form and perform specific operation on it, in order to get an enhanced image or to extract information from it [3]. Few applications are computer vision, remote sensing, feature extraction, face and finger-print detection etc.

Cancer is malignancy involving in abnormal growth of cell that spread to body parts and has uncontrolled growth [4]. The cancerous image representation process must tackle fundamental problems which cause the ineffectiveness in capturing the textural information and reduced discrimination capability of the features that result in low retrieval performance. Similarity measurement is one of the major tasks of a typical CBIR system that has larger impact on time and accuracy retrieval [5]. The project aims to overcome the problems such as "Which similarity measure is suitable for particular feature type and how to minimize the computation for similarity measurement?" and "Which texture feature is more representative and discriminative for defining the given query mammogram?" by using feature extraction methods, level set and clustering techniques [6–8].

2 Review of Literature

In this paper Turgut et al. [9], have worked on eight various algorithms related to machine learning. Practical application on eight algorithms with the data set and its results of classification were made note. Then in the second case, two different feature selection methods like RFE (Recursive Feature Elimination) and RLR (Randomized Logistic Regression) were applied on the microarray breast cancer dataset and 50 features were chosen as stop criterion. Again, the same eight machine learning algorithms were applied on the modified dataset. Every algorithm result would be compared with each other along with the first case outcome. The methods applied are Support Vector Machine (SVM), K Nearest Neighbor (KNN), Multi-Layer Perceptron (MLP), Decision Trees, Random Forest, Logistic Regression, Adaboost and Gradient Boosting Machines. The end result after applying two different feature selection methods, SVM gave better result in comparison with others. MLP is applied using different number of layers and neurons to examine the effect of the number of layers and neurons on the classification accuracy.

In this paper Varalatchoumy et al. [10], compared four novel approaches used for detecting ROI in Mammographic images based on database and real time images. In 1st approach histogram equalization and dynamic thresholding techniques are used for pre-processing. Region of Interest (ROI) was partitioned from the preprocessed image by using particle swarm optimization and k-means clustering methods. In 2nd approach pre-processing was done using diverse morphological operations. For the identification of ROI, a modified approach of watershed segmentation was used. In 3rd approach histogram uses equalization for pre-processing with an advanced level set approach for performing segmentation. 4th approach is most efficient as it uses different morphological operations and contrast limited adaptive histogram equalization for image pre-processing. A novel algorithm was developed for detection of ROI. Mammographic Image Analysis Society (MIAS) database images alone are applicable for 1st and 2nd approach. MIAS and Real time hospital images are applicable for approach 3 and 4. Graph depict the results of different algorithms comparative study of novel approach that exercised a novel algorithm for proving ROI detection that is considered to be most efficient, accurate and highly reliable approach that can be helpful in detecting tumors in MRI [10].

2.1 System Study

In the proposed scheme, Fuzzy-C-Means (FCM) clustering has been used for finding and segmentation in tumor region. By the same thing segmented region is completely analyzed by using the Multi-level Wavelet-PCA along with GLCM features based on it [11]. After these features are extracted and dataset processed trained also are classified by using the KNN algorithm of machine learning concept [12].

Raw features are the complex-valued responses of a set of multi-resolution Gabor filters were used on it. In that cross operator provides a simple approximation to the gradient magnitude and classical operator is detecting edges and their orientations. Have higher accuracy and less computational time. In the proposed system, classification after the extraction of texture features depends on the shapes of cancer cells in the region of interest and system also takes mean and variance measures into account for classification [13]. System will examine similarity measures and consider feature dataset reduction for similarity measurement improvement by using clustering and feature extraction techniques such as histogram of oriented gradients and speeded-up robust features [14, 2].

2.2 System Architecture

The system architecture is the conceptual model that defines the structure, behaviour, technology and other views of any system. Figure 1 shows system architecture of breast cancer prediction and tracking system. The system mainly consists of four processes. Once the image is acquitted, the system converts the image into gray scale image. By applying the suitable Image segmentation techniques, the system aims at extracting a meaningful object lying in the image. Clustering is one the powerful image segmentation technique involves grouping of data points to cluster, the system implements clustering technique using Fuzzy C-means (FCM). The segmented region is completely analyzed by using the Multi-level Discrete Wavelet Transform, Principal Component Analysis (PCA) along with Gray Level Co-occurrence Matrix (GLCM) features. Totally 13 features are extracted in the system and their pixel values in the form of matrix is stored in database that is in the

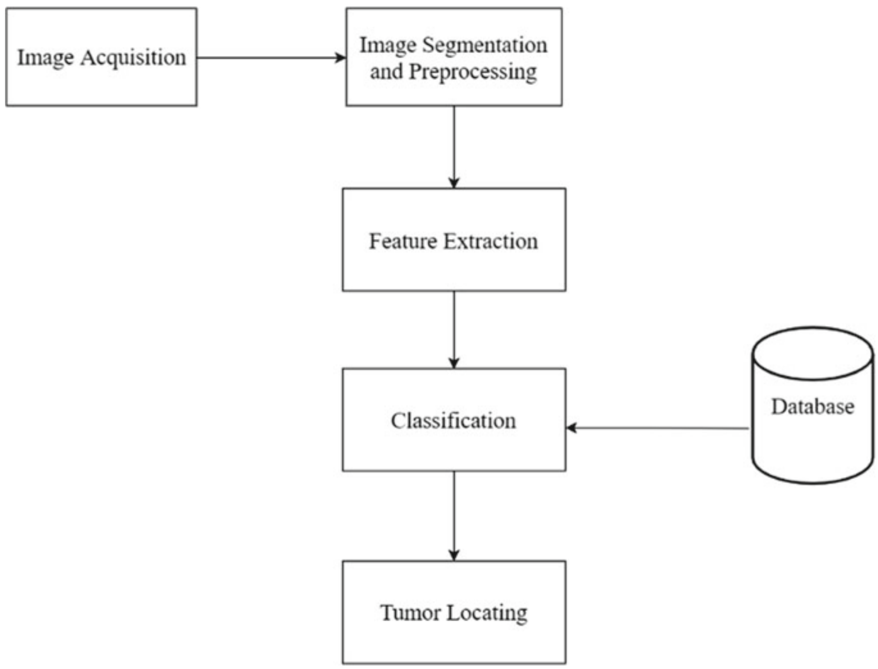


Fig. 1 System architecture of breast cancer prediction and tracking system

db.mat file, some of the features extracted by the system are mean, variance, entropy etc [15, 16, 4]. Then image undergoes classification process with respect to dataset in db.mat file and classifies the image into Benign, Malignant and Normal [17]. The system also performs morphological operations and calculates region properties of the image such as Area, Eccentricity and Euler number. If the image has cancer cells, then the tumor area is computed and displayed by the system along with the boundary detected image [3].

2.3 Activity Diagram

Activity diagram is an important behavioural diagram in UML to describe dynamic aspects of the system.

Figure 2 shows the Activity Diagram of Breast Cancer Prediction and Tracking System; once the image is loaded into the system, suitable image segmentation and pre-processing techniques are applied. Features from the image are extracted using feature extraction techniques. Classification is performed based on the trained dataset stored in the database and image is classified into Malignant or Benign. If image is classified as either Benign or Malignant then the boundary detected image along

Fig. 2 Activity diagram of breast cancer prediction and tracking system

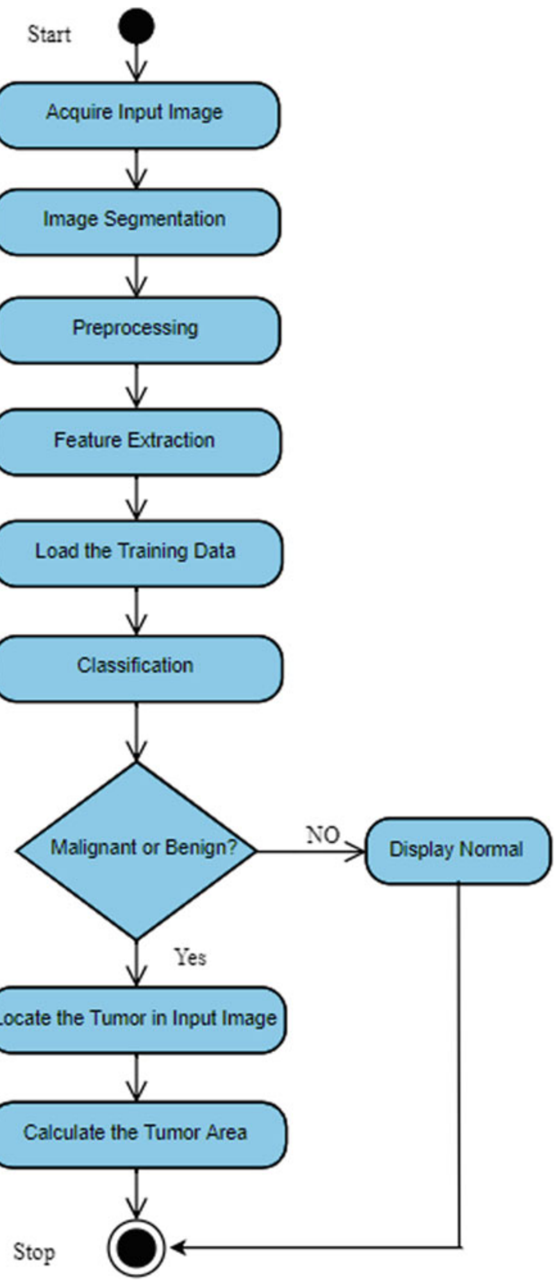


Table 1 Test case-classification

Step	Test data	Expected results	Observed results	Remarks
Step 1	Benign image	Benign classified	Benign classified	Pass
Step 2	Benign image	Benign classified	Malignant classified	Fail
Step 3	Benign image	Benign classified	Normal classified	Fail
Step 4	Malignant image	Malignant classified	Malignant classified	Pass
Step 5	Malignant image	Malignant classified	Benign classified	Fail
Step 6	Malignant image	Malignant classified	Normal classified	Fail
Step 7	Normal image	Normal classified	Normal classified	Pass
Step 8	Normal image	Normal classified	Benign classified	Fail
Step 9	Normal image	Normal classified	Malignant classified	Fail

with tumor area is displayed. If image is not classified as either Benign or Malignant then, the image is classified and displays the result as Normal.

2.4 Testing

Testing is an integral part of software development. Testing process certifies whether the product that is developed compiles with the standards that it was designed to. Testing process involves building of test cases against which the product has to be tested. Testing cases are as in Table 1.

2.5 Results

Figure 3 describes the GUI Page of the project where it contains four buttons, and when first image is pressed i.e., Browse the Image, using this button the image can be fetched from the folder.

Figure 4 shows the conversion of the fetched input image to the gray scale. If the input image is a 2D image then the gray scale image is same as input image, if the input image is a 3D image then it is transformed to 2D gray scale image.

Figure 5 shows the segmented image where it is a pre-processed image where all the noises are removed and it distinguishes between the normal cells and the suspected cells.

Figure 6 give you an idea about the classification of image, since the input image uploaded was Benign, hence the classification is done as Benign.

Figure 7 precise periphery of the cancer cells is detected and laid out.

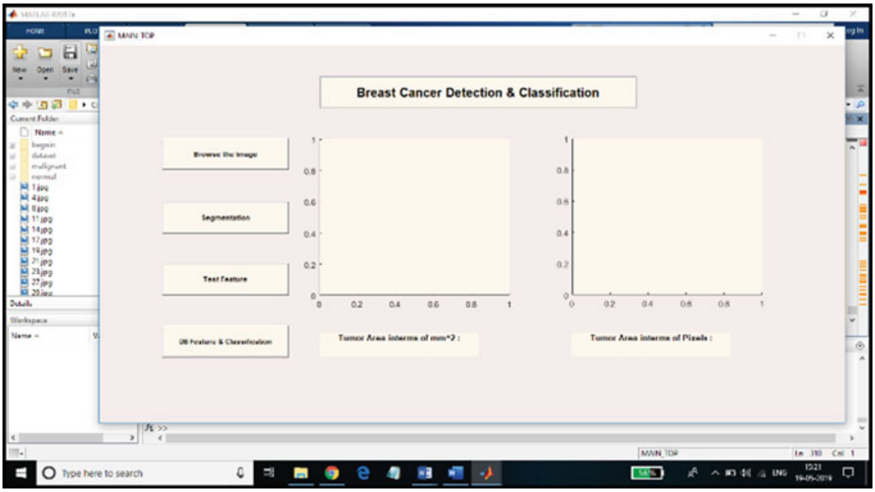


Fig. 3 GUI Page to accept Breast Cancer inputs to detect and classify

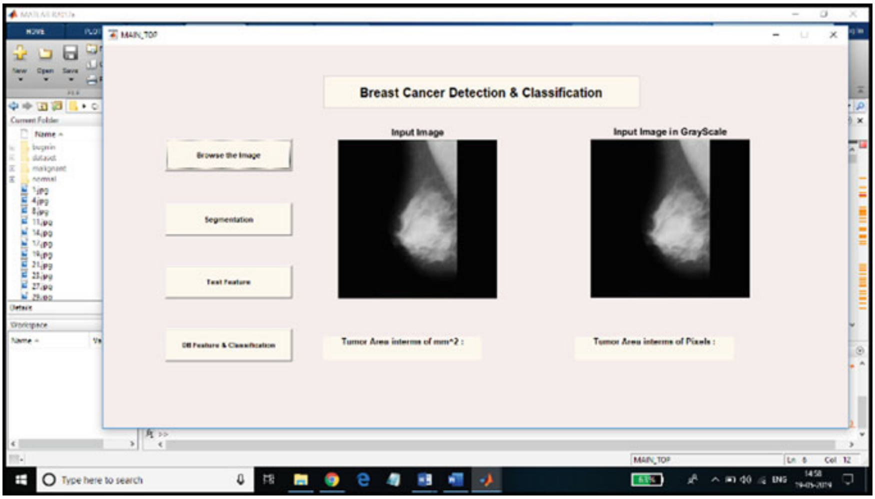


Fig. 4 Gray scale conversion image

3 Conclusion

The system helps to augment the performance of mammogram retrieval by selecting optimal features. The Fuzzy-C-Means (FCM) clustering has been used for Image segmentation. The segmented region is fully analyzed by using the Multi-level Discrete Wavelet Transform, Principal Component Analysis (PCA) along with Gray Level Co-occurrence Matrix (GLCM) features. Totally 13 features are extracted and

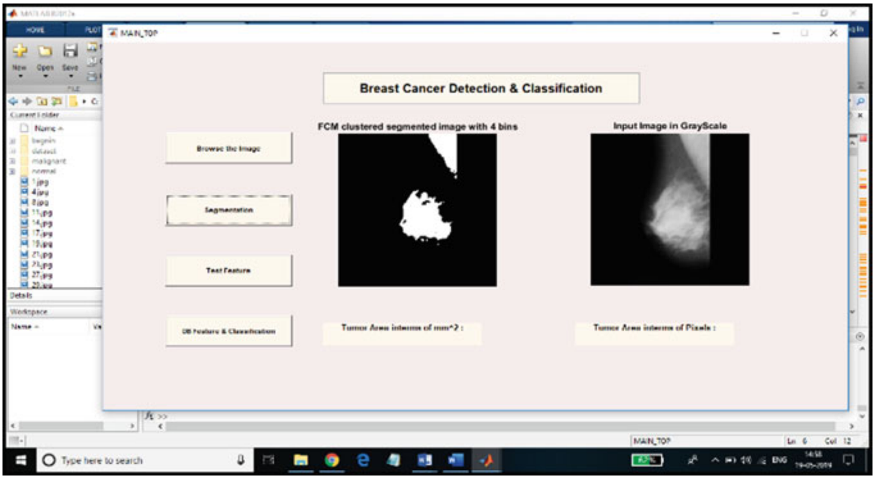


Fig. 5 FCM clustered segmented image with 4 bins

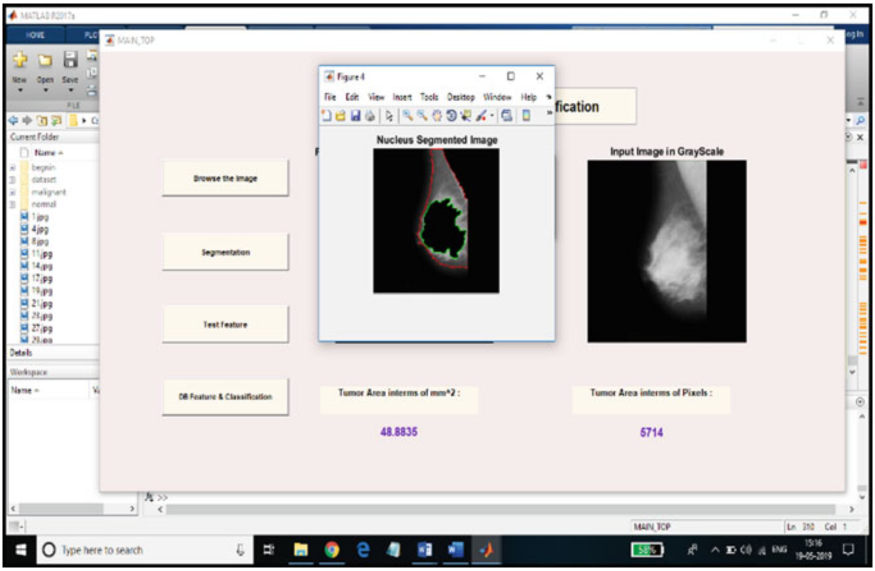


Fig. 6 Breast cancer detection and classification

their pixel values in the form of matrix is stored in database. After the features are extracted and completely trained the system classifies the image into Benign, Malignant and Normal using the KNN classifier technique which mainly depends on the

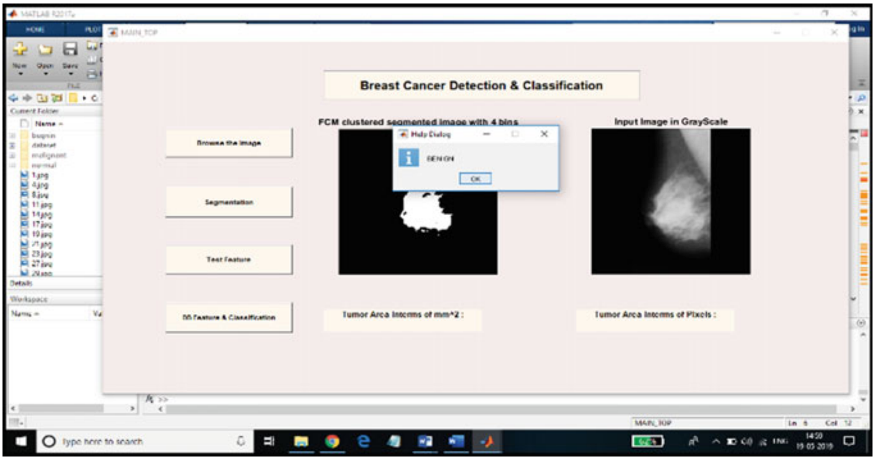


Fig. 7 Nucleus segmented image

shape of the cancer cells in the image. By performing suitable morphological operations, system computes the suitable region properties such as Area, Euler number etc., and displays the boundary detected image along with the tumor area.

References

1. Wadhwani S Classification of breast cancer detection using artificial neural networks. Current Res Eng Sci Technol (CREST) J
2. Elsawy N (2012) Band-limited histogram equalization for mammograms contrast enhancement. In: Cairo international biomedical engineering conference, Egypt. IEEE
3. Egmont-Petersen M, de Ridder D, Handels H (2002) Image processing with neural networks-a review. Pattern Recogn 35:2279–2301
4. Asad M (2011) Early stage breast cancer detection through mammographic feature analysis. IEEE
5. Pisano ED, Zong S (1998) Contrast limited adaptive histogram equalization processing to improve the detection of speculated masses in dense mammograms. J Digital Imaging 11
6. Karahaliou AN, Boniatis IS, Skiadopoulos SG, Sakellaropoulos FN, Arikidis NS, Likaki EA, Panayiotakis GS, Costaridou LI (2008) Breast cancer diagnosis: analyzing texture of tissue surrounding microcalcifications. IEEE 1089-7771
7. Ben Hamad N (2009) Wavelets investigation for computer aided detection of microcalcification in breast cancer. IEEE transactions
8. Ananth KR (2012) A geodesic active contour level set method for image segmentation. Int J Image Graph Signal Process
9. Turgut S, Dagtekin M, Ensari T (2018) Microarray breast cancer data classification using machine learning methods. In 2018 Electric electronics, computer science, biomedical engineerings' meeting (EBBT). IEEE 978-1-5386-5135-3
10. Varalatchoumy M, Ravishankar M (2017) Four novel approaches for detection of region of interest in mammograms—a comparative study. In: Proceedings of the international conference on intelligent sustainable systems (ICISS 2017). IEEE Xplore Compliant, Part Number: CFP17M19-ART, ISBN: 978-1-5386-1959-9

11. Ammu P K, Preeja V “Review on Feature Selection Techniques of DNA Microarray Data” International Journal of Computer Applications, Volume 61– No.12, January 2013.
12. Li BN, Chui CK, Chang S, Ong SH (2010) Integrating spatial fuzzy clustering with levelset methods for automated medical image segmentation. Elsevier Ltd. 0010-4825
13. Sampat PM, Markey MK, Bovik AC (2005) Computer-aided detection and diagnosis in mammography. In: Bovik AC (ed) Handbook of image and video processing, 2nd edn. Academic, New York, pp 1195–1217
14. Ameer Nisha S, Shajun Nisha S, Mohamed Sathik M (2017) A study on surf & hog descriptors for Alzheimer’s disease detection. IRJET. e-ISSN: 2395-0056, p-ISSN: 2395-0072
15. Kocur CM, Rogers SK, Myers LR, Burns T (1996) Using neural networks to select wavelet features for breast cancer diagnosis. IEEE Eng Med Biol 0739-5175
16. Mohanaiah P (2013) Image texture feature extraction using GLCM approach. Int J Sci Res Publ 3(5)
17. Singh S (2009) Breast cancer detection and classification using neural network. Int J Adv Eng Sci Technol

Enhanced and Verifiable Keyword Search over Encrypted Data

P. Rachana, Prerna Mohan, Kavitha Kumari, Priyam Shreyaskar, and D. Komala

Abstract As cloud computing is an attractive energizing technology for the reason that diverse resources are provided with more flexibility and cost effective to the end user, based on their demand. Considering the privacy issue of the cloud users and untrustworthiness of cloud servers the data has to be encrypted before deployed on the cloud. A series of problem arises with the encrypted data storage such as How searching is done on the outsourced data? How to notice user-side verifiability of search results to withstand malicious cloud servers? By what method the user will decide which is the required data that contains the keyword? Pointing to these challenges, in this paper, we introduce an enhanced verifiable keyword search technique called Trustworthy Keyword search Scheme over Encrypted data [TKSE] without interference of third party. Depending on the digital signature's authenticated information, the user is allowed to perform a search on the deployed encrypted data as well as the user can verify the outcome of the cloud data that accomplishes the previous search requirements. This approach specifically realizes server-side verifiability from being framed by malicious data owners in the data storage phase and protects honest cloud servers.

Keywords Blockchain · Cloud computing · Searchable encryption · Verifiability · Stemming · Keyword ranking

1 Introduction

The physical storage and the third party which is highly responsible for cloud maintenance that stores data off-site information in the cloud storage. By using the cloud method user is able to store data and access data globally with the help of internet availability. The cloud storage scheme is very well-known in all the storage system. The physical storage of data can travel all over the different servers by using the third party. Moreover with the security running, the reachability and availability of data by using the parties all the time [1].

P. Rachana · P. Mohan · K. Kumari · P. Shreyaskar · D. Komala (✉)
Department of Computer Science, Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_90

Fig. 1 Secure Blockchain based cloud framework

- The Cryptography allows us to get the solution for the problem. The reason behind the unexpected security issues even though privacy of data doesn't achieved by the existing systems in cloud storage [2]. At all the time, the user can get desired information from anywhere with the help of data decryption scheme (Fig. 1).
- Symmetric method is used for both decryption and the encryption of data. More adaptable technique is present in the advance of the asymmetric key encryption. Symmetric and asymmetric encryption system is available in the cryptographic technique [3]. In our work we are showing the fundamental method for the secure encryption plan by sharing the encrypted data to the different users. But, we are so forth to recognize the included gatherings in a different asymmetric key encryption approach applied on distinct cloud deployment model, a part of which each party in the attack-resistance cycle to understand the effect of safety [4].
- For making more effective and scalable security and key re-generation, trusted authority is responsible this can be done with user side verifiability and server side verifiability [5].

2 Literature Survey

2.1 Security Issues in Data Center with so Many Exceptions

All over the Internet cloud is only has the ability to store different information with so many powerful resources. Different resources is used for optimizing as well as sharing the data for utilization. In the above mentioned features inspiring the organizations for any specific users has to shift their applications to the cloud they can do by their own way [6]. The third party of the service provider gives the different kind of security threats. The administration which is outside the environment users collimated the security issues with the help of different data, information as well as applications [7]. In this analysis cloud computing has the different nature of security issues.

2.2 Policy of Attribute Based Encryption for Outsourcing Secure Data

The data owner as well as the attribution of user provides the secure way of sharing the data [8]. Some major issues such as data security, privacy and data sharing affects the rapid growth of cloud computing with the different hurdles or obstacle. In the attribute-based scheme policy is to suggest us an efficient way [9]. The stored files which is encrypted by integrated structure inside a single structure.

2.3 Various Techniques for Searching the Keyword on Encrypted Data

For decreasing the security and privacy issues, we need to store the data and information in different kind of storage servers in an encrypted form. We implying the functionality in a security purpose by using searching functions[10] it is described that how the storage server allow us to perform not only searching operation as well as to solve the every query without any loss of data. The techniques which we are using that is verifiable secure [11]: they provide us secrecy in verifiable manner in these type of encryption, the server give the control for searching results, they also supports those queries which are hidden, so the untrusted server doesn't allow to give any of the information related to the secret keyword to the user.

2.4 Efficient Way of Searching the Keyword with Public Key Encryption

In a study of the efficient searching of the keyword by using the public key encryption system. Consider Bob is a user who sends data by using mail to another user Alice used encryption method for the data which comes under the Alice's public key encryption. With the help of some gateway we can test keyword like \early is present or not in the specific mail or data [12]. For decrypting all the messages, alice doesn't have ability for the gateway. We are checking the keyword \early whether it is the same keyword or not in the mail by giving the different gateway to the alice. Here we are using the public key type of encryption for searching the keyword. In other example, we are considering different mail server which stores other messages for alice that is directly encrypted publicly by other. In this process alice can only send email to the specific server by using a secret key then that will identify all the messages [13]. The concepts which is related to search the keyword with several methods of encryption using public key.

2.5 Verifiable Symmetric Encryption for Multi-user in a Cloud Storage

In the searchable type of encryption which is having symmetric key in those schemes (or symmetric-key encryption with searching of different keywords), the security against any passive type of adversaries has been so far considered. In this work, worst dense case is security from any active opponent (i.e. including security and privacy) [3]. We will be using Unified Communications security in our work and using non-adaptive adversaries against next UC-security which is very similar to any of the privacy of our ivories as well as security [4].

3 Problem Statement

TKSE consists of entities that should incorporate a data owner, a data consumer, CSP (Cloud Service Provider) and also blockchain. It is important to concentrate on symmetric encryption which is searchable in TKSE. Data search process can be implemented with different approach. The request for search depends on this techniques and the search commitment is to be followed in this technique. Even payment needs to be considered in this approach. Then frame another technique to recognize verification and payment. At last, allow enduser to insist for compensation as a result of search payment in case of invalid search result at Data search level.

4 Existing System

In existing TKSE algorithm, enduser and Cloud service provider are of mutual distrust and can be malicious. Some of the behaviors to be considered for this case, it is as follows:

- The actual data of customers, cipher text data and secret key information may be learned by the malicious CSP.
- The CSP that are malicious may remit partial results or can create search results.
- By deploying cipher text data and index in data storage level to the CSP Malicious enduser can victimize CSP for compensation.
- Without giving the search results that fulfills the needs of enduser Malicious CSP desires to gain search fees from them.
- Similarly without settling the fee for the search, malicious enduser wishes to receive the authentic search results from CSP.

In consideration of all these issues, TKSE is redesigned with the objective as explained further.

- By considering the Privacy in TKSE it is designed to make it impractical for CSP to get any data based on the actual file and secret keys information.
- User-side Verifiability—It maintains reliability by which the CSP cannot produce any authentic search results or send partial results.
- Server-side Verifiability—The main objective is to secure the CSP from the malicious user against cheating.
- No TTP—Without any trusted third party (TTP) fair payment is accomplished.

5 Proposed System

TKSE should have the following four levels which are described below. The top three levels are mandatory but the user claim level is executed by the User when cloud service provider is malicious.

(1) **Initialization Level**

The parameters which has to be used in the subsequent levels are initialized by User and CSP, specifically their keys which are secret and misplaced transactions in the block chain.

(2) **Data Storage Level**

The outposting of the data storage is applied in this level. The four phases namely the data encryption, the index generation, the storage enforcement and the storage confirmation are successively performed as given below:

- *Data Encryption phase*: Consumer encrypts to get that outpost to cloud service provider for confidentiality.
- *Index Generation phase*: Consumer creates index I then conveys it along with the C in cloud service provider.
- *Storage Enforcement*: Firstly the cloud service provider assures later it saves I and C. Then, cloud service provider creates signature digitally according to the C and blockchain saves the signature. Later cloud service provider conveys such data to consumer which helps the consumer to acquire a signature from the blockchain.
- *Storage Confirmation phase*: While acquiring the digital signature from the blockchain, consumer assumes that cloud service provider stores the C.

(3) **Data Search Level**

Cloud service provider looks for her/his storage according to user specifications. It can be confirmed that cloud service provider gets fees for searching from the User only if the User gets correct outcomes. Also four consecutive sub phases are executed which includes search request, commitment search, payment commitment, validation and payment.

(4) **User Claim Level**

TKSE performs user claim level exclusively when cloud service provider breaks to show that the search outcomes reach the specifications of the user ahead runout time. User uses this to declare enough payments of cloud service provider without considering how the cloud service provider behaves.

5.1 System Architecture

The data storage phase has the greater significance as shown in Fig. 2 which includes encryption of data, keyword identification, index generation and storage confirmation.

The process of keyword identification and weighting is explained in the Eq. (1). The end result of this will be an array of the keywords and its frequency array. Both of these are uploaded to the server.

Another user end has to enter a searchable keyword which is encrypted. This is matched against the array of keywords and the relevant files are listed. The listing is ranked according to the frequency of the particular keyword present in a file. The file

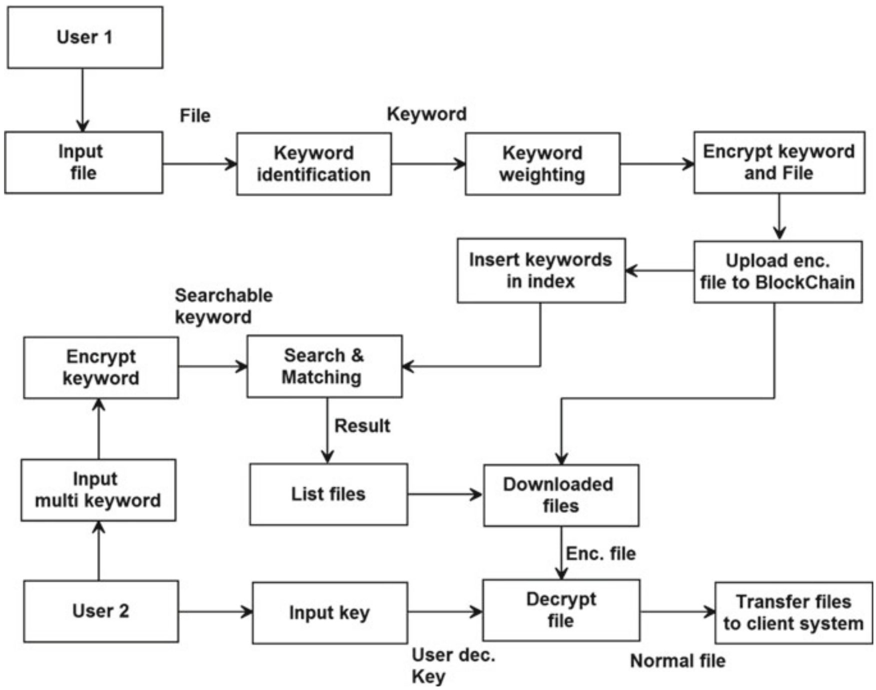


Fig. 2 System architecture

with the highest frequency of the entered keyword will be placed at the top. At last the user has an option to download the file of his choice to his system. The downloading process of the file includes the decryption of the file and then it is saved to the client system. The decryption key at each step has to be entered by the user so as to obtain the file and hence this ensures verifiable mechanism.

5.2 Keyword Weightage

The flowchart shown in the Fig. 3 shows **Keyword weightage** that performs the operation of taking out unnecessary words, stemming and storing in the array and search operations.

- *Keyword Stemming* [4] modifies various keywords by using variations. The variations can be in the form of prefixes/suffixes which can be appended to the root keyword. The root keyword is the base keyword whose variations are needed to be included. Hence this is a process of bringing down each word to its stem. As an example "included", "includes", "including" all become just the root word (stem) "include".
- After performing keyword stemming, the end results are the two arrays, one for the keywords and the other one for the corresponding weights.
- The mechanism is to check whether a word in the file is a keyword, if yes then increment the corresponding weight and if it is not then add it to the array.

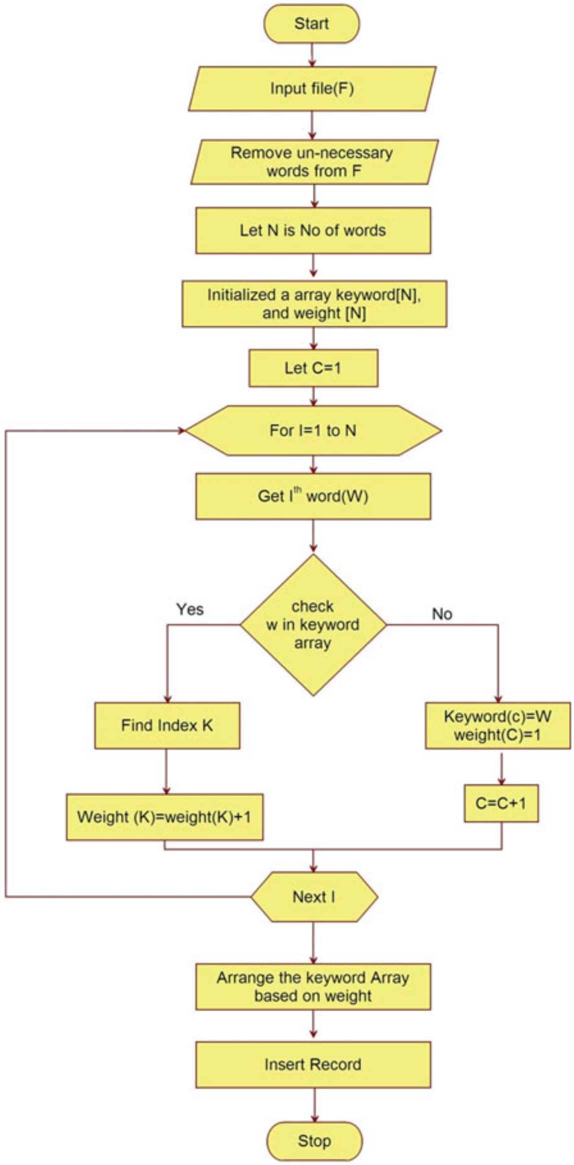
- *Keyword Ranking*—Eventually the outcome of the keyword search shows the list of all files that consist the specified keyword but having a ranked order [4]. The ranking is according to the keyword weight or density of the files. The density is calculated as shown in the below equation [5].

$$Keyword\ Density = \frac{Number\ of\ times\ you\ repeat\ Keywords}{Total\ words\ count\ of\ the\ article} \times 100 \quad (1)$$

6 Conclusion

The aim is realizing to search secure keyword over symmetric encryption of data against the malevolent users and CSP. So we came up with TKSE over encrypted cloud storage. In the TKSE, verifiability over server-side is allowed for only once in literature with the user-side verifiability and privacy, it shields genuine cloud servers by being framed in data storage phase by malicious owners of data. For an easy user experience keyword management algorithms are used. The analysis that we made concludes that performance evaluation and the security analysis which here indicates that TKSE is very efficient and secure. The number of transaction in data storage can also be improved with the search phase by enhanced techniques. The server

Fig. 3 Flowchart for keyword weightage



side enabled can be enhanced more by adopting a better hash function with better encryption methods which can protect the malicious operations.

References

1. Chen X, Li J, Susilo W (2012) Efficient fair conditional payments for outsourcing computations. IEEE Trans Inf Forensics Secur 7(6):1687–1694
2. Cai Z, Yan H, Li P, Huang Z-A, Gao C (2017) Towards secure and flexible EHR sharing in mobile health cloud under static assumptions. Cluster Comput 20(3):2415–2422
3. Kurosawa K, Ohtaki Y (2012) UC-secure searchable symmetric encryption. In: Financial cryptography, vol. 7397. Springer, Berlin, Germany, pp 285–298
4. Chai Q, Gong G (2012) Verifiable symmetric searchable encryption for semi-honest-but-curious cloud servers. In: Proceedings of IEEE International Conference on Communication (ICC), pp 917–922
5. Zhangj Y, Deng RH, Shu J, Yang K, Zheng D (2018) TKSE: trustworthy keyword search over encrypted data with two-side verifiability via blockchain. IEEE Trans 6:31077–31087
6. Ali M, Khan SU, Vasilakos AV (2015) Security in cloud computing: opportunities and challenges. Inf Sci 305:357–383
7. Karajeh H, Maqableh M, Masa'deh R (2016) Privacy and Security issues of cloud computing environment published
8. Saraswathi M, Bhuvaneswari T (2018) A trusted solution for secure outsourced data using modified key policy attribute based encryption. Int J Pure Appl Math 118(5):445–453
9. Wan Z, Liu J, Deng RH (2012) HASBE: a hierarchical attribute-based solution for flexible and scalable access control in cloud computing. IEEE Trans Inf Forensics Secur 7:743–754
10. Song DX, Wagner D, Perrig A (2000) Practical techniques for searches on encrypted data. In: Proceedings of IEEE Symposium on Security and Privacy, pp 44–55
11. Han P, Liu C, Fang B, Wang G, Song X, Wan L (2016) Revisiting the practically of search on encrypted data
12. Boneh D, Di Crescenzo G, Ostrovsky R, Persiano G (2004) Public key encryption with keyword search. Eurocrypt 3027:506–522
13. Hsu S-T, Yang CC, Hwang M-S (2013) A study of public key encryption with keyword search. Published

Camouflagge Technique Based Multi-functional Army Robot

N. Naveen, Rohan Raj Kumar Saini, Raseem Riswan, R. Sriram, and Shruthi Kumari

Abstract We are developing in the field of science in a rapid phase in order to create the technology which can make human life easier. Nowadays, many innovations and inventions are made in the field of defense to reduce the human lives. One such invention is Camouflage Robot that plays a vital role in reducing the damages that occur during disasters. This robot is designed to work on the principle of Chameleon's camouflaging technique which is one of the most natural and primitive methods for avoiding detection. The key objective of this paper is to implement a multifunctional army robot consisting of various sensors. Hence, the proposed system employing blue-tooth reduces errors at defense and protects the nation from threats.

Keywords IOT · Multifunctional · Internet of Things

1 Introduction

In the modern combat techniques employed by various militant forces across the globe, stealth and ability to manoeuvre inaccessible areas play a key role. The idea of the proposed system is to use robots which are capable of disguising itself in order to infiltrate the enemy campsite. The word robot refers to "A machine which is capable of performing complex series of actions automatically that is programmable by a computer." These robots used in defence are usually employed with the integrated system, including cameras, sensors and video screens.

SAR is an active remote sensor (i.e., it carries its own illumination and it is not dependent on sunlight) which makes it functional in all-weather and day-and-night operating conditions (OCs). Focused SAR images are different from their optical counterparts in many aspects. Some distinctive characteristics of SAR images include (1) a target size does not vary with the distance between the SAR sensor and the target, (2) the information about the imaged scene is carried in the magnitude of the radar backscatter (i.e., for a singlechannel SAR),1 (3) large specular reflections

N. Naveen (✉) · R. R. K. Saini · R. Riswan · R. Sriram · S. Kumari
Department of Computer Science and Engineering, Don Bosco Institute of Technology, Bangalore 560074, India
e-mail: naveensetty@dbit.co.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_91

pertaining to a microwave mirror-like behavior result from man-made scenes and some natural objects (e.g., rocks), and (4) high sensitivity to the very small changes in the target's pose and configuration for various reasons including: the shadowing effect, the interaction of the target's backscatter with the environment (e.g., clutter, adjacent targets, etc.), projection of the 3-D scene (including the target) onto a slant plane (i.e., SAR's line of sight (LOS)), and a noise-like phenomenon known as speckle due to the backscatter's dependence on the coherent combination of returns from points in the imaged scene.

The development of modern high-resolution synthetic aperture radar (SAR) was led by three key innovations. The first of these was pulse compression which enabled radars to range-resolve closely spaced targets. The most widely used pulse compression technique is linear frequency modulation (LFM), also known as chirp modulation, which was introduced in the early 1950s. Second, by 1951, it was possible to resolve closely spaced targets in angular position relative to the antenna beam center of side looking airborne radars.

Automatic Target Detection (ATR) deals with the information output from one (or more) sensor(s) aimed at a scene of interest. It generally refers to the use of computer processing capabilities to infer the classes of the targets in the sensory data, and to (optionally) characterize some attributes of interest such as articulation, orientation, occlusion, sub-class and so on, without human intervention. The term ATR originated in the military in the early 1980s under the Low Altitude Navigation and Targeting Infrared for Night (LANTRIN) program [1]. Today, ATR technology is important in both military and civilian applications. The ATR problem is a part of the general broad problem of machine vision; namely, how can computers be configured to do what humans do efficiently and naturally? Target, clutter and noise are three terms of military origins associated with ATR and are dependent on the application of interest. In the case of SAR imagery, target refers to object(s) of interest in the imaged scene. Clutter refers to either man-made (e.g., building, vehicles, etc.) or natural objects (e.g., trees, topological features, etc.) which tend to dominate the imaged scene. Noise refers to imperfections in the SAR image which are result of electronic noise in the SAR sensor as well as computational inaccuracies introduced by the SAR signal processor. In the literature, there is a spectrum of ATR problems ranging from classifying a pre-known signature in a well-characterized clutter to recognizing the source of signature that varies greatly with pose and state, and is located in a highly complex and probably occluded scene [2].

The key objective of Camouflage Robot is to reduce the human losses in terrorist attack or military operations. Many military organizations take the help of robots in the risk prone areas. Camouflage robot can play the role of a virtual spy, which can enter into enemy area collecting sensitive information via attached camera and sending them back to the controller. Robots can be made to interact and cooperate more closely with human beings by incorporating additional features such as robustness and autonomy. A versatile perception and recording of different parameters in this robot is accomplished using a multi sensor platform. In this system an interfacing module is incorporated to remotely sense the object parameters using IoT.

1.1 Related Literature

ATR deals with the information output from one (or more) sensor(s) aimed at a scene of interest. It generally refers to the use of computer processing capabilities to infer the classes of the targets in the sensory data, and to (optionally) characterize some attributes of interest such as articulation, orientation, occlusion, sub-class and so on, without human intervention. The term ATR originated in the military in the early 1980s under the Low Altitude Navigation and Targeting Infrared for Night (LANTRIN) program [3]. Today, ATR technology is important in both military and civilian applications. The ATR problem is a part of the general broad problem of machine vision; namely, how can computers be configured to do what humans do efficiently and naturally? Target, clutter and noise are three terms of military origins associated with ATR and are dependent on the application of interest.

In the case of SAR imagery, target refers to object(s) of interest in the imaged scene. Clutter refers to either man-made (e.g., building, vehicles, etc.) or natural objects (e.g., trees, topological features, etc.) which tend to dominate the imaged scene. Noise refers to imperfections in the SAR image which are result of electronic noise in the SAR sensor as well as computational inaccuracies introduced by the SAR signal processor. In the literature, there is a spectrum of ATR problems ranging from classifying a pre-known signature in a well-characterized clutter to recognizing the source of signature that varies greatly with pose and state, and is located in a highly complex and probably occluded scene [4].

Rohan Raj Kumar Saini, presented that Haar Cascade Classifiers is used to propose the new system to trace out the unknown and unauthorized persons and the robot will take the required action of its own [5].

Raseem Riswan proposed that the target should be detected automatically by PIR Sensors [6]. If object tries to cross the boundary, then PIR generates signal and gives to Arduino and it produces alert message. If object continues to go further then Arduino gives command to shoot that object. The main advantage is automatic target detection and shooting gun is wirelessly.

R. Sriram proposed that large scale border security system modeling and simulation can be done with the help of OPNET [7]. Shruti Kumari proposed that several mage recognition and tracking algorithms which can perform on numerous datasets. The proposed algorithm employs a wide range of datasets, ranging from hand gestures to shapes and objects to handwritten manuscript text. The performance of algorithms namely S.U.R.F Algorithm, Blob detection method, Template matching algorithm have been analyzed. Factors like invariance to rotation, processing speed, flexibility to use for various data sets, scale and illumination have also been analyzed. In addition, every image is identified using its unique set of features [8].

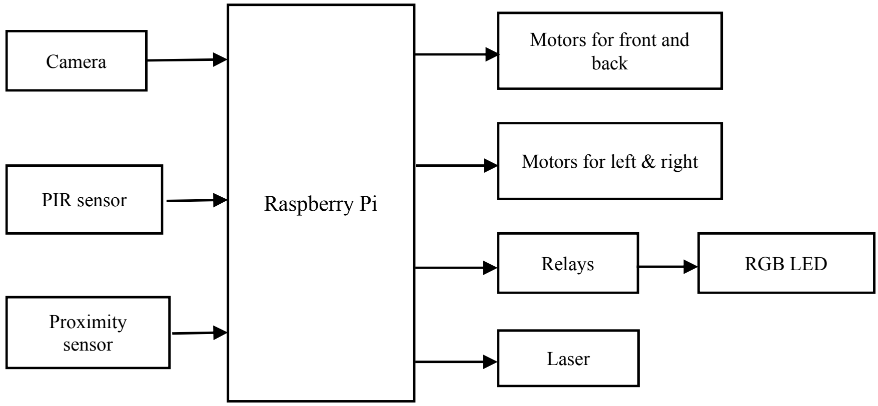


Fig. 1 Block diagram of multifunctional army robot

2 System Description

2.1 Block Diagram

See Fig. 1.

2.2 Proposed System

A motion tracking airsoft (or nerf gun) turret based on Raspberry Pi employed with autonomous motion detection, fires the gun when motion is detected. It also has an interactive mode, which allows the user for manual manually via a keyboard. Motion Detection employs computer vision and open CV for tracking moving objects. The system includes a colour sensor camera for surveillance purpose and camouflaging feature. This prevents the robot from easy detection by enemies. The proposed system uses Wireless transceiver for communication between transmitter and receiver. The robot can silently enter into enemy areas and transmit sensitive information via camera. The robot movement can be wirelessly controlled by a computer, and the robot is entirely battery operated. There are two modes: Interactive and Motion Detection. Interactive mode allows the user to control the turret remote and stream live video. Motion Detection uses computer vision and open CV for tracking moving targets in front of the camera. The system consists of one colour sensor camera as part of camouflaging feature and one for surveillance purpose. The robot changes the colour based on the surface, and because of this it makes it difficult for the robot to be identified by enemies. The system uses a wireless transceiver for communication between transmitter and receiver.

3 System Implementation

On robot side both the Input and Output is done whereas on PC side all the image processing is done. Robot has all these input devices like color sensing camera, video feeding camera and obstacle sensor to collect all the required data for processing. This data is then transmitted wirelessly to PC via Wi-Fi trans-receiver.

Computer then does the processing of the received data using various algorithms for image processing. It determines the color of background and transmits this data to robot. Through PC the user can also movement commands to the robot. One of the important functions of PC is to display the live video feed received from the robot. All the transmission is done serially using Wi-Fi trans-receiver.

Robot can output the received color by changing the color of LEDs covering the chasse. This is done by turning on one of the three relays present on the robot. The PC using which we will be able to handle the movement of the robot (Figs. 2 and 3).

Gun Targeting: A Raspberry Pi 3 based motion tracking airsoft (or nerf gun) turret has been built for gun targeting.

Camouflage Technique, the robot is being camouflaged and is controlled from afar an object.

Land mine Detection: The major problem faced by many countries around the world is the problem of Landmine removal, and the situation can be compounded by natural disasters or land development.

Rover Movement: Raspberry pi s used for two mode of action.

Image: A webcam is used for capturing real-tie images and sending over a computer network via Wi-Fi or Ethernet.

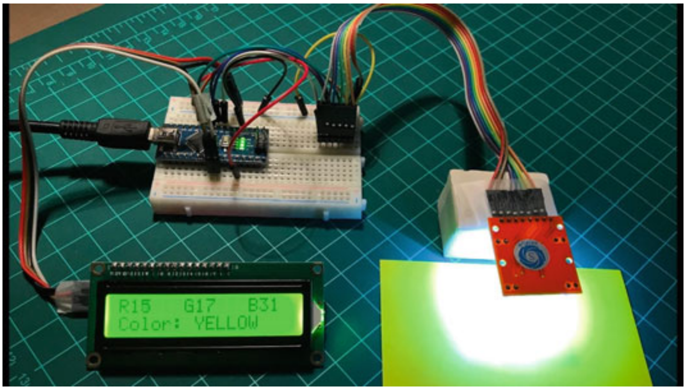


Fig. 2 Color sensor sensing yellow color

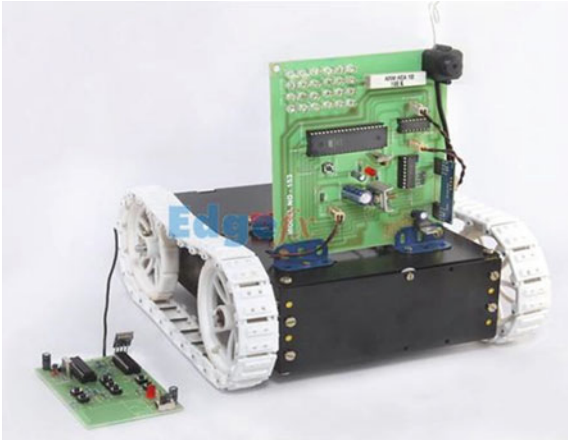


Fig. 3 Army robot

Fig. 4 Robot arm digging bomb

4 Results

It has basically three applications, first one is detecting bomb and disposing, second is to change colour according to input given to camera and third is to identify unknown person using R Sensor and shoot them (Figs. 4 and 5).

5 Conclusion

IoT devices will be more and more important in the future. Even though there may never be a foolproof malware detection solution, we always have to be consistent in our efforts to defend our systems. It is important that we, as cyber defenders, always maintain a lead against cyber attackers.

In this project, we implemented an Internet of Battlefield Things (IoBT) malware detection approach that s based on class-wise selection of OpCodes sequence as

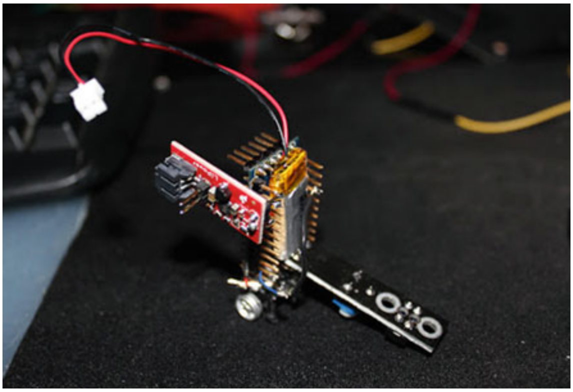


Fig. 5 Meta sensor

a feature for classification task. A graph of the important features of each sample was created and a deep Eigen space learning approach was used for classification of malware.

In the future, we plan to evaluate our approach against larger and broader datasets and also improve our system to detect even the most difficult and camouflaged malware.

References

1. Bhanu B (1986) Automatic target recognition: state of the art survey. IEEE Trans Aerosp Electron Syst AES-22(4):364–379. https://doi.org/10.1109/TAES.1986.310772
2. Alkhathami M, Alazzawi L, Elkateeb A (2017) Large scale border security systems modeling and simulation with OPNET 2017. In: IEEE 7th Annual computing and communication workshop and conference (CCWC). Las Vegas, pp 1–8. https://doi.org/10.1109/CCWC.2017.7868360
3. Padilla R et al (2012) Evaluation of Haar cascade classifiers designed for face detection. World Academy of Science, Engineering and Technology. Int J Comput Electr Autom Control Inf Eng 6:466–469
4. Jayanthi N, Indu S Comparison of image matching techniques. Int J Latest Trends Eng Technol 7(3):396–401. https://doi.org/10.21172/1.73.552
5. Shinde H, Sonawane K, Rane P, Pathak A, Chandak S (2018) Camouflage color changing robot. JRASET
6. Abdalla GOE, Veeramanikandasamy T (2017) Implementation of spy robot for a surveillance system using internet protocol of Raspberry Pi. In: 2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT). Bangalore, pp 86–89. https://doi.org/10.1109/RTEICT.2017.8256563
7. Schmidt A, Kasiński A (2007) The performance of the Haar cascade classifiers applied to the face and eyes detection. In: Kurzynski M, Puchala E, Wozniak M, Zolnierek A (eds) Computer recognition systems 2. Advances in soft computing, vol 45. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-75175-5_101
8. Fitriyani L, Yang C-K, Syafrudin M (2016) Real-time eye state detection system using Haar cascade classifier and circular hough transform Norma. In: 2016 IEEE 5th Global conference on consumer electronics

An Approach Towards E-Commerce for Agriculture with Modern Technologies

Athreya N. Patel, Afnan Ahmed, and B. R. Rohini

Abstract Agriculture plays an important role in the development of the economy of a country. It provides lots of opportunities to innovate and apply the innovation. Trading in agriculture from the farmer to the consumer goes through a lengthy process where farmers negotiate an amount for the generated yield with a middleman. The middleman has a set fee of brokerage on yield from farms based on quality. This process consumes significant time that can deteriorate quality while the farmer has no return of investment. In order to overcome these issues, an E-Commerce solution can be found which can be transparent enough to make sure farmer gets the return of investment with profits and is used without a middleman (broker) to sell the yield. There is no proper solution that exists for a transparent trading between farm and consumers. This paper highlights the importance of E-Commerce in Agriculture with less/no brokerage and that also helps a farmer to invest in ways that increase the yield.

Keywords Agriculture · E-Commerce · Transparent solution · Brokerage · Farmer-to-Consume · F2C · B2C

1 Introduction

- Agriculture plays an important role in the development of the economy of a country. It provides lots of opportunities to innovate and apply the innovation. Trading in agriculture from the farmer to the consumer goes through a lengthy process where farmers negotiate an amount for the generated yield with a middleman. The middleman has a set fee of brokerage on yield from farms based on quality. This process consumes significant time that can deteriorate quality while the farmer has no return of investment. In order to overcome these issues, an E-Commerce solution can be found which can be transparent enough to make sure farmer gets the return of investment with profits and is used without a middleman (broker) to sell the yield. There is no proper solution that exists for a transparent trading between farm and consumers. The objective of this concept

A. N. Patel (✉) · A. Ahmed · B. R. Rohini
Department of Computer Science, Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_92

is to explore E-Commerce in the field of Agriculture with less/no brokerage and that also helps a farmer to invest in ways that increase the yield [1].
- It is idealistic method to acknowledge precise requirements in the early stage of project. It is tough and costly to incorporate changes with great overhead and expense [2].
- There has been a major development in technology over the decade which has greater impacts in the arena of mankind. One of such change is in the field of Agriculture business via virtual market-Ecommerce. The predominant features of the Internet have supported the development of E-commerce of agricultural products. In the category of agricultural products, transportation and storage requirements are a high-risk task and hence opting a good supplier for the e-commerce of agricultural products is significant [3].
- The solution to E-Commerce illustrated in this paper focuses on how the farmer to consumer model will help in development of direct transactions between farmer and consumer with no brokerage and more profit to farmer. It also highlights on how the model is highly scalable and regional with the amount of features it offers using the modern technology. The solution should offer “easy to learn and return” interface which will encourage user to order or list the agro-based/agricultural yield. This model is not only applicable in India, but also the rest of the world.

2 Background

2.1 Current Scenario

Currently the agriculture commerce is mostly offline and the wide use of online solution for commerce is not thought of. Traditional commerce i.e. offline has many flaws which can cause low yield, less encouragement, high brokerage charges and less effective transportation. High brokerage and farmer to dealer and to lots of middleman will affect the cost increase. This leads to less profit to farmer. In India, farmer obtains maximum of Rs. 20/kg for high quality rice.

Net income shows what is available to pay for principal repayments, new farm investments, family living expenses and family off-farm investments. Net income with the current scenario is less. This might not encourage them to innovate and also leads to low profit margin [4, 5].

$$\text{Profit margin ratio}(\%) = \text{Net farm income}/\text{Gross farm income} * 100$$

2.2 E-Commerce Scenario

Currently the agriculture sector consists of farmers, brokers, wholesalers, retailers and secondary retailers. The transactions consist from mega to micro. The tariffs are collected from the farmer to consumer. This leads to cost increase in the sector.

In online processing cost is lowered by trading costs or by transfer costs. Trading costs lowers when the search by buyers and sellers is facilitated, when the costs of posted prices are lowered, when negotiations between geographically separate buyers and sellers are provided. When transactions involve goods that can be digitized, such as money, information products, or rights and obligations, transfer costs are saved on top of trading costs. Communication costs on the Internet are independent of data volume and distance between sender and receiver, geographic distance is not important in negotiation and finding out. Online medium has the prolifying impact on trading costs when information is digitized, e.g. when cattle or fresh product are marketed by digital video rather than by physical display [6].

E-Commerce readiness of agriculture [7] consists of participation from both ends i.e. buyers and sellers with access to the Internet. Currently, Agricultural products are listed from the retailers who have sorted out the product and have listing of the product in several E-Commerce companies. This occurs on both fresh products like vegetables, fruits and milk and essential products such as staples, etc.

The concerns of e-commerce in agriculture are: Which e-commerce business models are appropriate for agricultural markets? What is the impact of e-commerce on farms, agribusiness firms, markets, and rural communities? What is the role of government with e-commerce in agriculture? [8]

E-commerce transactions are classified according to the partners involved—consumers, business, and government. The classification is as follows: business-to-consumer (B2C) and business-to-business (B2B). Of the two, B2C e-commerce currently receives most public attention [9].

3 Proposed Solution

E-Commerce is where people search for a product and the product is picked or delivered after the transaction is completed.

The importance of user to accept, adopt and customize technologies to increase the potential of ICTs, with the environment are believed to assist the technology development of a society. Research on e-commerce have been widely cited, but research specific on B2B ecommerce readiness assessment for SME in agro based industry is rare and classified. The unique aspects of the SMEs in agro-food industry in Malaysia with regards to the adoption of B2B e-commerce have not been widely examined. Further, the current practical B2B e-commerce assessment models are largely based on the experience of developed countries which rather different from that in developing countries. B2B e-commerce implementation also is an enormous task due to the complexity of the process, thus, the simplified version on what are required and what are the possible challenges exist may fill up the technology gap [10].

Implementation of a solution where e-commerce is driven using modern technology and a farmer to consumer model for freshly sellable product/a product which are natively produced using machines. Also, the producer to end-user model for produced goods from primary supplier which also can be on-boarded as a seller. All these are technology driven and will help produce more income [11].

The proposed solution requires proper internet connection and a smart-phone/a computer. The internet literacy is also important. As a developing nation, India is developing its telecom infrastructure rapidly and over 51% are using smartphones [13]. The application should support native languages such as Hindi, etc. in India to provide easy to understand solution. The ease-to-use design should be integrated where it should be a PWA (Progressive Web Application) and have responsive and pure design [12].

Progressive Web Application is a solution which provides some features of a native application which combines the best of web and mobile apps for an enriched experience. These websites designed to be PWA as well as a website by itself should also follow guidelines on Performance, SEO, Protocols used and Responsive design.

1. Protocol i.e. the PWA should be served via HTTPS which is considered to be safe.
2. Caching—It should work offline. For this Service Workers are used to cache the data which are used often. This would help in rural where the proposed solution is being used.
3. Installation—A popup/manual installation should be possible in a click. These applications need not be installed via application stores. It is just a website which is responsive and has stored cache which performs basic functions similar to native applications.

Ease-of-use also consists of language support and responsive design. Responsive design is a way where in the website can fit any type of device and screen. It can be examples like desktop, mobiles and tablets.

With the use of this modern technology present in 2020, the e-commerce can provide information and also let farmer and consumer decide a better product to sell or buy. The use of information technology has lots of scope of advancement in agriculture including demand wise crop farming and data-based crop farming. The demand-based farming and data-based farming techniques can be presented in a way which can be understood by a farmer. These techniques potentially can help the growth of agriculture industry. The data can be collected from our solution based on number of orders of a specific agro product and also the quantity of the product bought can provide an idea on costs and the product quality. This will potentially empower farmer and encourage him to try modern tools like tractors, tillers, etc.

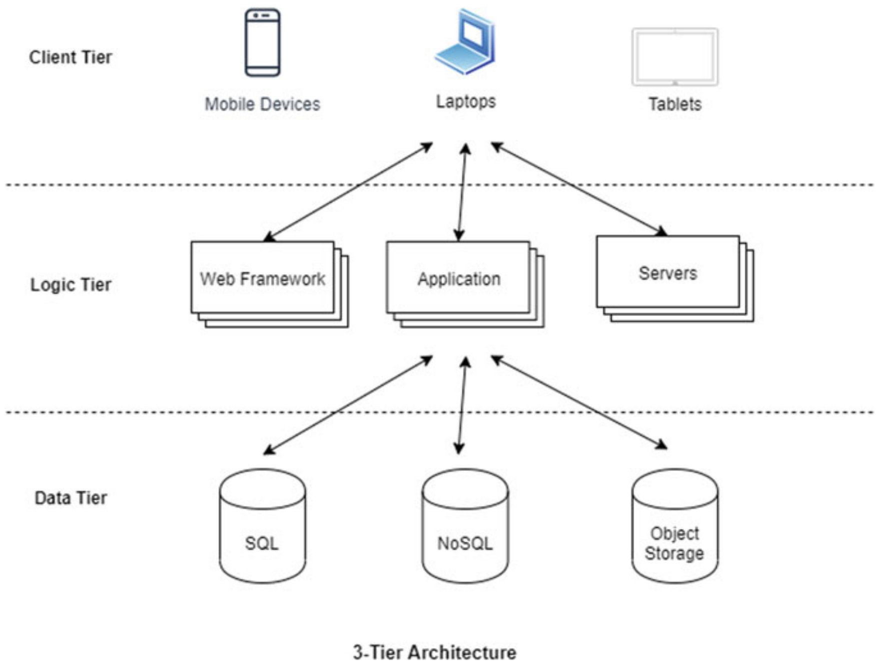


Fig. 1 Architecture of e-commerce for agriculture

4 Implementation

4.1 System Design of E-Commerce Architecture

The E-Commerce architecture includes a 3-tier architecture consisting of Client Tier, Logic Tier and Data Tier which are responsible for various tasks and it is depicted in Fig. 1.

4.2 Prototype of E-Commerce for Agriculture

The technology used in development of E-Commerce solution for Agriculture includes MongoDB for the database, React JavaScript Library for developing UI and NodeJS for building APIs. The implementation is responsive and also is a Progressive Web Application. React and MongoDB are interacting via APIs built with ExpressJS. These technologies used in the implementation are micro-services where each component can be deployed separately. The required component can scaled up or scale down based on the requirements. DevOps provides a way to deploy the application. Continuous Integration and Continuous Deployment which comes

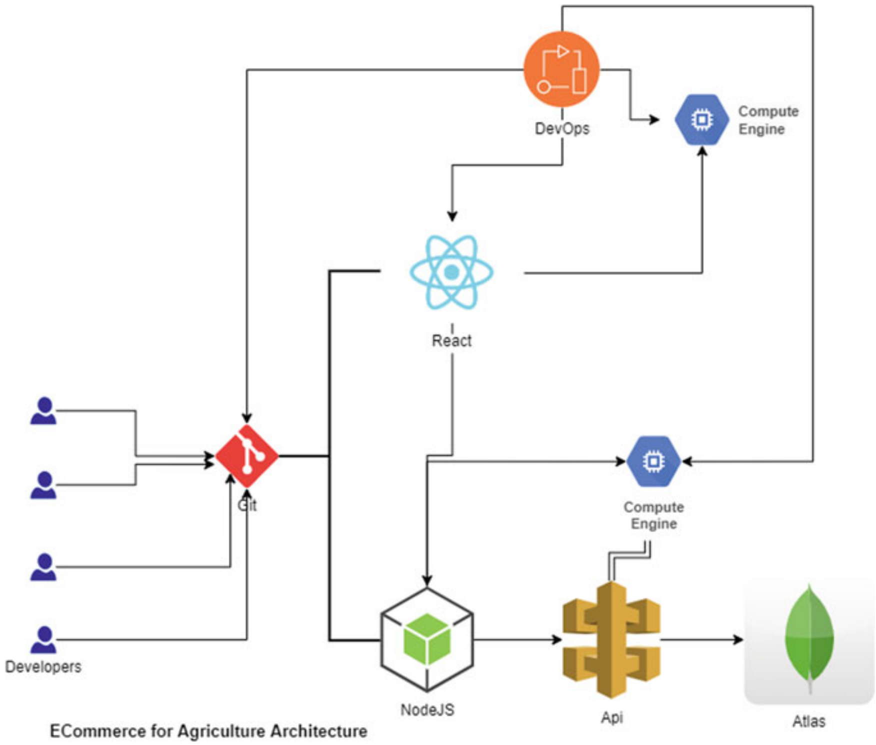
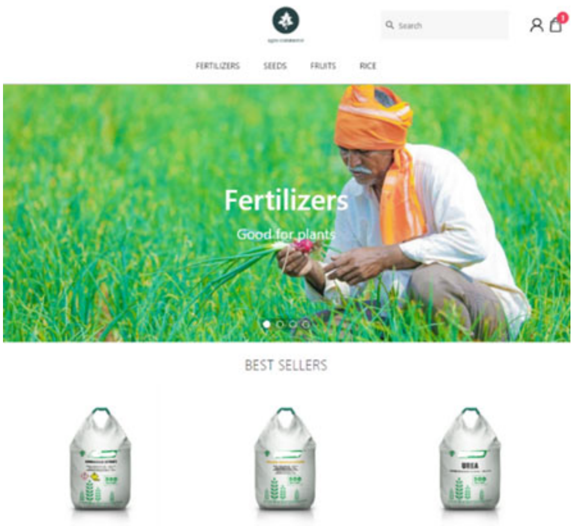


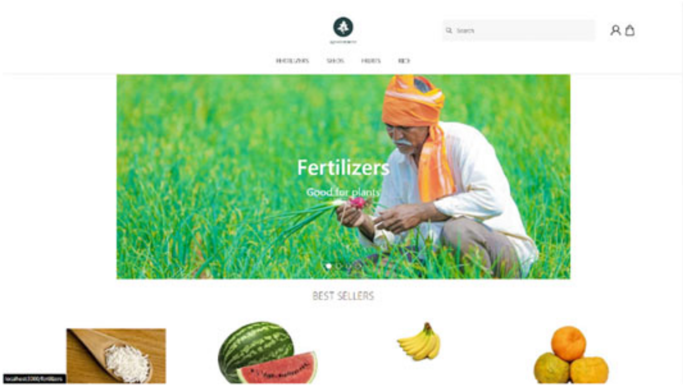
Fig. 2 Architecture of e-commerce for agriculture

under DevOps is used to deploy the tools using the pipeline. Jenkins which is a CICD tool consists of pipeline to deploy the micro-services. These pipelines can be used to have procedures which can be used to deploy the application to server/cloud. The present implementation uses Compute Engine from Google Cloud Platform (Fig. 2).

4.3 Home Page of E-Commerce for Agriculture



4.4 Customer Page



5 Results and Discussions

Using the E-commerce Agriculture application Farmers now will be more updated towards the current Web Technology. This application provides Independent E-Commerce Platform where Sales and Earnings can be documented to review whenever necessary. It contains systematically organized structure for buying and selling resources for farming. Sales are handled in a fast and easy method through Internet Technologies.

6 Conclusion

Agriculture is a practice from ancient times providing food to humans. The practice has changed from time to time based on tools, irrigation and climate. Modern technologies like internet is provides information and is wide spread. E-Commerce is widely used to sell products from seller to consumer and is extensively used because of internet. Modern E-Commerce solution especially for Agriculture can potentially provide a way to sell agriculture specific produce or products and the data analysis can provide insights on demand.

Acknowledgements This paper presents E-Commerce solution for Agriculture. Existing system for agriculture includes middleman to seller and an existing e-commerce solution which is built for selling all types of goods. In order to overcome the flaws, E-Commerce specific for agriculture with continuous feedback from the farmer and the customer can be implemented. E-Commerce solution provides a way to tackle problems that occur during trade between middleman and farmers. The data collected can help farmer to choose what crop should be grown.

References

1. Munassar AMN, Govardhan A (2010) A comparison between five models of software engineering. IJCSI Int J Comput Sci Issues 7(5)
2. Sun W, Chan Y, Xu HY (2015) A case-based distance approach to agricultural product's vendor evaluation in e-commerce. 978-1-4799-8375-9
3. Measures of farm profit, 9780643095168_CH11, CSIRO (Commonwealth Scientific and Industrial Research Organisation) Publishing, Australia
4. Mueller RAE (2000) Emergent e-commerce in agriculture. University of California, Number 14, AIC Issues Brief
5. Janom N, Zakaria MS, Ariffin NPM, The empirical evaluation of B2B e-commerce stages in agro based industry. In: 2012 IEEE colloquium on humanities, science & engineering research (CHUSER 2012). Kota Kinabalu, Sabah, Malaysia, 3–4 December 2012
6. www.publish.csiro.au
7. Bob Frankston, Harman (2018) Progressive web apps [Bits versus electrons]
8. IEEE https://ieeexplore.ieee.org/document/8287006
9. Megharani Building a template for intuitive virtual e-Commerce shopping site in India
10. Patil T, Rao MY 17-Nov-2017 MECS

11. Huang Y, Chai Y, Yi Architecture of next-generation e-commerce platform
12. Liu, Shen J (2019)
13. Almarabeh T, Majdalawi YKh (2018) Cloud computing of e-commerce

Classification and Identification of Dog Breed Using CNN

G. Santosh Kumar, R. Dhanush, B. M. Chirag, H. Chethan, and K. V. Hemanthkumar

Abstract This paper presents an image recognition problem using deep learning. This system identifies the breed of a dog by just taking one image as an input. This system makes use of deep learning and uses a convolutional neural network (CNN) to process the image. The proposed system identifies the dog breed by using a CNN algorithm where scanning of the image happens pixel wise. This model is trained and evaluated on the Stanford dataset. We use the pre-trained model and learned weighs to extract the feature from the data-set that contains the breed information of more than 120 dogs. Web application is used to display the output to the user. It contains a client and a server, which includes libraries for evaluating on a neural network in both offline and online systems. In this paper, we have come up with an idea to identify the dog breed and obtain its related breed information more efficiently and accurately using the Xception model architecture.

Keywords Convolutional neural networks · Deep learning · Flask server · Vue.js · Xception · Web application

1 Introduction

Convolutional neural networks (CNN) have been used to great effect in applications such as object classification, scene recognition, and other applications in image recognition. When the objects, the CNNs are trying to categorize share many similar features, such as the breeds of dogs, it becomes hard to imagine the specific features that CNNs must learn in order to categorize these dogs correctly. It is therefore interesting to see how well CNNs can perform on only dog breeds, compared to labels from all classes of objects in the regular ImageNet.

This system is able to determine the breed of a dog in an image provided by the user, and also displays detailed information about each recognized breed. Furthermore, there is low inter-breed and high intra-breed variation; in other words, there are relatively few differences between breeds and relatively large differences within

G. Santosh Kumar (✉) · R. Dhanush · B. M. Chirag · H. Chethan · K. V. Hemanthkumar
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_93

breeds, differing in size, shape, and color. In fact, dogs are both the most morphologically and genetically diverse species on Earth. The difficulties of identifying breeds because of diversity are compounded by the stylistic differences of photographs used in the dataset, which features dogs of the same breed in a variety of lightings and positions.

The main aim of the project is to build a high accuracy model capable of identifying the closest resembling dog breed by just looking at its image. The model accepts the image as an input and predicts the dog breed.

2 Related Work

There are some previous works related to the building of Convolutional Neural Networks, Image recognition, and deep learning systems.

Abdel-Hamid et al. [1] solve a similar problem which carries out speech recognition using traditional methods, making use of the size and position of each local part, and PCA, while Howard et al. [2] solved a problem by the CNN-based method for detecting dogs in complex images and considered the identification of the breed of dogs. Ráduly et al. [3] generalize R- CNN to detect different parts of an image and it is based on Resnet-50 Architecture using ImageNet database.

The current research paper is based on fine-tuning of CNNs using the Xception model and the results obtained are of high accuracy as the image classification happens pixel wise.

3 Design

3.1 System Architecture

This system is divided into three parts (Fig. 1):

- Image pre-processing
- Convolutional Neural Network
- Web Application.

3.2 Proposed System

3.2.1 Image Pre-processing

There are 20 thousand images to be worked on and all images have different sizes. Scaling the images is necessary as it would give a standard size to work with and

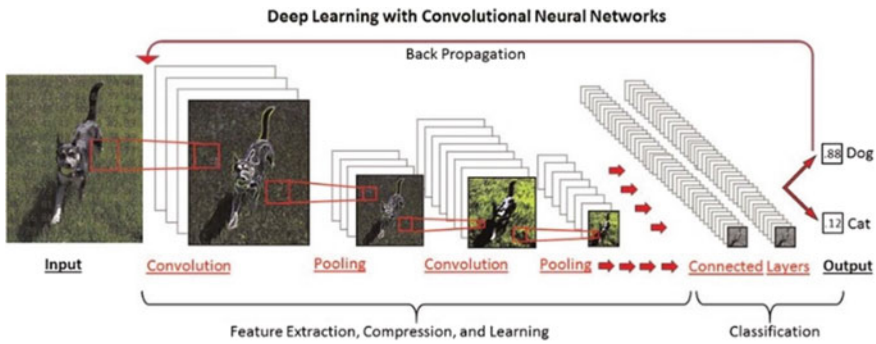


Fig. 1 System architecture

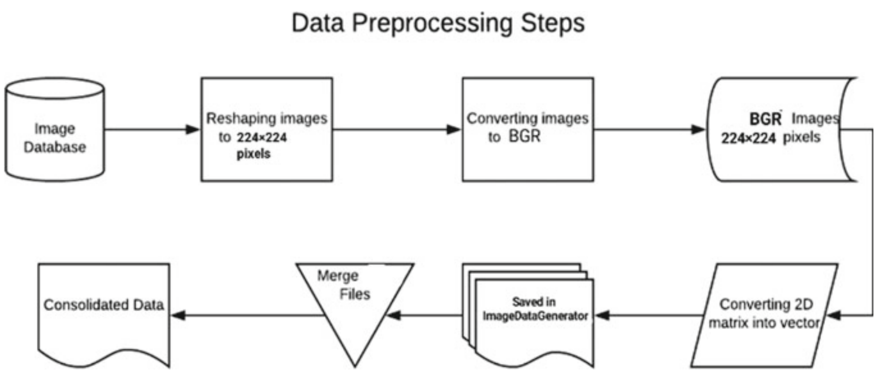


Fig. 2 Image pre-processing

the data would be uniform. So, we resized the images to 224 × 224 pixels. After resizing the images, the images are converted to BGR (Fig. 2).

Next, the images need to be stored in a format that would be suitable enough to train a model on. The images which were now in a 2D matrix format were then stored into a 1D array, basically, a flat-file that would consist of 784 attributes per image. This data is saved in the Image Data Generator file which is later directly called upon by the system and which will now be in a machine-readable format.

Then, 120 file data are merged into one single unit and the breeds will be stored in a dictionary. Now that the data is in a machine-readable format, the next stage is of the study which would be the modelling stage.

3.2.2 Convolutional Neural Network

The image after the pre-processing stage is fed into the ML model where the image is divided into several residual units (layers).

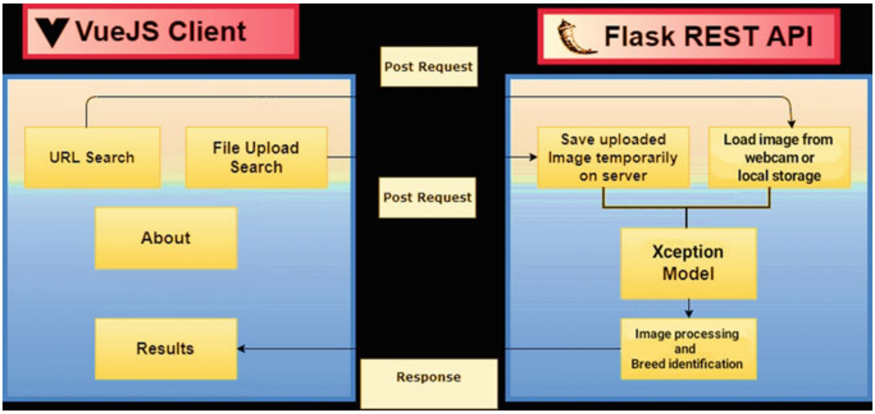


Fig. 3 Web application workflow

A CNN consists of an input layer, output layer, and hidden layers. Convolutional layers are present in the hidden layers. Hidden layers also consist of RELU layer which contains:

- Activation function
- Pooling layers
- Fully connected layers
- Normalization layers.

3.2.3 Web Application

The front end is implemented by the Vue.js framework which is used to create an interactive interface for the application (Fig. 3).

Back End Design—Flask framework is used to create a RESTful API.

4 Implementation

The main aim of image processing in this paper is to enhance the quality of the image and to carry out feature extraction and classification of the image (Figs. 4 and 5).

Tools used in developing this system:

- NumPy is a python library that consists of multidimensional array objects and various tools for processing arrays. NumPy can be used in processing powerful N-dimensional array objects, sophisticated functions, linear algebra, and random number capabilities.
- OpenCV is used to read and write images, capture and save videos, and process images (filter, transform).

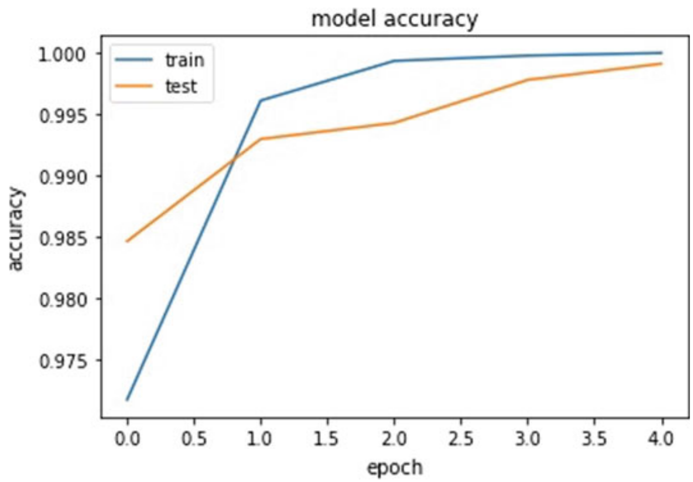


Fig. 4 Model accuracy in which training accuracy is 99% and the testing accuracy is 98%

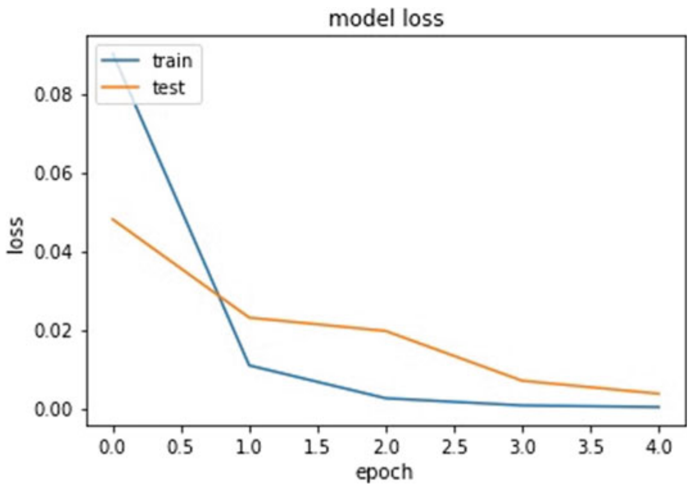


Fig. 5 Model loss

- Vue Router is the official router for Vue.js. It deeply integrates with Vue.js core to make building Single Page Applications with Vue.js easily.
- Bulma is a free, open-source CSS framework that has an extensive range of built-in features that leads to faster turnaround and lesser CSS code writing.
- Representational state transfer (REST) which defines a set of rules to be used for the creation of Web services.
- The Fetch API provides an interface for fetching resources.

- Axios is a JavaScript library used to make HTTP requests from Vue.js. Frontend to backend connection is achieved by Axios.
- The Web app is used to display the information obtained from the database to the user.

5 Result

This Web-based application provides the user a dog breed identifier which is more accurate and precise than the previous versions which were created using different algorithms and different technologies which were of low accuracy. The user gets a hassle-free user interface that assists the user in finding the dog breed and its related breed information where the input is given by the webcam or is uploaded from the local storage or any other camera device by connecting it to the system.

The system has been designed in such a way that if any other image like human, car, bike etc., is given as input other than a dog, then the user will be notified by the system to upload a dog image only. The user will also be able to adopt a dog by clicking an Adopt a puppy tab if they wish to do so.

6 Conclusion

This problem is a challenging one and very exciting. The solution can be used for other classification problems as well. In the real world, this kind of solution could be used in various studies which help scientists all over the world to save time and resources when conducting studies about the health and other species. Quick and accurate breed prediction helps veterinarians to identify the dog breed and get its related dog breed information during emergency situations.

Dog images are available in abundance and are the most loved creatures. Hence, we experimented on this topic. However, we would also like to expand our research on other aspects based on identifying an image and its prediction. This paper provides an understanding of the classification problems and helps scientists across disciplines all over the world.

References

1. Abdel-Hamid O, Mohamed AR, Jiang H, Deng L, Penn G, Yu D (2014) Convolutional neural networks for speech recognition. IEEE/ACM Trans Audio Speech Lang Process 22(10):1533–1545
2. Howard G, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobile nets: Efficient convolutional neural networks for mobile vision applications. CoRR, vol. abs/1704.04861

3. Ráduly Z, Sulyok C, Vadászi Z, Zölde A, Dog Breed Identification Using Deep Learning. In: SISY 2018 IEEE 16th international symposium on intelligent systems and informatics. Subotica, Serbia, 13–15 September 2018. https://doi.org/10.1109/SISY.2018.8524715

Empowering Integrity Auditing Based on Identity and Information Sharing for Secure Cloud Storage Using Delicate Data Concealment

S. V. Reshma, J. Ritisha, S. M. Pai, K. R. Sneha, and G. Shruthi

Abstract Users can store their Information in the cloud to reduce the cost of using regular storage by using the storage services provided by cloud and data management schemes that are provided. To ensure the integrity of information preserved in the Cloud, using several data integrity auditing. In cloud storage services, users can store their data remotely inside the cloud and can share it with any others. In this paper, an encryption method has been adapted to encrypt the information in accordance of conscious details of the documents and modifies these information file's signatures into reasonable ones for the encrypted file. To confirm the honesty of encrypted report within the segment of integrity auditing the signatures are used. The outcome results in this approach make the document saved within the cloud capable of proportion and utilized by others at the limit that the touchy facts is hidden. Along with that, identity-based cryptography is the basis of the proposed scheme, where it simplifies the complex certificate management. The proposed scheme is more secure and efficient than the older one is proved by the evaluation of performance and security analysis.

Keywords Cloud storage · Data integrity auditing · Data sharing · Sensitive information hiding

1 Introduction

Users are provided with privilege of using powerful and on demand supply of data storage facility. Users can save economy on maintenance of software and hardware for storing data by just using services provided by cloud [1], and it provides benefits to users. In this remote information integrity auditing schemes [2, 3], the records proprietor have to first of all generate signatures for the data factors earlier than importing them into the cloud. These signs are applied to confirm that the cloud honestly owns those information blocks at some stage in the integrity auditing phase.

S. V. Reshma · J. Ritisha · S. M. Pai · K. R. Sneha · G. Shruthi (✉)
Department of Computer Science and Engineering, Don Bosco Institute of Technology, Bengaluru, Karnataka, India
e-mail: shruthigcse@dbit.co.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_94

The data proprietor then uploads those data blocks at the side of their signs to the cloud. Data sharing is one of the maximum not unusual place capabilities of cloud storage, permitting many customers to share their information with many people. However, these shared facts saved inside the cloud may also include a few sensitive information. So if we upload the file directly to the cloud, the sensitive information might be unavoidably uncovered to the cloud. A prospective way to solving of a problem is to hide the entire shared document prior to uploading it into the cloud, and then produce the signatures used to verify the solidity of this encrypted document, This way can perceive the delicate data concealing since only the info holder or admin can decode the document. It will convert entire shared document unable to be used by others is essential to cover delicate records via encrypting the entire shared file. Thus, the way to realize records sharing with touchy data hiding in faraway records integrity auditing is incredibly essential and valuable.

2 Existing System

In cloud storage system, the document in cloud may have sensitive data, which shouldn't be revealed to other entities when the file being is uploaded, shared or downloaded. Encoding entire shared document can perceive sensitive information, this makes others unable to use the document. To repair this problem a foreign data integrity auditing scheme is proposed [4–9]. It identifies sensitive information data and hides the info with encryption. To realize data sharing with sensitive information hiding, consideration of usage of the thought within the encryptable signature to encrypt the sensitive information of the document by introducing a legitimate encryptor. It isn't feasible if the encryptable signature is legitimately utilized in remote information auditing.

The Signature development done on the basis of chameleon hashes. However, lots of chameleon hashes that show the key disclosure issue. To keep away from the security issue, the signature requires emphatically creatable chameleon hashes [10], which will definitely cause gigantic calculation/computation overhead. The Signature doesn't assist blockless-verifiability [1]. The verifier must download the entire information from the cloud to confirm integrity of data is implied, which can acquire gigantic correspondence overhead and exorbitant verification time in heavy data storage situation. The signature is predicated on the PKI, which experiences complex certificate management.

3 Proposed System

To address the above issues, a new efficient algorithm that uses signature is designed in the stage of signature generation. Signature based scheme which was designed supports blockless-verifiability, it permits the verifier to verify the integrity of information without downloading it entirely from the cloud. It's based on identity-based cryptography, in which complex certificate management is simplified. In the proposed scheme, the Private Key Generator produces the private key for the client as indicated by his ID. Verifying the private key can be done by user which is received by him. If the user is longing to transfer information to the cloud, to save the individual sensitive data of the primary document onto the encryptor, the user is required to utilize a binding element to bind data blocks relating to individual delicate data of the primary document. The point at which when it is essential, the user may recuperate the primary document from the blinded one by utilizing this blinding factor. Later on, the structured signature algorithm to produce signature for the blinded document is utilized by the user. These signatures are utilized to confirm the correctness of this blinded document. Besides, the user produces a document tag, to guarantee the accuracy of the document identifier name it is utilized and some verification values.

The client additionally figures a transformation value that is utilized to convert signatures for encryptors. The client sends the blinded document, its relating signatures, and the document tag alongside the transformation value to the encryptor. The encryptor encrypts the blinded information blocks into a uniform arrangement and furthermore encodes the information blocks with respect to the association's delicate data to secure the privacy of association when the user sends valid messages, and afterward converts their respective signature into substantial ones for encrypted document utilizing transformation value [4]. The encrypter uploads the encrypted document and the respective signatures to the cloud. The proof of auditing according to the challenge is generated by cloud from the TPA when the data auditing task is performed. It can be verified by TPA that the validity and safety of the encrypted document stored in cloud after evaluating if this auditing proof is true or not. Admin have privileges to create the users, during user creation we will send a User id and Password to their e-mail id and also admin will maintain the cloud server's configuration. Admin has the permission to ADD, EDIT or DELETE any number of users. User has to get the User id and Password through e-mail. User can able to login by using this User id and Password. Suppose if the user wants to download any file, first he has to select the file from the list and then get the key from Database, then decrypt the file and store into the local system [3, 4].

4 System Design and Methodology

This model involves different types of entities: (1) The cloud, (2) The user, (3) The encryptor, (4) The Private Key Generator (PKG) i.e. signature generator and (5) The Third Party Auditor (TPA), Fig. 1 represents the following.

The owner/admin blinds the data blocks relating to individual sensitive data in the document, and creates respective signatures. To guarantee the legitimacy of the document and check the solidity of file by using the signatures. The owner directs these blinded documents and its respective signatures to the sanitizer [2, 11]. The

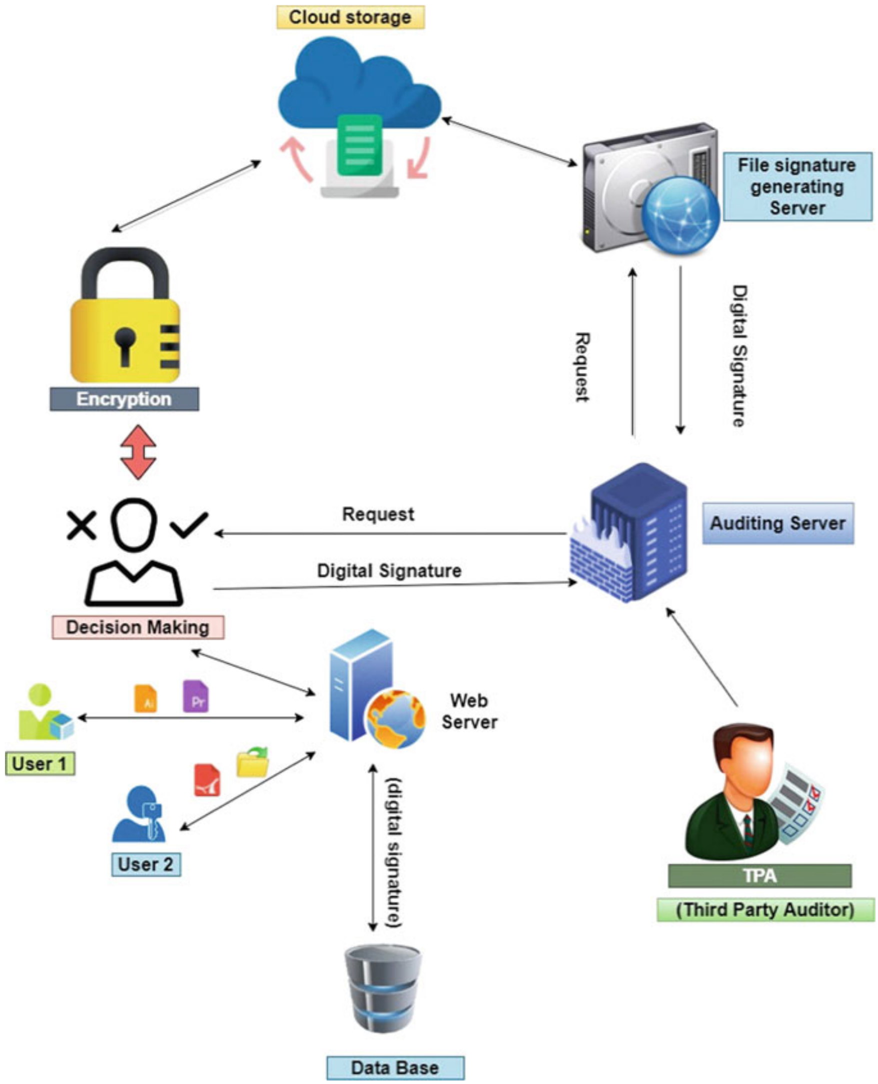


Fig. 1 System design

encryptor encrypts these data blocks and the blinded data blocks respective to the delicate information after receiving the message from the administrator, and then converts signatures of encrypted data blocks into legitimate ones for the encrypted document. Finally, the encryptor transfers this encrypted document and its respective signatures to the cloud. To check the solidity of the encrypted document in the stage of solidity audition the signatures are used. At the point when the TPA needs to confirm the solidity of the encrypted document preserved in the cloud, TPA transmits an inspecting challenge to the cloud. Afterward, the cloud reacts to him with an inspecting verification of information ownership. At long last, the TPA confirms the integrity of the encrypted document by checking whether this reviewing confirmation is right or not.

5 Results

We explore the best approach to accomplish data sharing to delicate data concealment using remote information auditing and suggesting a substitution idea called identity-based shared information auditing with delicate data concealment for secure distributed cloud storage. An encryptor is used to encrypt the data chunks just like the delicate data of the document/file. The complete proposal, initially, the user screens the data chunks like the private delicate data of its principal folder and then it produces the relating endorsements, then it is sent to an encryptor. An encryptor encodes these screened data chunks into a consistent design and encrypts the data chunks like the company's delicate data. This additionally changes the respective endorsements into encrypted document. This strategy understands the distant information integrity auditing as well as hold up the information sharing relying on to the prerequisite that delicate data is ensured in cloud storage. Also, our proposal is predicated on recognition-based cryptography, which disentangles the compound voucher management. We have used Drive HQ Cloud to store the files and algorithms are used to provide security to the files. User has to get the User id and Password through e-mail. User can able to login by using this User id and Password. Suppose if the user wants to download any file, first he has to select the file from the list and then get the key from Database, then decrypt the file and store into the local system (Figs. 2 and 3).

To adequately assess the exhibition in a few procedures, quantity of data chunks is set to 100 and subsequently the quantity of sanitized data chunks is set to 5 in the apparent procedure. As in Fig. 4, secret key generation and secret key check invest about a similar time, which is almost 0.31 s. The time expended by the signature generation is 1.476 s. The time of signature verification and that of sensitive data encryption correspondingly are 2.318 and 0.041 s. So we can infer that in these procedures, the signature heck spends the longest time and sensitive data encryption invests the briefest time.

The system analyses the data uploading rate by taking different sizes of files in two different sessions. Session 1 indicates the uploading rate with 100 mbps and session 2 indicates uploading rate with 80 mbps. Hence it is clear that the uploading rate always depends on the network speed.

This Fig. 5 depicts the sensitive text or context concealment after the file has been uploaded to cloud by the admin. It cannot be viewed or tampered with unless downloaded by authorized user.

Fig. 2 Sending OTP to the mail

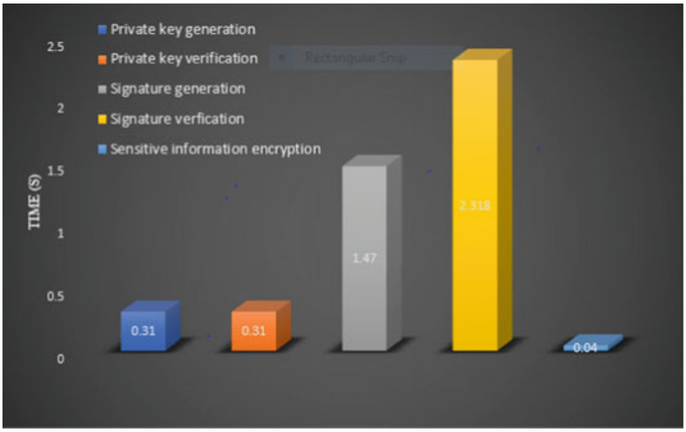


Fig. 3 Performance of different processes

6 Conclusion

In this paper, our system proposes an identity-based data integrity auditing scheme for storage of data in the cloud in a secure way, the data sharing is supported with sensitive data concealing. In this proposal, the saved document within the cloud is often communicated as well as employed by others on the restriction that the document is secured with sensitive information. Although, the remote data integrity auditing is efficiently executed. The security inspection of the suggested proposal and also performance justification by applications that currently exist. The outcome

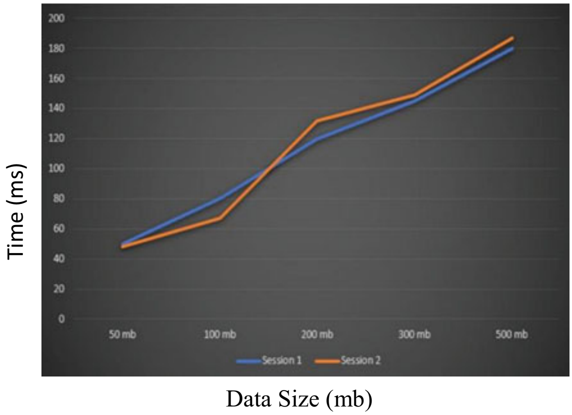


Fig. 4 Data upload rate comparison

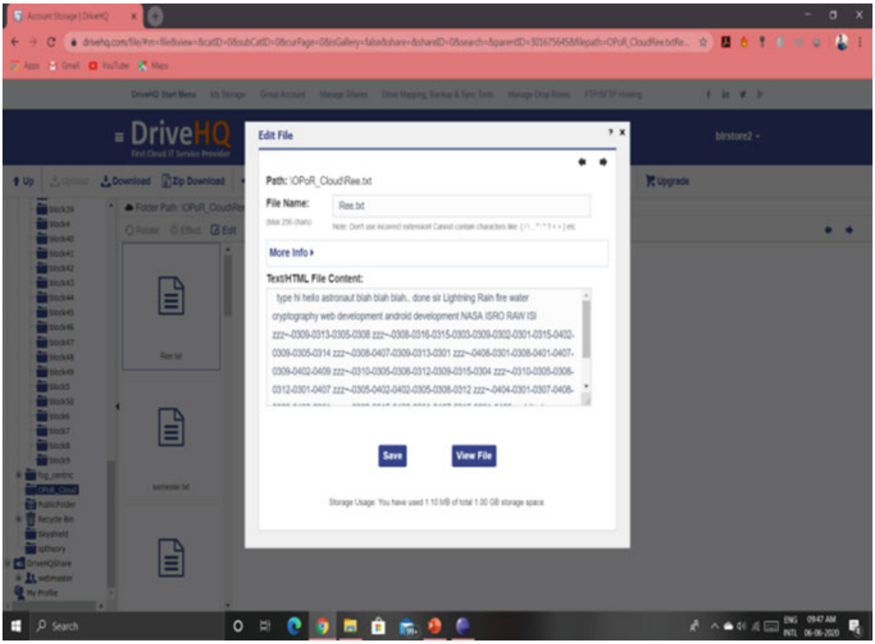


Fig. 5 Sensitive data concealment

clearly states that this suggested scheme is efficient in achieving proper security and efficiency.

References

1. Shen W et al (2018) Enabling identity-based integrity auditing and data sharing with sensitive information hiding for secure cloud storage. IEEE Trans Inf Forensics Secur 14(2):331–346
2. Ren K, Wang C, Wang Q (2012) Security challenges for the public cloud. IEEE Internet Comput 16(1):69–73
3. Ateniese G, Burns R, Curtmola R, Herring J, Kissner L, Peterson Z, Song D (2007) Provable data possession at untrusted stores. In: Proceedings of the 14th ACM conference on computer and communications security, ser. CCS '07, pp 598–609
4. Juels A, Kaliski BS (2007) Pors: proofs of retrievability for large files. In: Proceedings of the 168 14th ACM conference on computer and communications security, ser. CCS '07, pp 584–597
5. Shacham H, Waters B (2013) Compact proofs of retrievability. J. Cryptology 26(3):442–483
6. Wang C, Chow SSM, Wang Q, Ren K, Lou W (2013) Privacy-preserving public auditing for secure cloud storage. IEEE Trans Comput 62(2):362–375
7. Worku SG, Xu C, Zhao J, He X (2014) Secure and efficient privacy-preserving public auditing scheme for cloud storage. Comput Electr Eng 40(5):1703–1713
8. Guan C, Ren K, Zhang F, Kerschbaum F, Yu J (2015) Symmetric-key based proofs of retrievability supporting public verification. In: Computer security—ESORICS 2015. Springer International Publishing, Cham, pp 203–223
9. Shen W, Yu J, Xia H, Zhang H, Lu X, Hao R (2017) Light-weight and privacy-preserving secure cloud auditing scheme for group users via the third party medium. J Netw Comput Appl 82:56–64
10. Ateniese G, de Medeiros B (2005) On the key exposure problem in chameleon hashes. In: Security in communication networks. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 165–179.
11. Ateniese G, Chou DH, de Medeiros B, Tsudik G (2005) Sanitizable signatures. In: Proceedings of the 10th European conference on research in computer security, ser. ESORICS'05. Springer-Verlag, Berlin, Heidelberg, pp 159–177

Machine Learning Techniques to Predict Diabetes Mellitus

B. M. Yashaswini, Y. Kavya, M. S. Akshatha, R. Bhavya, and Arunima Chanda

Abstract Diabetes occurs when our body is not able to respond to insulin with proper function, where the insulin is playing very major role in human body. Diabetes regulates the sugar level in the body and maintain the excess of sugar to normal range. It helps to keep effective of sugar on body organs as kidney, eyes, heart, foot etc. Sugar can cause critics like blindness, heart stroke foot amputations, kidney failure etc. Due to disturbance of carbohydrate, fat and protein metabolism. It is named as hypoglycaemia in daily routine whatever we serve from the source of fibre or any other nutritional value, our body extracts sugar and uses it for workouts. Insulin is most need to control sugar even though there is no diabetes cure, diabetes can be treated and controlled and some people goes into remission. To control diabetes manage the sugar level in body.

Keywords KNN · SVM · Stochastic gradient descent · Adaboost algorithm · Naïve Bayes · Decision tree

1 Introduction

Diabetes is an incessant sickness or gathering of metabolic ailments where an individual experience a huge amount of blood glucose in the body, which is either the insulin creation is insufficient, or on the grounds that the body's cells don't react appropriately to insulin. The steady hyperglycaemia of diabetes is identified with long stretch mischief, brokenness, and disappointment of different organs. The ordinary distinguishing process is that patients need to visit an analytic focus, counsel their PCP, and hold on for a day or more to get their reports Human services industry contains significantly and cosmically tremendous and touchy information and should be taken care of carefully. Diabetes Mellitus is one of the becoming cosmically deadly maladies everywhere throughout the world. Clinical experts need a dependable expectation framework to analyse Diabetes. Applying AI and information mining strategies in research is a key way to deal with using cosmically tremendous volumes

B. M. Yashaswini (✉) · Y. Kavya · M. S. Akshatha · R. Bhavya · A. Chanda
Department of Computer Science and Engineering, DBIT, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_95

of accessible diabetes-related information for removing savviness. Apparently, thus, AI and information mining approaches are of incredible concern with regards to determination, the executives and other related clinical organization perspectives.

Distinctive AI procedures are auxiliary for analysing the information from different points of view and combining accessibility of cosmically colossal measures of information will have the option to give us auxiliary discernment if certain information mining strategies are applied on it. The primary objective is to decide early examples and afterward to decipher these examples to appropriate central and auxiliary data for the clients. In this way, mining the diabetes information in a productive way is a vital concern. The target of the dataset is to symptomatically foresee whether a patient has diabetes dependent on certain symptomatic estimations remembered for the dataset. A few limitations were set on the determination of these cases from a bigger database. The dataset is examined what's more, investigated to fabricate viable model that predicts and analyse the diabetes sickness. In this investigation we mean to apply the bootstrapping resampling strategy to upgrade the exactness and afterward applying Naïve Bayes, Decision Trees and k closest Neighbors (KNN) and analyse their performance.

2 Literature Review

Diabetes through Machine Learning is paper [1], they return to the information of the San Antonio Heart Study, Foreseeing and utilize machine figuring out how to foresee the future advancement of type-2 diabetes. To manufacture the expectation system, they utilized with help of vector machines and 10 highlights that are notable in the writing as solid indicators of diabetes. Because of uneven idea of the dataset in the class names, they utilized 10-overlap cross-approval for preparing the system and a wait to approve it. The after effects of this examination is an approval exactness of 84.1% with a review pace of 81.1% found the middle value of more than 100 emphases. The examination contained 5158 men and non-pregnant ladies of Mexican-American also, non-Hispanic white eth-comfort, matured somewhere in the range of 25 and 64 years and dwelling in San Antonio, Texas [2]. The point of this paper is to devise a AI plot that can distinguish sound subjects that are at an expanded danger of creating type-2 diabetes. To this, the information utilized here is unit of the SAHS that incorporates the OGTT information of 1496 solid subjects at pattern, out of which 171 were named as diabetic at the subsequent evaluation and 1281 kept up their sound status [3]. To decide the execution of our pre- word usage models, we use precision, review and particularity of the models. While the preparation, they accentuated on boosting the review of the classifier which in different names, amplifies the distinguishing proof pace of high-chance diabetes. Utilizing the system portrayed in the past area, they show the exhibition output that are arrived at the midpoint of more than 100 emphases. The use of just one significant calculation for assurance of diabetes is fused in this study.

3 Design

The proposed framework centres around the forecast of diabetes utilizing Adaboost Algorithm. Different base classifier mix is gone after for Adaboost calculation. The framework is executed utilizing four stages. This incorporate assortment and arrangement of worldwide and nearby dataset. Here preparing is finished with worldwide information and approval is nearby data. Figure 1 tells the square outline of the proposed framework. The underlying stage fuses the social affair of the worldwide dataset and nearby dataset. A rate of 75 is for the worldwide dataset. Here 25% is utilized to testing. For approval, the nearby dataset is gathered through leading review to better places in Kerala. The missed qualities in nearby datasets were filled with using the relating mean qualities as s part of the worldwide dataset. Second stage incorporates the exchanging of the worldwide dataset. Second stage incorporates the Preparing of the worldwide dataset with Adaboost calculation by thinking about different base classifier. The base classifier viable are bolster vector machine, innocent Bayes, choice tree also, choice stump, third stage incorporates the approval of the neighbourhood dataset with the abovementioned prepared classifiers. Last stage incorporates the precision check for Adaboost calculation with base classifiers. The starting stage joins the social occasion and worldwide dataset also the neighbourhood dataset. Preparing not accomplished to the worldwide dataset. A rate split of 75 is given for the worldwide dataset. 25% is utilized to testing. For approval, the neighbourhood dataset was gathered by directing overview from better places in Kerala. The missing qualities in nearby dataset are filled by using the relating mean qualities as s part of the worldwide dataset. Second stage incorporates the exchanging of the worldwide dataset. Second stage incorporates the Preparing of the worldwide dataset with Adaboost calculation by thinking about different base classifier. The base classifier viable are bolster vector machine, gullible Bayes, choice tree also, choice stump, third stage incorporates the approval of the nearby dataset with the

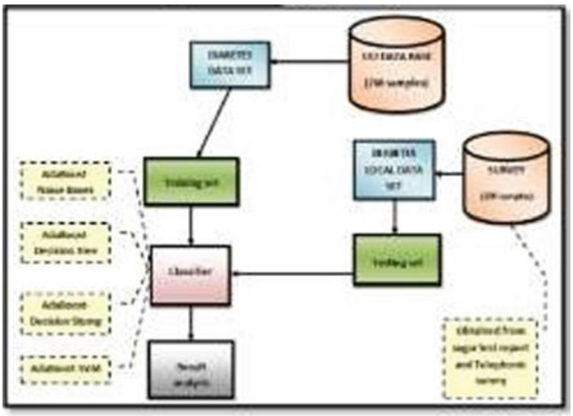


Fig. 1 Block diagram of proposed system

abovementioned prepared classifiers. Last stage incorporates the precision confirmation of every one of the Adaboost calculation with base classifiers. MATLAB-Weka interface is utilized for the exactness confirmation.

4 Implementation

Procedure: We have isolated the dataset dependent on the class level. At that point apply the grouping calculation with Pearson Correlation Similarity. Subsequent to applying the bunching, from each group we have picked the qualities with greatest divergence. In light of the chose highlights, the information is prepared with the characterization calculation. Normally, Pearson relationship is utilized to gauge separation (or similitude) before actualizing a grouping calculation. Since Pearson Correlation is very vulnerable to exceptions. Pearson Correlation metric estimates how exceptionally associated are two factors and is estimated from -1 to $+1$. Pearson Correlation Coefficient of 1 demonstrates that the information objects are entirely connected at the same time, for this situation, a score of -1 implies that the information objects are not associated. At the end of the day, the Pearson Correlation score evaluates how well two information objects fit a line. The advantages of utilizing this separation strategy is that the precision of the score increments when information isn't standardized. In the wake of getting the similitude between the highlights, grouping calculation is applied. K-means and K-medoids are two mainstream calculations for grouping. Here, we are examining both the calculation individually. A characterization model endeavours to make some determination from watched esteems. Given at least one data source a grouping model will attempt to foresee the estimation of at least one results. Results are names that can be applied to a dataset.

Here we have utilized numerous order calculations:

Bolster Vector Machine
Gullible Biased
K Nearest Neighbors

Setting: The target of the dataset is to symptomatically anticipate whether a understanding has diabetes, in view of certain demonstrative estimations remembered for the dataset. A few imperatives were set on the choice of these examples from a bigger database. Specifically, all patients here are females at any rate 21 years of age of Pima Indian legacy. Content: The datasets comprise of a few clinical indicator factors and one objective variable, Result. Indicator factors incorporate the quantity of pregnancies the patient has had, their BMI, insulin level, age, etc.

Algorithm: Algorithm choice for anticipating the best outcomes. Generally Data Scientists utilize unique sorts of Machine Learning calculations to the enormous informational indexes. In any case, at significant level every one of those unique calculations can be ordered in two gatherings: Administered learning and Unsupervised learning.

Administered learning: Supervised learning is a kind of framework where both information and wanted yield information are given. Information and yield information are named for grouping to give a learning reason for future information preparing. Administered learning issues can be additionally assembled into Regression and Classification issues. A relapse issue is the point at which the yield variable is a genuine or ceaseless worth, for example, "pay" or "weight". An order issue is when the yield variable is a classification like sifting messages "spam" or "not spam".

Unaided Learning: Unsupervised learning is the calculation utilizing data that is neither grouped nor named and permitting the calculation to follow up on that data without direction. In our dataset we have the result variable or Dependent variable for example Y having just two arrangement of qualities, either M (Malign) or B (Benign). In this way, we will utilize Classification calculation of regulated learning.

1. **Stochastic Gradient Descent**
   The word 'stochastic' signifies a framework or a procedure that is connected with an arbitrary likelihood. Henceforth, in Stochastic Gradient Descent, a couple of tests are chosen arbitrarily rather than the entire informational index for every cycle. In Gradient Descent, there is a term called "bunch" which indicates the all out number of tests from a dataset that is utilized for computing the slope for every emphasis. In normal Gradient Descent improvement, similar to Batch Gradient Descent, the bunch is taken to be the entire dataset. In spite of the fact that, utilizing the entire dataset is extremely helpful for getting to the minima in a less boisterous or less arbitrary way, yet the issue emerges when our datasets get extremely colossal.
2. **Closest Neighbor**
   K-Nearest Neighbors (KNN) is perhaps the least difficult calculation utilized in Machine Learning for relapse and arrangement issue. KNN calculations utilize an information and group new information focuses in view of a likeness measures (for example separation work). Characterization is finished by a dominant part vote to its Neighbors.
3. **Support Vector Machines**
   "Support Vector Machine" (SVM) is a managed AI calculation which is utilized for either characterization or relapse difficulties. Notwithstanding, it is for the most part utilized in characterization issues. In this calculation, plotting all the information as a point in n-dimensional space to the estimation of each of element that is being estimated with the specific organize. During this we perform the grouping using hyper plane. Support Vectors are just the co-ordinates of individual perception. Bolster Vector Machine is an outskirt which best isolates the two classes (hyper- plane/line).
4. **Random Forest Classification**
   Arbitrary Forest has almost had an similar hyper parameters as a sacking classifier. You don't need to consolidate a sacking classifier with a packing classifier and can just effectively utilize the Random Forest. Like I previously stated, with Random Forest, we can manage Regression undertakings by using the

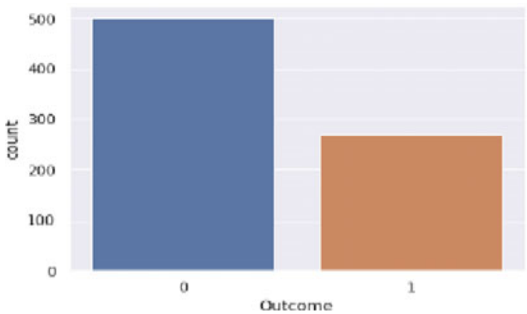
Random Forest regressor.in general Random timberland assembles different sacking classifiers.

5. **Naïve Bayes**
Bayes hypothesis named after Rev. Thomas Bayes. It uses a similar method to predict the probability of different class upon attributes. This is most used in text classification which have problems with multiple classes. Every pair of datasets being classified is independent of one another.

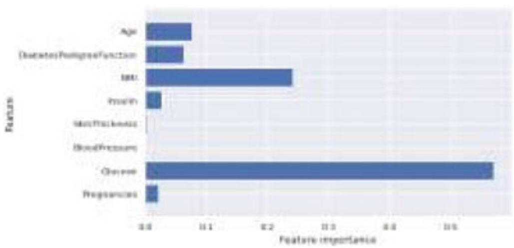
6. **Decision Tree**
Choice Trees are a sort of Supervised Machine Learning where the information is ceaselessly part as per a specific boundary. The tree can be clarified by two substances, specifically choice hubs and leaves. The leaves are the choices or the ultimate results. Also, the choice hubs are the place the information is part.

5 Results

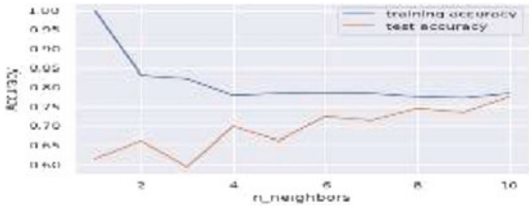
Outcome



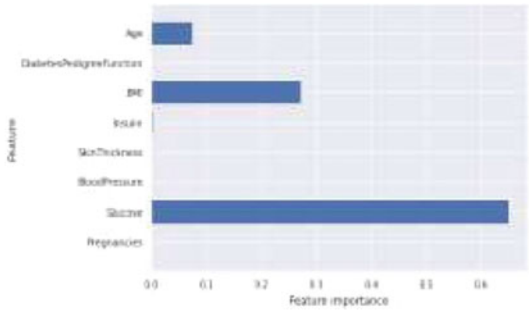
Gradient Boosting Algorithm



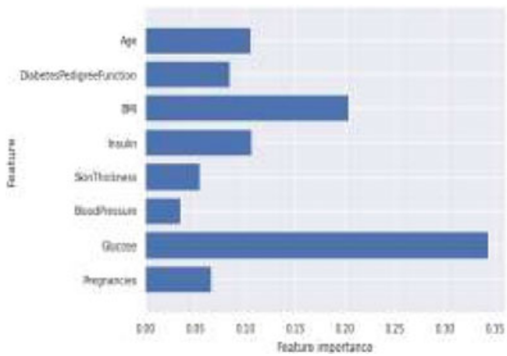
K-NN Algorithm



Decision Tree Algorithm



Random Forest Algorithm



Accuracy of K-NN

Training set: 0.78 Test set: 0.78.

6 Conclusion

AI has the extraordinary capacity to alter the diabetes chance expectation with the help of cutting-edge computational strategies and accessibility of enormous measure of epidemiological also, hereditary diabetes chance dataset. Location of diabetes in its beginning phases is the key for treatment. This work is to depict an AI way to deal with anticipating diabetes levels. The procedure may likewise assist specialists with developing a precise and viable instrument that will reach at the table of clinicians to assist them with settling on better choices about the ailment status.

References

1. Juan Camilo Ram'irez Faculted de ingenier'ia de sistemas university, Antonio Narino Bogota, Colombia
2. Department of Electrical and Computer Engineering, Texas A and M university at Qarter, Doha, Quatar 23874
3. Department of Industrial and Systems Engineering, Texas A and M University, College Station, TX 77843-3128
4. Alic L, Abbas HT, Rios M, Abdul Ghani M Qaraqe K
5. Magnetic detection and imaging group. Faculty of Science and Technology, University of Twente, Enschede, The Netherlands
6. Division of Diabetes, University of Texas Health Science Center at San Antanio, Texas

Managing Database for Satellite Health

N. S. Chandrashekar, B. G. Chethan, Ashith V. Shibu, and F. Flinders Samuel Asher

Abstract As there is a rapid climb of satellite data sets over the years, the volume of information which is required to be stored, a thorough management is required to preserve such enormous amount of information. Time to retrieve the data and process the data is tedious and time-consuming and there is a lack of option for analyzing the data. Designing and developing database architecture understandable and effective. This architecture will take less processing time and efforts will be reduced as it is need to handle and manage only one system. It also generates plots and reports for review.

Keywords MySQL database · Raw data generator (RDG) · Partitioning · Compression

1 Introduction

The most important aim of a DBMS system is to produce the simplest way to accumulate and fetch information most conveniently and economically [1]. By data, we tend to mean facts that may be recorded which have embedded meaning.

A RDBMS framework is a DBMS framework structured explicitly for relative databases. In this manner, RDBMS is developments of a lot of DBMS. A social database alludes to a database that it stores information in an organized configuration misusing lines and segments. The information has developed by the executive firm in the course of recent decades. Initially the first relative information is the executive's framework (RDBMS) innovation [2]. Still nowadays, the main part of back-end frameworks generally are RDBMS for online computerized data, monetary arrangement, clinical, transport, protection offices, and media transmission business. As of the absolute information gathered and investigated in tries have increased numerous folds in volume, assortment, and speed of age and utilization [3].

Time to retrieve the data and process the data is tedious and time-consuming and there is a lack of choices for analyzing the data. A database architecture is planned

N. S. Chandrashekar (✉) · B. G. Chethan · A. V. Shibu · F. F. S. Asher
Don Bosco Institute of Technology, Kumbalgodu, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_96

and developed that is apprehensive and effective. The data working on is the data of that of the satellite. The main operation in satellite communication is orbital monitoring, the altitude of the satellite, monitoring, and controlling of many other subsystems. Since the data from the satellite which arrives is nearly continuous and will be recorded as a brand-new entry every time, it is considered as the database architecture to be of a time-series database. All parameters from the subsystem have been viewed concerning time.

2 Proposed Methods

The raw data which is present in the archival area will be passed through the RDG and then the processed data will be added to the database which is created. The data should be automatic in such a way that there should be no gaps between the present and the future data. The GUI is interfaced to the database through external software to retrieve data and plot graphs for the same.

2.1 Block Diagram and Overall Description

Figure 1 shows the block diagram for the analysis of data from the satellite. In this work are receiving and analyzing the data from the satellite is done. The data from the satellite is received at the ground station during the visibility period and at a lower sampling rate during the non-visibility period. The data received from the satellite is raw data. The data received will go through a set of downlink subsystems that will perform de-modulation and de-multiplexing. This data will be stored on mission-critical servers. The archival are stores all the raw data from the mission-critical servers.

The Raw Data Generator (RDG) is used to access the raw data from the archival areas and gives the processed data as output to the scientists. This data is used for analysis and is the existing method. In the proposed method the raw data from the archival area is taken and the processed data is extracted from the RDG and gets populated in the MySQL database. This process of populating the database is a continuous process in real-time.

The main objective here is, the time taken to get the information (d2) from the MySQL database for analysis is faster compared to the time taken to access the data (d1) from the RDG for analysis. Hence, it can be said that d2 < d1. The architecture for this MySQL database is designed in such a way that it increases the scalability and flexibility.

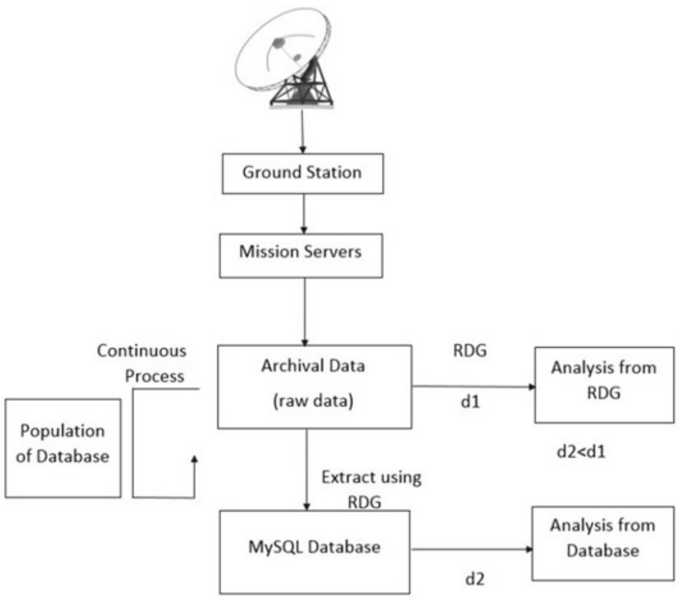


Fig. 1 Block diagram for analysis of data from the satellite

2.2 *Visibility Period*

This is the period during which the satellite will be visible to the ground station. In this case, the ground station has a visibility period of ten minutes, during this time the data is received by the ground station at the full rate of 4 kbps, but during the non-visibility period, the data is stored after a set rate of sampling which is lower than the real-time data.

2.3 *Mission-critical Servers*

Mission-critical server operation sounds straight forward—something that is essential to the overall operations of a process within an organization. Essentially, something that is critical to the mission but mission-critical is far more complex than that. Mission-critical computing is a secure, reliable and scalable computing that supports the organizations process and operations [4]. The processes and operations are mission-critical because they are the heart to the organizations' goal and, if they fail it might cause reputational damage to the organization and have an impact on national security. Hence the data coming in from the satellite is secured with no single point of failure (Fig. 2).

Fig. 2 Mission-critical servers

2.4 Archival Storage

Archival storage is a type of storage designated for the data that may not currently need but must be saved so it can be accessed in the future, if necessary. Archiving is used to hold data for the long-term in the event it needs to be retrieved.

2.5 Raw and Processed Data

The data that will be received from the satellite is a raw data, should be converted into processed data for use. For example, the data coming from the satellite accumulates huge amounts of raw data each day but that data does not yield much information until it is processed [5].

Here the processed data is obtained by running the raw data through the RDG. Hence it is required to process the data from the satellite which is readable by the scientists for analyzing the same (Fig. 3).

2.6 Raw Data Generator (RDG)

The Raw Data Generator is used to access the raw data from the archival storage and gives the processed data output. The RDG is a utility designed to manipulate

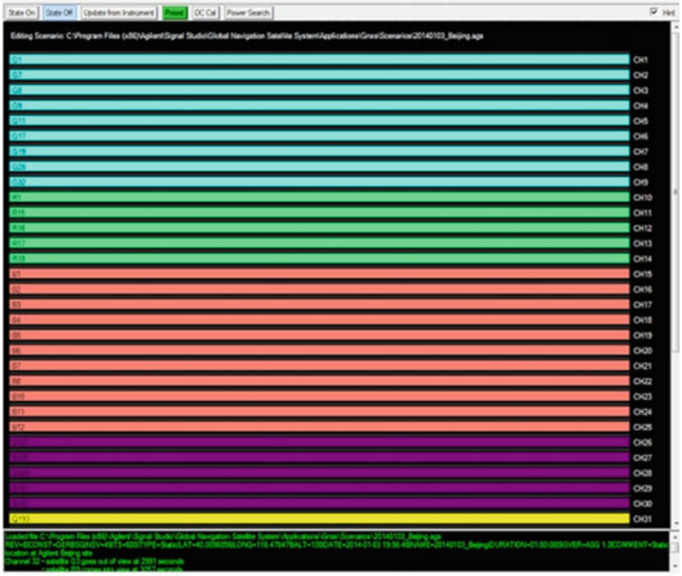


Fig. 3 Raw data

the raw data in the archival area. This utility can be used to access the raw data and process it in any format needed and access the required parameters. The RDG has many different options while retrieving the data. While storing the data in the MySQL database, since only processed data is stored, the raw data goes through the RDG, and only then it is populated into the database.

2.7 Requirement Specifications

The server part is coded with the following specifications.

- RDBMS
- MySQL Community Server v8.0 Popular Open source platform Backed by Oracle
- Operating System Linux Distribution
- Red Hat Enterprise Linux.

2.8 Methodology

The design given in Fig. 4 is the database architecture developed to store large amounts of data. The database designed here is 3-Dimensional. This method is incorporated to handle enormous parameters to analyze from the satellite. 2-Dimensional

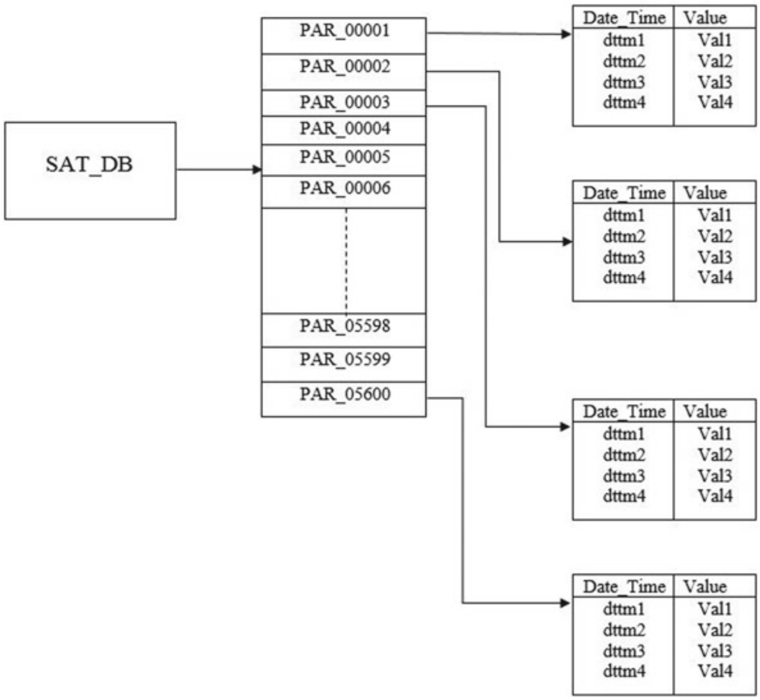


Fig. 4 Database architecture

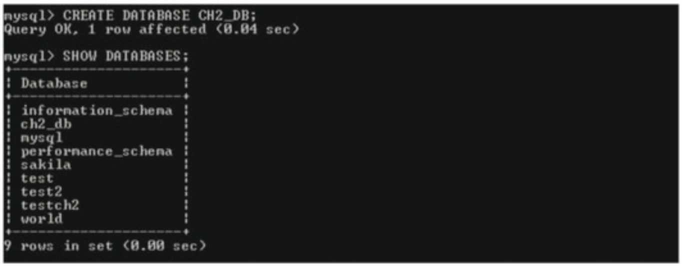
was impractical due to the large number of parameters and MySQL has a restriction to the number of columns to be used in a particular table. Making the table 3-D makes it easier, more efficient and user friendly to access and analyze the data. Since the database is 3-Dimensional it is highly scalable and has increased flexibility [7].

2.9 *Create Database*

Here it can be seen a number of databases created in Fig. 5.

2.10 *Table Creation*

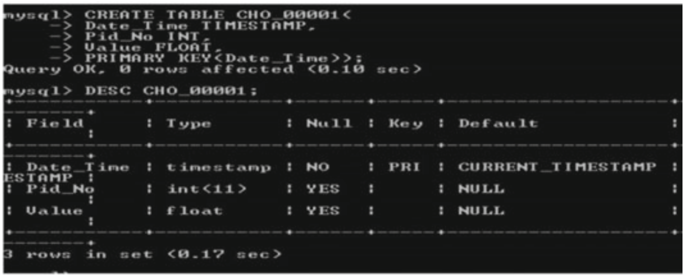
In the Fig. 6, it is made the column Data Time as PRIMARY because to avoid duplicate present within the system. Now when the table is created, script is put in a loop using bash in the Linux command line and add all the tables needed for all



```
mysql> CREATE DATABASE CH2_DB;
Query OK, 1 row affected (0.04 sec)

mysql> SHOW DATABASES;
+--------------------+
| Database           |
+--------------------+
| information_schema |
| ch2_db             |
| mysql              |
| performance_schema |
| sakila             |
| test               |
| test2              |
| testch2            |
| world              |
+--------------------+
9 rows in set (0.00 sec)
```

Fig. 5 Database creation



```
mysql> CREATE TABLE CHO_00001(
    -> Date_Time TIMESTAMP,
    -> Pid_No INT,
    -> Value FLOAT,
    -> PRIMARY KEY(Date_Time));
Query OK, 0 rows affected (0.10 sec)

mysql> DESC CHO_00001;
+-----------+-----------+------+-----+-------------------+
| Field     | Type      | Null | Key | Default           |
+-----------+-----------+------+-----+-------------------+
| Date_Time | timestamp | NO   | PRI | CURRENT_TIMESTAMP |
| Pid_No    | int(11)   | YES  |     | NULL              |
| Value     | float     | YES  |     | NULL              |
+-----------+-----------+------+-----+-------------------+
3 rows in set (0.17 sec)
```

Fig.6 Table created

the parameters. Now that all the tables are present, the data will be added into the database continuously within the respective parameters IDs.

2.11 Load Data

In MySQL, the LOAD DATA explanation peruses columns from a book record at a rapid. To compose information from a table into the database the SELECT INTO OUTFILE is utilized [6]. To peruse a document over, utilize the LOAD DATA INFILE [8]. In MySQL usually, one file at a time can be added using the MySQL user interface (Fig. 7).

Bash scripting is used to load multiple text files into the database which is much easier than using external software through a connector.

Fig. 7 Loading data using the SOURCE command

iii. LOAD DATA

```
Syntax: Source <drive_name:/location of
.sql file
Example:
Source e:/MySQL/textfiles/loaddata.sql
```

Table 1 Comparison test between RDG and MySQL database

Amount of data	Time taken by PVG (s)	Time taken by MySQL database (s)
• 5 months data: 1 Oct–28 Feb	138.572	27.25
• 20 day's data: 15 Oct–5 Nov	10.06	8.53
• 20 day's data: 20 Dec–10 Jan	11.107	9.54
• 15 day's data: 5 Jan–20 Jan	7.392	2.48
• 20 day's data: 5 Nov–25 Nov	9.803	4.57
• 15 day's data: 25 Jan–10 Feb	60.783	3.45
• 20 day's data: 20 Oct–10 Nov	10.277	5.26
• 15 day's data: 5 Jan–20 Jan	7.392	2.40
• 15 day's data: 10 Dec–25 Dec	7.141	3.10

2.12 Partitioning

The reason why we are incorporating partitioning is because we can have more flexibility between the large amounts of data in the database while retrieving the data at higher speeds (Table 1).

3 Conclusion

The proposed database architecture design ensures an efficient way of storing a large amount of satellite data at high speeds and scalable. The database architecture is based on time and additionally by making the database. Dimensional the data is well organized and easily accessible. Partitioning is introduced into the database to maximize the retrieval performance, and the results have proven to be positive when compared to the previous method. By incorporating these methods, it is possible to achieve greater processing speed for retrieving and analyzing the data, scalable database architecture, increased flexibility for data analysis, and handling of null values in a better way.

Acknowledgements It is acknowledged that Mr. Amith Kumar Singh, Mr. Karthik D V, Mr. Viraj Tilara have extended their support and gave valuable comments.

References

1. Palanisamy S, Suvitha Vani P (2020) A survey on RDBMS and NoSQL databases MySQL vs MongoDB. IEEE
2. Patil MM, Hanni A, Tejeshwar CH, Patil P (2017) A qualitative analysis of the performance of MongoDB vs MySQL database based on insertion and retrieval operations using a web/android application to explore load balancing—sharding in MongoDB and its advantages. IEEE
3. Sunny Kumar A (2016) Performance analysis of MySQL partition, hive partition-bucketing and Apache Pig. IEEE
4. Tongkaw S, Tongkaw A (2016) A comparison of database performance of MariaDB and MySQL with OLTP workload. IEEE
5. Győrödi C, Győrödi R, Pecherle G, Olah A (2015) A comparative study: MongoDB vs. MySQL. IEEE
6. Li D, Han L, Ding Y (2010) SQL query optimization methods of relational database system. IEEE
7. Vikram Phaneendra S, Madhusudhana Reddy E Big data—solutions for RDBMS problems—a survey. IJARCCE
8. Schwartz B, Zaitsev P, Tkachenko V High performance MySQL: optimization, backups, and replication

Multi-Terrain Rover Based on Rocker-Bogie Mechanism

Jai Prakash Prasad, Virupaksha V. Sahukar, M. Suresh, K. R. Shreyas Shetty, and N. Vinay Kumar

Abstract During this paper we have got projected to create in all-terrain mobile automation based Rocker Bogie Mechanism with Arduino as microcontroller which will move across alternative planet surfaces to seek out data and to require samples. The rovers of planetary motion have important part in exploring the areas of mars or any other planet when landed and they should work as planned for future missions. This Rover could be a sensible, reconfigurable all-terrain multi mission small Rover it was made to deal with problems of exploration of planet. Rover stands for remotely operated video increased receiver. Rocker bogie mechanism could be a mechanism primarily employed in the mars rovers to beat the rough terrains whereas maintaining stability. The look consists of a spring free suspension-based difference drive system that enables the bogie to maneuver over rocks, pebbles with ease.

Keywords Rocker Bogie mechanism · Multi-terrain · Motor driver · Rover · Arduino · Bluetooth module · Wireless camera · Suspension free

1 Introduction

There is increasing want for mobile robots that are ready to work in unknown environments and extremely uneven parcel [1, 2]. These are measure principally used for tasks that common man unable to do and that don't seem to be safe. So as to realize these kinds of missions, any transportable automaton has to have an acceptable mobile system per every scenario. Among these mobile systems, it's the rocker-bogie suspension that was first used for the Mars Rover resident and it's presently NASA's favored style for rover wheel suspension. The rocker bogie suspension could be a mechanism that allows a six-wheeled vehicle to inactively stay up with the latest with a surface even once driving on seriously lopsided bundle. So as to overcome back vertical obstacle faces, the wheels which are at front, are forced on the object using

J. P. Prasad (✉)
Department of ECE, Don Bosco Institute of Technology, Bengaluru, India

V. V. Sahukar · M. Suresh · K. R. S. Shetty · N. V. Kumar
Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_97

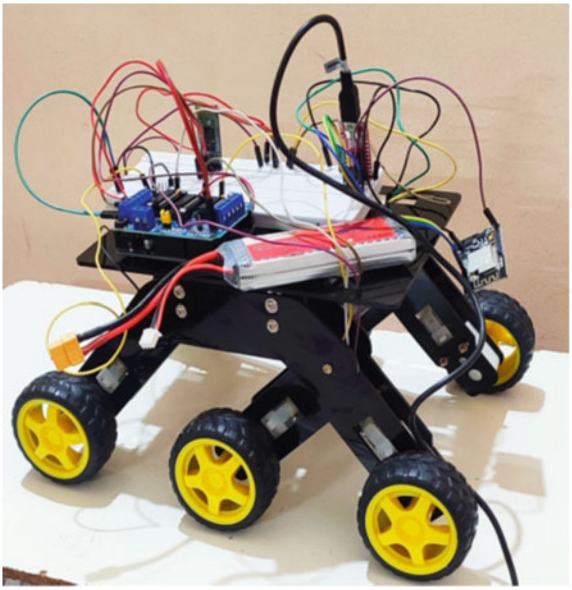


Fig. 1 Rocker bogie mechanism

the Centre and rear wheels that generate most needed torsion [3]. The pivot of the forward wheel at that point lifts the forward of the vehicle over-top the obstruction and hindrance overwhelmed. Those wheels that stay in the center, is then pressed against the deterrent by the back haggles against the hindrance by the front until the time it is upraised up and over. At last, the wheel which is at backside is force over the object at the front 2 wheels thanks to applying pull force. During every wheel's traversal of the obstacle, forward progress of the vehicle is slowed or utterly halted that finally maintain vehicles Centre of gravity as Refer to Fig. 1.

2 Literature Survey

In Ref. [4] this describes a unified live of stability of a Rocker Bogie vehicle that accounts for the tendency to slip, tip over, or lose contact with the bottom considering each static equilibrium and dynamic effects.

In Ref. [5] we tend to use mechanical managemental model sensing element system. It was developed to observe and control the sensing element performance of the rover. The mechanical man application developed for Smartphone has capability to modify on/off sensing element device over Wi-Fi. Figure 1 shows Rocker Bogie Mechanism.

In Ref. [6] we tend to describe a technique for collision rejection among multiple robots in plate like environments, that models sensible inter robot interaction and

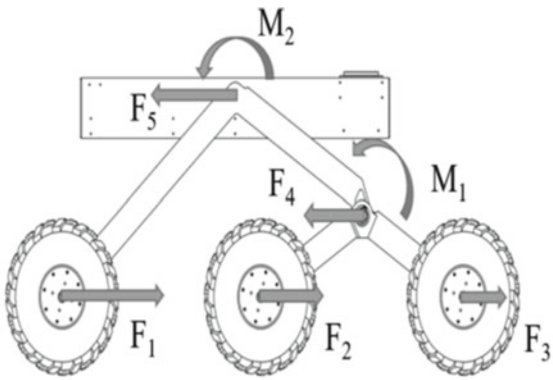


Fig. 2 The forces acting during turning the robot

higher cognitive process, this technique additionally helps in permitting sleek transmission and safety navigation in unknown surroundings of the rover.

In Ref. [7] we have a hybrid design, composed of a centralized planner and native managemental that gives motion control and network routing so as to finish a task for a multirobot team in illustrious environments, was projected.

In Ref. [8] we've got thought of the matter of coming up with the trajectories of an automaton once its measurements an maliciously compromised by an aggressor, this helps in security and dominant of the rover in non-adversarial parcel.

3 Problem Formulation

The major limitation of current rocker-bogie is its inability to climb steps over doubly the wheel diameter. Having only one movable joint with 2 tyres, the peak climbed is proscribed. The uneven parcel of land keeps on displacing the cg of car. Hence the design ought to be therefore created that the cg, even once displaced, remains underneath the crucial worth maintaining the steadiness. When the front tyres face associate obstacle the second tyre pushes it forward but the third tyre pulls it back off unless the link of the front tyre makes angle of over 90° with the vertical. For rocker-bogie there's a limitation in angle once that the vehicle is upturned. Generally, the angle for economical traversing is 45° however with some style modifications, the rocker will go uphill of over 50° (Fig. 2).

4 Methodology

The basic Schematic figure of the project is shown in Fig. 3. Mainly this block diagram consists of the following essential blocks. Arduino Uno, L293D Motor Driver, ESP32

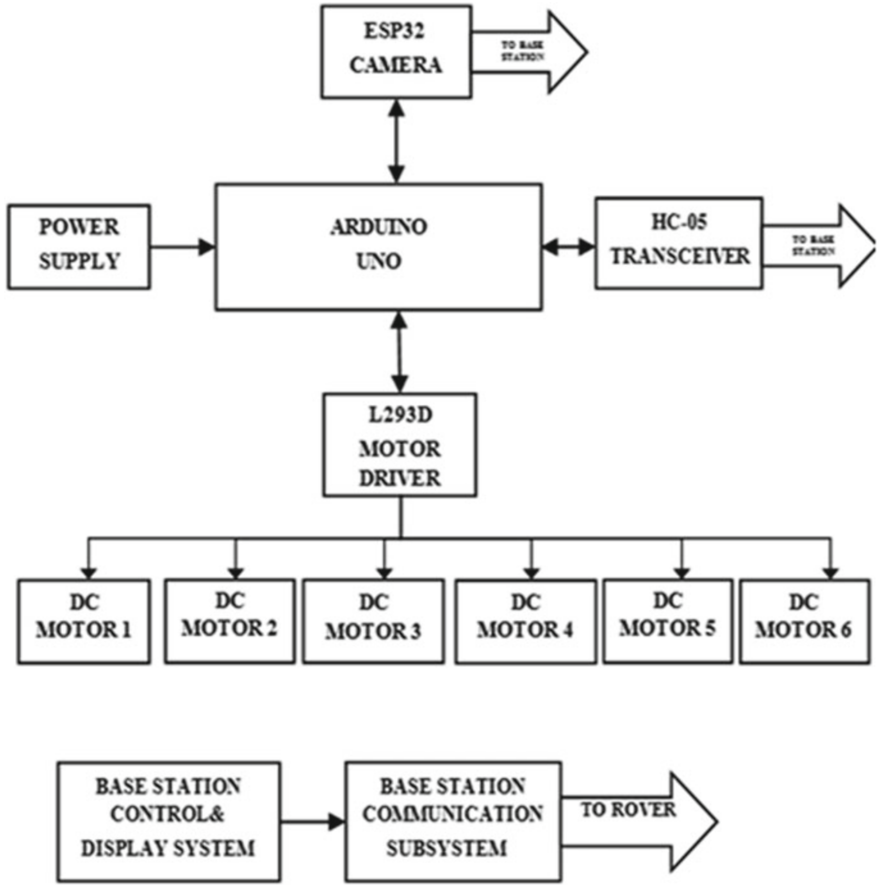
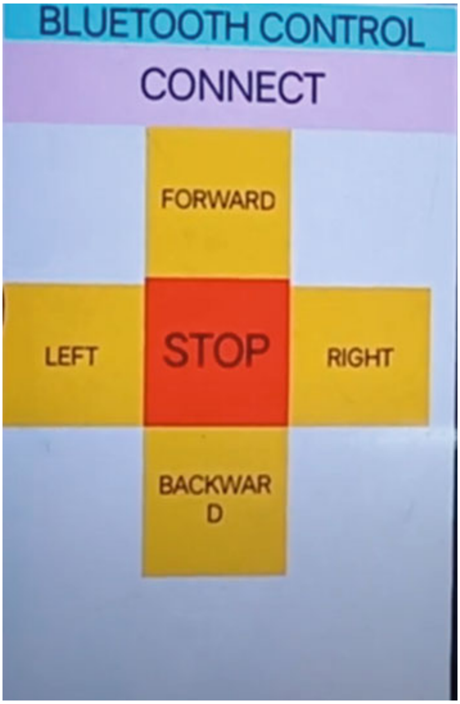


Fig. 3 Schematic figure of the Rover

Camera, HC-05 Transceiver Bluetooth module, Power supply, DC Motors, Base Station Control & Display System, Base Station Communication System.

In order to design, we used 3 sections input section, output section and circuit section. In input section it contains power supply, here we are using 9 V battery to power up the circuit. Output section contains ESP32 Camera for live video recording and image capturing, The HC-05 Transceiver is used for data transferring in the serial communication, And L293D Motor Driver for the direction and speed control of the DC Motors. Circuit section is one of the main and important sections of the project. It mainly consists of all the circuits which are processing the instruction that are received from the sensors and input section which then produce the respective output.

Fig. 4 Android application



5 Control System

The signals which are required to control the movement of rover is sent to microcontroller through wireless Bluetooth module. By using android or IoS application we can connect our Bluetooth module to our smart phone. User interface of Bluetooth app is shown in Fig. 4. When we select required command in app it sends the data to microcontroller which in turn controls movement and direction of wheels [16].

Here Fig. 4 is the Application Interface used in this project Also, we use a camera which is based on ESP 8266 Module ESP-32 wireless camera module with ov2640 camera is used to receive video and images captured by rover. ESP-32 is a Wi-Fi enabled device which connects to Wi-Fi network. Using IP address, we can view live stream or capture images through website opened by corresponding IP address, which is as shown in Fig. 5.

6 Applications

- The rover is also utilized for studying the geology and Environment of selected area [9].

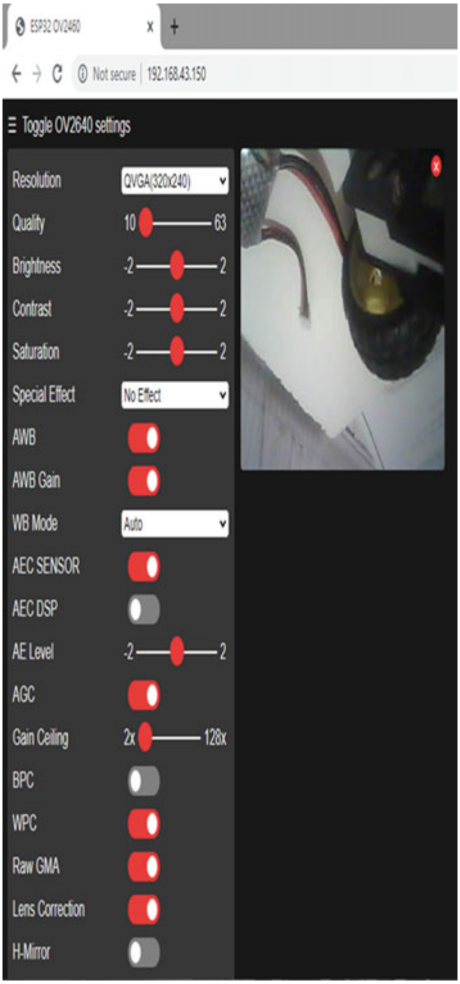


Fig. 5 ESP-32 camera live stream on internet through an IP address

- The rover can be used for the Military and defense Application, e.g. Metal detector.
- With certain advancements like appending arms to the meanderer it very well may be made valuable for the Bomb Diffusing Squad to such an extent that it tends to have the option to cut the wires for diffusing the bomb [1, 10].
- It can also be converted it into a Wheel Chair [1, 9]. It is also utilized as stair case climbing in houses where patients and old folks may face difficulty climbing. The rover can be used in Industrial Application, e.g. Load Carriers.
- If the size is increased then we could use the rover to carry large military weapons [17].

7 Advantages

The most significant advantage of created system is reduction of the suspension mass [11]. Without steering wheels rover has 6 motors less than classic rocker-bogie based robots. Less mass and less amount of motors allows to design very simple and durable solutions based e.g. on aluminum profiles. Differential steering method unfortunately increases friction especially on high CoF surfaces. Solution showed in this paper creates an opportunity to make cheap 6 wheeled all terrain platform without difficult to design wheel turning mechanism [15].

8 Result

The design and made model will go up the angle of 45°. Also we have a tendency to tested for the net cam with recording mounted on rocker bogie system and located satisfactorily performance obtains throughout this take a look at camera has revolved around 360°. During stair rising take a look at for length but 375 mm (15 inch) system cannot climb the support. It will be potential to develop new models of rocker bogie which might climb the steps having low lengths.

9 Conclusion

The growth of science and technology makes human life easier. The main theme of this project is to build Multi-function rover based on Rocker Bogie mechanism [12]. The idea, one rocker-bogie system with two modes of operation, is useful to increase the effectiveness and efficiency of the all-terrain mobile robots. This auxiliary improvement can make the rocker-bogie more versatile and productive in circumstances that require rapid crossing or managing surfaces that need a heartier exhibition over extreme impediments. This is utilized for used space exploration, this designed architecture is suitable for all terrain like rigid and smooth surface. A rover can take an advantage of multi-legs to walk easily, with increasing advancement in rover technologies for extra-terrestrial exploration. Its real time video surveillance and feedback feature helps in detecting various threats in place where human reach is not possible [2]. This show rocker bogie framework works at various surfaces. As stated by the distinctive weight following up on interface decides force working on it [1]. By accepting exact step measurements, precisely dimensioned rocker bogie can climb the step with extraordinary steadiness. The structure and made model can climb the edge up to 45°. Also tested for the cam with AV recording mounted on rocker bogie system and found satisfactorily performance obtains during this test camera is utilized for capturing the video and can be streamed using an IP address [13, 14]. With increasing advancements in Rover technologies for extra-terrestrial exploration, this robotic innovation could be seen as a potential application on earth [2]. It can likewise be utilized in military and guard field as a result of its reduced plan. Meanderers in not so distant future may supplant the need of human necessity for recon, looking over, danger giving, etc. There is boundless potential in this innovation to be utilized in different applications.

Acknowledgements This work was supported by Don Bosco Institute of Technology, under the guidance of Dr. Jai Prakash Prasad, Associate Professor, Dept of ECE, Bengaluru, India.

References

1. www.ijetajournal.org
2. www.slideshare.net
3. www.irjet.net
4. Mann MP, Shiller Z (2016) Dynamic stability of a Rocker Bogie vehicle: longitudinal motion Q. In: International conference on robotics and automation
5. Kim D, Hong H, Kim HS, Kim JW (2011) Optimal design and kinetic analysis of a stair-climbing mobile robot with rocker-bogie mechanism. Elsevier
6. Javier-Alonso-Mora, Beardsley P, Siegwart R (2018) Cooperative collision avoidance for nonholonomic robots. IEEE Sens J 1552-3098
7. Stephan J, Fink J, Kumar V, Ribeiro A (2017) Concurrent control of mobility and communication in multi-robot systems. IEEE Sens J 15523098
8. Nádvorník J, Smutný P (2014) Remote control robot using Android mobile device. IEEE conference
9. Harish Chandu K, Hari Narayana P, Charan Teja KC, Sai B, Murali Mohan Y (2018) Fabrication of rough terrain vehicle using Rocker Bogie mechanism. IJETA
10. Bankar ND, Faiz D, Kulkarni T, Kulkarni Y, Jagtap S (2020) Design and fabrication of stair climbing mechanism. IRJET
11. Tomaszuk P, Łukowska A, Re Vko M, Dzier Xek K, Straszy Wski P (2019) Active wheel speed control to avoid lifting the swingarms in rocker-bogie suspension. IEEE
12. www.hindawi.com
13. Babu B (2018) Design and fabrication of rocker bogie mechanism geo survey rover. Int J Sci Dev Res 154–159
14. www.ijert.org
15. Thueer T, Siegwart R (2010) Mobility evaluation of wheeled all-terrain robots. Robot Auton Syst 58:508–519
16. https://www.arduino.cc/en/guide/introduction
17. www.ijsdr.org

Comparative Study of Identifying Biomarkers for ASD Classification Using a Genetic Database

Ayesha Uzma Khan, M. Shalini, L. Shweta Bai, B. Sindhu, and B. S. Roopa

Abstract Autism Spectrum disorder is a neurodevelopmental disorder. This disorder is different compared to a normal person given his/her interaction with other people in terms of poor eye contact, speech deficiency (i.e., Abnormal tone in voice, blabbering, neologism) delay in learning to speak, repetitive behavior such as hand flapping or body rocking. In the medical field, genetics research is an interesting area. In recent years, various new techniques have been developed for detecting the wide range of diseases associated with genes. Identifying the genes which are prone to ASD from a vast amount of genetic data is a challenging task. In this paper, different methods have been used to find the generalized way of obtaining biomarkers which in-turn plays a very important role in classifying ASD over non-ASD. These methods include supervised learning and unsupervised learning methods which consist of a combination of statistical filters to select the most prominent genes, kinematic analysis, clustering method, and various machine learning classifiers. The main aim here is to find and justify which dataset is best to classify ASD over non-ASD individuals. After comparing all the methods used, it indicates that the most significant genes which emerge as biomarkers help in the better classification of ASD over non-ASD individuals based on the gene expressions

Keywords Autism spectrum disorder (ASD) · Genes · Genetic data · Gene expression · Machine learning

1 Introduction

Autism, also referred to as Autism Spectrum Disorder (ASD), could be a set of mental illness of the neuro-developmental type. Autistic individual frequently experiences troubles with social interaction and communication and exhibits repetitive patterns in interest, behavior. These autistic symptoms can be observed whose age groups are from 1 to 2. The uncertainty in autism is because of various risk factors including

A. Uzma Khan · M. Shalini · L. Shweta Bai · B. Sindhu · B. S. Roopa (✉)
Department of Electronics and Communication Engineering, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_98

certain genetic conditions as autism contains a strong genetic base. The risk factor is evaluated based on the diagnosed case history of autism [1]. The diagnosis relies on symptoms. The treatment for autism includes various behavioral therapies [2] for each individual and includes the teaching of various skills to be acquired by each autistic individual

Compared to men, women are diagnosed four times less usually. The word "spectrum" in autism spectrum disorder refers to the variance within the complexity of symptoms [3]. Those individuals with severe symptoms have to depend on others daily and those who have few symptoms can manage themselves and turn out to be independent

2 Literature Survey

2.1 Selection and Classification of Gene Expression in ASD Using a Combination of Statistical Filters and GBPSO-SVM Algorithm

In this paper [4], the prominent genes are derived which contribute more to classify the ASD over non-ASD. This is achieved by using two methods. First is the statistical filter and another is wrapper-based filter. In this paper, the method used is GBPSO-SVM algorithm (Geometric Binary Particle Swarm Optimization-Support Vector Machine). The dataset is taken from NCBI repository having 54,613 genes from 146 samples.

2.2 Hybrid Feature Selection Method

Machine learning techniques can be efficiently and effectively used for medical diagnosis. Single nucleotide polymorphism is found to be one of the most important sources in human genome variability. ASD SNP microarray dataset [5] from NCBI GEO is used.

To examine the data of SNP and to differentiate between healthy and affected samples, the SNPs dataset has been applied to various techniques to separate or to isolate the normal ones from the affected samples [6]. One among them is a hybrid feature selection method [7]. The optimal subset of SNPs is extracted on the application of the Hybrid feature selection model [8]. The proposed method's performance has been measured by applying the extracted SNPs into different classifiers.

2.3 A Comparison of Machine Learning Algorithms

In this paper, machine learning algorithms were applied to the data collected from a single site which gathered the data for educational evaluations and medical purpose which is not publicly available. Supervised learning algorithms were applied to the data gathered to diagnose whether children meet the factors of ASD or not [9]. So, all the eight algorithms' performances are evaluated over ten random train-test splits of the data by considering the classification accuracy of all eight algorithms to compare their potential use for monitoring [10]. The false negative is produced more by NB-SVM than false positive, but the random forest didn't. All the algorithms were compared and accuracy was tabulated. Among all the applied algorithms, NB-SVM was found to be more accurate.

2.4 Identifying Autism in Children Using Kinematic Analysis

In this paper, a conceptual study of simple upper-limb movement is demonstrated to verify whether it is useful to classify children with ASD whose age groups are between 2 and 4. Here, in this study, a supervised machine learning method is proposed to classify 15 children with ASD over 15 normally developing children using kinematic analysis by executing a simple reach-to-drop task. This method of kinematic analysis gives a classification accuracy of 96.7% by taking the features of the movements related to the goal of the task. The results show that the motor deformity associated with autism would be suitable and efficient to determine ASD.

2.5 Identifying Autism Using the Features of Restriction

Usually, to identify ASD, the assessment of an individual's behavior plays a very important role. In this paper, the restricted behavior or restrictedness in an individual can be very useful to classify ASD over non-ASD. Here, the classification is performed using machine learning classifiers and the results exhibited that the KNN algorithm gives the utmost classification accuracy of 88.37%. This study shows that the kinematic features related to restriction could be effective to distinguish between ASD and non-ASD. Also, by applying machine learning for classification results in an efficient diagnosis of ASD.

2.6 Unsupervised Machine Learning

Autism spectrum disorder (ASD) is basically a heterogeneous disorder. Research has been able to identify various subgroups of ASD which are genetically and behaviorally distinct, but research is yet to use the machine. Autism spectrum disorder (ASD) is basically a heterogeneous disorder. Research has been able to identify various subgroups of ASD which are genetically and behaviorally distinct, but research is yet to use the machine.

2.7 Emotion Recognition System for Autism Children Using Non-verbal Communication

The facial expression, prosody, bodily expressions, or language are various ways to express emotions for human beings. Autism Spectrum Disorder (ASD) is actually a neurodevelopmental disorder, where the individual suffering is inefficient in communication skills as well as social skills. Here in this particular paper, the various emotions of children have been identified with the help of their body movement. It is difficult to identify the emotions of an ASD individual. To detect the type of emotions an ASD child is going through, various body movements of the child were used.

3 Methodology

The methodology in this paper consists of seven different methods applied to find a generalized way of obtaining biomarkers which comprises of both supervised and unsupervised machine learning methods.

3.1 Method 1

(1) In this step, similar genes are removed by calculating the mean and median ratio of gene expression. Those genes whose ratio is greater than or equal to 0.95 are removed. Therefore, in this step, the 54,613 genes are reduced to 9454 [4].

(2) The 9454 genes are divided into two parts. The first 85% is used for model training and testing. The remaining 15% noninvolved set was used as a new real-world dataset over the gene classification based on the predefined model. Three filters namely TT (t-test), COR (Feature Correlation), and WRS (Wilcoxon Rank Sum test) were applied and this was repeated for 10-folds which resulted in the selection of 200 most prominent /discriminative genes.

(3) In this step, the GBPSO-SVM algorithm is applied and most discriminative subsets of genes are selected and classification is performed on resulting genes. The three subsets of genes generated are finalized into a merge subset of genes based on frequencies of appearance of genes in these three subsets of genes.

3.2 Method 2

The framework of the hybrid feature selection method consists of three stages:

(a) Pre-processing stage
(b) A hybrid feature selection stage
(c) A classification stage.

Further analysis of each stage is explained as:

(a) Pre-processing stage: This stage consisting of data transformation and data refinement.
Data Transformation: SNPs get transformed into numerical form.
Data Refinement: It follows 2 steps, removal of redundant SNPs, replacing by missing values by making use of attribute mode.
(b) A hybrid feature selection stage: This stage is involved in the selection of the most informative subset of features.
(c) A Classification stage: Various classifiers such as Naive Bayes (NB), SVM, k-NN, and Linear Discriminant Analysis (LDA) and were used to measure the performance of the proposed hybrid feature selection method.

3.3 Method 3

In this paper, they have used 8 Difference Algorithms such as latent semantic analysis (LSA) [11], latent Dirichlet allocation (LDA) [12, 13], support vector machine (SVM) with a linear kernel [14], multinomial naive Bayes (MNB) [15], interpolated Naive Bayes-SVM (NB-SVM) [16, 17], and two neural networks [18].

3.4 Method 4

Here, the classification of ASD over non-ASD is performed using a supervised machine learning model where 15 children with ASD and 15 normally developing children are taken to execute a kinematic task i.e., a simple reach-to-drop task [19]. This method of kinematic analysis gives a classification accuracy of 96.7% by taking the features of the movements related to the goal of the task.

3.5 Method 5

In this method, machine learning plays a key role in classifying ASD over non-ASD using the features of restriction. Here, 20 children with ASD and 23 normally developing children are taken and they were made to execute a kinematic task which required the maximum variance in their movements [20].

The key elements evaluated from the movements are amplitude, velocity, and acceleration. Then, entropy and range were calculated for each of these key elements followed by applying machine learning classifiers which includes support vector machine (SVM), random forest (RF), decision tree (DT), linear discriminant analysis (LDA), and k-nearest neighbor (KNN) for classification.

3.6 Method 6

Unsupervised machine learning was applied to the ASD model subgroups [21]. The regression method was applied later for analysis of treatment response for each subgroup.

After the GMM was applied, 16 subgroups were identified. Later on, from the application of Hierarchical Agglomerative Clustering, the subgroups were further grouped under 2 categories which had variations in the development of the skills. Thus, once the subgroups were identified, treatment could be applied accordingly.

3.7 Method 7

Here, 12-dimensional body movement features like acceleration, velocity, distance, and angle from the left hand, head, and right hand are used to identify the various emotions of ASD children using the various movements in their body parts like left hand, head, and right hand [22]. The dataset used here includes comprised videos of children of ASD whose age groups varied from 5 to 10 years and the total number of samples is 10 in number. The data obtained was further given to classifiers like Random Forest (RF) and Support Vector Machine (SVM) to identify the emotions of the ASD children. Quantitative analysis can be used to calculate the performance. Thus, the identification of various emotions like anger, fear, happy, sad, and neutral within the ASD individual was easier to predict.

4 Comparison of Results

See Table 1.

Table 1 Comparison of all methods used

Methods used	Outcome	Inference
Method 1	NCBI repository having 54,613 genes146 samples. A combination of statistical filters results in the extraction of most prominent genes prone to ASD	Classification accuracy of 92.1%. Each filter method identifies the most discriminative genes among them
Method 2	ASD SNP microarray dataset from NCBI GEO repository. To examine the data of SNP and to differentiate between healthy and affected samples	Classification accuracy of 89% is achieved The method SVM-RFE was found to be very useful to eliminate unwanted genes
Method 3	A single site which gathered the data for educational evaluations and medical purpose. The performance of machine learning algorithms was evaluated	Among all the applied algorithms, NB-SVM was found to be more accurate
Method 4	15 children with ASD and 15 normally developing children. This method results in a classification by considering the features of the movements related to the goal of the task	Classification accuracy of 96.7%. Machine learning classifiers play a vital role in classification
Method 5	20 children with ASD and 23 normally developing children were taken. Applying machine learning for classification results in an efficient diagnosis of ASD	Classification accuracy of 88.37%. Here, kinematic features related to restriction could be effective to distinguish between ASD and non-ASD
Method 6	Total of 2400 samples. 16 clusters (subgroups) were identified on basis of evaluation on 8 skills	Due to limitations in unsupervised machine learning, this method was found to be inefficient
Method 7	Comprised videos of children of ASD body movement features were used for recognizing the emotions of autistic children	SVM shows a good accuracy rate of 96.3%

5 Conclusion

From all the above methods used in various papers referred here, the genetic dataset is feasible and efficient to classify ASD over non-ASD individuals compared to all other datasets like kinematic data, videotapes, data obtained after application of ABA treatment using various questionnaires. Any changes in a specific gene will result in the identification of ASD in an individual. If a kinematic or video dataset has been used, there can be chances of misdiagnosis as a non-ASD person too can have a physical disability similar to an ASD individual or vice versa. In recent years, various new techniques have been developed for detecting a wide range of diseases associated with genes which can be of great use in the field of medicine. In the medical field, the biomarkers act as a measurable indicator to identify various diseases. These biomarkers help in early diagnosis, drug target identification, drug reaction, disease preclusion, etc., as ASD is a spectral disorder that has many symptoms. The biomarkers play a very important role in the identification of ASD in an individual. Among the seven different methods applied, a combination of statistical filters resulted in a classification accuracy of 92.1%, hybrid feature selection method resulted in a classification accuracy of 89%. Similarly, kinematic data analysis resulted in a classification accuracy of 96.7, 88.37, and 96.3% (SVM) respectively. In the meanwhile, the most significant genes emerge as biomarkers which help in the better classification of ASD over non-ASD individuals based on the gene expression.

References

1. Tick B, Bolton P, Happé F, Rutter M, Rijsdijk F (2016) Heritability of autism spectrum disorders: a meta-analysis of twin studies. J Child Psychol Psychiatry 57(5):585–595. PMC 4996332. PMID 26709141. https://doi.org/10.1111/jcpp.12499
2. Fleuret F (2004) Fast binary feature selection with conditional mutual information. J Mach Learn Res 5(Nov):1531–1555
3. Lai MC, Lombardo MV, Chakrabarti B, Baron-Cohen S (2013) Subgrouping the autism "spectrum": reflections on DSM-5. PLOS Biol 11:4. PMC: 3635864. PMID: 23630456. https://doi.org/10.1371/journal.pbio.1001544
4. Hameed SS, Hassan R, Muhammad FF (2017) Selection and classification of gene expression in autism disorder: use of a combination of statistical filters and a GBPSO-SVM algorithm. PLoS ONE 12(11):e0187371. https://doi.org/10.1371/journal.pone.0187371
5. Waddell M, Page D, Shaughnessy Jr J (2005) Predicting cancer susceptibility from single-nucleotide polymorphism data: a case study in multiple myeloma. In: Proceedings of the 5th international workshop on bioinformatics. ACM, pp 21–28
6. American Psychiatric Association (2013) Autism spectrum disorder. 299.00 (F84.0). In: Diagnostic and statistical manual of mental disorders, (DSM-5), 5th edn. American Psychiatric Publishing, Arlington, VA, pp 50–59. ISBN 978-0-89042-559-6. https://doi.org/10.1176/appi.books.9780890425596
7. Saeys Y, Inza I, Larrãnaga P (2007) A review of feature selection techniques in bioinformatics. Bioinformatics 23(19):2507–2517
8. Alzubi R, Ramzan N, Alzoubi H (2017) Hybrid feature selection method for autism spectrum disorder SNPs, SBN information. INSPEC Accession Number: 17240499, IEEE, https://doi.org/10.1109/CIBCB.2017.8058526
9. Roopa BS, Manjunatha Prasad R (2019) Concatenating framework in ASD analysis towards research progress. In: 1st International conference on advanced technologies in intelligent control, environment, computing & communication engineering (ICATIECE). Bangalore, India, pp 269–271. https://doi.org/10.1109/ICATIECE45860.2019.906
10. Lee SH, Maenner MJ, Heilig CM (2019) A comparison of machine learning algorithms for the surveillance of autism spectrum. PLoS ONE. https://doi.org/10.1371/journal.pone.0222907
11. Dumais ST, Furnas GW, Landauer TK, Deerwester S, Harshman R (1988) Using latent semantic analysis to improve access to textual information. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, pp 281–285
12. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 993–1022
13. Ramage D, Hall D, Nallapati R, Manning CD (2009) Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. In: Proceedings of the 2009 conference on empirical methods in "natural language processing", vol 1. Association for Computational Linguistics, pp 248–256

14. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
15. Rennie JD, Shih L, Teevan J, Karger DR (2003) Tackling the poor assumptions of Naive Bayes text classifier. In: Proceedings of the 20th international conference on machine learning (ICML-03), pp 616–623
16. Wang S, Manning CD (2012) Baselines and bigrams: simple, good sentiment and topic classification. In: Proceedings of the 50th Annual meeting of the association for computational linguistics, short papers, vol 2, pp 90–94
17. Mesnil G, Mikolov T, Ranzato MA, Bengio Y (2014). Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews. arXiv preprint arXiv:1412.5335
18. Joulin A, Grave E, Bojanowski P, Mikolov T (2016) Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759
19. 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 Dev Disord 45:2146–2156. https://doi.org/10.1007/s10803-015-2379-8
20. Zhao Z, Zhang X, Li W, Hu X, Qu X, Cao X, Liu Y, Lu J (2019) Applying machine learning to identify autism with restricted kinematic features. IEEE Access 7:157614–157622. https://doi.org/10.1109/ACCESS.2019.2950030
21. Stevensa E, Dixonb DR, Novackb MN, Granpeeshehb D, Smithc T, Linstead E (2019) Identification and analysis of behavioural phenotypes in autism spectrum disorder via unsupervised machine learning. Int J Med Inf 129:29–36
22. Santhoshkumar R, Geetha M (2019) Emotion recognition system for autism children using nonverbal communication. Int J Innov Technol Explor Eng (IJITEE) 8(8). ISSN: 2278-3075

Unmanned Marine Robot

G. H. Sangamkumar, N. Dinesh Kumar, Satyam Shreeh, B. M. Suhas, and Vivek Ambi

Abstract It is a review of floating robot vehicle which has multi uses. It is formulated on unmanned systems which is on high demand I surveillance field. The developed SGT (Spatial Grasp Technology), SGT is been used in this system, which is able to pleasant these needs can be briefed. In addition to their combinations, SGT uses high-level holistic challenge scenarios that self-navigate and cover the entire systems in a high-quality virus mode. SGT can successfully assist gradual transition to computerized as much as absolutely robot systems below the unified command and manipulate. At once operating with virtual, physical, and executive spaces, this brings pinnacle operations, data, preference common sense, and simple command and control to the allotted resources at run time, providing flexibility, ubiquity, and functionality of self-restoration in solving complex problems, in particular the ones requiring quick reaction on unpredictable conditions. Specific scenarios of completing paper and dealing with robot collectives at particular conceptual levels in a unique language might be provided.

Keywords Arduino mega · Wireless camera · Conveyor belt · SGT (Spatial Grasp Technology) · Cloud-based surveillance system · Bluetooth module · WSN (wireless sensor network) · IoT(Internet of things)

1 Introduction

Surveillance is major target when the main focus is to secure an object as it is tedious job. It is risky to keep track around the surrounding objects, therefore developed model continuously monitors the object in a target area. The robot continuously sends the live streaming videos through wireless camera to control station. Practical implementation of this paper can be replaced human to surveillance robots. Due to small size this particular robot is capable enough to enter into the tunnels. It can also enter easily in mines and small holes of building. it is successful to live on in harsh and hard climatic situations for existence long with out inflicting any damage.

G. H. Sangamkumar (✉) · N. D. Kumar · S. Shreeh · B. M. Suhas · V. Ambi
Department of ECE, Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_99

these varieties of robots were designed from previous few many years and making the difference in army.

Today's Most of the system use a mobile robot with a wireless camera to keep eyes on activities. The camera which is set on the top of robot can move to different directions. Because of the movable digital camera those kinds of robots are more bendy. On this form of robots are flexible for motion of the robotic vehicle so wheel-based robots are more suitable for flat platform. With the improvement in wireless communication and internet, the films captured via wheel-based robotic may be visible remotely on laptop or computer. The sensors and the complex operations are however disadvantage of this robotic vehicle.

2 Literature Survey

U.S. Army Considers Replacing Thousands of Soldiers with Robots‖ [1] Aniket. A. More, Akshay. V. Jadhav & Sweta Patil presented the Surveillance Robot for Defense Environment, This paper presents gives a practical present-time approach for surveillance robots at remote locations and enemy territories using a remote controller based robotic vehicle on wireless technology that can be used for defense and military applications. The sensors and camera are used to detect and identify human, objects etc. This vehicle is designed to work in limited area with better efficiency for example In Armed forces. They can use these types of robot vehicles in hostage situations to determine the number of terrorists in the building, types of weapons used, bombs etc. The processing unit used in proposed system is Raspberry pi working on Raspbian operating system. The Pi board controls the movement, gather information using sensors and camera that is used to stream the real time video of surrounding to the operator.

Ackerman [2] V. Choudhary presented the IOT based Weather Monitoring and Reporting System, This paper supplied The tool proposed is a complicated answer for climate tracking that makes use of IoT to make its actual time data with out difficulty available over a totally big range. The tool deals with monitoring climate and climate changes like temperature, humidity, wind velocity, moisture, the use of more than one sensors. The information uploaded to the internet internet page can with out issues be to be had from anywhere in the international & moreover can be used for destiny references. ithe ones sensors ship the data to the internet net internet page and the sensor information is plotted as graphical information. The undertaking even consists of an app that sends notifications as an powerful alert system to warn humans approximately surprising and drastic adjustments inside the weather. Due to the compact format and much less shifting factors this layout requires a lot less safety. The additives on this undertaking don't eat an entire lot strength. Gadget is lots less luxurious and price powerful.

JMEST [3] Sarmad Nozad_Mahmood presented the proposed layout of climate tracking device the usage of Arduino. This paper particularly combines among have a look at fields-primarily based completely definitely manage structures and facts acquisition method, to create a database machine depending at the employed attributes to generate the furnished statistics. The number one attributes were decided on based totally completely totally on the sensors used to assemble the gadget so that you can create an effective climate station challenge. Real database introduction technology is taken into consideration the number one mission of this artwork, which gives an possibility to mine the records, recorded within the past. The proposed sensors used to degree and store Temperature, Humidity, and Wind pace information. The received information may be displayed in techniques identified as direct and oblique due to periodic statistics have a look at and storing the information as real database tool respectively. Furthermore, the entire tool supervises and governs locations domestically based totally totally on the periodic trade that occurs within the weather conditions, which will preserve the proposed places in preferred climate situations. Ultimately, moderate sensing module is protected with the module to offer climate station tool through the usage of the information regarding day / night time time instances based totally absolutely absolutely mild intensity.

IJIR [4] Trupti_Bhondve and R. Satyanarayan, r presented Cellular Rescue robotic for Human frame Detection of disaster. This paper gives Many areas of worldwide are becoming affected because of natural calamity. failures are fantastic & unstoppable activities which might be each guy made or herbal, at the side of terrorist assaults, earthquakes, wirelesses and floods and so on. Failures create emergency situations to provide fundamental services to the sufferers ought to be coordinated rapid. regularly, we look at that many humans die via trapping in the ones disasters but the human beings furthermore die on large scale certainly due to the reality they didn't get assist at immediately time or the help furnished to them is overdue. This paper proposes a mobile robot primarily based absolutely absolutely WiFi Sensor community (WSN) that is designed for human life & detection in an unmanned vicinity may be done wi-fi through an automatic system.

3 Methodology

The robot is constructed use of an Arduino board processor to which an ultrasonic sensor, temperature & humidity sensors, Bluetooth module, pH sensor, and air pressure sensor are connected to it (Fig. 1).

Arduino mega: Arduino Mega2560 is a microcontroller chip board based on the ATmega2560. The outputs of the sensors are the inputs to the Arduino. The outputs of the Arduino are the control signals to the motors and camera.

pH sensor: pH sensor is used to determine acidic, neutral and basic nature of the water that's within the pH range of zero to 14, where zero to seven indicates the acidic content of the water, seven indicates the water is neutral and 7–14 indicates the primary nature of the water.

Air pressure sensor: The BMP180 barometric pressure sensor is used to predict the weather by measuring parameters such as air pressure, temperature and humidity.

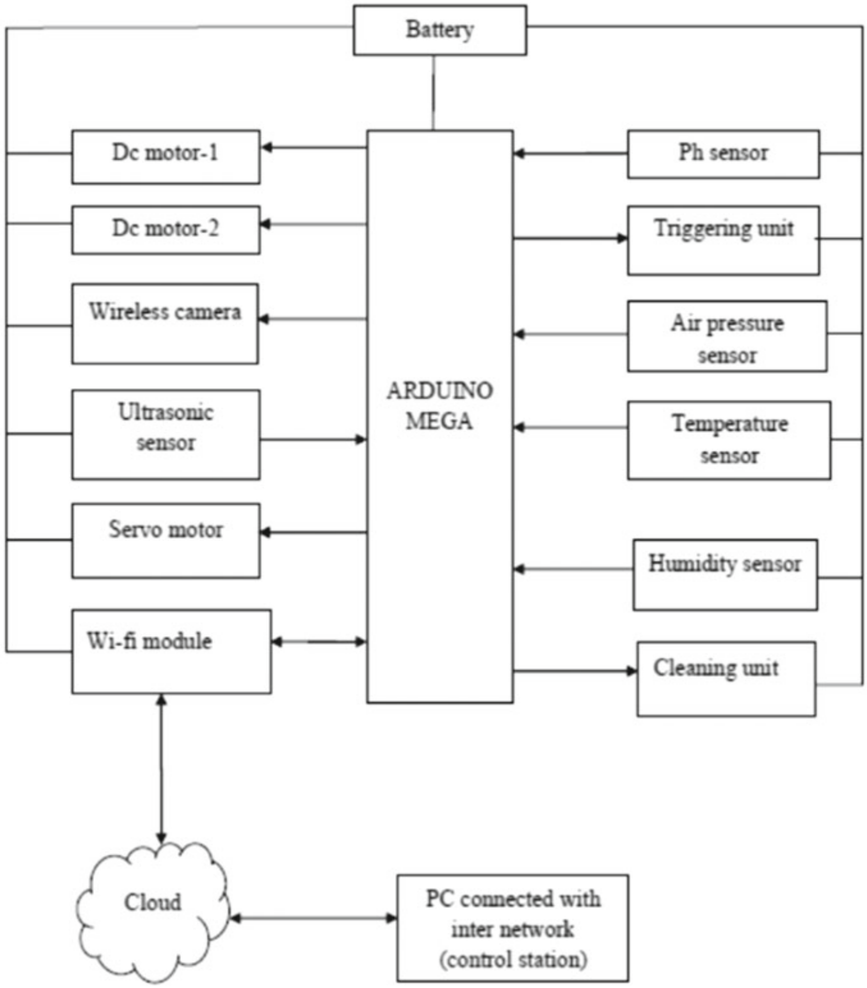


Fig. 1 Block diagram of unmanned marine robot

Ultrasonic sensor: RADAR system is set up through an ultrasonic sensor and servo motor, which will reveal the presence of enemies or unauthorized objects and send the information to the control unit.

Servo Motors: This motor is a rotary actuator or a motor that lets in for specific control in phrases of linear position, acceleration, and speed.

DC motors: A direct current or DC motor, converts electrical energy into mechanical energy. It is one of two basic types of motors: the other type is the alternating current or AC motor. Among DC motors, there are shunt-wound, series-wound, compound-wound and permanent magnet motors. Its used for robot movement and conveyer belt.

Wi-Fi Module: The WIFI module is a self-contained SOC (system On Chip) module advanced with the aid of Espressif with integrated TCP/IP protocol stack that can deliver a micro-controller access to the WIFI. it is used for development of IoT (internet of things).

Wireless Camera: When the ultrasonic sensor is triggered, the controller sends a signal to the Wireless digicam unit for capturing a live video coverage recordings & sends to the control unit. The later information can be uploaded to the cloud if required.

Cloud: It makes use of the internet to store the data such as values of temperature, humidity, pH and air pressure sensors.

Triggering unit: When the ultrasonic sensor detects any enemies, the wireless camera is turned on and if we come to know that it is a spy or an enemy object the triggering unit is turned on and shoots the object.

Cleaning unit: Nylon mesh is rolled as a conveyer belt to collect the plastic materials floating on the water and puts it in the collector.

Control Station: In this station the distance of the objects, the values of the pH, air pressure, temperature and humidity sensors are monitored.

Advantages

- We can monitor/destroy our enemies from remote places at marine border
- We will view live coverage videos from distant places
- We will monitor the environmental situations, so that it will useful for climate broadcasting
- We can use the robot to clean the river/lake/sea
- We can use the robot to test the water Ph.
- All measured records might be uploaded to the cloud, that is useful for further evaluation
- We are able to save lives.

Applications

- We can use this system for navy
- We can use this system for lake/river cleaning
- We can use the system for weather broadcasting
- We can use the floating robot in flood affected area for Rescue operation (Figs. 2 and 3).

4 Acknowledgements

This work was supported by our guide Mr. Sangam Kumar G.H and entire Dept of Electronics and Communication, DBIT.

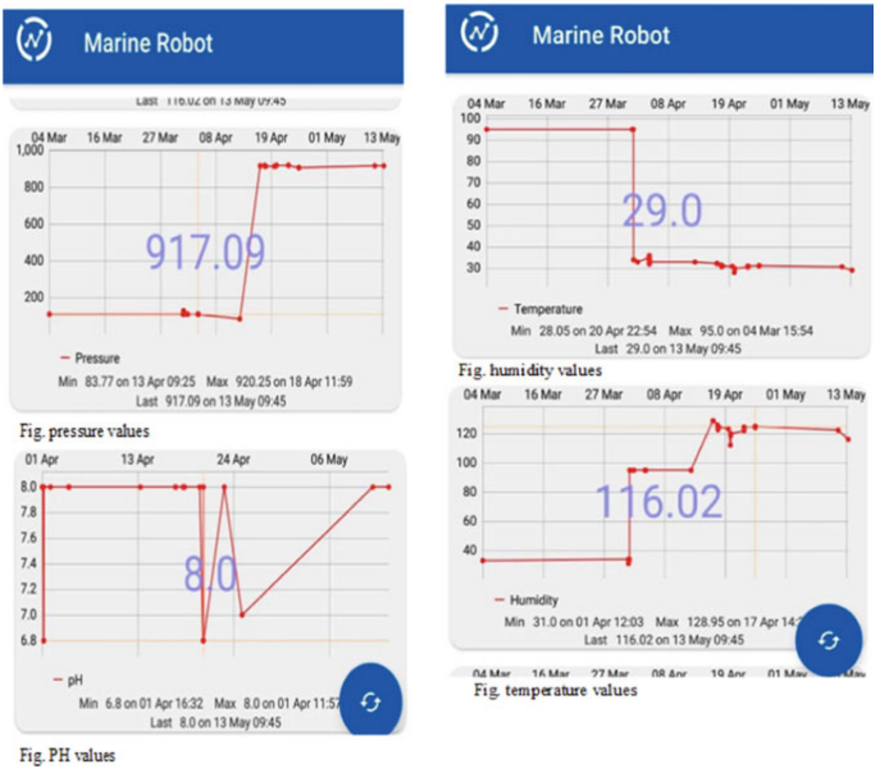


Fig. 2 Readings of pressure sensor, PH sensor, humidity sensor, temperature sensor

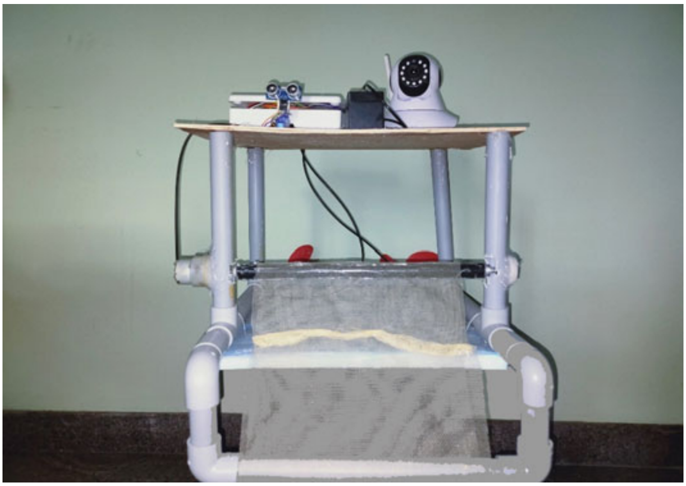


Fig. 3 Actual Model of unmanned marine robot

5 Conclusion

There are various security techniques to guard our coastal areas, but implementation of android based robotic for coastal place safety may be more helpful… This technique reduces the risk within the lives of our soldiers. It reduces human involvement in the coastal place.

This marine robot includes ultrasonic sensor and servo motor which both acts as a Radar machine. Radar device facilitates in figuring out the movement of unauthorized person or target and servo motors enables inside the motion of the ultrasonic sensor. Alert message "target detected" is dispatched to control station, and the surroundings are monitored the usage of wireless digicam. With the usage of live streams captured, similarly actions like blasting can be accomplished.

We can also use this system for river/lake water quality testing and cleaning floating waste from river. Which will helpful for maintain the hygiene in water. Which helps to build healthy society.

References

1. U.S. Army Considers Replacing Thousands of Soldiers with Robots||, U.S.S. Enterprise, IEEE Starship U.S.S Enterprise Section, 2017
2. E. Ackerman, U.S. Army Considers Replacing Thousands of Soldiers with Robots, IEEE spectrum,22 Jan 2017
3. J Multi Eng Sci Technol (JMEST) ISSN: 2458–9403 4(4) Apr 2017
4. [IJIR] Imperial J Interdisciplinary Res 3(4) 2017

Raspberry Pi Based Smart Refrigerator to Recognize Fruits and Vegetables

Santosh M. Nejakar, K. R. Nataraj, K. R. Rekha, S. Sheela, P. Pooja, and K. S. Nafeesa

Abstract Recently, IoT has become one of the most vital technologies of 21$^{\mathrm{St}}$ century. The internet of things (IoT) describes the network of physical devices that interfaced with the sensors, software, microcontroller and other technologies. It serves the purpose of connecting and exchanging of data with other devices and systems over the internet and it also provides various services in the variety of domains and applications. Times are changing and so do the lifestyle of people. Earlier were the day's people spent more time being at home, but today, is trend is on declined. The reason is the people are overwhelmed with work. So they will be left with very less time to visit the market in order to fulfill their daily essential requirements. It is very complicated and time consuming. So to overcome from this problem we implement an informative refrigerator. The refrigerator which we have developed transfigures any preexisting refrigerator into a smarter one using various sensors and microcontroller. It also presents an art method for vegetables and fruits recognition. When we are away from home and when we can't access the data of the refrigerator this smart refrigerator develops the diagnostics solution around a conception of smart fridge facilitated with an onboard camera and informs us through a developed android application on the mobile phone about the quality and quantity of various fruits and vegetables. By that we can know which vegetables and fruits are available or not. And by using camera we capture the image and sends to user mail id. Hence, we come up with a concept of smart refrigerator using Raspberry Pi, load cell, RFID, and various sensors to do their respective jobs.

Keywords Raspberry Pi · Refrigerator · Raspberry Pi · Load cell · Gas sensor · IR sensor · Temperature sensor · RFID scanner · Camera · Power supply · Android application

S. M. Nejakar (✉) · K. R. Nataraj · K. R. Rekha · S. Sheela · P. Pooja · K. S. Nafeesa
Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_100

1 Introduction

Refrigerator is most commonly and frequently used on kitchen. Nowadays the people do not have time to monitoring the validity and availability of fruits and vegetables inside the refrigerator. A smart refrigerator converted old refrigerator into the new one Technology used refrigerator to smarter. Smart appliances with multimedia capability have been used into our daily life. Now we are using IoT based smart fridge [1]. The smart refrigerator gives the, what is the status of the fruits and vegetables like weight, quantity, quality, captured image and so on. Smart refrigerator is capable of sensing and continuous monitoring of fruits and vegetables and also provides various useful features. Smart refrigerator is notifications send through the user about the product via Wi-Fi module on user mobile android applications [2]. Significantly we avoid the food waste, reducing disease and make the healthier lifestyle in today. Internet of things (IOT) is the most common and useful upcoming domain which is already in the world. We are using various sensors to recognization of fruits and vegetables. When we are away from home and when we can't access the fruits and vegetables inside the refrigerator this smart refrigerator develops a solution for avoiding the wastage of fruits and vegetables in smart fridge [3, 4]. Onboard web camera is present inside the refrigerator. This camera will capture the image of various fruits and vegetables inside the refrigerator and sends the captured image to the registered user email Id. In this project, the main component is raspberry pi to which various sensors, RFID em18 reader module, and camera are connected [5]. The main aim of this project is to provide a smooth communication between the user and the refrigerator. This project consists of two parts, in the first part al the hardware components are interfaced with the raspberry pi. In the second part, the required blynk application to notify the user about his/her refrigerator is developed [6].

2 Related Work

"Recognition of Edible Vegetables and Fruits For Smart Home Appliance".

In this project they introduced the various neural technologies and deep learning using this technologies they proposed divergent interpretation. They focused mainly on classification of consumable fruits, mushrooms and plants. Using all this technologies the refrigerator will be able to perform various activities like recommending recipes to the user based on the available contents, evaluate the preservation condition, proposing the items to purchase that are in low in stock [7, 8].

"RFID Based Smart Fridge".

In this paper RFID based smart fridge had been proposed. They used RFID technology to detect a particular grocery items. RFID technology includes an antenna, tag id, containing information about the item, RFID reader. The RFID antenna will be employed on top of the fridge. Each rack or bag will be given an RFID tag and it will be pre-decided what kind of item will be placed into that particular shelf/bag. Each time an item that placed in the rack/bag will be scanned using RFID reader and information is collected. The collected information will be stored in the database using a serial cable and it will be updated in the computer [9, 10].

3 Methodology

3.1 System Architecture

Figure 1 shows the outline of the device that will eventually be interfaced inside a refrigerator that shows each component of the system, how the system works and the flow of the system and so on. This is basically amalgamation of regular refrigerator and a microcontroller, a smart mobile to fabricate an inventory monitoring to track the stock of a fridge with an absolutely wireless approach using a smart mobile application. It deals with the issue of food wastage and efficient food distribution through a connected refrigerator. This network can be created through a microcontroller and sensors and a wireless mobile application. This process was discussed to determine the fruits and vegetables below the defined quantity or threshold limit, determining the product expiration date and eventually placing the order for purchase of fruits and vegetables online. It also includes the connection of an IR sensor to an IOT platform

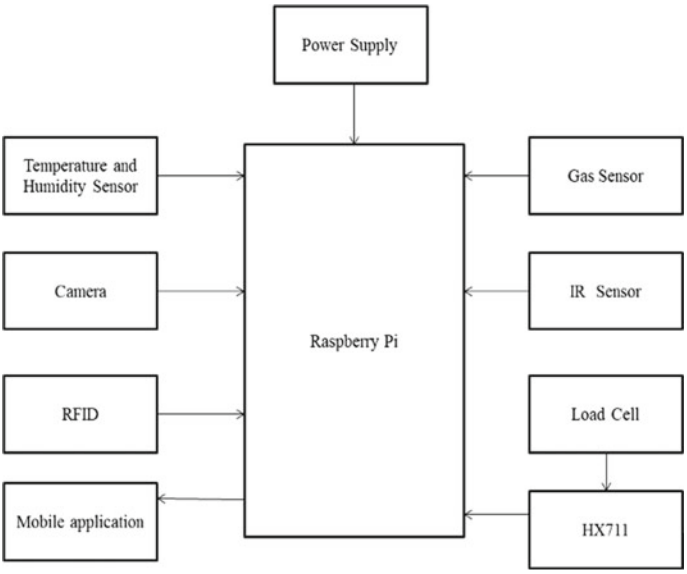


Fig. 1 Block diagram

that can allow the exchange of information on the positioning of fruits and vegetables inside the refrigerator. It includes RFID technology which recognizes expiry date of a milk can, yogurt or any beverages. It also has a Gas sensor which recognizes any gases or smoke which leads to explosion of refrigerator and also detects any bad odors resulted from spoilage of vegetables and fruits inside the fridge. We have also utilized a weight sensor which detects the weight of the fruits and vegetables kept in the refrigerator.

3.2 Refrigerator

Refrigerator is indispensable storing approach that is being used universally. The rudimentary idea behind the refrigeration is to prevent the rapid growth of bacteria; in turn it reduces the festering of fruits and vegetables. Ideal temperature range to keep the food fresh is 3–5 °C [11].

3.3 Raspberry Pi

The Raspberry pi can open opportunities for users to create their own home automation projects. Raspberry Pi is a minicomputer that connects to a computer and utilizes input devices like keyboard and mouse. Raspberry Pi has inbuilt SD card slot, and has memory of 1GB RAM. Python is one of the most generally used software in Raspberry pi.

Technical specification: Operating voltage: 3.3 V, Clock Frequency: 1.2GHz, Raw voltage input: 5V, 2A power source [12].

3.4 Load Cell

A load cell is a type of transducer which is utilized to generate an electrical impulse. The generated electrical signals magnitude is proportional to the applied force that is being measured. Load cells are available in different sizes, shapes and ranges in market. For this we used 10Kg load cell.

Technical specification: Full Scale Output: 3.0 mV, Safe overload: 200% RC Material: alloy steel, nickel plated

3.5 Hx711

HX711 is a load cell driver which is 24-bit high precision and accurate analog to digital converter. HX711 amplifies the low electrical output of the load cell obtained from the applied mechanical energy transforms into digital output.

Technical specification: Operating supply voltage: 2.6–5.5V, Operating Temperature range: −40–85℃

3.6 Gas Sensor

There are various types of gas sensors but MQ series gas sensors are mostly used and widely popular gas sensors. A gas sensor is a device that recognizes the presence or concentration of various gases in the atmosphere. It comprises of a sensing element which is exposed to current through connecting terminals. This current is sensed as heating current.

Technical specification: Operating voltage: 5V, Sensing Resistance: 10K–60K ohm

3.7 Temperature and Humidity Sensor

DHT11 is a basic and cost effective digital temperature and humidity sensor. It is mandatory that DHT11 is implemented away from other devices that generates heat. No analog input pins are needed.

Technical specification: Operating voltage: 3.3 or 5V DC, Measurement Range: 20–95% RH; 0-50 degree Celsius, Resolution: 8 bit (temperature and humidity) [13].

3.8 Raspbain OS

Raspbain OS is a Debain- based operating system for Raspberry Pi. It is free and open source software.

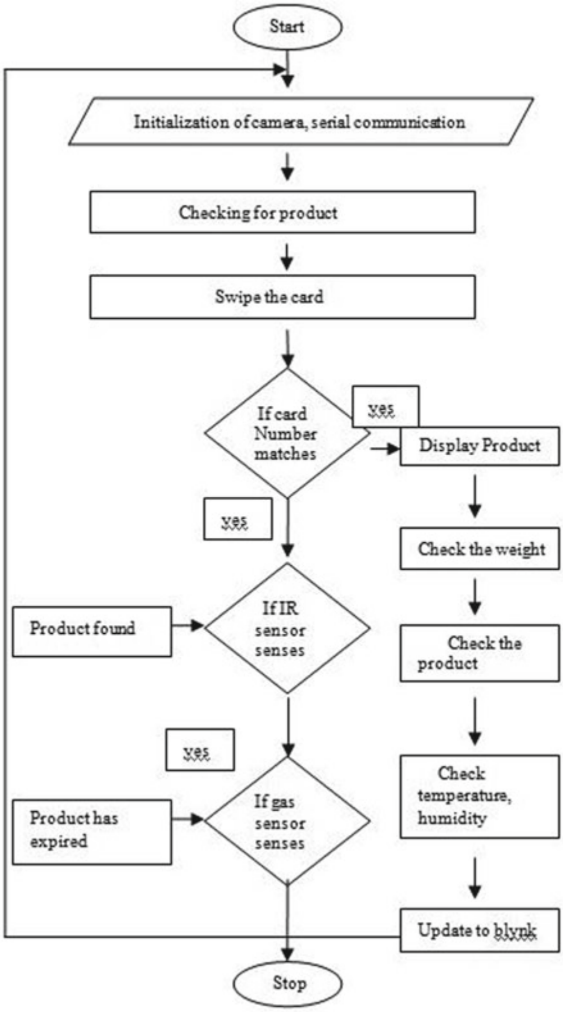
3.9 Blynk Application

Blynk serves as a platform for the users to develop a hardware project which interfaces hardware and software components and facilitates the user to control and monitoring purposes.

4 Flow Chart

Figure 2 shows the flowchart of the refrigerator and it explain the step by step procedure for working of refrigerator. After turn on the module it initializes the camera and makes serial communication. Once serial communication between all the devices completed, it checks for the products (vegetables & fruits). After receiving message from the authenticated or registered mobile number (swipe the card stage) camera will get turn on and it takes the images inside the refrigerator and it will update in blink application, which is available in user mobile phone. When the user opens the app it gives the images captured by camera which is placed inside the refrigerator.

Fig. 2 Flowchart



By using gas sensors user can able to identify the quality of the vegetables and fruits. After some days vegetables starts to decay at that stage those will produces some gases, gas sensors used in our project will identifies and sends messages to the registered mobile numbers. And also by using weight sensors we can able to identify the quantity of the vegetables & fruits inside the refrigerators by sitting anywhere.

5 Results

In this we have presented the overall connection that is all the components such as sensors, voltage regulator, Raspberry Pi are connected together. This makes a demo device that cannot be implemented directly in to the fridge.

First step in this user must scan the RFID card which contains the information about the vegetables and fruits that he wants to keep and store inside the refrigerator. The RFID card will be scanned and the information of which type of fruit or vegetable it is will be stored in the database (Figs. 3, 4, 5 and 6).

Secondly, the IR sensor detects the fruits or vegetables that are kept in front of it. And informs the user about the presence of the particular fruits and vegetables to the user through the developed Android application.

Thirdly, the weight and expiry date will be sensed using the weight sensor and gas sensor. To check the weight user should measure the vegetable bowl or bag on top of the load cell. To check and determine the expiry date of that particular vegetable or fruit a gas sensor has been installed (Figs. 7 and 8).

The figure shows picture of the Blynk app in the User's mobile showing the fruits and vegetables inside the fridge. It shows the weight and the name of fruit and vegetables that user has kept inside the refrigerator. And also information about temperature and humidity of the refrigerator are also known.

Fig. 3 Experimental set up

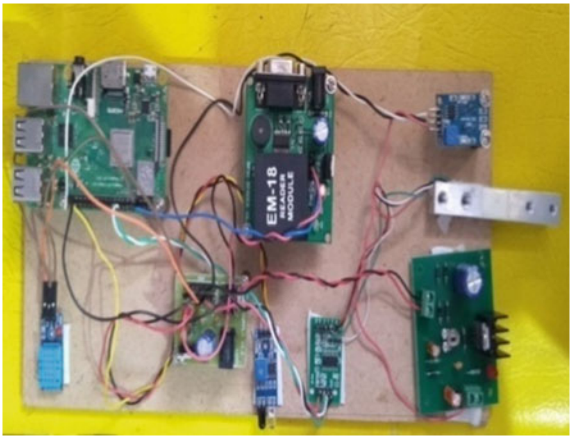


Fig. 4 User scanning the RFID card

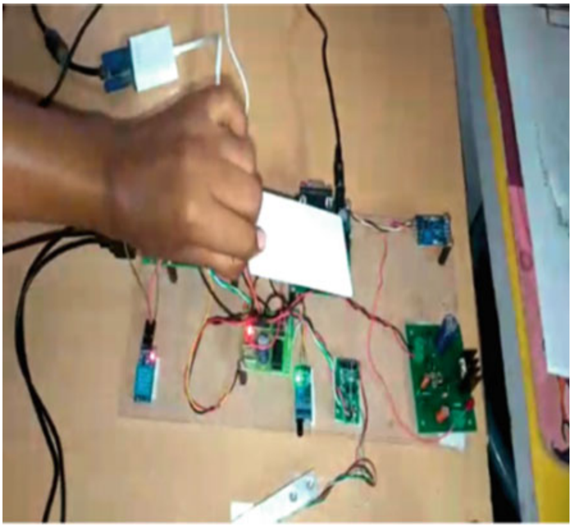


Fig. 5 IR sensor detecting the vegetable

To the check the image of the fridge the email-ID of the user will be given to the database and to this-email.

6 Conclusion and Future Work

The project raspberry pi based smart refrigerator provides a purpose of monitoring the fruits and vegetables in the fridge. What is special, apart from, through these intuitive refrigerator users will be able to spend less money with minimum effort. The refrigerator that we have developed in this project is going to be a necessary tool

Fig. 6 Measuring bowl of vegetables using load cell and capturing the image

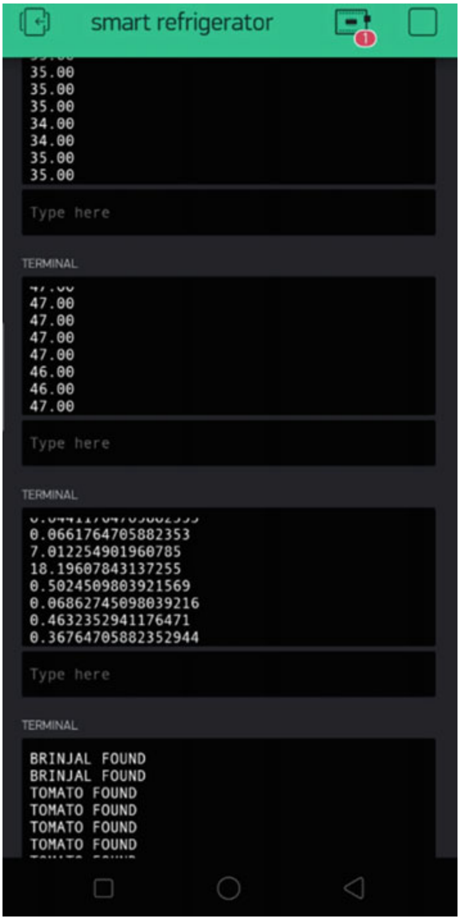


Fig. 7 A screenshot of the App showing the details of the fruits and vegetables

Fig. 8 Showing the picture that is sent to the user's email

in upcoming modern homes. This project has provided us an opportunity to operate on Raspberry Pi3, Raspbian and Python. Furthermore, this project has enabled us to enlighten ourselves about sensors and their working. This smart refrigerator is innovative and user friendly. We definitely hope this innovative refrigerator would become budget friendly too.

This project can be enhanced further in future, which is mentioned below:

- Refrigerator can be developed such that it can be able to provide the nutritional values of the fruits and vegetables or the available food products in it. It can be very helpful for the user due to the fact that he/she can get an accurate idea on what he is eating.
- User can be provided with all the information required for a recipes with the available contents in his refrigerator.
- We are in the generation of Internet of Things, so it is difficult to believe a gadget cannot communicate with the another gadget. For instance the smart refrigerator might be able to interact with an electric cooker to recommend to some rice.

References

1. Buzzelli M, Belotti F, Schettini R (2018) "Recognition of edible vegetables and fruits for smart home appliances" IEEE
2. Nejakar SM, Pruthvi S (2013) Titled "Wireless infrared remote controller for multiple home appliances" In Int J Electric Electron Res (IJEER), 2, 25–35, Month: January–March 2014, Available at www.researchpublish.com

3. Meghana KC, Nataraj KR (2016) IOT based intelligent bin for smart cities. Int J Recent Innovation Trends Comput Commun 4(5):225–229
4. Edward M, Karyono K, Meidia H (2017) “Smart fridge design using nodemcu and home server based on raspberry Pi 3”, IEEE
5. Miniaoui S, Atalla S, Hashim KFB (2019) “Introducing innovative item management process towards providing smart fridges”, IEEE
6. Hachani A, Barouni I, Said ZB, Espritech LA (2016) “RFID based smart fridge”, IEEE
7. Rezwan S, Ahmed W, Mahim MA, Islam MR (2018) “IoT based smart inventory management system kitchen using weight sensors, ldr, led, arduino mega, nodemcu, wi-fi module with website and app”, IEEE
8. Durani H, Vaghasia M, Sheth M, Kotech S (2018) “Smart automated home application using iot with blynk app”, IEEE
9. Ragesh N, Giridhar B, Lingeshwaran D, Siddharth P, Peeyush KP (2019) “Deep learning based automated billing cart”, IEEE
10. Kumar S, Balyan A, Chawla M (2017) “Object detection and recognition in images”. Int J Eng Dev Res
11. Xie L, Sheng B, Yin Y, Lu S, Lu X (2013) “An intelligent fridge for food management depending on rfid generation”, Zurich, Switzerland
12. Meghana KC, Nataraj KR (2016) “IOT based intelligent bin for smart cities”. Int J Recent Innovation Trends Comput Commun 4(5): 225–229, Hsin-Han wu Yung-Ting Chuang “Low-cost Smart Refrigerator” 2017, International Conference on Edge Computing
13. Sheela S, Shivaram KR et al (2016) Int J Innovative Res Sci Eng Technol. Proceedings of international conference on “Innovative technology for smart roads by using iot devices”, with ISSN(online): 2319–8753, ISSN (print): 2347–6710, in IJIRSET, India, Special Issue 10th May, 2016, 5:190–194

Smart Mirror Based on Raspberry Pi

G. Shruthi, N. Gowtham, Mohammed Junedh, M. Sahana, and S. K. Sahana

Abstract Intelligent systems are not only used on mobiles, tabs, computers; but can also be used on mirrors. An intelligent mirror based on Raspberry pi is set for home. The system of raspberry pi is linked to net through Wi-Fi and obtain the data about weather forecast, daily news etc. The Raspberry pi is programmed with the help python & connects to a monitor with speaker so as to provide onscreen interface. This adds an additional features of face recognition for security & smart unravel process. This also includes the emotional recognition of a face, which identifies the different emotions on a face. The details are displayed on a LED monitor camouflaged with a mirror. The mirror provides basic conveniences like weather of the city, time, news details and also we can recognize the user's face and verify the person.

Keywords Raspberry pi · Smart mirror · Python

1 Introduction

Smart Mirror will be wall seated mirror that shows weather, time, news and many more things. Things began to become agile, mobile phones changed to sensible phones and more over net is associated with a range of devices and idea can be called as 'Internet of Things' [1]. Our project design is to explore the technology in an exceedingly mirror. Another favor of this device is to add a face recognition, which we have done using Open CV. This helps the enjoyer with surveillance. Our mirror would facilitate in establishing sensible homes by artificial intelligence [2] and eventually recommendation an area in industries. It can do all the work which we command it to, as it has a personal assistant which will listen to your commands and will response according to it. The smart mirror has microphone through which it will listen to the commands, and there is speaker from which it will give its response. Smart mirror have the some capabilities like showing date and time which was done before when it was developed, but objective of this project is as follows: It must

---

G. Shruthi (✉) · N. Gowtham · M. Junedh · M. Sahana · S. K. Sahana
Department of Electronics and Communication, Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_101

be capable of showing all the real time data, the Smart Mirror must be capable of showing the information on the screen correctly, the Smart Mirror must take the voice command as the input and give the required response to the user.

2 System Overview

The Smart mirror consists of associate electronic display along with two-way mirror for the information shown in Fig. 1. The organic structure consists of 2-way mirror placed on the show monitor, raspberry pi, speaker, webcam. It permits the rays of the light from one side of the mirror while leaving the rays from the display to pass through. On the outside, the hardware is encapsulated inside a wood frame. On the front, a 2-way mirror is placed in before an associate LCD touch screen monitor. This way, the system will act as a mirror once in use. The wood frame contains a bezel on the front which the mirror and LCD panel area unit ironed against

The software of this mirror is established using the programming communication called Python. It contains some segment which knob some activity of the arrangement, like data of the weather, name of the city, date and time, news and interactive media presentation.

This system principally includes 3 elements a two-way reflector, Liquid Colour Display monitor and Raspberry pi. The two-way reflector is contemplative on one aspect and clear on the opposite aspect. The Liquid Colour Display monitor are going to be linked to the Raspberry pi. The Raspberry pi would be used for computing of discrete widgets using Python. The reflector will be turn on employing a voice

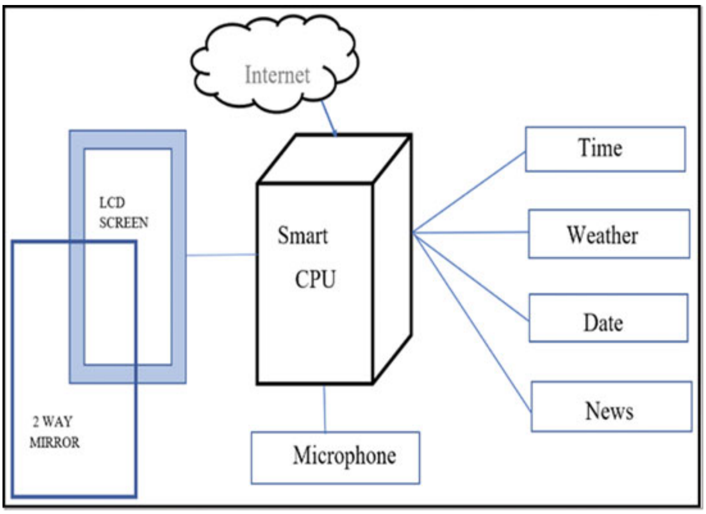


Fig. 1 The overview of smart mirror

compel like “Hello Mirror!”, “Good morning mirror” or the alternative keyword. The Sensible mirror will also give speech and document response like acknowledgement the enjoyer or provide few respects as response, for which the system will use system suitable mike and announcer.

3 System Design and Implementation

The model uses a face recognition authentication. Smart mirrors are modern extension to the smart household peoples that has been getting more consideration currently by both economical manufactures and academic. The use of creating a sensible reflectors is to save the time of humans and also to save time and enabling them to modernize with current news.

A. **Raspberry Pi**

The Raspberry pi [3] is having other category but in the plotting this medical reflector that a writers uses the Raspberry pi 3 version. It contains the both WIFI module and Bluetooth. Raspberry Pi is a smaller lower power simple board minicomputer it is helpful in functioning an operating systems like Linux. We are adopting model B raspberry pi. This board is having 512MB SDRAM and it requires 700mA. It have 40 General Purpose Input Output pins. To achieve this, the Raspberry Pi is linked to the Internet through Wi-Fi module in the Raspberry Pi. Raspberry pi 3 consists of Broadcom BCM2387 ARM Cortex-A53 Quad Core Processor and provide GPU Video core IV. This also gets an input from the linked mike from the external USB sound card, which grants to a user to do more functions like web search, setting up of reminders and even engaging in two way conversations with the voice assistant, with a help of the Google’s voice assistant API.

B. **Face Recognition**

Face recognition or identification technique acquire the image matching technique to smudge people. The flow diagram for face recognition is shown in Fig. 2. It will determine folks from the video sources or digital pictures. Whereas many face identification algorithms, that often worked in examination elected to extract the facial attributes from the obtained pictures with the faces within the dataset. Smart Mirror work will administer face identification by use of Open CV library. This will bear facial spotting, facial coaching, and facial forecast. The identification of faces in the picture frame is the procedure of face spotting.

The Facial training is the process of identifying the some images of a same individual person to identify a characteristics determined to an independent. With a help of an Open CV, the two modules, namely user enrolment and users certification, they have evolved in the substantiate substructure.

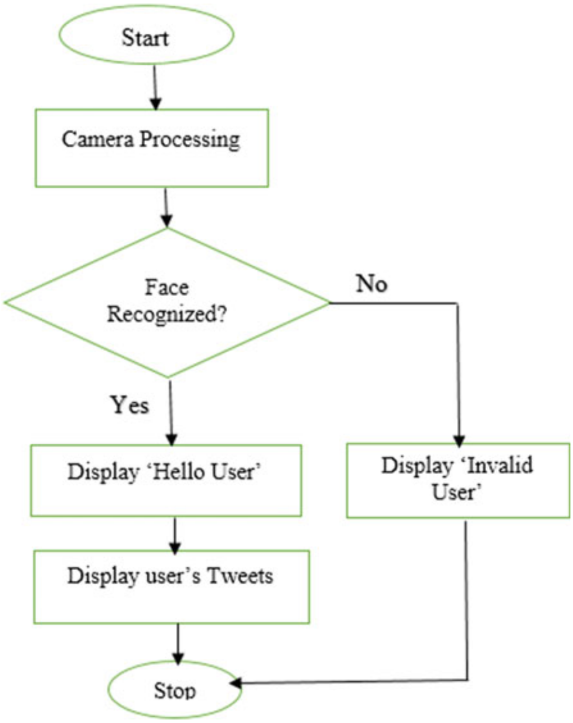


Fig. 2 Flow diagram of face recognition

- Users Enrolment—facial observation and facial coaching

This is will be executed by proceeding the pictures with the help of Raspberry camera. Then the pictures are sent to an Open CV cascade classifier to recognize human faces. A cascading classifier has the imitation that will be pre-coaching with the others some many face images, called positive examples, and negative examples arbitrary images of the same size.

- User Authentication—spotting and prediction [4].

Whenever the user wishes to log in to the structure they have to be verified. Then the verification policy will include both the facial observation and facial identification [5]. This will compare the picture with the representation set generated previously and gives back a special identity of a user's (from an imitation set) though face attributes are adjacent to a present picture.

C. Voice Based Command

Users can control all the functions using voice commands, including authentication. This system helps in the interaction between user and system. Google Speech to Text Conversion helps the users to switch from audio to text by implementing the neural network models. The API (Application Programming Interface) acknowledges 120 languages and alternatives to assist a wide-range users prop [6]. This would be helped to enable voice commanding and controlling, and more by processing real-time surge or pre-recorded audio.

D. Emotional Recognition

The Facial expression will include the muscle tightening features and the facial parts and the main focus is on the emotional detection. It supports as a base for stress detection. People's mood can be recognized through their facial expressions. The Facial emotional recognition is globally used in artificial intelligence and the automated systems that uses the vision based enumerate [7] [8]. Proper exposure of exposition in the identified faces can even boost the effectiveness of Human-Computer Interaction (HCI) system to more extent and it supports machines to conceive humans in a better way.

4 Result

We have concentrated on an interactive and user-friendly home appliance that is beneficial to the user. A user's personalized data such as calendar, news feeds, and other information relevant to their lifestyle have been implemented as shown in Fig. 3 and how the results are displayed on desktop is shown in Fig. 4.

5 Conclusion

The Smart Mirror is having the scope in the field of Internet of Things and Home Automation. The Smart Mirror will be a bridge to the home gadgets, mobile devices, etc., which may expand the practicality of the mirror. Adding security means unknown won't be able to attempt to access sensitive information that perhaps displayed on the mirror via a use of APIs. Alexa can also be implemented through this mirror by using Amazon web services. This intelligent mirror can also be used like a smart-phones by using various applications like you-tube, Google maps, accessing emails, Instagram etc.

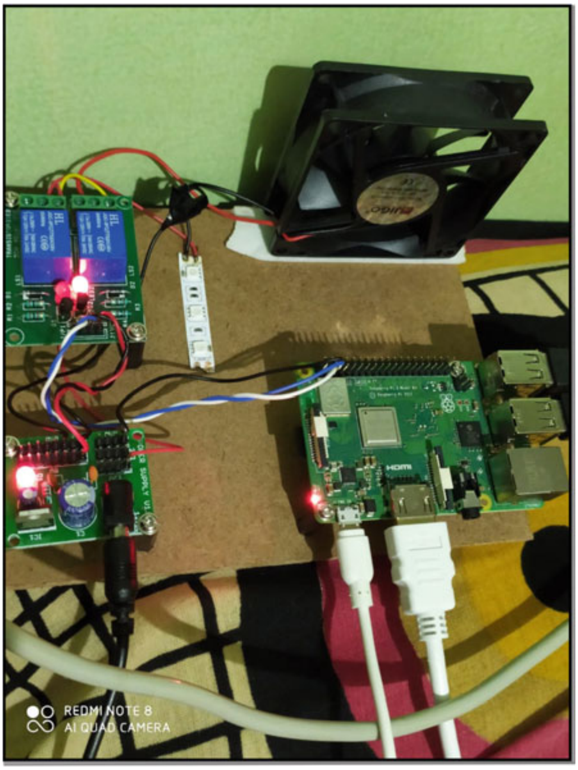


Fig. 3 Hardware circuit

Fig. 4 Mirror output on desktop

References

1. "Smart interactive mirror display", 2019 International conference on machine learning, big data, cloud and parallel computing, India
2. Mr. Mathivanan Sakthivel A, Anbarasan G (2019) "Home automation using smart mirror"
3. Yong S, Liqing G, Ke D (2018) "Design of smart mirror based on raspberry pi". International conference on intelligent transportation, big data & smart city
4. Wahi RN, Jafri S, Johri A, Dr. Pandey D (2018) "Smart mirror: a time-saving and affordable assistant", 4th International conference on computing communication and automation (ICCCA)
5. Francis F, Athira S, Raphel R, Sachin N, Porinchu S, Ms. Francis S, "Smart mirror: a novel framework for interactive display"
6. Yong S, Geng L, Dan K (2018) "Design of a smart mirror based on raspberry pi". International conference on intelligent transportation, big data & smart city
7. Gomez-Carmona O, Casado-Mansilla D (2017) "SmiWork: an interactive smart mirror platform for workplace health promotion". 2nd international multidisciplinary conference on computer and energy science
8. Kaur A, Jasuja A (2017) "Health monitoring based on IoT using raspberry pi". International conference on computing, communication and automation (ICCCA)

Application of Swarm Robotics Systems to Marine Environmental Monitoring

H. S. Suresha, D. N. Sumithra, J. N. Renuka, B. N. Deepika, and N. B. Meghana

Abstract Mechanized natural observing in marine ecological is at present done either by little scopes mechanical autonomy frameworks, made out of one or hardly any robots, or static sensor systems. In this apply autonomy frameworks to complete marine ecological observing missions. In swarm apply autonomy frame works; every individual unit is moderately basic and reasonable. The robots depend on decentralized control and nearby correspondence, permitting the multitude to scale to several units and to cover enormous territories. We study the utilization of a multitude of sea-going robots to ecological observing undertakings. In the initial segment of the investigation, we incorporate multitude Control for a temperature observing strategic approves our outcomes with areal multitude mechanical autonomy frameworks. At that point, we direct a reproduction based assessment of the robots execution over enormous zones and with huge multitude measures, and exhibit the multitude a strength to shortcomings. Our outcomes show that swarm mechanical autonomy frameworks are appropriate for ecological checking undertakings by effectively covering an objective zone, taking into consideration repetition in the information assortment Process, and enduring individual robot is use.

1 Introduction

Since a solitary robot usage can't satisfy the human fulfillment because of the expanding of undertaking necessities, numerous researchers created multi robots to achieve an entangled assignment, either in the homogenous [1, 2] or heterogeneous [3–6] structures. For the homogenous structures, specialists typically utilized a gathering of self-ruling robots that takes care of an issue utilizing a disseminated approach. These robots are typically called as a multitude robot. Multitude has been applied in an assortment of undertakings, for example, conglomeration, running, searching, object grouping and arranging, route, way development, sending, community oriented control and assignment distribution issues [7]. In other word, a multitude

H. S. Suresha (✉) · D. N. Sumithra · J. N. Renuka · B. N. Deepika · N. B. Meghana
Department of Electronics and Communication Engineering, Don Bosco Institute of Technology, Bangalore-74, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_102

utilizes some straightforward operators to perform various sorts of errands to arrive at the objective.

In its application, the individual operators o1f multitude are not educated about the worldwide status of the settlement [8]. There is no pioneer that directs all of individual specialists of multitude to arrive at their objectives. The information on the multitude is appropriated all through all the specialists. The objective might have the option to be reached by utilizing agreeable conduct among singular operators. In this manner, the people won't have the option to achieve their undertakings without the assistance or collaboration of the people with different operators.

In achieving morose confinement errands, swarm robots ought to have a unique expertise for moving in their condition, interfacing with different robots, seeing and preparing the data of their surroundings [9]. The multitude shares data about the earth and the individual specialists collaborate with one another, in this manner a qualification between the detecting and the correspondence organize is made [10]. Because of no focal coordination inside their assignment execution, their agreeable conduct turns into a basic part in swarm research. The viability of the coordination among the individual operators of multitude can be accomplished utilizing data sharing.

This data sharing can lessen the length of time looking. To control robots through correspondence isn't a simple way. There are a few components ought to be focused [11], for example, correspondence design, the measure of calculation, and correspondence data transfer capacity [12]. Consequently, a few analysts create circulated mechanical sensor arrange so as to accomplish a vigorous correspondence with low calculation and low correspondence trouble [13, 14]. This paper applies a dispersed automated framework so as to accomplish a decent correspondence execution in looking, following, and discovering scent source.

2 Literature Survey

In this survey we utilize two scientific classifications: strategies and aggregate practices, In Sect. 2, we investigate strategies to structure and examine swarm mechanical autonomy frameworks. In Sect. 3, we dissect a portion of the conceivable aggregate practices a multitude mechanical autonomy framework can display. By aggregate practices we mean practices of the multitude considered all in all. Such aggregate practices can be utilized as building hinders for applications, for example, rummaging or development. In Sect. 4, we finish up the paper with a conversation of the open issues in swarm mechanical autonomy and multitude designing.

Past surveys proposed scientific classifications that contrast from those that we propose here. Dudek et al. (1993) picked swarm size, correspondence go, correspondence geography, correspondence transmission capacity, swarm reconfigurability and multitude unit preparing capacity to order the writing [15]. Cao et al. (1997)

utilized: bunch design, asset clashes, causes of collaboration, learning, and mathematical issues [16]. Iocchi et al. (2001) received a various leveled scientific classification: in the principal level they considered mindful versus ignorant collaboration. The mindful class is partitioned into unequivocally planned, pitifully organized and not-composed frameworks. Works identified with emphatically organized frameworks are partitioned into unequivocally brought together, pitifully incorporated and dispersed. A different area is devoted to uses of multi-robot frameworks. Gazi and Fidan (2007) decided to isolate the Past overviews proposed logical arrangements that differentiate from those that we propose here [17]. Dudek et al. (1993) picked swarm size, correspondence go, correspondence geology, correspondence bandwidth, swarm reconfigurability and large number unit planning ability to arrange the composition. Cao et al. (1997) used: pack structure, resource conflicts, reasons for joint effort, learning, and scientific issues. Iocchi et al. (2001) got a different leveled logical grouping: in the chief level they considered careful versus uninformed cooperation [18]. The careful class is apportioned into unequivocally arranged, sadly sorted out and not-made structures.

3 Proposed Methodology

The objective of this area is to order the articles distributed in the multitude apply autonomy writing as indicated by the techniques used to plan or to examine swarm mechanical autonomy frameworks. In Sect. 2.1, we present the most widely recognized plan strategies used to create aggregate practices for multitudes of robots. In Sect. 2.2, we present the most widely recognized strategies used to comprehend, foresee, and examine the aggregate conduct of a multitude. 2.1 Design techniques Design is the stage where a framework is arranged and created beginning from the underlying details and necessities. Sadly, in swarm mechanical autonomy there are still no formal or exact approaches to structure singular level practices that produce the ideal aggregate conduct. The instinct of the human architect is as yet the primary fixing in the advancement of multitude mechanical technology frameworks. We isolate the plan techniques into two classifications: conduct based structure and programmed plan. Conduct based plan is the most well-known approach to build up a multitude mechanical autonomy framework. In an iterative manner, the individual conduct of every robot is executed, considered, and improved until the ideal aggregate conduct is gotten. In conduct based structure, motivation is frequently taken from the perception of the practices of social creatures. This may facilitate the plan procedure as, now and then, the subtleties of a specific conduct are as of now comprehended and numerical models are accessible. Another approach to create swarm apply autonomy frameworks is by means of programmed plan techniques. Programmed structure strategies can be utilized to diminish the exertion of the engineers in making an aggregate conduct. We order programmed plan strategies in two classifications: transformative mechanical technology and multi-robot

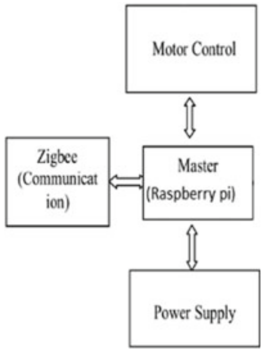


Fig. 1 Master block diagram

support learning. In the accompanying, we talk about conduct based structure strategies and programmed plan techniques by depicting the overall standards and the relative favorable circumstances and weaknesses. It is denoted in Fig. 1.

4 Description

We are actualizing an ace robot which will check the matrix zone and send orders to the particular slave robot.

The above square graph speaks to the ace robot. The core of the ace is Raspberry pi. ZigBee convention is utilized for correspondence between the pontoons. It is denoted in Fig. 2 robots. There are numerous instances of all things considered complex conduct for instance social bugs, for example, insect provinces, termites, honey bees, wasps … and so on. So, a multitude insight framework comprises commonly of a populace of moderately basic specialists collaborating just locally with themselves and with their condition, without having a worldwide information about their own state and of the condition of the world [4–9]. The multitude robots have following properties.

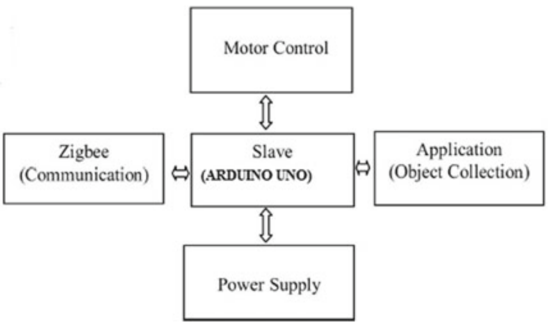


Fig. 2 Slave block diagram

- Autonomy—people that make the multitude mechanical framework are self-sufficient robots. They are autonomous and can cooperate with one another and nature.
- Large number—They are in huge number so they can help out one another.
- Scalability and strength—another unit can be effectively added to the framework so the framework is effectively versatile. More number of units improve the presentation of the framework. The framework is very vigorous to the loosing a few units as there still exists a few units left to perform. In spite of the fact that the frame work won't perform up to its greatest abilities.
- Decentralized coordination—The robots speak with one another and with condition to take a ultimate choice.
- Flexibility-It requires the multitude mechanical framework to be able to produce modularized answers for various assignments.

5 Swarm Robotics System

In a multitude mechanical autonomy framework, the robots depend on decentralized control. Each automated unit is self-ruling and settles on choices dependent on tangible readings and data got from different robots in its prompt region. Along these lines, singular robots can progressively react to occasions in the earth and help out neighbors based on nearby prompts. The robots can consolidate the tactile readings into the dynamic process, in request to follow ecological slopes, track sea, life, etc. It is as shown in Fig. 3.

Decentralization of control prompts various key properties [3, 11–13] that make swarm mechanical autonomy frameworks especially appropriate for marine conditions, to be specific:

Robustness to individual faults:

Fig. 3 Swarm robotics system

Given the decentralized idea of the robot control, there is no main issue of disappointment in a multitude mechanical autonomy framework. In this way, the swarm is strong against the disappointment of individual robots and, since there is repetition inside the multitude, issues don't bargain the finishing of the mission. Such strength is particularly important for long haul missions in marine situations, as waves, wind, and trash can cause sudden disappointments in the robots.

Scalability:

Multitudes can scale progressively to tens or many robots [14], as the robots just connect with different robots in their quick region. Such adaptability is basic for observing enormous waterways, as it empowers information examining at a few places at the same time.

Flexibility:

Robots in a multitude apply autonomy framework can show various practices in light of changes in nature and tangible contributions, rather than depending on predetermined crucial. The utilization of multitude mechanical technology frameworks to marine situations is worthwhile in assignments where enormous and dynamic conditions must be observed. Having various unique estimating focuses empowers a high spatial and fleeting goal of the accumulated information, which is especially important in conditions where the highlights being estimated change all through reality [19].

6 Results and Discussion

The Fig. 4 gives the Results & output of swarm robotics system. Here Master will control the slave. Using Zigbee Technology we can control the slave and we can find application or object collection.

Communication in Swarm Robots

1. For working in swarm, the ace and slave needs to speak with one another.
2. For this we use transmitter and collector.

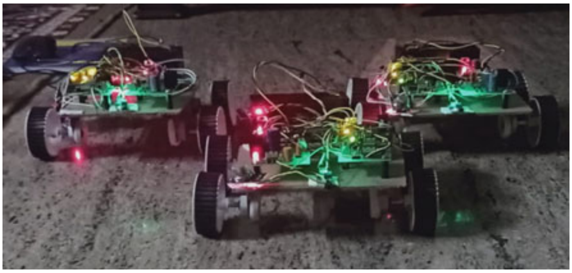


Fig. 4 Output of swarm robotics system

3. The kind of transmitter and collector rely up upon the scope of correspondence.
4. In the correspondence procedure, microcontroller assumes a key job.

7 Applications and Advantages

Applications: Possible applications for swarm apply autonomy are for sure tremendous. It incorporates assignments that interest for scaling down (Nano mechanical autonomy, microbotics), like disseminated detecting undertakings in smaller scale apparatus or the human body. One of the most encouraging employments of multitude apply autonomy is in catastrophe salvage missions. Multitudes of robots of various sizes could be sent to places salvage laborers can't reach securely to identify the nearness of life through infra-red sensors. Then again, swarm apply autonomy can be fit to assignments that request modest plans, for example mining errands or rural scavenging undertakings. Additionally, a few specialists use swarm mechanical procedures to acknowledge new types of intelligent workmanship.

- A few possible proposition of multitude apply autonomy which are entirely appropriate are portrayed beneath.
- Undertakings which spread huge territory.
- Tasks risky to robot.
- Tasks which require scaling populace.
- Tasks which require repetition.
- Post-debacle help tasks
- Target looking
- Military application.

8 Disadvantages

Interference: robots in a gathering can meddle between them, due to collisions, impediments, etc.

(ii) Uncertainty concerning other robots" expectations: coordination requires recognizing what different robots are doing. On the off chance that this isn't clear robots can contend rather than coordinate.

9 Future Scope and Conclusion

Multitudes of robots performing together to do exchanges could give new opportunities to people to interface the intensity of machines. The capacity to control robot multitudes could demonstrate enormously valuable in a scope of settings, from military to clinical. The robots can likewise amass themselves into a solitary bunch in the wake of being scattered over a room, and shape themselves by request of need. For a bigger scope, they could have an impact in military, or search and salvage activities, acting together in regions where it would be excessively hazardous or unreasonable for people to go. In industry as well, robot multitudes could be put to utilize, improving assembling procedures and work environment wellbeing. The bionic air vehicles roused from swarm knowledge innovation will get relevant in a couple of years. It very well may be anticipated that machine honey bees or cockroaches with examination hardware and rockets will conceivably appear in future war.

10 Conclusion

Utilizing robots in genuine applications, for example, space investigation and search-and-salvage is very fascinating for specialists these tasks require elevated level of power and variation. One approach to expand the strength is developing a conveyed framework, for this situation utilizing multi robot with disseminated control. The advantage of utilizing appropriated framework is having less reliance to a disappointment of a specific piece of the entire framework. Another significant component of SAR robots is flexibility. During a strategic would face to unpredicted circumstances and it ought to have the option to deal with that circumstance. One way to deal with have adaptability and variation is self-reconfigurable robots. This component empowers the robot to change its structure rely upon nature. Multitude mechanical technology is another methodology for SAR. In this methodology every robot is a self-sufficient specialist with some ability and sensors. Singular robots are able to self-governingly and afterward team up with different robots to accomplish more perplexing undertakings, for example, moving over a gap, moving item, and so forth. Multitude Bot opens up another exploration field. The framework worked here is equipped for playing out the unpredictable assignment of strategic administration in a simpler manner. The dissemination of the errand brought about partitioning the work lastly accomplishing the outcome.

References

1. Liu E, Wong E (Jun 1996) Emergency ambulance services. Research and library services division legislative council secretariat
2. Kerber RE, Becker LB, Bourland JD, Cummins RO, Hallstrom AP, Michos MB, Nichol G, Ornato JP, Thies WH, White RD, Zuckerman BD (1997) Automatic external defibrillators for public access defibrillation. Circulation (American Heart Association) 95:1677–1682
3. Arif M, Samani H, Yang CY, Chen YY (May 2013) "Adaptation of mobile robot to intelligent vehicles," IEEE 4th proceedings of international conference on software engineering and service science, 550–553, 23–25
4. Casper J, Murphy RR (June 2003) Human-robot interactions during the robotassisted urban search and rescue response at the World Trade Center. Syst Man Cybernetics Part B: Cybernetics IEEE Trans 33(3):367–385

5. Matsuno F, Tadokoro S (Aug. 2004) “Rescue robots and systems in Japan,” Robotics and Biomimetics. ROBIO 2004. IEEE International Conference on, 12(20):22–26
6. Murphy R, Stover S, Pratt K, Griffin C (Oct 2006) “Cooperative damage inspection with unmanned surface vehicle and micro unmanned aerial vehicle at hurricane wilma,” Intell Robots Syst 2006 IEEE/RSJ International Conference on, 9:9
7. Ho M, Fall K (2004) “Poster: delay tolerant networking for sensor networks.” Proc IEEE Conference on Sensor and Ad Hoc Communications and Networks
8. Ohno K, Kawatsuma S, Okada T, Takeuchi E, Higashi K, Tadokoro S (Nov 2011) “Robotic control vehicle for measuring radiation in Fukushima Daiichi Nuclear Power Plant,” Safety, Security, and Rescue Robotics (SSRR), 2011 IEEE International Symposium on, 38(43):1–5
9. Hayashi T, Kawamoto H, Sankai Y (Aug 2005) “Control method of robot suit HAL working as operator’s muscle using biological and dynamical information,” Intell Robots Syst (IROS 2005). 2005 IEEE/RSJ International Conference on 3063(3068):2–6
10. Jezernik S et al (2003) “Robotic orthosis lokomat: a rehabilitation and research tool.” Neuromodulation: Technology at the Neural Interface 6(2):108–115
11. Clement RGE, Bugler KE, Oliver CW (2011) Bionic prosthetic hands: a review of present technology and future aspirations. The Surgeon 9(6):336–340
12. HA Samani Book: Lovotics, Loving robots, ISBN: 3659155411, 2012
13. Mukai T et al (2011) Realization and safety measures of patient transfer by nursing-care assistant robot riba with tactile sensors. J Robot Mechatron 23(3):360–369
14. Degen JW et al (2005) CyberKnife stereotactic radiosurgical treatment of spinal tumors for pain control and quality of life. J Neurosurgery: Spine 2(5):540–54
15. Dudek G, Jenkinj M, Milios E, Wilkest D (1993) A taxonomy for swarm robots. In: Proceeding of the 1993 IEEE/RSJ International conference on Intelligent Robots & Systems yokahama, Japan, 26-30 July 1993, pp 441–447
16. Cao et al (1997) An emerging view of scientific collaboration: scientists’ perspectives on collaboration and factors that impact collaboration. J Am Soc Inform Sci Technol 54(10):952–965, 2003
17. Gazi & Fidan (2007) Coordination and control of multi-agent dynamic systems: models & approaches. In: Conference: International Workshop on Swarm Robotics**, pp** 71–95
18. Iocchi et al (31 Aug 2001) Coordination in multi-agent RoboCup teams. Robo Auton Syst 36(2–3):67–86
19. Wada K, Shibata T (2007) Living with seal robots—its sociopsychological and physiological influences on the elderly at a care house. Robotics, IEEE Transactions on 23(5):972–980

Static Simulation of Star Images

M. L. Tejaswini, N. S. Vaarunya, Surabhi S. Sugur, Sushma B. Byahatti, and M. Spoorthi

Abstract A star tracker is an optical device used to track and measure the star positions. A star tracker, when placed on a satellite or spacecraft, can be used to determine the attitude or orientation of the satellite/spacecraft with respect to the position of stars. For this to be done, the tracker must obtain images of the stars, calculate the position with reference to the field of view of the spacecraft and identify the stars by relating their current position with respect to their absolute position by using a star catalog. This project simulates the acquired image and motion of the stars according to the field of view of the star tracker placed on a spacecraft orbiting a planet. An interface for the user was made including input parameters like the field of view (FOV), magnitudes, direction cosines etc. After these parameters, having been defined by the user, are given as input, the corresponding simulation output is displayed.

Keywords Direction cosines · Field of view · Star catalog

1 Introduction

Star tracker is an electro-optics system that plays an important role in highly accurate attitude measurement. Many satellites and spacecraft are equipped with star trackers and rely on them for this information for missions like Earth observation or deep space exploration. It takes an image from a set of stars and compares it with the star catalogue, determines the attitude of the spacecraft and thus the deviation is controlled [1].

For example, to obtain an image with high-resolution during earth observation for a particular geographical position, it is very important to precisely control the attitude of the remote sensing satellite using high-precision attitude detail from a star tracker [2]. In order to test these devices before being deployed, testing would have to be done using the night sky, but factors like the weather would interfere with the testing. Additionally, the cost of testing is also a concern. Considering all

M. L. Tejaswini (✉) · N. S. Vaarunya · S. S. Sugur · S. B. Byahatti · M. Spoorthi
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_103

of these issues, testing would be better suited in a software simulation capable of simulating real space environments according to the input parameters which vary with the specifics of the mission. A simulation gives the benefit of getting clear results of testing without interfering factors so that the test results can be used to reduce or prevent the chances of failure of the device [1].

Simulating a star map reliably is a vital source for testing the performance of a star tracker. Design input alternatives, including wide Field-of-View optics and other parameters may vary to suit specific mission needs [3]. The programming was done in MATLAB.

2 Input Parameters

- **Field Of View (FOV)**: It is the measure of the angular coverage area of a sensor to electromagnetic radiation or alternatively, can be defined as the extent of the area that is observable at any point in time. It is defined in X and Y directions in degrees.
- **Rows and Columns**: The parameters used to represent the X-axis and Y-axis of window size in pixels.
- **Star Catalog**: It is a catalog that contains a list of stars with various attributes like magnitude etc. Depending on the catalog, the coordinates of the stars are in direction cosines, RA and DEC, etc.
- **Magnitude**: The magnitude of the star represents the brightness of the stars.
- **Focal length**: The focal length of the optics used for collimating the dots displayed on the screen. Used to determine the angular resolution of pixels forming the display.
- **Look angle**: The angle pointing to a particular part of the sky. It is specified in direction cosines. It must be normalized.
- **Direction Cosine**: They are the cosines of the angles between the vector and the three coordinate axes X, Y. Z. They are used for constructing matrices that express one set of orthonormal vectors in terms of another set, or for expressing a known vector on a different basis.

The values for these input parameters are entered by the user through the Graphical User Interface after which the algorithm will process these values and give the corresponding output.

3 Algorithm

In order to obtain the desired output of the star-field image simulation, the user must give the input values in the Graphical User Interface for the different parameters like Field Of View, Look Angle which will be calculated using the degree of shift value entered by the user, Resolution specified in rows and columns (in pixels) for defining window size, magnitudes of the stars which are to be extracted from a star catalog.

The focal length of a lens defines the lens angular FOV. Equation to find focal length from FOV is given below:

$$fx = 1/\tan\left(\frac{fovx}{2}\right) \tag{1}$$

$$fy = 1/\tan(\frac{fovy}{2}) \tag{2}$$

where "*fx*" is the focal length for FOV in the x-direction and "*fy*" is the FOV in the y-direction. By changing the FOV parameter the angular resolution of the display pixels per degree is changed. The various details of the stars like their coordinate positions in space and magnitude are accessed from a star catalogue.

The angle between the look angle direction cosine and the star catalogue direction cosines will be calculated. If the angle between the two is less than half of the FOV, then the star is taken into account to be within the FOV and therefore, the remainder of the stars are not counted. Thus, the number of stars lying in the field of view is obtained.[4]. Angle between the look angle direction cosine and the star catalog direction cosine can be calculated as:

$$\theta = \cos^{-1}(a.b) \tag{3}$$

where a is the look angle direction cosine and b is the star catalog direction cosine.

After obtaining stars in FOV, coordinate conversion transforms will be applied to convert from three-dimensional world coordinates to two-dimensional pixel coordinates. First, the world coordinates will be converted into camera plane coordinates and then a 3D to 2D conversion to the image plane and finally to pixel plane. Thus, the simulation will display an image according to the user values to the input parameters [5].

$$u = f\frac{X}{Z} + ox \tag{4}$$

$$v = f\frac{Y}{Z} + oy \tag{5}$$

where u and v are pixel coordinates, f is focal length, X, Y, Z are image coordinates, ox, oy are image center. Thus, the simulation will display an image according to the user values to the input parameters [5] (Fig. 1).

Fig. 1 Flow chart for the process of Static Star Image Simulation

4 Simulation Results

The proposed algorithm is designed using the MATLAB software and is simulated to obtain the static star images. Different output results are obtained for different input parameters. An increase in the FOV results in increase in the number of the stars. Also, a change in the position of the star tracker's camera results in the shift of the position of stars which can be simulated using the proposed algorithm.

We can compare Figs. 2 and 3. and observe the shift in the stars position for a change in the position of the star trackers' camera.

As the FoV is increased, the number of stars captured by the star sensor also increases. This kind of result can be observed by comparing Fig. 4 and 5, where the number of stars in Fig. 5 is greater than the number of stars in Fig. 4. Also, the numbers of stars in a given FoV are directly proportional to the magnitude given by

Fig. 2 A static star image with FOV 30 × 30 with 0 degree shift

Fig. 3 A static star image with FOV 30 × 30 and a 5 degree shift

the user. When the code is simulated through GUI, then the user inputs the necessary parameters and clicks on the pushbutton which simulates and displays the required static star image (Fig. 6 and Table 1).

Table 1 gives a comparison between the numbers of stars present in different Field of View. It can be observed that as the FOV of the camera lens is increased, the number of stars get added and the count of stars in the FOV also increases. As

Fig. 4 A static star image with FOV 40 × 40

Fig. 5 A static star image with FOV 25 × 25

the value of magnitude entered in GUI decreases, number of stars in corresponding output images will also decrease.

5 Conclusion and Future Enhancement

The star image simulations for various combinations of input parameter values were obtained and the magnitude values were found to have an effect on the number of

Fig. 6 Graphical user interface of the simulation

Table 1 Variation of number of stars with FOV

FOV	Stars within the given magnitude	Stars in FOV
25 × 25	57	3
30 × 30	57	5
40 × 40	57	6

stars being displayed. The shift of the stars could also be observed with changing look angle values. Finally the GUI was created using MATLAB's GUIDE to enable the user to input values.

The future scope of the project includes the dynamic simulation of the star image using the quaternions that is, simulating the high rate and acceleration of the star image such that when seen by the star sensor, the movement of stars is being observed and smearing of stars can also be observed.

References

1. Ardi NS, Saifudin M, Eko-Poetro R, Fathurrohim L (2018) "Development of star image simulator for star sensor algorithm validation". J Phys: Conf Series 1130
2. Iwata T (11 Jan 2005) "Precision attitude and position determination for the advanced land observing satellite (ALOS)," Proc SPIE 5659, Enabling Sensor and Platform Technologies for Spaceborne Remote Sens
3. Kazemi L, Enright J (2019) "Accurate star tracker simulation with on-orbit data verification," IEEE aerospace conference, Big Sky, MT, USA, pp 1–8
4. Kandiyil R (2010) "Attitude Determination software for a star sensor", Master Theses, Lulea University of Technology, Space Science and Technology, Department of Space Science, Kiruna
5. Smith CG (2017) "Development and implementation of star tracker based attitude determination", Masters Theses, thesis number T 11238, Missouri University of Science and Technology

Power Quality Improvement by Photovoltaic Integrated-UPQC-S with Three Level Neutral Point Clamped Diode Inverter Using Modified PQ Theory

R. Anguraja and W. M. Sivakumar

Abstract The paper presents a modified active and reactive power based control of a solar PV array integrated unified power quality conditioner (PV-UPQC-S) with three level neutral point clamping inverter (NPCI). The system achieves improvement in power quality and generation of clean energy and this in turn improves feature of the system. The fundamental-frequency positive-sequence (FFPS) of voltage components at the point of common-coupling (PCC) are obtained using the generalized cascaded delay signal cancellation technique (GCDSC), and are used in reactive power based control to approximate reference signal for the PV-UPQC-S. This alteration in reactive power based theory allows its implementation for control of PV-UPQC-S for distorted state of PCC voltage. The voltage source converter of series type for PV-UPQC-S gives a portion of VAR(reactive power) to load even during any condition of grid. The photovoltaic array is unified at UPQC dc bus which contributes a proportion of load active power and this reduces the demand on supply. Simulation using MATLAB-Simulink is done to verify the system performance for steady state and dynamic conditions.

Keywords PV-UPQC-S · PCC · Voltage source converter · Fundamental-frequency positive-sequence · FFPS · MATLAB-Simulink · Steady state performance · Dynamic performance · Solar energy · Wind energy · VAR

1 Introduction

A collective necessity for renewable energy systems which has additional features especially where there is low voltage distribution system. If the quality of the power is low then there will be high power loss and it may lead to displeasing conduct with adjacent communication lines [1]. This paper emphasizes on unified power quality conditioner is basically from APF family. In UPQC shunt and series APF is combined to increase the quality of the power so the power quality can be improved

R. Anguraja (✉) · W. M. Sivakumar
Department of Electrical and Electronics Engineering, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_104

at distribution level. UPQC categorization can be done as (1) current or voltage source converter (2) two-wire, three-wire, three-phase (3) new configurations for three-phase system [2]. Although semiconductor device loads are energy competent, and injects harmonics to the grid and this leads to disturbance that may occur at PCC especially in a weak grid system. Sensitivity of power electronic loads to voltages may also cause distortion and there may be fluctuations in voltage in distribution system because of periodic energy sources like wind and solar energy.

The fluctuation of voltage has impact on sensitive power semiconductor loads for example lighting systems it may chances of tripping the electrical system which will increase maintenance costs. A PV-UPQC-S photovoltaic (PV)-array-integrated system will generate clean energy and quality of the power is also improved, thus making the system efficient [1]. PLL is a 3 phase cascaded signal is used for selective harmonic detection and unwanted harmonics is totally removed which would result in steady-state error detection. Multiple harmonics can be tracked by parallel arrangement of numerous CDSC operators is proposed [3].

In distribution systems of low-voltage there is a high demand for renewable energy systems. The solution is an UPQC-S proposed in this paper, shunt VSC has an advantage of improved load voltage regulation and it also provides improved grid current quality. This paper proposes the improved PQ theory algorithm and this algorithm of PQ theory is based on GCDSC. It allows PV-UPQC-S operation below adulterated voltage conditions.

PV array along with PV-UPQC dc link reduces demand on supply system load, maintenance cost in turn quality of power is improved and also frequent of power tripping is reduced.

2 PV-UPQC-S Modelling

PV-UPQC-S with three levels NPC is illustrated in Fig. 1. It contains two part series VSC and shunt VSC, which are connected using a DC bus as shown in the Fig. 1. Interfacing inductors are connected to grid of VSC's. Switching harmonics are filtered using Ripple filters. Voltage is injected through series transformer in VSC. Reverse block diode is used to connect DC bus and SPV array.

3 Pv-Upqc-S Control

3.1 Structure for Control of Shunt VSC

The Fig. 2 shows the simulation of Shunt VSC model using Simulink. The problems related to Power quality like unbalanced load and inaccurate current are minimized by using shunt VSC. Shunt VSC also supplies real power from the PV arrays to

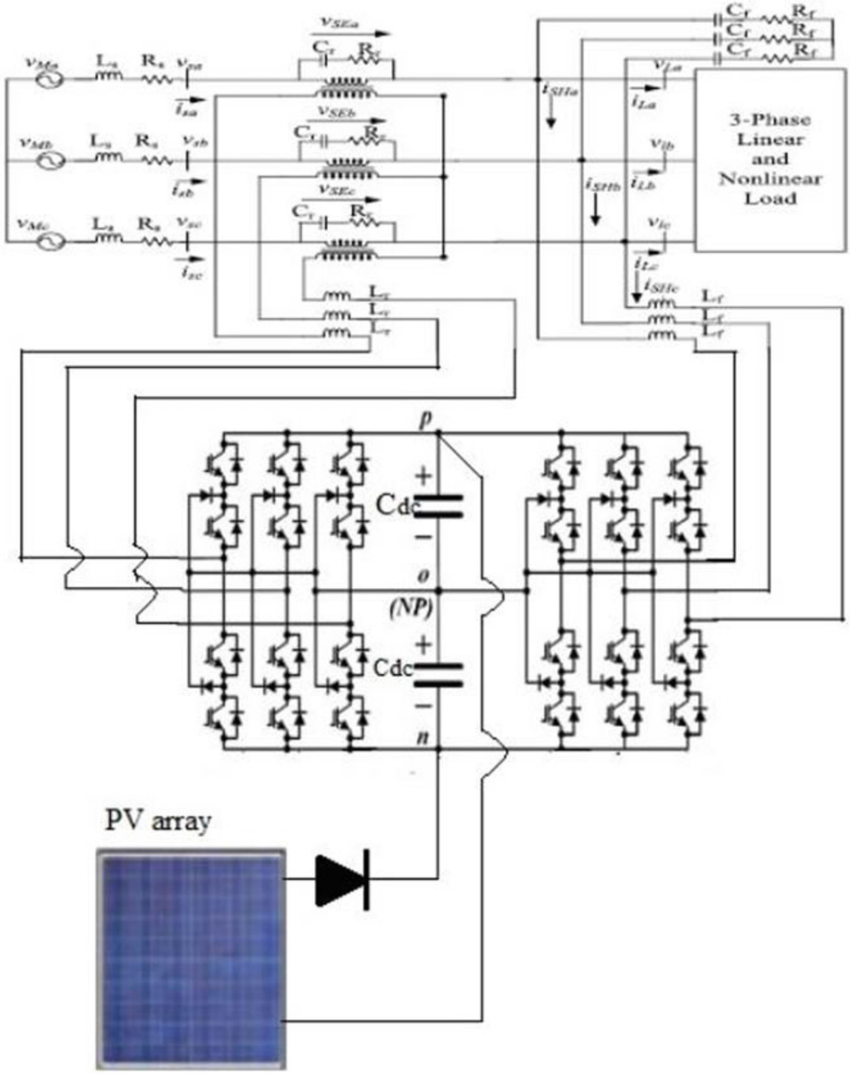


Fig. 1 Three level NPC PV-UPQC-S Circuit diagram

the grid and dc bus voltage is maintained at PV array MPP voltage. The MPPT algorithm provides the dc bus voltage reference value. The PV array maximum power is obtained from P&O algorithm. The measured dc bus voltage (Vdc) is sent to LPF and compared with dc reference bus voltage (V∗dc). The error in V∗dc and Vdc is sent to PI controller, which gives the loss component (Ploss). Power (Pref) drawn from grid is given by,

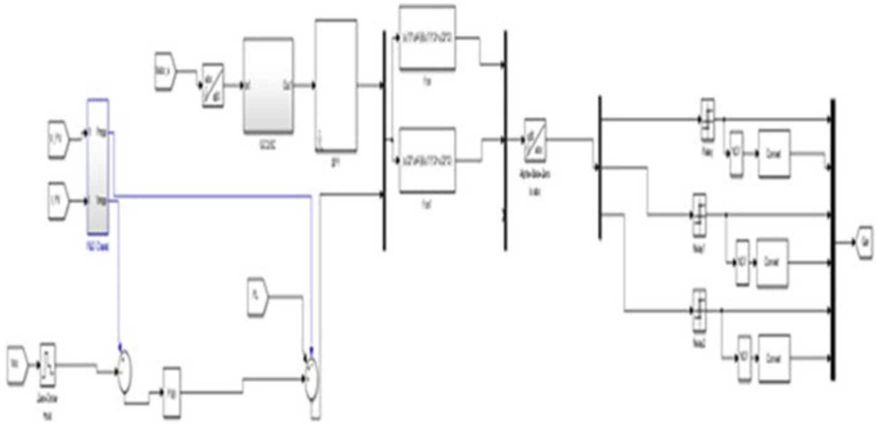


Fig. 2 Shunt VSC model using Simulink

$$\text{Pref} = P_L + P_{loss} - P_{pv}$$

GCDSC block gives vs1α and vs1β.

3.2 Photovoltaic Array Desgin

The PV array works on solar photovoltaic principle. The design parameter of FIRST SOLAR FS-272

Open circuit voltage (Voc) = 94.57 V
Maximum voltage (Vmax) = 70.88 V
Short circuit current (Isc) = 1.18 A
Maximum current (Imax) = 1.010 A
Twelve number of series connected module per string and twenty four number of parallel string are used to construct PV array

4 Results

The MATLAB-simulink tool is used to simulate the steady state performance & dynamic performance of PV-UPQC-S system.

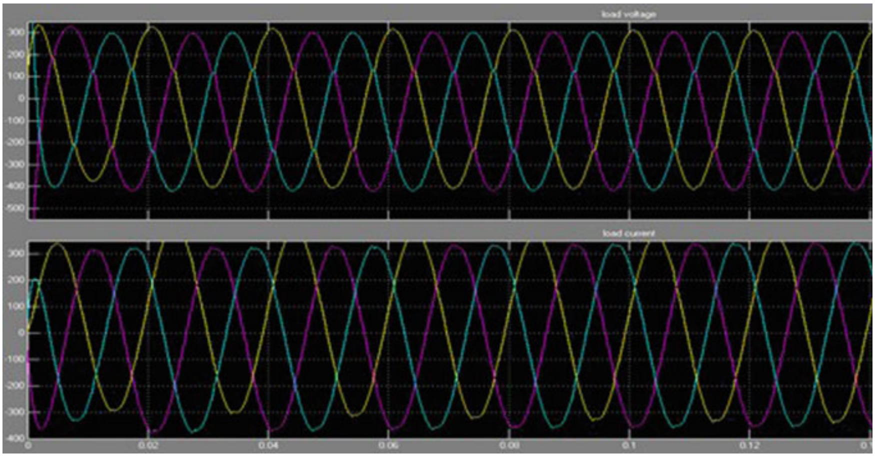


Fig. 3 PV-UPQC-S Steady state performance

4.1 Steady State Condition

The system performance during steady state is illustrated in Fig. 3. The nonlinear load causes the total harmonic distortion at supply side and load side because current is not sinusoidal with respect to voltage. In order to reduce the %THD the PV integrated UPQC-S is connected to PCC.

It can be observed that up to 0.15S initial disturbance after that the load current is sinusoidal, by connecting PV-UPQC-S the disturbance due to nonlinear load is mitigated and also the FFT analysis in the MATLAB Simulink at 50Hz fundamental frequency the load current distortion is 1.30%, which is within the limit of IEEE-519.

4.2 Dynamic State Condition

The PV-UPQC-S is illustrated in fig. 4. It is observed that the shunt VSC keeps the balanced grid current and the load current is settled within 0.2 Sec.

And also the FFT analysis shows at 50Hz fundamental frequency and 1.49 % load current harmonic distortion which is within the limit as shown in fig. 5.

Irradiation change PV-UPQC-S performance is illustrated in fig 7. UPQC dc bus is integrated with PV array; hence demand on supply system is reduced. It is evident that the integrated photovoltaic shares active power to the load which is illustrated in the Fig. 6.

Due to the discontinuous nature of solar energy, fluctuations increases based on load demand, during this condition PV- UPQC-S compensates for the fluctuations in voltage and current of load. The FFT analysis is performed and %THD is 1.31% as shown in Fig. 8.

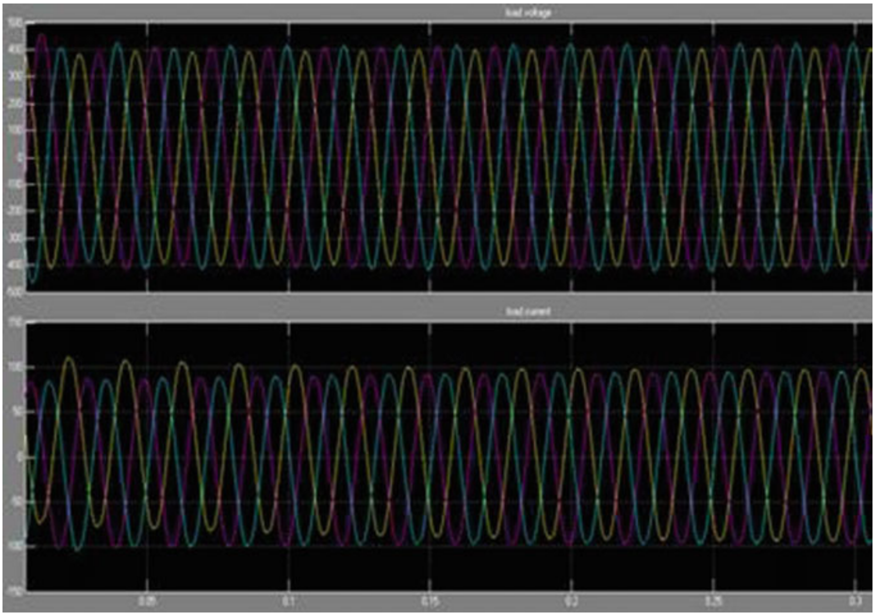


Fig. 4 UPQC Performance during load unbalanced

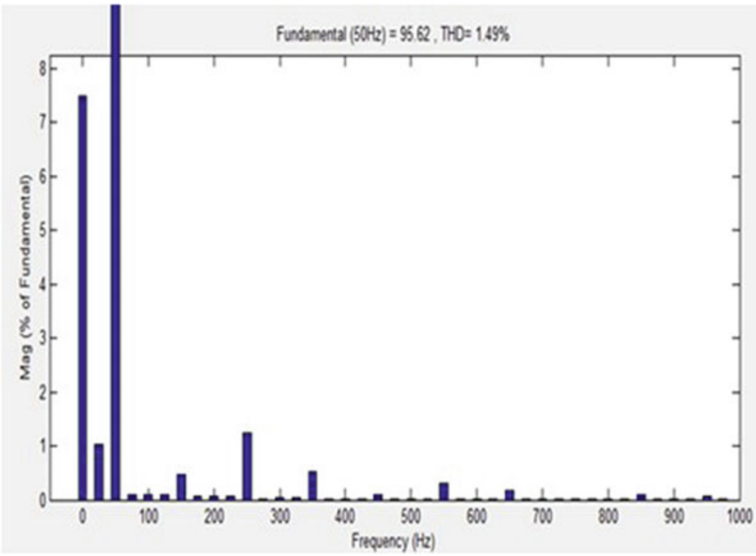


Fig. 5 THD in FFT analysis during load unbalanced

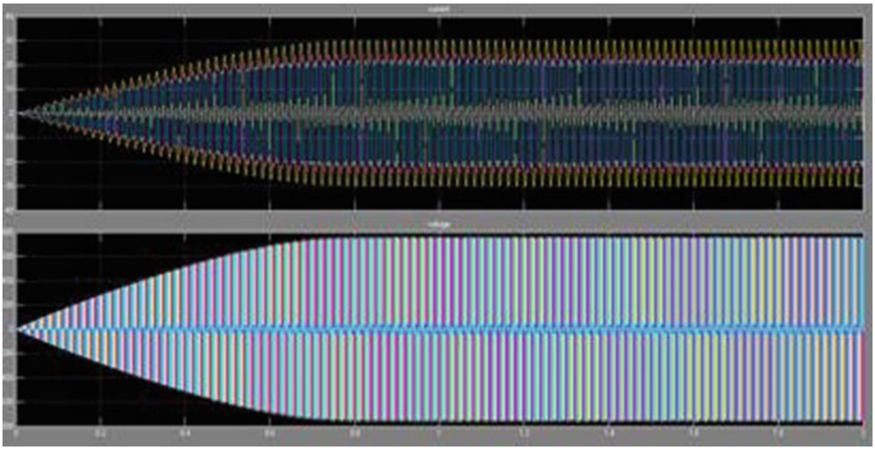


Fig. 6 Current and voltage supplied by PV array

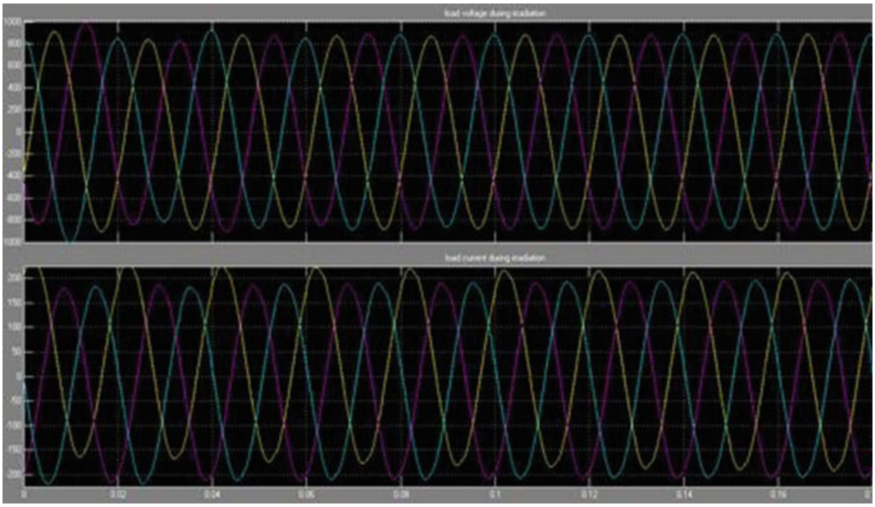


Fig. 7 Load current and voltage of PV-UPQC-S during irradiation change

Under Ideal PCC voltage distortion which is taking place with respect to grid voltage varies from 0.7 PU to 1 PU corresponding to the time from 0.5 to 1.5 seconds, the PV-UPQC-S injects or absorbs a voltage, at 0.7 PU the voltage sag is compensated by UPQC by adding 0.3 PU voltage and at 1.3 PU results a voltage swell is compensated by series voltage source converter by absorbing the voltage.

The grid voltage distortions and the Load voltage distortion compensated by PV-UPQ is shown in Figs. 9 and 10.

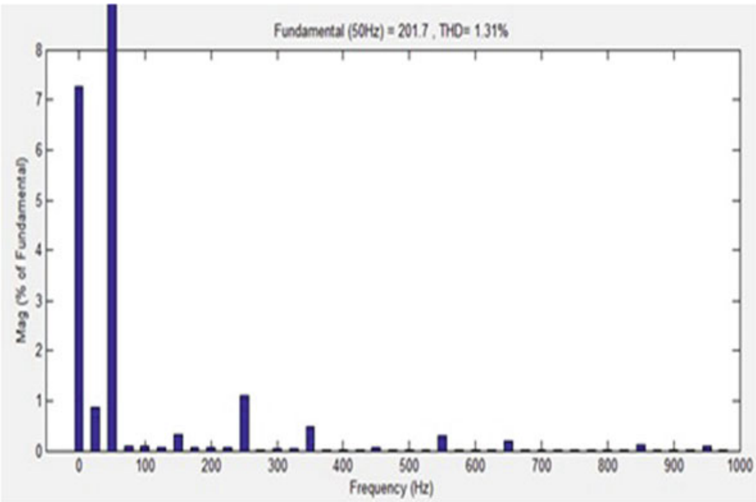


Fig. 8 THD in FFT analysis during irradiation change

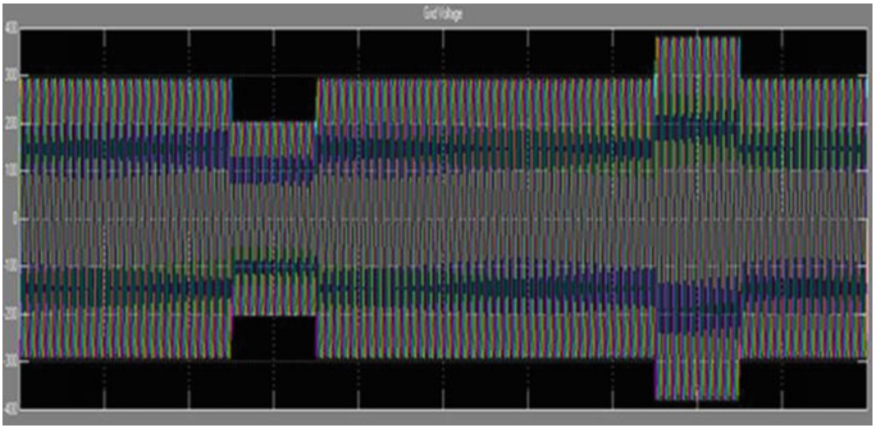


Fig. 9 Grid voltage distortions

5 Conclusion

The performance of proposed system is analyzed using MATLAB-simulink. The system provides the generation of green energy and improvement in quality of power which increases the utility of system. The distorted grid voltage of FFPS is extracted from GDSC block that can be utilized for the control of PV-UPQC-S. Dynamic performance of proposed systems is analyzed for irradiation's sudden change and grid voltage fluctuations. The proposed control technique boosts the performance

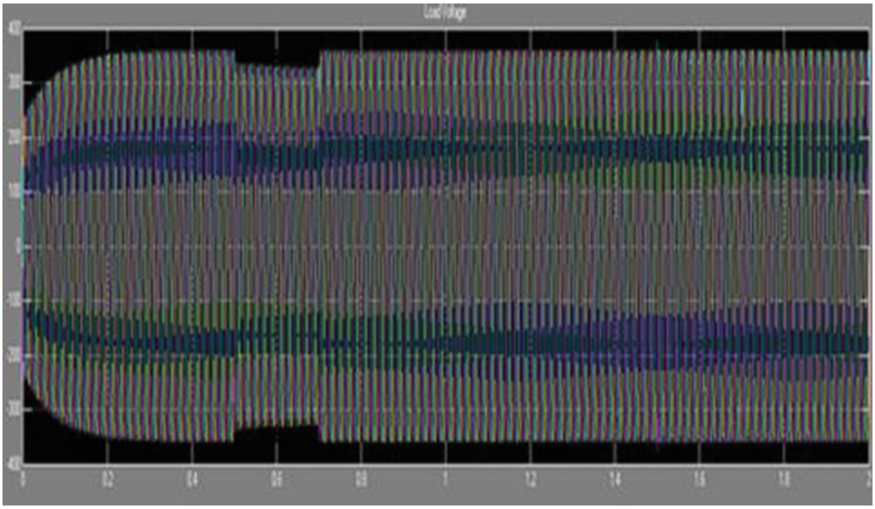


Fig. 10 Load voltage distortion compensated by PV-UPQ

of series VSC by compensating VAR to the load during nominal conditions of the grid. In the proposed work, the series VSC utilization is increased by supplying the reactive power to the load.

Acknowledgements The authors are thankful to Don Bosco Institute of Technology for providing required facilities & continuous encouragement for research work.

References

1. Devassy S (2017) Modified pq-theory-based control of solar-pv-integrated upqc-s. IEEE Trans Ind Electron 53(5):5031–5040
2. Khadkikar V (2009) A novel structure for three-phase four-wire distribution system utilizing unified power quality conditioner (UPQC). Proc IEEE ISIE 53(5):1897–1902
3. Wang YF (2013) Three-phase cascaded delayed signal cancellation pll for fast selective harmonic detection. IEEE Trans Power Syst 60(4):1452–1463
4. de Oliveira Costa LF (2016) Electrical power quality and the challenges faced by power assembly's applications in petrochemical industry. IEEE Trans. Power Electron. 30(1):387–397
5. Kumar S, Verma AK, Hussain I, Singh B, Jain C (2017) Better control for a solar energy system: Using improved enhanced phase-locked loopbased control under variable solar intensity. IEEE Ind Appl Mag 23(2):24–36
6. Rauf AM, Khadkikar V (2015) Integrated photovoltaic and dynamic voltage restorer system configuration. IEEE Trans Sustain Energy 6(2):400–410
7. Khadkikar V, Chandra A (2009) A novel structure for three-phase four-wire distribution system utilizing unified power quality conditioner UPQC. IEEE Trans Ind Appl 45(5):1897–1902
8. Khadkikar V (2012) Enhancing electric power quality using UPQC: a comprehensive overview. IEEE Trans Power Electron 27(5):2284–2297

Electric Field Distribution of 800 kV OIP Transformer Bushing

R. Anguraja and Pradipkumar Dixit

Abstract Electrical bushings are the basic equipment of HV power transformer and their successful operation depends on its reliability. Most of HV power transformer bushings are condenser oil impregnated paper (OIP) type. High voltage insulated bushing are needed to draw the electrical energy through earthed barrier. This paper deals with improving the electric field distribution inside a 800 kV OIP condenser bushing by varying number of foils and their thickness using finite element method. Simulation results show that 18 foils, 4 mm thickness each and 6 mm gap between the foils is optimum for better distribution of the field.

Index Terms Electrical field intensity · Finite element method · High voltage bushing · Oil impregnated paper

1 Introduction

The porcelain bushing of non-condenser type for low voltage transformers mainly consists of four types of dielectrics namely Bakelite tube/PVC/paper tape, transformer oil, porcelain and air, because of their good dielectric property and mechanical strength [1]. Bushings mechanically support external conductors and provide insulation from earth as the conductors are fed into the transformer tank [2]. The cost of bushing is small portion of the total power transformer cost but their failure is often catastrophic.

The study of HV bushings electric field distribution is of substantial importance for enhancement of HV transformers reliability. The electric field distribution is mainly controlled by the geometry of bushing, permittivity, insulating materials volume resistivity and the surface resistivity due to contamination of the surface. Voltage

R. Anguraja (✉)
Department of Electrical and Electronics Engineering, Don Bosco Institute of Technology, Bangalore, India

P. Dixit
Department of Electrical and Electronics Engineering, MSRIT, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_105

distribution control mainly is the main problem in metal flange section and solution for that has to be found.

Capacitance graded bushings are usually made of paper insulation. The metallic foils are placed between the paper insulation at calculated radial distances. The potential is equally distributed because the capacitance between is equally distributed. In this paper, transformer employs (OIP) as the main insulator. The main advantage of OIP over other type of insulators is its economical cost, better thermal performance and dielectric property [3].

2 Modeling of Bushing

A. Dimensional details of 800 kV bushing

The dimensions details of 800kV bushing are taken according to Central Board of Irrigation and Power.

(i) The bushing total length is 9655 mm
(ii) The inner conductor diameter is 30 mm
(iii) The radial distance between the conductor and the porcelain is 350 mm.

To get uniform electric field, Aluminum foils are used in between paper layer.

The model is built for different number of aluminum foils within it for different thickness of OIP and different thickness of foils as discussed in section III. Figures 1 and 2 shows the 800 kV bushings model and meshed model developed according to CBIP for FEM analysis. The parameters assigned for the analysis are

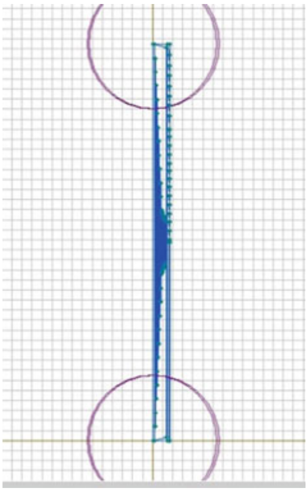


Fig. 1 Foils OIP bushing model built in quick field

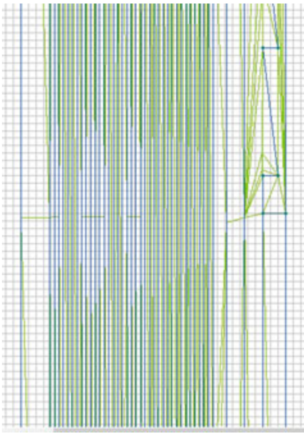


Fig. 2 Meshed model of 800 kV

(i) Input voltage = (800*(1000)/ $\sqrt{3}$))
(ii) Relative permittivity of air = 1
(iii) Permittivity of conductor = 100,000
(iv) Permittivity of foils = 10,000
(v) Relative permittivity of oil = 2.2
(vi) Relative permittivity of OIP = 4.

Electric field analysis for different foils and thicknesses has been carried out and details of which are discussed in next section.

3 Estimation of Electric Field

As mentioned earlier, the main objective of the paper is to obtain optimum electric field distribution inside and voltage distribution along the oil reservoir side of bushing 800kV, OIP bushing. The estimation of electric field has been carried out with different number of foils and thicknesses without changing the dimensions of the bushings. The study has been made with 9 foils, 12 foils and 14 foils. The different combinations are studied.

The FEM simulation has been carried out for 12 foils 15 foils and 18 foils for all 12 cases mentioned above.

Table 1 gives the maximum electric field estimated in each of the cases.

Table 1 Maximum electric field estimated

Case	Description	Maximum electric field kV/mm
1	12 foils 4 mm foils and 4 mm OIP	1.176
2	12 foils 4 mm foils and 6 mm OIP	7.896
3	12 foils 6 mm foils and 4 mm OIP	12.236
4	12 foils 6 mm foils and 6 mm OIP	7.758
5	15 foils 4 mm foils and 4 mm OIP	9.848
6	15 foils 4 mm foils and 6 mm OIP	6.473
7	15 foils 6 mm foils and 4 mm OIP	10.390
8	15 foils 6 mm foils and 6 mm OIP	6.796
9	18 foils 4 mm foils and 4 mm OIP	9.010
10	18 foils 4 mm foils and 6 mm OIP	5.666
11	18 foils 6 mm foils and 4 mm OIP	9.403
12	18 foils 6 mm foils and 6 mm OIP	5.969

4 Results and Discussions

Figure 3a and b shows respectively the complete model and zoomed out sketch of electric field strength distribution inside the OIP bushing with 12 foils of 4 mm thickness and OIP thickness of 4 mm.

To verify the distribution of the field, electric field has been plotted as function of radial distance inside the bushing for 12, 15 and 18 foils and shown in the Figs. 5, 6 and 7 respectively.

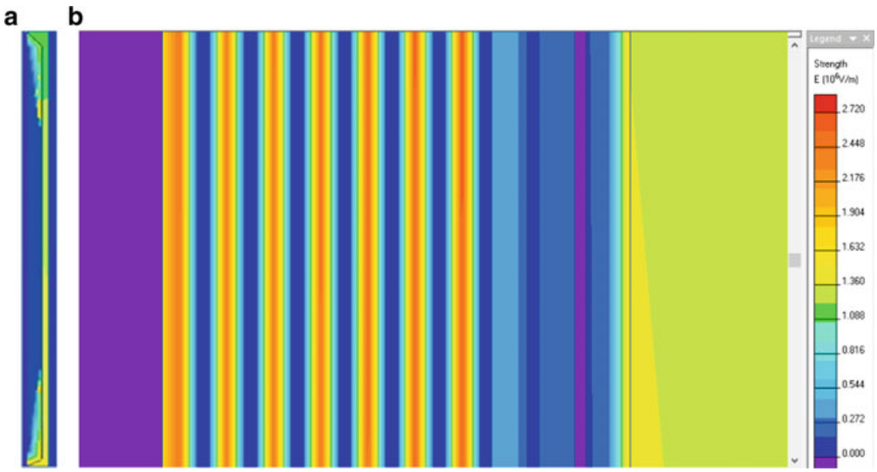


Fig. 3 **a** Electric field distribution & of 12 foils with 4 mm thickness OIP 4 mm thickness. **b** Electric field distribution for 12foils with 4 mm thickness and OIP with 4 mm thickness

The Fig. 4 shows field distribution of 12 foils
Where

a = 4 mm foils and 4 mm OIP (Graded);
b = 4 mm foils and 6 mm OIP (Graded)
c = 6 mm foils and 4 mm OIP (Graded);
d = 6 mm foils and 6 mm OIP (Graded).

From Fig. 4, it is observed that the least electric field strength 7.758kV/mm is obtained from 12 foils which are graded with 6mm thickness and 6mm thickness of OIP.

The Fig. 5 shows the representation of 15 foils results
Where,

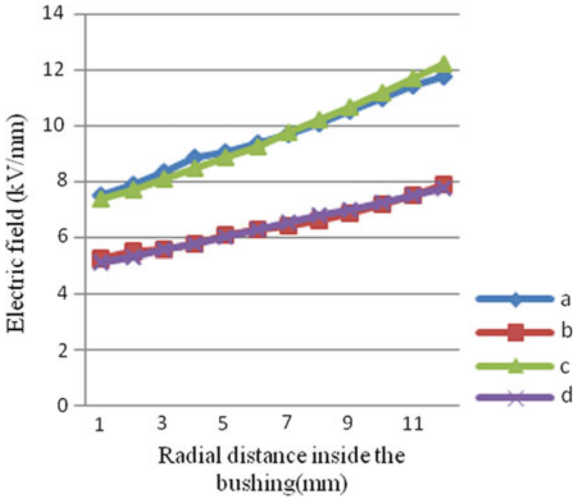


Fig. 4 Electric field distribution inside 12 foils results

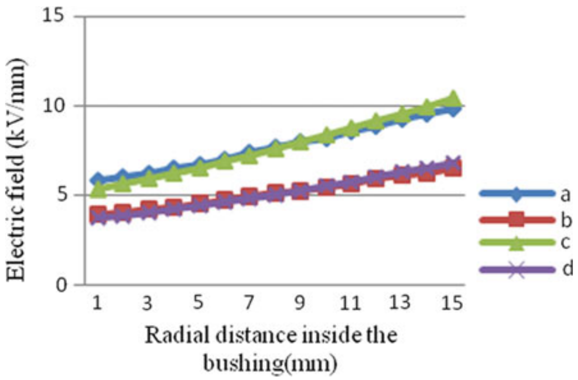


Fig. 5 Electric field distribution inside 15 foils results

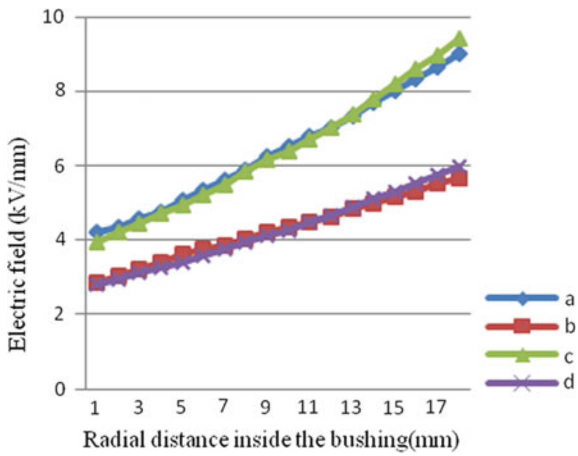


Fig. 6 Electric field distribution inside 18 foils results

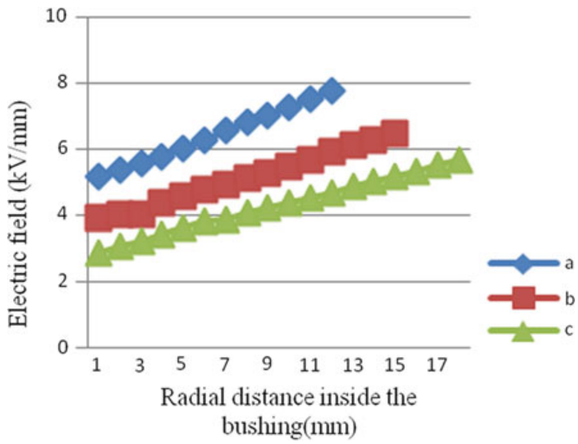


Fig. 7 Uniform and least Electric field in 9 foils, 12foils and 18foils respectively

i = 4 mm foils and 4 mm OIP (Graded)
ii = 4 mm foils and 6 mm OIP (Graded)
iii = 6 mm foils and 4 mm OIP (Graded)
iv = 6 mm foils and 6 mm OIP (Graded).

From the Fig. 6, it is observed that the least electric field strength 6.473kV/mm is obtained from 15 foils which are graded with 4mm thickness and 6mm thickness of OIP.

Figure 6 shows the representation of 18 foils results

Where,

i = 4 mm foils and 4 mm OIP (Graded)
ii = 4 mm foils and 6 mm OIP (Graded)
iii = 6 mm foils and 4 mm OIP (Graded)
iv = 6 mm foils and 6 mm OIP (Graded).

From the Fig. 6, it is observed that the maximum electric field strength 5.666 kV/mm is obtained from 18 foils which are with 4mm thickness and 6mm thickness of OIP.

The Fig. 7 shows uniform field distribution in 9 foils, 12 foils and 18 foils respectively.

Where,

a = 6 mm foils and 6 mm OIP in 9 foils (Graded)
b = 4 mm foils and 6 mm OIP in 12 foils (Graded)
c = 4 mm foils and 6 mm OIP in 14 foils (Graded)

From Figs. 5, 6 and 7, it could be observed that, case (10) of 4mm thickness with OIP of 6mm 18 foils the least electric field recorded is 5.666kV/mm and it is optimum for better distribution of Electric Field in 800kV OIP bushing. It is perhaps the least value to be found in comparison for all other readings taken in all other cases. This apparently suggests that the field is largely uniform in the above case. Consequently, the electric stress is relatively less and the possibility of insulation failure would be least, and hence it is optimum for 800 kV OIP bushing.

Figure 8 shows the field distribution along the oil reservoir (upper) side of bushing for 12, 15 and 18 foils.

Where,

a = 6 mm foils and 4 mm OIP in 12 foils (Graded)
b = 6 mm foils and 4 mm OIP in 15 foils (Graded)
c = 4 mm foils and 6 mm OIP in 18 foils (Graded)

Fig. 8 Field distribution along the oil reservoir (upper) side of bushing for 12, 15 and 18 foils

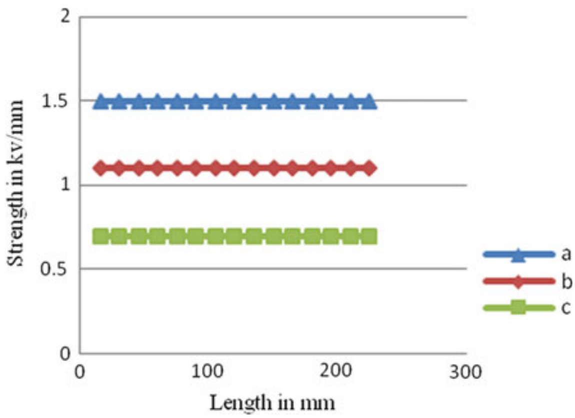
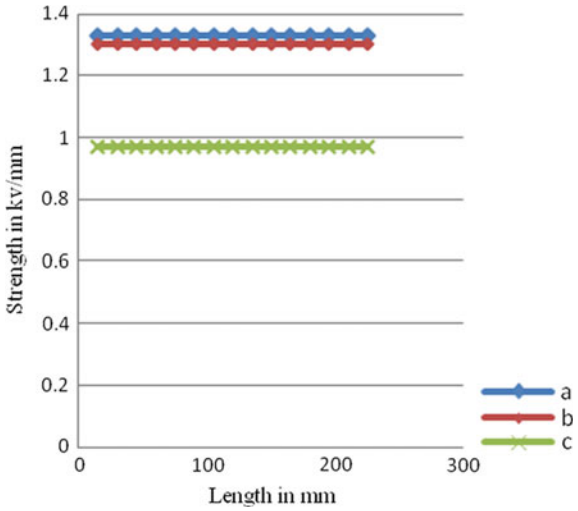


Fig. 9 Field distribution along the oil reservoir (down) side of bushing for 12, 15 and 18 foils



From Fig. 8 it is observed that 4mm foils and 6mm OIP in 18 foils, the electric field strength is recorded close to the conductor 0.786 kV/mm which is optimum for better distribution along the oil reservoir & top of the foils in 800 kV OIP bushing.

Figure 9 shows the field distribution along the oil reservoir (down) side of bushing for 12, 15 and 18 foils.

Where,

a = 6 mm foils and 4 mm OIP in 12 foils (Graded)
b = 6 mm foils and 4 mm OIP in 15 foils (Graded)
c = 4 mm foils and 6 mm OIP in 18 foils (Graded)

From Fig. 9 it is observed that 4mm foils and 6mm OIP in 18 foils, the electric field strength is recorded close to the conductor 0.974 kV/mm which is optimum for better distribution along the oil reservoir & bottom of foils in 800 kV OIP bushing.

5 Conclusion

The present paper deals with the estimation of electric field inside the 800kV OIP bushing using FEM. The study involves with different no of foils and thickness of OIP. The study shows that 18 foils which are with 4mm thickness and 6mm OIP is found to be well suited for optimum distribution of electric field inside the bushing.

Acknowledgements The authors are thankful to M.S.Ramaiah Institute of technology and Don Bosco Institute of Technology for providing necessary facilities and support.

References

1. Ganga S, Kanyakumari M, Shivakumara Aradhya RS "FEM approach to design functionally graded transformer bushing "CPRI", Bengaluru, India
2. Smith DJ, McMeekin SG, Stewart BG, Wallace PA "Transformer bushings-modelling of electric field and potential distribution within oil impregnated paper with single and multiple spherical cavities".school of enginnering & computing Glasgow Caledonian university Scotland UK
3. Jyothi NS, Ramu TS, Mandlik M (2010) "Temperature distribution in resin impregnated paper insulation for transformer bushings", IEEE Trans Dielectrics Elect Insulation

Automatic Engine Locking System for Drunken Driving

Padmashree V. Kulkarni, N. Manu, Meenaz Sadaf, Shadab Khan, and Rani

Abstract Today every country is facing serious problems with drunken drivers. Drunk driving is the reason behind most of the car accidents. So we are designing a system that will detect level of alcohol content in driver's breathe. The system is also equipped with an automatic braking system for any obstacle on its way. This system will make the vehicle driving more safe and secure when compared to current system. We are using alcohol content in driver‘s breathe to control the access the vehicle. In this system whenever the alcohol consumed by the driver exceeds the permissible limit the speed of the vehicle will be reduced and the ignition of the vehicle will be turned off. The Global Positioning System (GPS) unit will trace the current location of the vehicle and through IoT communication an Alert message will be sent to police or family members or to respective authorities. This paper deals with reduction on number accidents caused due to drunken driving.

Keywords Alcohol detection · MQ3 sensor · Arduino uno · ATmega328P · Engine locking · Internet of Things (IoT) · GPS · Drunken driving

1 Introduction

In the current scenario of the world Road safety have been decreasing considerably due to drunken driving, especially in India more accidents cases are being recorded due to drunken driving. Drunken driving is also emerging as the major problem in public safety and it should be strictly monitored and rectified as soon as possible. So we are proposing our project to minimize the cases of drunken driving and to boost up the road safety.

In our project we have found a solution for this by alcohol detecting system that monitor and controls the vehicle ignition whenever the level of alcohol consumed crosses the permissible limit. The alcohol molecule or contents are determined by the driver's breath using an alcohol detector placed in the steering in front or by a

P. V. Kulkarni (✉) · N. Manu · M. Sadaf · S. Khan · Rani
Department of Electrical and Electronics Engineering, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_106

detector place in the seat belt of the driver. When exceeding level of alcohol content in the blood is detected Arduino turns off the engine ignition and a buzzer alarm is initiated along with an indication "Alcohol Detected" is displayed in the LCD display. The Global Positioning System (GPS) tracks the current location of the vehicle and by using IoT communication system a distress message will be sent to respective authorities and control room. By using this system in vehicles the access of vehicle to drunk drivers are reduced and the number of accident cases due to it can be reduced which boosts up the public road safety and enhances safety measures for both public and driver.

2 Literature Survey

Sahu et al. [1] In this paper the author has designed a method to control the drunken driving. The author has fabricated an alcohol detecting system using a Atmega8 microcontroller and MQ-3 sensor. Here he uses his alcohol detection system to analyze the condition of the driver and monitors the ignition of the vehicle. When the alcohol consumed by the driver exceeds the permissible limit the vehicle ignition will be turned off and by using GPS and GSM units he is sharing the location of the vehicle to a registered number.

Dai et al. [2] An more efficient alcohol detecting unit, GPS and GSM units are used in this paper for controlling the speed of the vehicle when the driver is drunk, By using the units of this project more accurate results are obtained even though the cost increase than the previous paper.

Dhivya et al. [3] A Multiple monitoring systems have been designed in this paper which makes its system more complicated. The level of complex design have affected the project by giving the false alarm indication in there model which turns out to be the major drawback of this system. And we have adapted the necessary techniques used in this paper for our model and we have designed a more simplified and cost effective system.

James et al. [4]. The papers talks about the alcohol-detecting unit designed by using microcontroller PIC16F877A. The design proposed in this paper is outdated and expensive. So we are designing our system using an Arduino by reducing the cost and also making it easy to add more suitable units to it in future. My proposing our design components we are reducing the overall cost of the unit to make it affordable to everyone.

Prashanth et al. [5] Here the author has designed a alcohol detecting system using PIC microcontroller and MQ-2 sensor. This design is not used future because its false alarm and inaccurate data. so we have proposed the manufacture of our system using MQ-3 senor and Arduino by minimizing the false alarm and derived more accurate data. By using the Arduino in our model we have given an additional feature to add more systems along with our system.

3 Proposed Methodology

We have proposed our design using Arduino Uno R3 and MQ-3 alcohol sensor for detecting the level of alcohol consumption by the driver. And our designs also have control of the speed and ignition of the vehicle‘s engine. The level of alcohol consumption is derived by analyzing the level of alcohol content present in drivers breathe and further action are designed accordingly. We are using different alarms and indication system to communicate the threat to nearby public along with the respective authority.

At the Initial stage the alcohol content is measured and if it is under the permissible limit the driver is set free to access the vehicle. But, as soon as our system detects that the alcohol consumption of driver has exceeded the limit an immediate buzzer indication is given and the speed of the vehicle is reduced and the ignition of the vehicle is turned off. A message will be displayed on LCD screen displaying that “Alcohol have been detected”. The location of the vehicle will be tracked and shared to registered numbers and authorities. The flow chart of Methodology is as shown in Fig 1.

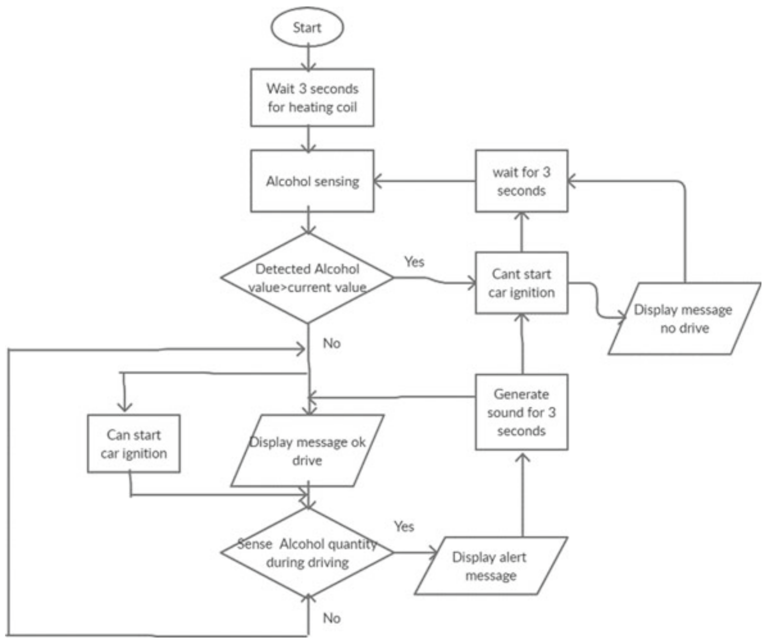


Fig. 1 Proposed methodology

Fig. 2 Arduino uno

4 Hardware

4.1 Arduino Uno

Arduino Uno is the widely used in designing monitoring and controlling system. ATMega328 microcontroller is embedded on this Arduino board. It is a 20 pin IC out of which 14 are digital pins and the other 6 pins are analog pins. These can be easily programmed by using Arduino IDE software which is free source software. These are cost effective and more reliable than conventional microcontroller and can handle more number of operations effectively. The Arduino board is shown in Fig. 2.

4.2 MQ-3 Sensor

MQ-3 sensor is a gas sensor, which is used to detect the alcohol content. Thses sensor are more accurate and sensitive than MQ-2 Sensor. It can detect the alcohol content upto 2 m range. The additional benefit of this sensor is that we can adjust the sensitivity of it which makes it more flexible to use.

4.3 16X2 LCD Display

An LCD display is the widely use indicator in various systems, A 16X2 LCD display is used in our project to display the status of status of the measurement of level of alcohol content and status of engine monitoring unit. It most commonly used than over seven segment display because these are easy to program. The 16x2 LCD display is shown in Fig. 3.

Fig. 3 LCD display

Fig. 4 Buzzer

4.4 Buzzer

A buzzer is also an alert system used in our model which indicates the status of the alcohol detecting unit. Whenever our system detects the exceeded level of alcohol buzzer alerts the nearby people by making a beep sound. A buzzer indication is widely used and effective indicator method to alert the nearby people. And it's efficient working and low cost makes it more convenient. Figure 4 shows the image of a buzzer.

4.5 GPS Module

Global Positioning System (GPS) is the unit which is used to track the location of vehicle. The GPS can track the exact location of the vehicle at any instant time unobstructed line of sight to four or more GPS satellites. Obstacles such as mountains and buildings block the relatively weak GPS signals. Along with Geo location GPS can also track the time and can share this information to any point on the earth. The GPS module used is shown in Fig. 5.

Fig. 5 GPS module

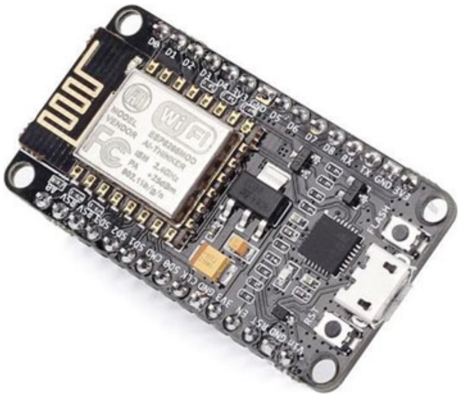


Fig. 6 Node MCU

4.6 Node MCU

Node MCU is a free open source IoT product. Node MCU is an open source Wi-fi Module which helps in communication through the wireless system. It can be easily programmed and many IDE are available to code this module according to needs. IT can be integrated to online chatting system such as Telegram, Blynk etc. Node MCU unit is shown in Fig. 6.

5 Software

5.1 Arduino IDE

Arduino IDE is an open souce arduino programming software system. We can write the working program and conditions for the arduino by using Embedded C or C++

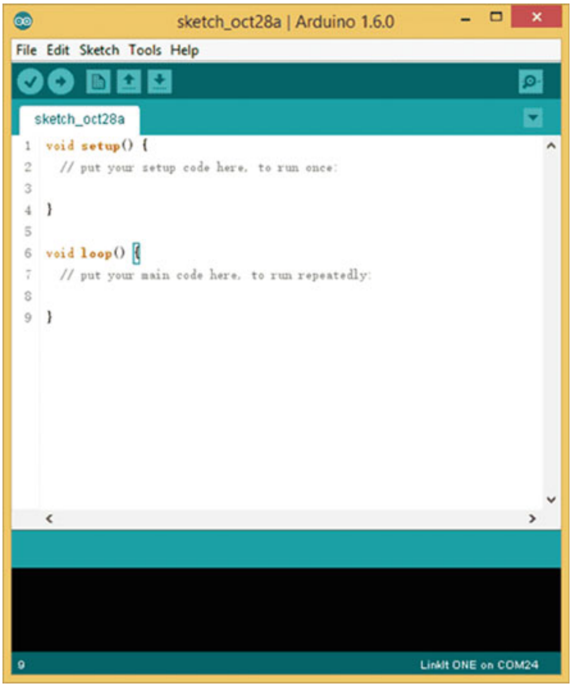


Fig. 7 Arduino IDE

coding language. This software is also used to burn the code into arduino‘s microcontroller and this alsoenables us to edit the program as many time as we want making it mire convinient. Ardunio software is have plenty inbuilt programs which gives more optio to code the microcontroller. The interface of the Arduino IDE software is shown in Fig. 7.

6 Result

When level of alcohol consumption by the driver exceeds the permissible limit an immediate buzzer alert is given to the driver and nearby people. The speed of the vehicle will be reduced and with a specified amount of time delay the ignition of vehicle is turned off. A distress message will be sent to registered number along with vehicle‘s location. An LCD display indicates the status of the alcohol detecting unit and shows a message “Alcohol Detected”. Figure 8 represents our prototype model for our proposed project and methodology. The distress message indicating the status of the driver is communicated through the message as shown in Fig. 9. It also shares the location of vehicle through a GPS link where one can access the location by clicking on it.

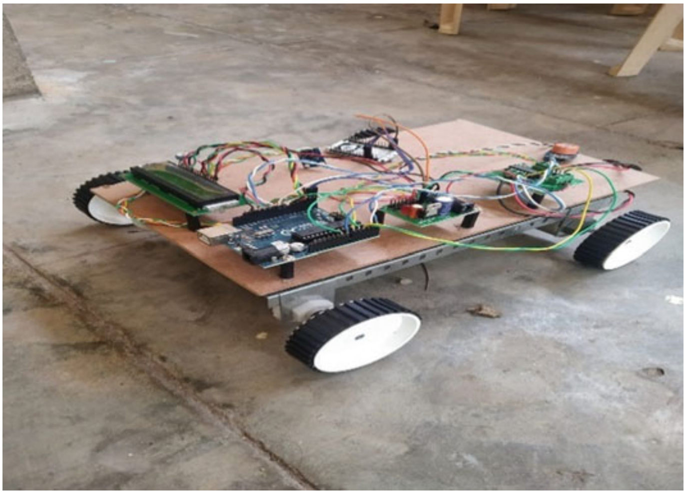
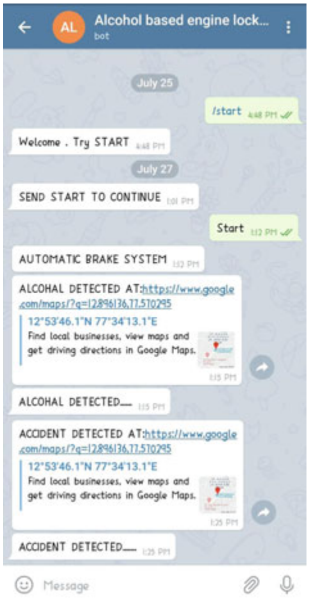


Fig. 8 Model

Fig. 9 Message panel



7 Conclusion

In our Project we have designed an alcohol detecting system which compares the level of driver's alcohol consumption to the permissible limit, if the permissible limit is crossed the speed and ignition of the vehicle will we regulated and turned off.

Our system is enabled with several indicators to alert nearby people and respective authorities. The Geo location of the vehicle will be tracked and sent through wireless communication using the IoT. This system prevents the driver from accessing the vehicle after being drunk and this helps in reducing the accidents cause by drunken driving. Our system increases the safety measure to both the driver and people making the road safety more safer than before. When it is equipped and implemented in vehicle in future it reduces the cases of accidents and saves many lives.

References

1. Sahu P, Dixit S, Mishra S, Srivastava S (2016) Alcohol detection based engine locking system using MQ-3 sensor, B. Tech students, Department of Electronics and Communication, IMS Engineering College, Ghaziabad-201009, Uttar Pradesh
2. Dai GJ, Teng J, Bai X, Shen ZH, Xuan D Mobile phone based drunk driving detection. The Ohio State University Columbus, Ohio, USA
3. Dhivya M, Kathiravan S (Feb 2015) Driver authentication and accident avoidance system for vehicles. [Smart Computing Review, vol. 5, February 2015]. Dept. of ECE, Kalaignar Karunanidhi Institute of Technology
4. James N, Aparna C, John TP (Jan 2014) Alcohol detection system. Int J Res Comput Commun Technol 3(1)
5. Prashanth KP, Padiyar K, Naveen Kumar PH, Santhosh Kumar K (Oct 2014) Road accident avoiding system using drunken sensing technique. [International Journal of Engineering Research & Technology (IJERT) ISSN: 2278–0181, IJERTV3IS100754, 3(10) October 2014]. Dept of Mechanical Engineering, East-West Institute of Technology, Bangalore, India

Transformerless Bidirectional Converter Fed Hybrid Power System

Prasanth Venkatareddy, Nagendra Prasad, and Ramesh Kumar

Abstract There is a very huge demand for the electricity hence the generation is also forced to be increased to meet the demand but the generation due to the fossil fuel will lead to pollution and global warming in the atmosphere hence a solution which is an alternate and a clean source is proposed in this paper. A new transformer less bidirectional buck boost converter with SEPIC and MPPT connected to grid is proposed It consist of a simple generic architecture consisting of low components, and experiences low voltage stress on the transistor and has a wide voltage gain. The voltage variation is done by buck boost converter. It can either boost the voltage generated by the PV to be greater than the grid voltage or buck depending on the demand to maximize more power from PV the design includes MPPT. Battery is employed for uninterruptible power supply, synchronous rectification to increase the efficiency. when the model is in continuous conduction mode, study of stable operation, model with moderate signal, component propose, and converter efficiency is obtained.

Keywords BBB · Converters · SEPIC · MPPT · Renewable energy · ESU · Grid

1 Introduction

In modern days there is large scarcity for electric power. This has forced to increase the generation of electricity. But increase in generation cannot be done with any non-renewable energy resource due to the main disadvantage such as energy exhausting, pollution, and global warming. An alternate method is to generate the energy using renewable energy resource. By integrating the renewable energy with grid more amount of energy can be generated [1]. Also, by connecting a battery in parallel excess of energy can be stored. A solar has a photo voltaic component to generate

P. Venkatareddy (✉)
Nitte Meenakshi Institute of Technology, Yelahanka, Bangalore, India
e-mail: prashv143@dbit.co.in

N. Prasad · R. Kumar
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_107

energy. Along with the solar circuit an MPPT and SEPIC is connected to increase the efficiency. Renewable energy source has a very large demand which can produce the electrical energy, but the main concern is the environment, to which pollution hazard must be prevented. Hence the solution is the development of micro grid integrating with renewable energy [2]. But when the generation is connected with the renewable source there is lot of interruption in the power grid hence the solution lies with connecting a energy storage unit to remove the fluctuation. It helps in maintaining the balance among consumption and power generation.

1.1 Photovoltaic

Photovoltaic or PV is a renewable energy converter that use sunlight as an energy source into electrical energy [2]. PV consists of a series of two or more layers of semiconductors. The principle is related with the photoelectric principle. Sunlight contain energy form photons. When these photons on the surface of the PV, the excess electrons from the n-type semiconductor will move toward the hole on the p-type semiconductor. Because of the flow of electrons and holes is then formed an electric field. PV module is a collection of several PV cells. This research uses Solaria module MSX-60 [3]. Module consists 36 cells that can generate a maximum power of 60W with light intensity conditions ($I_{radiance}$) of 1000W/m^2 and 25 °C in temperature.

Characteristics of power which can be generated versus the magnitude of light intensity on the PV surface as in Figs. 1 and 2. The array is a connection of PV modules that connected to one another both serial and parallel. Serial installation will cause voltage to increase according to the number of serially connected modules, while parallel installation causes the increase of the resulted current.

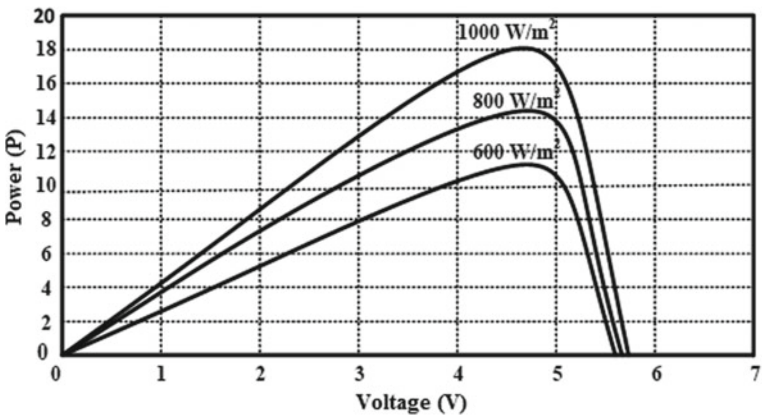


Fig. 1 Characteristics for different variation due to power and temperature

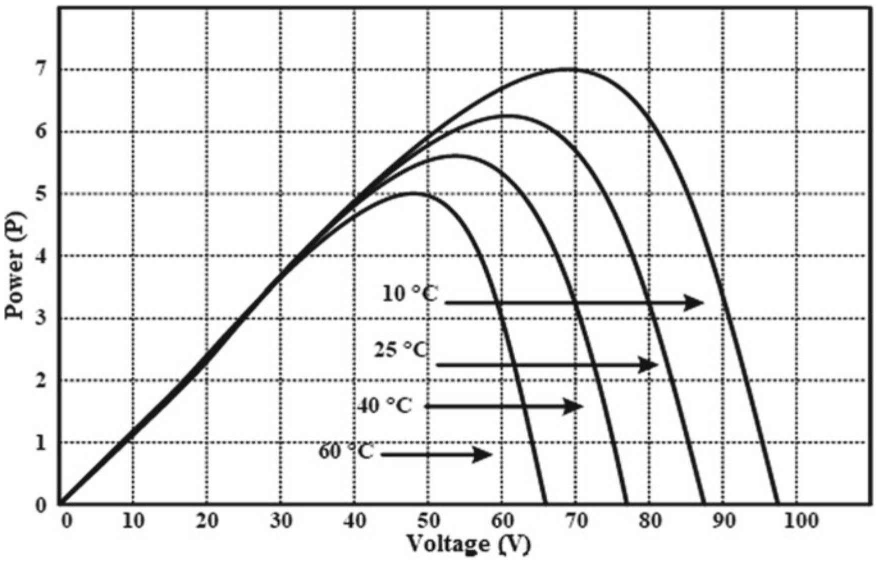
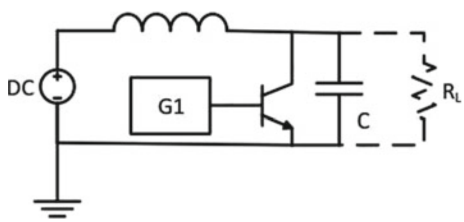


Fig. 2 Characteristics for different variation due to power and temperature

1.2 BoostConverter

Boost converter is one of DC–DC converter that used to increase DC voltage as shown in Fig. 3. The output voltage of boost converter is always greater than the input voltage. The topology of boost converter generally consists of inductor, diode, and switch. Inductor and diode on boost converter are serially connected towards the voltage source, while the switch is parallelly connected. The inductor is used as a current and voltage source to limit the ripple of output current. The Capacitor on boost converter is used to limit the ripple of the output voltage. The value of the output voltage of boost converter is determined by setting duty cycle value.

Fig. 3 Boost converter



1.3 Transformer Less Inverter with GridConnected

Inverter is an electronics circuit which converts DC input voltage to AC output voltage with the controllable amplitude and frequency. The output voltage can be obtained by control the switching of the inverter. The output voltage is ideally sinusoidal form and contains harmonics [4]. Transformer less inverter is a type of bridge inverter which can result in sinusoidal voltage. This inverter is designed to directly connected to grid without transformer as a linker. The result of the transformer less inverter is better than inverter which still uses a transformer to connect to the grid. Power flow process in the system connected to grid needs an inductor in the transmission process.

1.4 Maximum Power PointTracker

Perturb method and Observe (P & O) is one of methods MPPT [5]. P & O method is easy to use and costs less than the existing methods. This method has a simple feedback and multiple parameter measurements. This method operates in stages (such as addition or subtraction). When there is a change on duty cycle, it compares the output power of the PV with the increase in the previous cycle. If the disruption tends to show an increase (or decrease) of power array, the next disruption will be formed in the same direction (or the opposite). In this condition, the peak power tracker will seek continuous peak power conditions.

2 Transformer Less Boost Inverter System

The main components used in the system consist of solar modules (PV), boost converters, transformer less inverter, inductors, and controller. In general, a PV absorbs sunlight and generates voltage. The output of a PV is direct current. The output voltage of the PV is then stepped-up by the boost converter. A PI controller controls the converter output at a specified reference voltage. This allows the PV to work at a lower voltage level than the fixed network without causing over- modulation on the inverter. The output of the converter will be stored in two capacitors in turns based on the frequency of the network. Two capacitors output the voltage to the network through an inverter to get the amount of single-phase AC voltage. Inverter will adjust the output voltage through an inverter modulation index obtained from the MPPT Perturb and Observe process (P & O). The inverter output voltage phase.

2.1 Phase 1 Structure and Operating

Synchronous generator is an electromechanical machine which works on the principle of Faraday's law of electromagnetic induction. It generates three phase AC current from stator winding. The stator consists of armature winding which are electrically displaced by 120 degree each. An AC current generated by the stator is the output of the SG. The rotor which rotates is surrounded by the magnetic field. The field is either produced by a permanent magnet or by magnet wound on the field coil.

Wind turbine has certain drawbacks such as constant monitoring, removal of carbon filth etc. Hence alternate method suggested is brushless excitation by using permanent magnets instead of electromagnets. Electromagnet being replaced by stable magnet, is placed on external part of rotor hence the field excited by this. Also, when compared to electromagnet the excitation is easy by using permanent magnet. The permanent magnet efficiency is increased by reducing the gap between stator and rotor. The permanent magnet has an advantage of more no of poles and its durability. The disadvantage is that there is no control of the rotor flux hence efficiency is attained only at pre-defined wind speed. Synchronous machine has very a smaller number of poles compared to other machines, but these poles rotates at very high speed. The wind speed determines the rotor speed. The speed control variation is done by the gear box, also it maintains generator speed constant when speed of turbine changes. This action is effectively done by matching the generator speed to variable speed of blades. The major disadvantage comes when maintenance must be done also when wind turbine weight increases then generator noise also increases, power losses is high and system efficiency is lost completely as the gear box cannot be driven easily.

2.2 Full Wave Three Phase Rectification

The rectification of three phase AC supply circuit is constructed by combining both half-wave rectifier circuits which consist of six diodes.The supply to the rectifier circuit can be given from three phase three wire delta connected system. At the output there is a filter which results in low ripple (Fig. 4).

In circuit each phase is connected between pair of diodes. The load is connected between positive and negative. Working of circuit are in three phases. Diodes one, two, three and four are between phase first and second phases. Diodes third, fourth, fifth and sixth diodes are between second and third phases. Diodes fifth, sixth, first and second are between third and first phases. Diodes one, three and five are connected to anode terminal and diodes two, four and six are connected to negative terminal. Hence the diodes conduct that conduct in pairs are $D_{1\text{-}2}$, $D_{1\text{-}6}$, $D_{3\text{-}6}$, $D_{3\text{-}6}$, $D_{3\text{-}4}$, $D_{5\text{-}4}$, $D_{5\text{-}2}$ and $D_{1\text{-}2}$ to give a sinusoidal path for the load current .

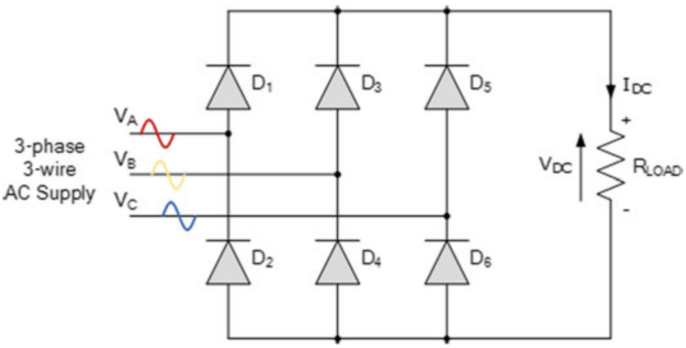


Fig. 4 Rectifier circuit

2.3 Full-Wave Three-Phase Rectifier Waveform

The circuit is fed by a three-phase transformer secondary. The connection from the transformer can either be star or delta. Diodes conduct in pair with a phase difference of 120 degree each. In every cycle pair of diodes conduct for part of cycle every time (Fig. 5).

The V $_{DC}$ avg value from output voltage waveform is: $V_{DC} = \frac{3\sqrt{3}}{\pi} V$ s, $= 1.65 *$ V_S.

Where: $V_S = V_{L\,(PEAK)7***} \div \sqrt{3}$

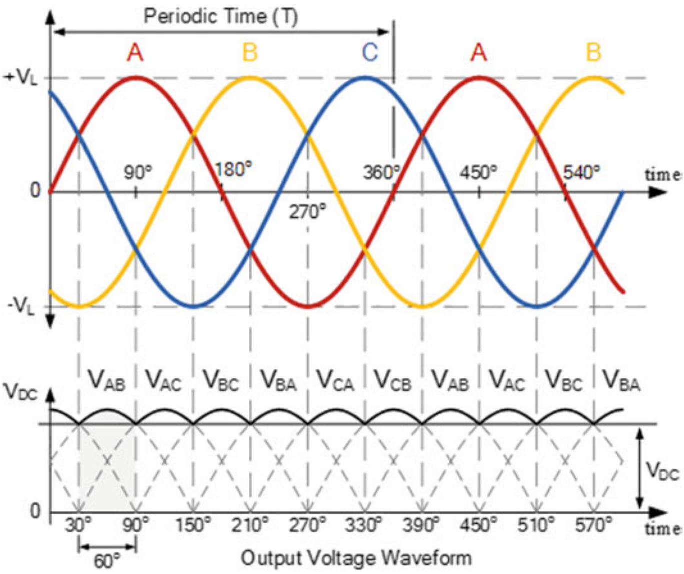


Fig. 5 Rectifier output

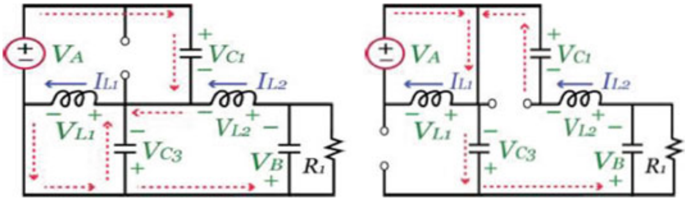


Fig. 6 Circuit of power flow from A to B

$V_{L\ (PEAK)}$ is the maximum line-to-line voltage (V_L*1.414).

3 Phase II Structure and Operation

3.1 A to B Power Flow Operation

In the circuit voltage source is connected to port A and resistance R1 is connected as load at port B. Switch Q1 is the main transistor. Turn ON and OFF state is represented as 0 and 1. Q2 and Q3 are synchronous rectifiers. According to the waveform of proposed converter when G1 is turned ON G2 and G3 are in off position hence no signal. This period between the raise of G1 and off of G2 and G3 is called as dead time. Similarly, ON of G2 and G3, and OFF of G1 is another dead time. During this dead period there is a flow of current in the diodes connected across Q2 and Q3. Voltage passing through D and S of transistor will become nil. This provides zero-voltage switching (ZVS) during turn off Q2 and Q3. Due to this converter efficiency is increased (Fig. 6).

3.2 State 0, G1, G2 and G3 is 100

In this state G1 is turned on and G2 and G3 is turned off. This period is represented by product of duty cycle periodic switching time. d_1T. Above Figure shows the flow of current path. By applying Kirchhoff voltage and current rule the following voltage and current equations can be derived.

$$V_{L1} = V_A \tag{1}$$

$$V_{L2} = V_{C1} - V_B + V_{C3} \tag{2}$$

$$\mathrm{I}_{\mathrm{C1}} = \mathrm{I}_{\mathrm{C3}} = -\mathrm{I}_{\mathrm{L2}} \tag{3}$$

$$\mathrm{I}_{\mathrm{C4}} = \mathrm{IL2} - \frac{VB}{R1} \tag{4}$$

3.3 State 1, G1, G2 and G3 is 011

In this state G1 is turned OFF and G2 and G3 is turned ON. The current direction is reversed. This period is represented as d_2T and d_3T here d2 and d3 are duty cycle of Q2 and Q3. Based on current flow the voltage and current equations are as follows.

$$\mathrm{V}_{\mathrm{L1}} = -\mathrm{V}_{\mathrm{C3}} = \mathrm{V}_{\mathrm{A}} - \mathrm{V}_{\mathrm{CI}} \tag{5}$$

$$\mathrm{V}_{\mathrm{L1}} = -\mathrm{V}_{\mathrm{B}} + \mathrm{V}_{\mathrm{C3}} \tag{6}$$

$$\mathrm{I}_{\mathrm{C1}} = \mathrm{I}_{\mathrm{C3}} = \frac{IL1 - IL2}{2} \tag{7}$$

$$\mathrm{I}_{\mathrm{C4}} = \mathrm{I}_{\mathrm{L2}} - \frac{VB}{R1} \quad \mathrm{I}_{\mathrm{C4}} = \mathrm{I}_{\mathrm{L2}} \tag{8}$$

During this stage, the equations for voltage across capacitor and current across inductor is as follows

$$\mathrm{V}_{\mathrm{C1}} = \frac{VA}{1-d1}, \mathrm{V}_{\mathrm{C3}} = \frac{d1}{1-d1}\mathrm{V}_{\mathrm{A}}, \mathrm{V}_{\mathrm{B}} = \frac{2d1}{1-d1}\mathrm{V}_{\mathrm{A}}.$$

Hence voltage gain can be written as

$$\mathrm{M}_{\mathrm{AB}} = \frac{VB}{VA} = \frac{2d1}{1-d1}, \mathrm{I}_{\mathrm{L1}} = \frac{1+d1}{1-d1}\mathrm{I}_{\mathrm{L2}}.$$

$$\mathrm{I}_{\mathrm{L2}} = \frac{VB}{R1} = \frac{2d1}{(1-d1)R1}\mathrm{V}_{\mathrm{A}}.$$

3.4 B to A Power Flow Operation

Circuit represents voltage source V_{B} is connected to port B and load is connected to port A which is a resistor (Fig. 7).

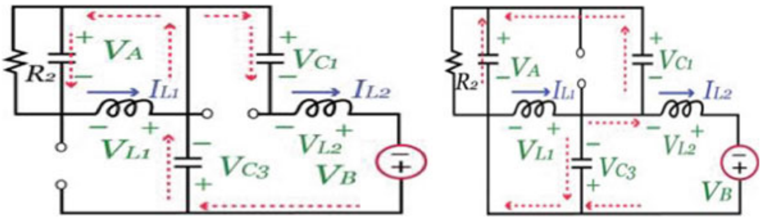


Fig. 7 Circuit of power flow from A to B

Here Q1 and Q2 are switches and Q1 is the rectifier when G1 is turned OFF G2 and G3 are in ON position hence receives gate signal. This period between the raise of G2 and G3 and off G1 is called as dead time. Similarly, ON of G2 and G3, and OFF G1 is another dead time. Hence zero-voltage switching (ZVS) is provided during turn off and on of Q1. The capacitor voltage, inductor current and voltage gain of the circuit is as shown below.

$$\mathrm{M_{AB}} = \frac{VA}{VB} = \frac{d2}{2(1-d2)}\mathrm{V_{C1}} = \frac{VB}{2(1-d2)} \quad \mathrm{V_{C3}} = \frac{VB}{2}$$

$$\mathrm{V_A} = \frac{d2}{2(1-d2)}\mathrm{V_B I_{L1}} = \frac{2-d2}{d2}\mathrm{I_{L2}}.$$

$$\mathrm{I_{L2}} = \frac{VA}{R2}\mathrm{MAB} = \left(\frac{d2}{2(1-d2)}\right)\frac{VB}{R2}$$

4 Result and Analysis

All the three system are integrated and Simulated. In this work there are three phases which integrates three systems. PV of solar forms the phase one. Here PV panel is integrated with the SEPIC converter along with MPPT. The signal received from PV panel is given to MPPT. This MPPT produces the PWM signal. This pulse produces the triggering pulse for the MOSFET which is present in the SEPIC converter.

The SEPIC converter gets the pulse signal from MPPT and triggers the circuit. The SEPIC converter boosts the voltage and feeds to the inverter to convert from pulsating DC to Stabilizes DC. In phase two wind energy is integrated. The wind turbine is connected to PMSG. The output of PMSG is AC later fed to the inverter to convert it DC. The last phase is the ESU also called as battery. The output is connected to BDC whose operation is either buck or boost. The output voltage is again connected to converter to get a stabilized DC. All the output is integrated and connected to grid (Fig. 8).

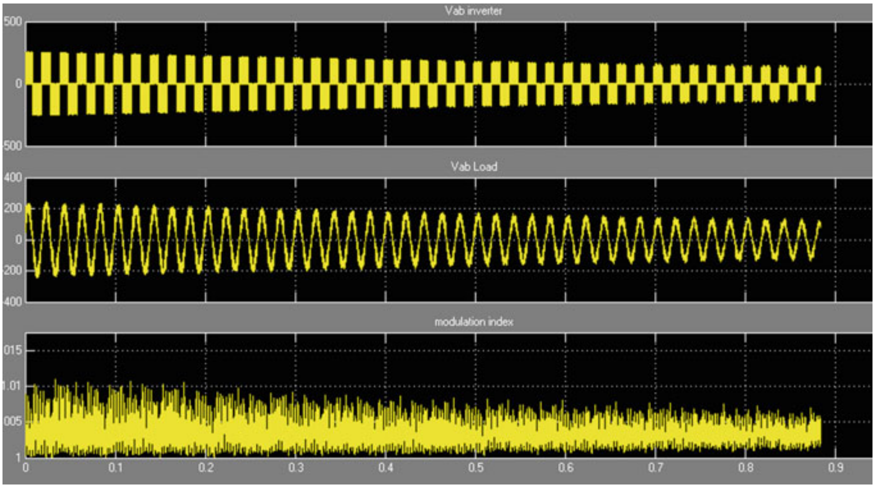


Fig. 8 Result of hybrid system

5 Conclusion and Future Scope of Work

This work concludes with an uncomplicated circuit and simple architecture of three integrated power sources and is converted into DC for storage purpose. All the three sources are integrated to load with wide voltage gain. A new BBB converter which has flow of power in each direction is shown. Stress on switches is also comparatively low. The BBB has a dynamic performance. Efficiency is improved due to synchronous rectification. Hence it can be concluded that the proposed project has a good advantage for applications such as bidirectional power flow, wide voltage gain and good energy system. In future there are many other natural renewable energy sources like water which can be combined with the hybrid energy sources and can be analyzed with Mat-lab for future scope of work.

References

1. Thirugnanam K, Kerk SK, Yuen C, Liu N, Zhang M (Aug. 2018) Energy management for renewable microgrid in reducing diesel generators usage with multiple types of battery. IEEE Trans Ind Electron 65(8):6772–6786
2. Dasgupta S, Mohan SN, Sahoo SK, Panda SK (Mar. 2013) Application of four-switch-based three-phase grid-connected inverter to connect renewable energy source to a generalized unbalanced microgrid system. IEEE Trans Ind Electron 60(3):1204–1215
3. Refaat SS, Abu-Rub H, Sanfilippo AP, Mohamed A (Feb. 2018) Impact of grid-tied large-scale photovoltaic system on dynamic voltage stability of electric power grids. IET Ren Pow Gen 12(2):157–164

4. Bragard M, Soltau N, Thomas S, Doncker RW (Dec. 2010) The balance of renewable sources and user demands in grids: power electronics for modular battery energy storage systems. IEEE Trans Pow Electron 25(12):3049–3056
5. Pegueroles-Queralt J, Bianchi F, Gomis-Bellmunt O (Jan. 2015) A power smoothing system based on supercapacitors for renewable distributed generation. IEEE Trans Ind Electron 62(1):343–350

Monitoring and Controlling of Domestic Appliance Using IoT

P. S. Rajath Shankar, Sunil Davis, P. J. Deepika, Megha Sandesh, and Dinesha

Abstract IoT (Internet of Things) is used to send data over network without human interference. IoT has made automation easier. In the designed system we have used the concept of IoT to create an automated monitoring and controlling of domestic appliance. The purpose of the system design presented in the paper is, to create awareness on current usage. In this generation people are exposed to many types of electronic appliances. The people need to be aware of how much current is needed by each appliance they own. The designed system concentrates on letting people to know how much current is consumed by each appliance. The proposed system monitors the power passing through each appliance and shuts it off on high power because when a low power device meets a high power a short circuit occurs. This helps people to understand and reduce the use of power.

Keywords IoT · Node MCU · Arduino uno · Relay switch · Voltage sensor · Current sensor

1 Introduction

IoT (Internet of Things) [8–10] is used to send data over network without human interference. IoT has made automation easier. In the designed system we have used the concept of IoT to create an automated monitoring and controlling of domestic appliance [1, 2]. This application we have used IoT to create awareness about domestic appliance among the user. IoT makes the designed system easy to access at any place or time. People use many appliances without the knowledge of power required and being consumed by that appliance. The power supplied to the appliance is sometimes more than the required quantity which may lead to short circuit.

Short circuit is an electrical circuit that allows very low or no current to flow through an unintended path, this result in excessive current flow through the circuit. Such condition is harmful for the appliance that has consequences such as burning of

P. S. R. Shankar (✉) · S. Davis · P. J. Deepika · M. Sandesh · Dinesha
Department of Electrical & Electronics Engineering, Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_108

wires, internal breakdown and in worst cases causes fire accidents. In this paper we have come up with an idea that could reduce issues of short circuiting. The proposed system is an IOT model that learns the current and voltage flowing in a circuit and if the rate of power is higher than required it shuts the appliance off. The proposed system contains a circuit of voltage and current sensors to calculate voltage and current passing through the device. With the help of this the power is calculated. If there is a fluctuation in the power then the devices are shut down. This functionality prevents short circuiting. Along with this the other functionality implemented is that the user can monitor the power flow and can observe the rate of power that affects the appliance. This will help the user to reduce the current flow through the circuit

2 System Design

The aim of this project is to make the user aware of power consumption by each appliance and to reduce short circuit. This is implemented with the following physical components:

(A) Relay Switch
(B) Voltage Sensor
(C) Current Sensor
(D) Arduino UNO
(E) Node MCU.

A. **Relay Switch**: The relay switch acts as a basic on and off switch. The relay switch is grounded to the switch and the appliance. Once the user wants to turn on the circuit to check the power flow, he can turn it on using the mobile device. It is controlled by the voltage passing through it Fig. 1.

B. **Voltage Sensor**: The voltage sensor is a hardware component that monitors and calculates the amount of voltage supply. It determines whether the voltage supply is AC or DC voltage. The input to this sensor is voltage whereas the output is in analog signals. The measurements can vary with the voltage divider. There are two pins the positive and negative pin to which the device can be

Fig. 1 Relay switch

Fig. 2 Voltage sensor

connected. The usage in our circuit is to measure and calculate the voltage flow through the appliance Fig. 2.

C. **Current Sensor**: The current sensor measures the current passing through the wire. The current is measured using magnetic flux. Due to the flowing AC current there is a magnetic flux induced in the coils. As an opposite reaction to this flux a secondary magnetic flux is induced in the secondary coil to cancel out the primary magnetic flux. A proportional AC current is induced to this flux. The secondary current flows through the shunt resistor causing a voltage difference at its both ends. The voltage difference is proportional to the current flowing through measurement conductors Fig. 3.

D. **Arduino UNO**: The Arduino is used to constructing and programming in electronics. It consists of a microcontroller, analog and digital pins, reset button, USB port, external power supply and In-Circuit programmer in general [4].

- Microcontroller- is to send and receive signals, also command that needs to be sent. All the other components of the Arduino support the microcontroller
- Analog and Digital pins- The analog pins vary from A0 to A5 and are often used to read analog sensors. The digital pins are labeled from 0 to 13 which can act as input or output.
- Reset Button- this button resets the Arduino, the current program running will be executed from beginning.

Fig. 3 Current sensor

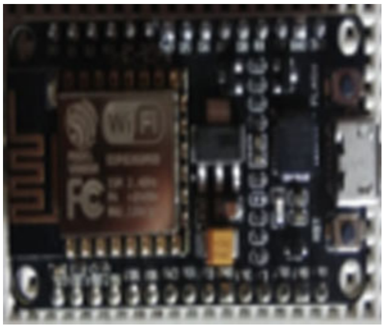
Fig. 4 Arduino UNO

- USB Port- this port is used to upload a program on the microcontroller. It usually works at 5v supply.
- External power supply- This is only used to power the board and has a regulated voltage of 9 to 12 V, mostly if the USB plug does not provide sufficient power for the program.
- In-Circuit Programmer- it is another source to upload program done by Tx = 1 output and Rx = 0 input. Here T refers to transmit and R refers to receive Fig. 4.

E. Node MCU:

A Node MCU is a combination of a node and microcontroller unit. It is an open platform for edit/modify. It is permanent software with read only memory and is useful in prototyping. The ESP8266 is a low-cost Wi-Fi chip developed by Espress if Systems with TCP/IP protocol. Node MCU is similar to an Arduino with ESP8266 [3] as its major component. It has Programmable pins and built in Wi-Fi [5]. It can get power through micro-USB port. Its cost is low and can be programmed through multiple programming environments Fig. 5.

Fig. 5 Node MC



3 Implementation

- **Circuit Design**

The circuit design is done using components Arduino UNO, Relay switch, Node MCU, resistors, capacitors, voltage sensors, current sensors and breadboard. The connections are made by connecting these components. The relay is grounded to the switch socket. The voltage and current sensors are grounded to relay switch, a shunt resistor of 30Ω is used to measure current and a capacitor of 10microF is used to make the current flow stabilized. The output pins of the voltage and current sensor are given as analog input to the Arduino. The Arduino helps in conversion of two analog inputs to one digital output Fig. 6.

The Arduino acts as the Slave in the "Master Slave design". This Master Slave circuit is done by connecting the digital output of Arduino to Node MCU. The Node MCU is the Master of this circuit. The Arduino stores digital power signals which is fetched to the Node MCU on command from the Node MCU. The signals received by the Node MCU is stored in data base and delivered to user device Fig. 7.

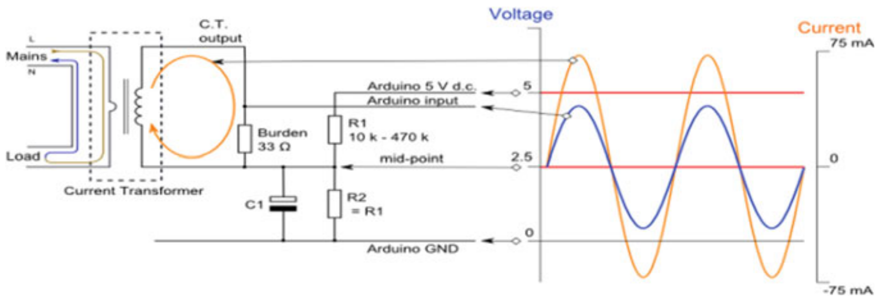


Fig. 6 Circuit diagram

Fig. 7 Implemented circuit

- **Application Design**

The programming is done in C language. The programming consists of code to turn on/off the circuit, shutting the appliance in case of high power and storage database. The system is programmed in such a way that once the user turns on the system, the turn on request is sent through the server to the circuit. The request signal is processed by the few components of circuit to turn it on. Once the circuit is on, the reading of current and voltage are used to calculate the power. The power is a digital output; this is stored in Arduino and fetched to Node MCU. Signals in the Node MCU are externally stored in a private database. The digital signals are converted to understandable readings of power and transferred to user device through sever. The main feature of this project is that the circuit shuts down the appliance when there is a high power being delivered. This is done by code snippet where the power flow is given its highest limit for a particular appliance. When the power through the device exceeds this limit the program sends the circuit signal to break the circuit. This shuts down the appliance. As the programming hardware used are IoT [6–8] modules that is Arduino and Node MCU it allows to turn off the circuit at/in any time or place. This makes the designed system compatible and convenient.

4 Working

The circuit is turned on by the user using his mobile application. The users send a request to turn on the circuit from its server. This request is processed by the server and the command is sent to the circuit to turn it on. The on receiving the signal relay switch turns the circuit on. As the voltage and current sensors are connected to the relay, when the circuit is turned on the sensors begin to measure and calculate voltage and current flow. The voltage sensor measures the voltage supply through it. Also differentiates AC and DC supply, the output is in analog signals. Similarly, the current sensor measures the current flow using magnetic flux. The output of current sensors is also analog signal. The analog outputs are fetched as input to the Arduino. The Arduino helps to convert the two analog inputs into one digital output. As there is a Master Slave connection implemented between the Arduino and Node MCU, the signals are fetched to the Node MCU only on request. The Node MCU is programmed such that it sends a request after every 10 s of delay. On receiving the request, the Arduino sends the signals to the Node MCU. The Node MCU stores these signals in the database for easy access. The signals are digital making it understandable. The user can view the power being consumed by the appliance. The power flowing through the circuit is set to a certain limit. When the power through the appliance exceeds this limit, the circuit shuts down the appliance preventing short circuit. This limit varies from appliance to appliance Fig. 8.

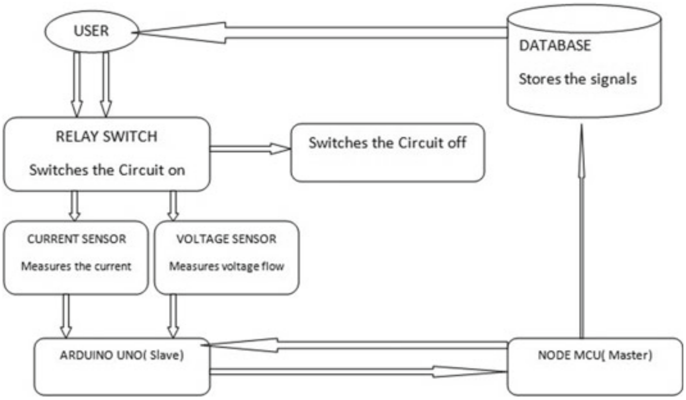


Fig. 8 System model

5 Result

The main objective of the proposed scheme is to bring awareness among the users about the power consumption. This is achieved using the system design. The user could turn on or off the circuit at any point and any place. Once the circuit is on it automatically reads the current and voltage through the appliance and calculates power. This power as set a limit to 120 V when exceeds or experiences high fluctuation shuts the appliance, which prevents short circuit. Else the signals are recorded and stored in the database for easy access. These signals are digital which makes it accessible. With the mobile application the user can monitor the power the flow Fig. 9.

Appliance	ON	OFF	Current	Voltage	Power consumed
Bulb	24.07.2020 12:04:59 AM	24.07.2020 12:09:50 AM	0.07 mA	217.23 V	1.10244225WH
Fan	24.08.2020 1:04:59 PM	24.08.2020 1:50:50 PM	2A	210 V	350WH
Refrigerator	24.08.2020 12:00:59 PM	24.08.2020 2:00:50 PM	10A	220V	4400WH

Fig. 9 Time usage of power consumption

6 Conclusion

The design is efficient as it is designed to a mobile application, which makes it easy and convenient to access at any place and any time. If the user forgets to turn the circuit off, he could turn it off from any place at any time. This makes is more accessible, convenient and reliable system. The circuit shuts down the appliance when there is higher power flowing through it than the specified limit. This prevents short circuit which in turn avoid the major accidents caused due to short circuit. As the user is able to view the power flow it creates awareness on usage of the appliance. People can choose the appliance suitable for the power flow and reduce usage of inappropriate power appliances. Hence the designed is efficient, reliable, convenient and manageable.

References

1. Dwivedi AK (ME Student) Department of Electrical Engineering NITTTR, Chandigarh Mrs. Shimi S. L (Assistant Professor) Department of Electrical Engineering NITTTR, Chandigarh" Home Automation and Energy Management using Android App" Int J Eng Res Technol (IJERT)
2. Azeez A, Sivanagaraju N, Amilesh M Assistant Professors, ECE Dept., VKR, VNB & AGK College of Engineering, Gudivada, A.P, India "GSM based home appliances automation and energy auditing"
3. Parihar YS Scientist D and district informatics officer national informatics centre, Mahoba(U.P.), India "Internet of things and node mcu a review of use of node MCU ESP8266 in IoT products"
4. Badamasi YA Nigerian Turkish Nile University Abuja, Nigeria "The working principle of an arduino"
5. Nair AG, Alias E, John K, Mathew MP, Rini Varghese P, Joy JS "Remote controlled home automation using android application via wifi connectivity"
6. Dey S, Roy A, Das S (2016) Home automation using internet of thing. IRJET 2(3):1965–1970
7. Piyare R, Lee SR (2013) Smart home-control and monitoring system using smart phone, The 1st International Conference on Convergence and its Application, 84:83–86
8. Gunge VS, Yalagi PS (2016) "Smart home automation: a literature review", National Seminar on Recent Trends in Data Mining- RTDM
9. Gaikwad PP, Gabhane JP, Golait SS (2015) "A survey based on Smart Homes system using internet-of things", Computation of Power Energy Information and Commuincation (ICCPEIC) 2015 International Conference on, pp 0330–0335
10. Asadullah M, Raza A (2016) 'An overview of home automation system'

Smart Garbage Separation Using Robotic Arm with Image Processing

R. Raveendra, N. Shwetha, S. Sahana, M. Veeresh, and P. Bharath Kumar

Abstract Waste production in India is around 36.5 million tons i.e., 1 lakh tons per day approx. This is a huge amount of waste being produced. Waste contains medical, household and industrial wastes etc. Though there are many measures taken to reduce the production of waste it is harder to reduce production of waste in such a vast country with large population. Hence many other measures are taken to reduce waste that is being produced. Key to efficient waste management is segregation. Segregation till today is being done manually. The humans involved in this process face lot of issues and infected by various diseases. To help segregation of waste and humans being affected, in the proposed system we have come up with an automated system to segregate waste. This method can efficient and reduce risk of human lives.

Keywords Efficient waste management · Segregation · Automated system · Reduce risk of human lives

1 Introduction

India is a vast country with huge population, in proportion with the population the waste being produced is also in large quantities. Almost 36.5 million tons of waste is produced per annum which is approximately 1 lakh tons every day. There are many waste production control measures been taken, still it is hard to reduce the waste being produced when the population is large. Control measures are taken such as compost, reuse and recycle. Key of waste management is segregation before dumping [1].

Segregation is process of separating wet and dry waste so they can be recycled or reused. Till day segregation is done manually. Waste is essentially segregated as wet and dry waste. Such that wet waste can be composted and dry wastes can be either reused or recycled. In a recent study done in Kerala, India it is observed that these workers are more likely to get eye diseases, respiratory diseases and dermatological

R. Raveendra (✉) · N. Shwetha · S. Sahana · M. Veeresh · P. Bharath Kumar
Department of Electrical & Electronics Engineering, Don Bosco Institute of Technology, Banglore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_109

problems and nail infections [2]. Apart from being infected by diseases, these workers are physically hurt during segregation [3]. It has always been a threat to human life.

In the proposed system, we have come up with an automated method to segregate waste. In this method we have ensured that the waste segregation does not need any human intervention [4]. We have involved a robotic arm to pick the wastes and conveyor to segregate them. The waste is segregated as wet and dry waste into two different bins [5]. These have sonars with which once the bin is filled the worker is sent a message to empty the bin so that he would not check on it regularly.

2 Types of Waste and Problems Caused in Segregating Them

Household waste: Household waste contributes a big portion of total waste produced. These wastes contain contaminated food, which attracts vermin, also they can cause parasite, lung and skin infections during the disposal or segregation. Household waste also contains plastics, few medical wastes and sanitary wastes. These wastes also cause various skin disorders like psoriasis [6].

Medical waste: Waste that needs a proper treatment and disposal. Medical waste can be dangerous when they are disposed without any treatment. These wastes can cause HIV, pneumonia, tuberculosis, Hepatitis B and C etc. Few of the diseases are not curable and for others medications are not affordable by all. Though medical wastes are taken extra care and treated before disposal, there can be few cases where they have not been treated which is still dangerous.

Industrial waste: Toxins are a part of industries because of chemicals used in production units. These toxins cause immune suppression, acute poisoning and reproductive failure. Also few infectious diseases like typhoid other like diarrhea, gastroenteritis and kidney problems may occur.

3 Implementation

Proposed system is implemented using IoT and Artificial Intelligence. IoT helps the information to be passed over network without human intervention and AI is used to segregate waste with images of the waste [7]. To accomplish this we have used the hardware components as follows:

a. **Arduino**: It is used to constructing and programming in electronics. It consists of a microcontroller, analog and digital pins, reset button, USB port, external power supply and In-Circuit programmer in general.

i Microcontroller—is to send and receive signals, also command that needs to be sent. All the other components of the Arduino support the microcontroller.

ii Analog and Digital pins—Analog pins vary from A0 to A5 and are often used to read analog sensors. The digital pins are labeled from 0 to 13, which can act as input or output.

iii Reset Button—This button resets the Arduino, the current program running will be executed from beginning.

iv USB Port—This port is used to upload program on the micro-controller. It usually works at 5v supply.

v External Power Supply—Used to power the board with regulated voltage of 9–12 V, mostly if sufficient power cannot be given by USB plug for the program.

vi In Circuit Programmer—uploading program by another source is done by Tx = 1 output and Rx = 0 input. Here T refers to transmit and R refers to receive.

Arduino is as shown in Fig. 1 is used to control the robotic arm, to pick the objects and drop on conveyor belt.

b. **Servo motor**: A servo motor as shown in Fig 2 is a rotary actuator or a motor that consents for a accurate control in terms of the angular position. acceleration, and velocity. There are many types of servomotor such as AC, DC, brushless, positional rotation, continuous rotation and linear servomotor. It basically has three wires ground, power and control. We user servo motor to move the object towards the bin it belongs to [5].

c. **Robotic arm:** The arm functions similar to human arm and is programmable. It is linked such that there can be linear, transitional or rotational motions. The robotic arm as shown in Fig. 3 is used to pick the waste and drop it on the conveyor belt.

Fig. 1 Arduino

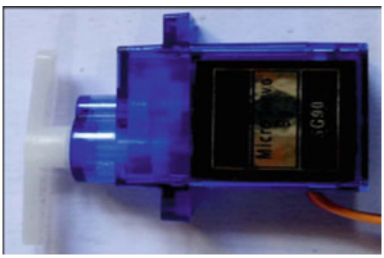


Fig. 2 Servomotor

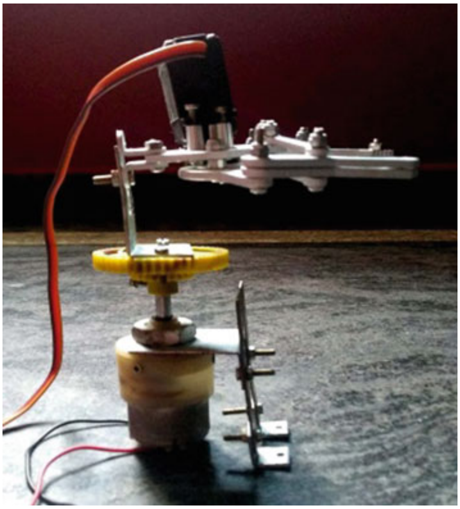


Fig. 3 Robotic arm

d. **IR Sensor:** sensor that uses infrared waves to find any object. There are two types active and passive IR sensor. We use an active sensor, which emits infrared rays through a LED. When object is near to sensor, the waves are reflected and receiver senses the waves indicating an obstacle [8]. The IR sensor as shown in Fig. 4 is used to detect any waste placed over the belt.

e. **Sonar:** HC-SR04 sonar as shown in Fig. 5 has been used. It consists of ultrasonic transmitter, receiver and control circuit. The HC-SR04 ultrasonic sensor uses SONAR to fix the distance of an object just like the bats do. It offers excellent non-contact range detection with high precision and stable readings in an easy-to-use package from 2 to 400 cm or 1″ to 13 feet. The sonar is used to track how much waste is filled in the segregated bins.

Fig. 4 IR sensor

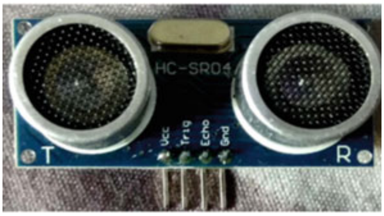


Fig. 5 Sonar

Fig. 6 Node MCU

f. **Node MCU:** A Node MCU as in Fig. 6 is a combination of a node and micro-controller unit. It is an open platform for edit/modify. It is permanent software with read only memory and is useful in prototyping. The ESP8266 is a low-cost Wi-Fi chip developed by Espressif Systems with TCP/.IP protocol. Node MCU is similar to an Arduino with ESP8266 as its major component. It has Programmable pins with built in Wi-Fi. Power is channelized through micro-USB port. Its cost is low and can. be programmed through multiple programming environments. There is various functionality of Node MCU in the proposed system like it receives signal of what kind of waste and sends the signal to servo motor, receives a signal from sonar when the bin is filled and message is sent to the worker that the bin is filled.

g. **Bluetooth**: Bluetooth module as shown in Fig 7 is a PCBA board, which integrated Bluetooth functions. Bluetooth module can be used in short-distance wireless communication, which can be separated into the Bluetooth module and Bluetooth voice module according to its usage. Bluetooth module is a elementary circuit set of chips which combined Bluetooth functions and which can be used in wireless network transmission. Here Bluetooth is used for moving the arm in convenient way.

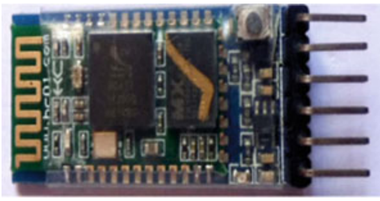


Fig. 7 Bluetooth

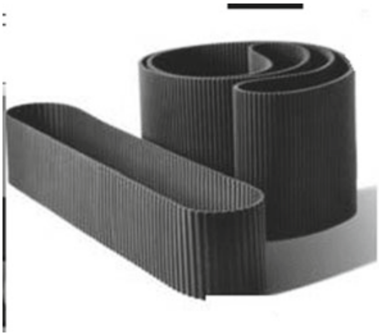


Fig. 8 Conveyor belt

h. **Conveyor Belt:** It is used to transport or carry an object. It uses two motorized pulleys to rotate in a fixed path. The motors in the pulley operate in such a way that the belt continues same path and in same speed [9]. There are two types of conveyor belt lightweight and heavyweight. In this system we use lightweight conveyor belt. The function of the belt as shown in Fig. 8 is to carry the waste that is placed on it till the bins.

The implementation of the design is explained below:

The proposed system as shown in Fig. 9 is designed such that there is no manual segregation of waste [4]. As mentioned above the components are combined to form an IoT model. Robotic arm picks the waste and drops it on the conveyor belt. The conveyor belt moves with help of motorized pulleys. As the conveyor belt moves near to the IR sensor it identifies the obstacle on the path and stops the conveyor belt from moving forward [8].

As soon as the conveyor belt is stopped the image of the object is captured and uploaded in the server for identification. In the server with AI the image is classified

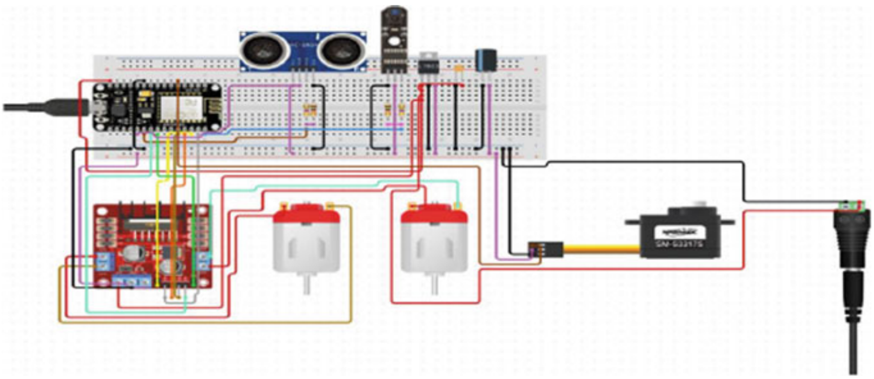


Fig. 9 Circuit diagram

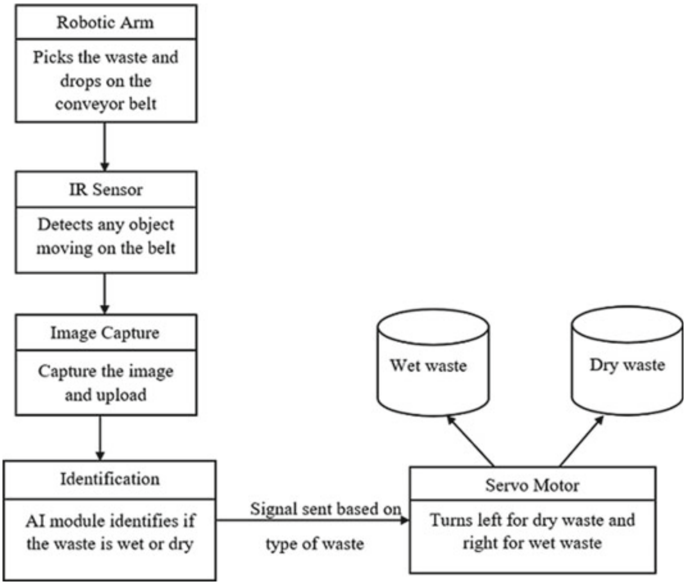


Fig. 10 Flow chart

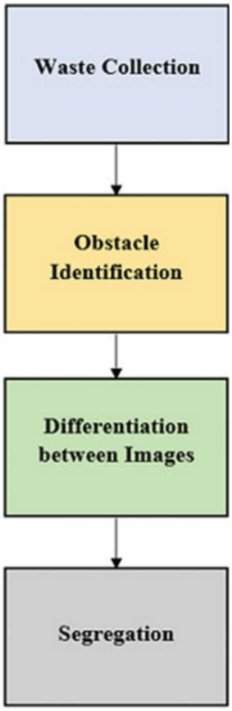
as wet or dry waste. The machine is trained in prior with a sets of wet and dry waste images such that it is capable of differentiating between any image being fetched to it. Once the image is classified as wet or dry waste a signal is sent to through the Node MCU to the servo motor and the conveyor belts motion is recovered. The servo motor turns to left if the object is wet else to right. The dry waste is collected in the right bin and wet waste on the left bin. The bins contain sonar, which is used to send a signal to the worker once the bins are filled. The sonar keeps a track of how far the bin is filled. Once the waste reaches the specified level it generates and sends signal to the Node MCU, which sends a message to the user that the bin is filled. The whole idea of this proposed system as shown in Fig. 10 is to avoid any kind of human intervention.

4 Working

The proposed system as shown in Fig. 11 works in four steps:

Waste collection: This is the first step of the process. The waste is picked one after other by the robotic arm. This object is place on the conveyor belt. The robotic arm is handled by Bluetooth. This can be done using phone and other mobile devices. This is implemented with help of Arduino, Bluetooth module and dc motor. The robotic arm moves in linear left to right and transitional up and down. This step makes the

Fig. 11 Working flow



identification and classification of object easier. The conveyor belt carries the object while it moves rotationally [10].

Obstacle identification: We use an IR sensor to identify any obstacle moving in front of it. The IR sensor used is an active sensor that emits infrared waves from a LED. When an obstacle moves in front of it this the emitted infrared waves are reflected. The reflected waves are absorbed by a receiver. This now transmits a signal that an obstacle is found. As soon as this signal is received, the conveyor belt is stopped. After the belt is stopped an image of the obstacle is taken. This image is uploaded on a server and to the trained machine.

Differentiation between images: This is the most important step of the process. This step is based on AI and fully automated. The machine is trained using CNN algorithm. A Convolutional Neural Network is based on Deep Learning; the input is an image to which importance is assigned considering various aspects of image. CNNs are used for image classification and recognition because of its high accuracy. Like a funnel CNN follows a hierarchical model, which is used to build a network, where all the neurons are connected to each other and output is processed by forming a fully connected layer. With the help of CNN the machine is able to differentiate between wet and dry waste. Once the image is differentiated a signal is sent based on type waste to the Node MCU. The belt starts moving again.

Segregation: The last step of the process. This step the Node MCU receives a signal on what kind of waste is being sent on the belt. The Node MCU sends a signal

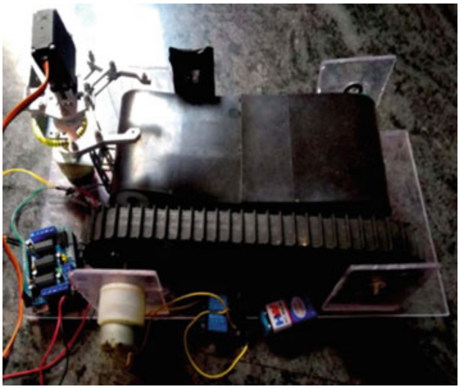


Fig. 12 Image of the implemented device setup

to the servo motor. If the object is wet the motor turns right else left. This makes the object reach left bin if it's wet and right if it's dry waste. After segregation, to make the system have less human intervention we have used sonars that will help notify when the bin is filled.

5 Result

As per the implementation and working of the device built it is automated Fig. 12. The robotic arm is handled by Bluetooth. This can be done using phone and other mobile devices. It picks the object and drops it on the conveyor belt it's done without human intervention. The object is identified and segregated with the other IoT components used. IR sensors are used such that all objects being placed are segregated. With the help of CNN the machine is able to differentiate between wet and dry waste. Once the image is differentiated a signal is sent based on type waste to the Node MCU. The CNN algorithm is highly accurate making the proposed system accurate. The sonars in the bins send automated messages to the workers when the bin is filled.

6 Conclusion

The main objective of the proposed system is to design a completely automated waste segregation device. This is implemented with usage of robotic arm and the segregation process. This will reduce human intervention in the process of segregation and help in waste reduction. Also it will reduce risk of human lives during the process of segregation. This will be much faster than the traditional method. The proposed system is reliable, efficient and cost effective. The sonars used sends automated messages when the bin is filled which does not need to regularly check the bin.

References

1. Jayson M, Hiremath S, Lakshmi HR (2018) Smart bin-automatic waste segregation and collection. In: Second International Conference on Advances in Electronics, Computers and Communications (ICAECC) (2018)
2. Thayyil J (2013) Occupational health problems of municipal solid waste workers in India. Int J Environmental Health Eng, June
3. Nandhini S, Sharma M. Naveen S, Balachandran K, Suryanarayana DS, Ram H (2019) Electronically assisted automatic waste segregation. In: 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (2019)
4. Bashir A, Banday AS, Khan R, Shafi M (2013) Concept, design and implementation of automatic waste management system. Int J Recent Innov Trends Comput Commun 1(7):604–609
5. Archana Babu S, Arunima SJ, Athira J, Chandran B, Naveen S (2016) An economic automatic waste segregator using Arduino. Int J Res Advent Tech 4(7):2321–9637
6. Nwazor NO (2019) Development of a smart, automated waste management system. Eur J Eng Technol 7(2)
7. Sivasankari, Shri B, BevishJinila Y (2017) "Smart Waste Management" Using WSN and IoT. ICIIECS
8. Pavithra (2014) Smart trash system: an application USING ZigBee. IJISET 1 (8)
9. Pushpa MK, Gupta A, Shaikh SM, Stuti J, Suchitra V (2015) Microcontroller based automatic waste segregator. Int J Innovative Res Electrical, Electronics, Instrumentation Control Eng 3(5)
10. Anagnostopoulos T, Zaslavsky A, Medvedev A, Khoruzhnicov S (2016) Top–k query based Dynamic Scheduling for IoT-enabled Smart City Waste Collection. In: IEEE

Battery and Its Management for E-Rickshaw

R. Santhosh Kumar, U. R. Geethashree, P. Kirthi, and B. C. Ranjeeta

Abstract The battery operated vehicles have entered in almost all the transport sector. One of the main transportation vehicles in cities is E-Rickshaw. E-Rickshaw—is an alternative option for the fossil fuel Auto-Rickshaw. The main electrical components in any electrical vehicle are Battery and motors. These two are designed in such a way that it should provide desired output. The battery and its management system is designed and implemented for an E-Rickshaw to carry maximum persons of four. Apart from endowing to clean, healthy and safe environment, it can also help the people to empower from lower social and financially weaker section. EV has attained importance and acceptance because of economic mode of transport. The renewable energy from Solar or wind energy is used to charge the battery.

Keywords E-Rickshaw · Renewable energy · Charge controller · LiFePO4 battery · Lead-acid battery · DC motor · Batteries · Lithium-Ion batteries

1 Introduction

E-rickshaws are three wheeled vehicles to carry a maximum people of four [1]. The E-Rickshaw uses electric power from battery to run which can be charged from Solar and Wind energy or external electrical supply [2, 3]. Electric motor acts as an engine which is powered by batteries and is suitable for smaller distance transportation for goods carrying & human transportation. They are suitable for the usage on narrow streets because of its size. The main advantage of E-Rickshaw over fuel vehicle is that, its operating cost is low and negligible or pollution free [4, 5]. They are pollution free and eco friendly transport system. E-rickshaws are one of the preferable modes of transport because of their low fuel and maintenance cost, eco friendly, zero noise pollution, simple in driving and e-rickshaw is a boon for a common man [6]. This vehicle runs at three wheels with a differential mechanism adopted for the two rear wheels. These vehicles chassis is made up of mild steel tubular type which makes

R. S. Kumar (✉) · U. R. Geethashree · P. Kirthi · B. C. Ranjeeta
Department of Electrical & Electronics Engineering, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_110

Table 1: Comparison of various three-wheeled vehicles based on energy consumption and cost [4]

Vehicles	Specific energy consumption (kJ/passenger-km)	Passenger capacity	Specific cost (INR/passenger-km)
LPG Auto rickshaw	362	4	0.6
Diesel Auto rickshaw	286	4	0.4
E-rickshaw	53.76	4	0.096

the system light. The most commonly used type of motor is series DC motor. The comparison of various three-wheeled vehicles based on energy consumption and cost is as listed in Table 1.

From the above table we can observe that the specific energy consumption is very less for E-rickshaw compared to Diesel Auto rickshaw and LPG Auto rickshaw. Even the cost per passenger for km is also very less for E-rickshaw. Hence the E-rickshaw is very much suitable for city transportation system of passengers.

2 Main Components

2.1 Battery

The EV mainly uses the following battery types [7].

- Li-Ion Batteries
- Ni Metal Hydride Batteries
- Lead acid Batteries

Out of these batteries, Lithium-Ion Batteries are used for many applications. The advantage of Lithium ion battery is their energy / unit mass is high, power/weight ratio is also high, energy efficiency is also high, better performance at higher temperature, and its discharge is also low. Another advantage of lithium ion battery is most of its components are recyclable and it is shown in Fig. 1.

2.2 Battery Management System (BMS)

A BMS manages the rechargeable battery by providing protection to the battery [8]. It also monitors the battery state, calculates and reports the secondary data etc. the BMS is a very important block in an EV. It is required to keep battery stable. The flow diagram of BMS is illustrated in Fig. 2.

Fig. 1 Lithium Ion battery

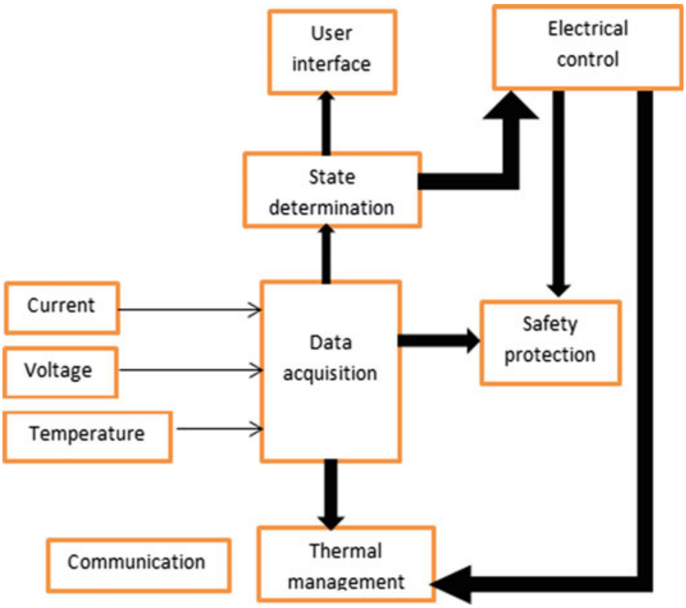


Fig. 2 Flow diagram of BMS

The BMS is mainly used to monitor the battery. The battery monitoring requires three important parameters such as the temperature, current, and voltage of each cell of battery. The other use of BMS is estimate the state of the Battery. This is done by measurement of state of charge (SOC) and state of health (SOH). SOC can be found using the voltage and current of cell, charging & discharging profile. SOH can be found from number of cycle of charge and battery performance.

3 Design of Battery Ratings

Gross weight = 210 kg

Maximum velocity required = V= 40km/hr=11.11m/s

Wheel size = 0.110 m

Area of vehicle A= length × breadth =2.625× 1.3 = 3.41 m2

Linear distance travelled = 2πr = 0.6911 m

Speed in RPM= 40000/ (0.6911×60) = 965 rpm

Power = [(mass in kg) × (g) × (V m/s) × rolling resistance] + [air density× co-efficient of drag× area of vehicle× V3]

= [210×9.81×11.11×0.01] + [0.64×0.88×3.41×1371] =2889 watts = 3000 W

Battery calculations

Distance per charge = 60 km at 40km/hr velocity

Power output/km= 3000/40 = 75 Wh/km

Let the voltage of motor be 48V

AH used per km = 75/48= 1.56 AH/km

Total AH required = 1.56 × 60 × 1.25 = 117 AH = 120 AH battery

From the above calculations we can choose 48V, 120 AH battery for the proposed work.

4 Methodology

The motor is powered by the battery (Fig. 3). Instead of using a single type of battery, here both lead-acid batteries and Li-ion batteries are used. The lead-acid batteries are charged with the help of Solar or Wind energy and lithium ion battery is charges using mains. The charge controller is used to control the current drawn from the batteries. The switch is provided to switch between the batteries based on the availability of charges in the battery. The forward or reverse switch is included and the motor drive is used to control the various parameters of the motor like direction and speed. The E-Rickshaw works on DC series motor, battery & suspension system which is contrasting with respect to the conventional auto rickshaws. The electrical battery charges on 12V. Some models made up of fibre are also used because of their strength which results in less maintenance. The commonly used battery is Lead acid battery which has a life of 12 to 24 months. Batteries with deep discharge specially made for Electrical vehicles are commonly used (Figs. 4, 5, and 6).

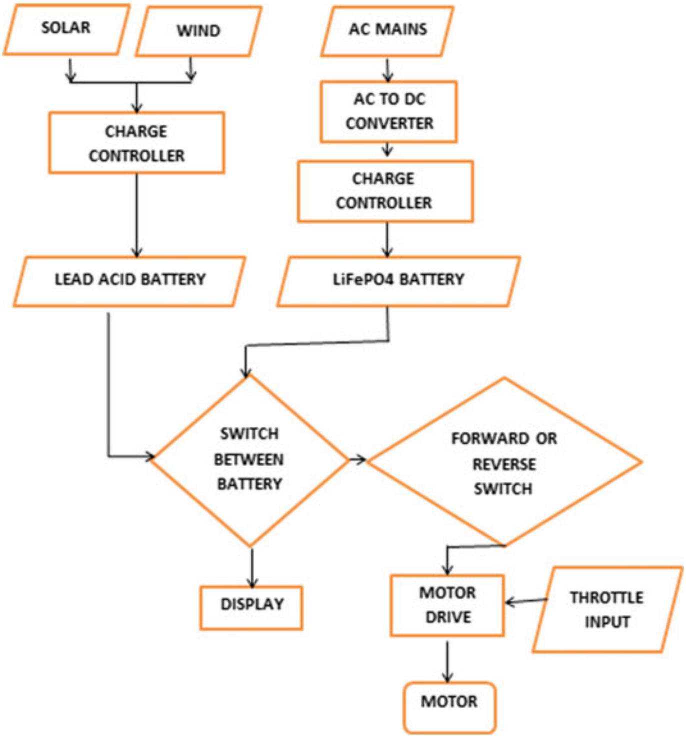


Fig. 3 The flow diagram of the E-rickshaw

Fig. 4 Front steering interface of E-Rickshaw

4.1 *Advantages of E-rickshaw Over Conventional Rickshaw*

- Air pollution can be controlled.
- Cost per distance is reduced.
- Can utilize the renewable energies like solar and wind.

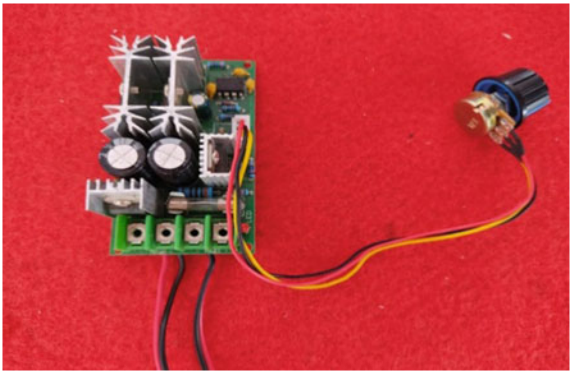


Fig. 5 Motor driver used for E-Rickshaw

Fig. 6 Battery charge controller or battery management system used for E-Rickshaw

5 Results & Discussions

The E- rickshaw powered by solar panels is assembled and the final project is as shown in the Fig. 7.

Fig. 7 Front portion of the E-rickshaw

Figure 7 shows the front portion of the E-rickshaw. It mainly consists of accelerator, speedometer, indicator, horn etc.

The rear portion of E- rickshaw is illustrated in Fig. 8.

This mainly consists of two wheels coupled with two DC series motors, Lead acid battery charged by solar panel and also from supply.

The solar panel used for the proposed project is as shown in Fig. 9.

The final project is as shown in Fig. 10.

Fig. 8 Rear portion of the E-rickshaw

Fig. 9 Solar panel

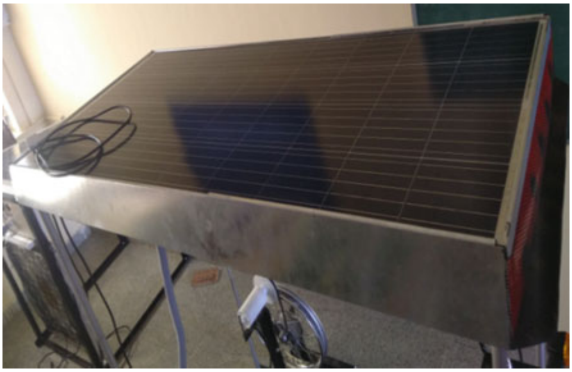


Fig. 10 Final product

6 Conclusion

This project "Battery and its management for E-rickshaw" is a cost effective, practical, eco-friendly and the safest way to transport and also to save energy. It is best suitable for urban areas. The technology helps in reducing the complexity. This system can be extended to drive the vehicle for a larger distance.

References

1. Srdjan ML, Priscilla M (2007) Usage pattern development for three-wheel auto rickshaw taxis in India. In: IEEE VPPC. https://doi.org/10.1109/VPPC.2007.4544195
2. Jain J, Khare S (2015) Multi modal public transportation system-Indore Case Study. IJSER 3(9):47–50
3. Tushar J (2007) Renewable energy and environment. Geographical Review of India, January
4. Ramachandra TV (2009) Shwetmala "Emissions from India's transport sector: statewise synthesis." Elsevier, Atmospheric Environment
5. Majumdar D, Jash T (2015) Merits and challenges of E-Rickshaw as an alternative form of public road transport system: a case study in the state of West Bengal in India. In: 2015 International

Conference on Alternative Energy in Developing Countries and Emerging Economies, Elsevier, pp 307–314
6. Kumar Singh S (2006) Future mobility in India: implications for energy demand and CO_2 emission. Transport Policy, Elsevier.
7. Iclodean C, Varga B, Burnete N, Cimerdean D, Jurchiș B (2017) Comparison of different battery types for electric vehicles. In: IOP Conference Series: Materials Science and Engineering (2017). http://doi.org/1088/1757-899X/252/1/012058
8. Hariprasad A, Sandeep PR, Ravi V, Shekar O (2020) Battery management system in electric vehicles. IJERT 9(5)

Water Pollution Monitoring System Using IoT

R. S. Sharmila, R. Sushma, M. U. Mahanth, A. Chandana, and C. Sunil

Abstract For the issue of water pollution we need a critical implementation i.e., water pollution monitoring system by using IoT with the increase in the development of technology. The system monitors the parameter through processing, storing the data, and transmission. This system uses an Arduino board with few sensors, a Wi-Fi system, and a PC. The proposed system monitoring with sensor interfacing device using IoT and programmed with high speed embedded C programming language. Finally, the various sensors monitor the data and send it to the cloud can be viewed in the IoT platform using the Wi-Fi system.

Keywords Arduino · Sensors · IoT platform · Wi-Fi system · Personal Computer (PC)

1 Introduction

Water is abundant on the earth but the amount of potable water available is in a tiny fraction (2.7%) of the total amount of water in the world. The vast majority of the world water is in the ocean (75.2%) and remaining are ground water. water. Water pollution is one of the important issues that needs monitoring of water regularly. Pollution causes many diseases and that is a reason for the death of approximately 15,000 people monthly [1]. The important use of water is consumption, agriculture, travel, industrial use which affects the quality of water.

We need to corroborate the supply of potable water quality to be monitored. The existing paper which Water pollution occurs when the water bodies are contaminated with several chemical parameters like ammonia and chloride ion amount, dissolved oxygen [2] and pollution occurs also due to the waste products like sewage and liquid waste of households, agriculture lands, and industries which diffuse pesticides, nitrogen, phosphorous load, nutrients, and hydro-morphological impacts.

R. S. Sharmila (✉) · R. Sushma · M. U. Mahanth · A. Chandana · C. Sunil
Department of Electrical and Electronics Engineering, Don Bosco Institute of Technology, Banglore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_111

The existing system using GPRS or GSM modules uses 3G cellular service ability with different equipment. There is the reason for improving the existing system to provide protection for public health and to develop an operational response in real-time. The proposed system uses Embedded circuits with a built-in internet connectivity and different sensors are used to monitor the quality of water effectively and it is economical, less complex, and less man-power required. Improve and expand monitoring of the aquatic nature for the purpose of statistical robustness [3].

2 Related Work

Paper explores the sensors' cloud domain and also the assessment of water monitoring techniques, information dissolution, network operator, influence of government and villages in provide proper information. Improving the water quality involuntary is not feasible, methodical use of technology and economic practices can help, improvise water quality and create consciousness.

We need to corroborate the supply of potable water quality to be monitored. The existing paper which processed by a micro-controller and Zig-bee protocol. Online quality of water monitoring system is relaying on GSM or GPRS. This module gathers and the data perception is sent to the monitoring centre through GPRS. It is an artificial technique and therefore collection of data and other processes will be slow. Method of maintaining the water quality in Ganga and Yamuna rivers are checked by using sensors as these two rivers are the most in our country. Therefore, surveying the water standards using this modus-operandi is costlier.

The module designed in this paper uses an Ethernet-shield host server with a database hold on to all the sensors input and evaluates the sensor input with some initial values that are present in the database already. The outcomes of the specimen can be distinguished through a browser that is connected to the corresponding network as the embedded C programming when the assigned IP address to the Arduino is retrieved.

3 Methodology

The Arduino IDE (1.8.12) software is used to upload the code [3]. We have drafted a water quality observing system forecast using embedded C programming, sensors, ADC (Analog-to-Digital-Converter). This IoT project allows the people to interconnect with real world by enabling the communication system which is embedded in the network which enables monitor to interconnect with the real-world from anywhere [3]. This system rearranges the value of water quality data by interfacing the device through wireless transmission (Fig. 1).

The hardware components which is used in our project play a very important role. The Arduino board is furnished with Analog and digital I/O pins for interconnecting

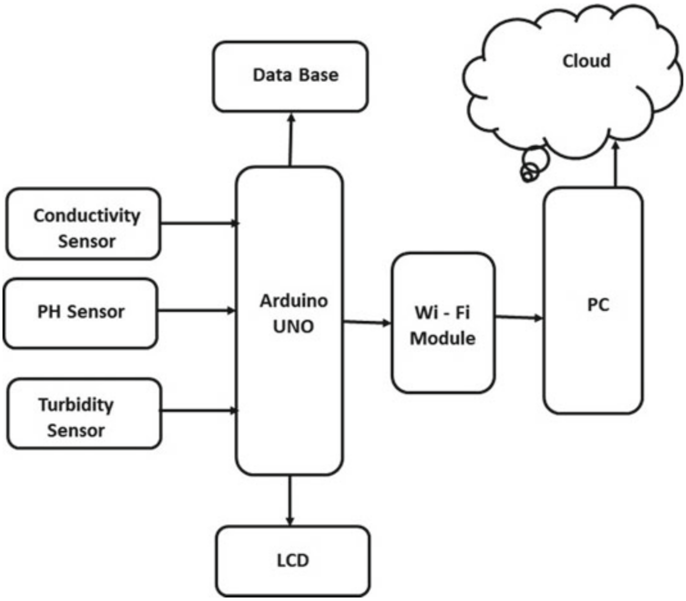


Fig. 1 Proposed system

the various boards like LCD, Wi-fi system to broadcast the data. The proposed system monitors the three criterion of water such as pH, conductivity, turbidity with high speed from various sensors using thingspeak.

This monitoring system mainly has three parts:

1. Sensors part
2. Cloud part
3. User part

Following are the Hardware components of water monitoring system: -

1. Conductivity Sensor

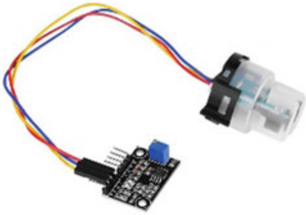
The conductivity measurement is important for detecting the impurities in the water. The conductivity temperature ranges from 28 to 158°. Farad. Conductivity present in the water means there are many ions present in the water immersed in a solution (Fig. 2).

2. Turbidity Sensor

Turbidity is a measure of the cloudiness of water when loses its transparency. Turbidity blocks the suspended particles near the surface, turbidity value must be less than 0.1NTU for drinking water and for other purpose it should not be more than 4NTU. Increase in turbidity decreases the amount of Photosynthesis which is harmful for aquatic plants (Fig. 3).

Fig. 2 Conductivity sensor

Fig. 3 Turbidity sensor



3. pH Sensor

The caustic or basic content in the water is measured using pH sensor. The pure water has a 7-pH value, less than 7 pH is caustic and more than 7 pH is a base [4]. For the drinking purpose, it should be around 6.5 to 8.5 pH therefore the span of pH is lies between 0 and 14 pH [4]. It operates on a 5 V power supply.

The temperature ranges from −5 to + 95 °C and it has a sensitivity of 0.002 pH and stability of 0.02 pH per 24 h. pH sensor indicates the chemical change in the water (Fig. 4).

4. Arduino UNO

Arduino board has 14 digital input–output pins, 6 analog inputs, a power jack, a USB connection, and a reset button [5]. The Arduino operates with the baud-rate of 9600.software IDE is the reference version of Arduino with an embedded C programming. The operating voltage is 5 V with a clock speed of 16 MHz (Fig. 5).

Fig. 4 pH sensor

Fig. 5 Arduino UNO

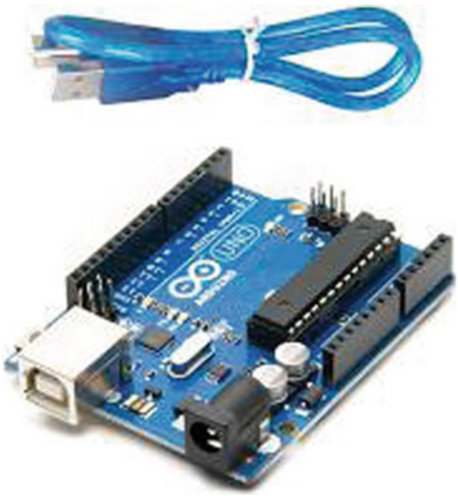


Fig. 6 Wi-Fi module

5. Wi-Fi Module

In this proposed system make use ESP8266 module because it uses TCP/UDP protocol. This Wi-Fi module is extremely cost-effective with high-speed processing (Fig. 6).

6. LCD

The outline size of the LCD is 16 $\times$ 2 mm with a maximum thickness of 13.2 mm and it has a built-in controller ST7066. The power consumption is lesser and it is widely used in the screen industry. LCD is interfaced with embedded C programming (Fig. 7).

4 Implementation

Firstly, the Arduino establishes the digital interfacing between the controller and network then the sensor monitors the quality of water and displays the value of

Fig. 7 LCD

the sensor on LCD and also uploads on the cloud. If the sensor value exceeds the preliminary value it displays as pollute on LCD and once again it uploads on the cloud. The uploaded data on the cloud can be viewed by the user by login into the cloud (Thingspeak).

The uploaded data on the cloud can be seen by everyone by using API key credentials. The proposed system monitors the source, transmits, stores, and processes wirelessly, and collects the sensor values. This creates slow development in an operational response (Fig. 8).

5 Experimental Results

Three sample were tested for the impurities if any in the water. The result for highly impure among the three samples are listed below.

Figure 9 shows the pH sensor reading on LCD indicating acidic level to be more than the specified limits.

Figure 10 shows the turbidity sensor reading on LCD indicating that the water contains suspended solid matter consisting of particle level to be more than the specified limits.

Figure 11 shows the conductivity sensor reading on LCD.

In the water monitoring, ADC will convert the designated sensor reading analog assessment into a digital representation, and from that value evaluated with the related environment parameters [3]. This system checks the water quality parameters and furnish the information on water safety and also give the knowledge regarding continuous monitoring of the water. The user can visualize the data in terms of graph using IoT platform as shown in Fig. 12.

6 Conclusion

The proposed water pollution system is a wireless monitoring system which uses the IoT and also uses Arduino with sensor network. This system is economical in determining the water quality of a reservoir and it protects the environment and

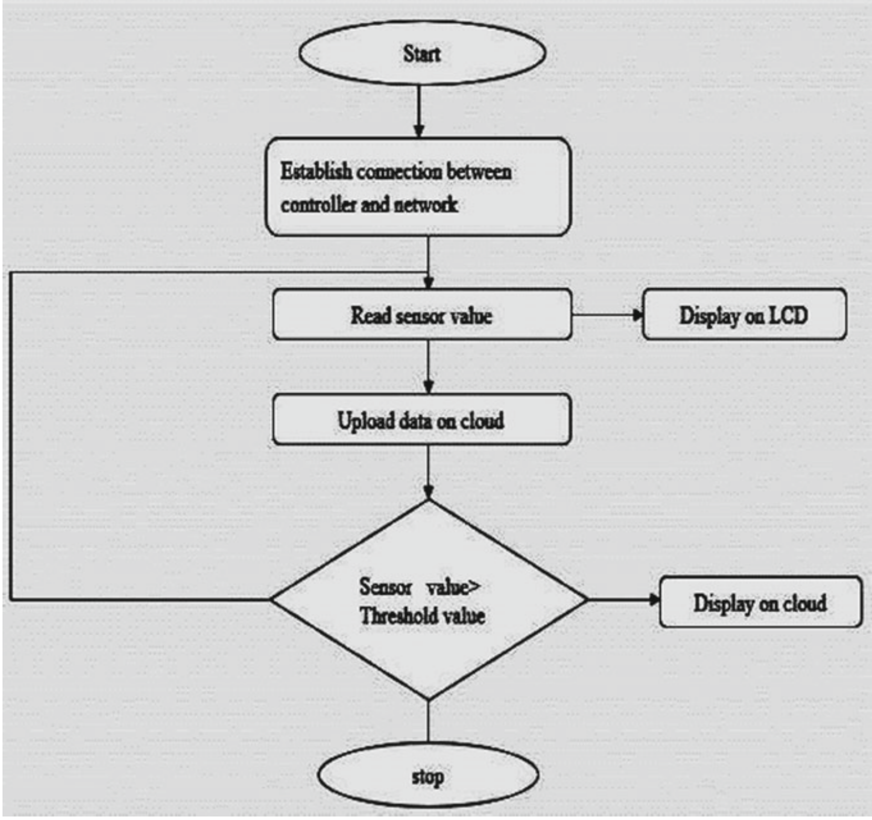


Fig. 8 Flow chart

Fig. 9 pH sensor reading

Fig. 10 Turbidity sensor reading

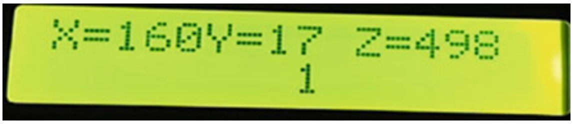


Fig. 11 Conductivity sensor reading

Fig. 12 (**a**) shows the conductivity sensor scaling on the on ThingSpeak. (**b**) shows the turbidity sensor scaling on the on ThingSpeak. (**c**) shows the turbidity sensor scaling on the on ThingSpeak

human life from the danger. For the actual monitoring process, the system is reliable and feasible and inherits high execution speed [3]. The operation is simple and has widespread application. By using the embedded and sensor devices for monitoring protects the environment and helps to interact with the other objects through the network.

References

1. Kedia N (2015) Water quality monitoring for rural areas—A sensor cloud-based economical project. In 1st International Conference on Next Generation Computing Technologies (NGCT-2015) Dehradun, India, 4–5 September
2. Water Pollution Monitoring using IOT. International J Innovative Technology and Exploring Eng, 2019
3. Rajender, Dushyantha G, Khadri MNS, Jagadevi N, Kalshetty, International Journal of Innovations in Engineering and Technology for Smart Water Quality Monitoring System Using IoT Environment by Nikhil
4. Vaishnavi V, Daigavane, Gaikwad MA, Advances in Wireless and Mobile Communications. ISSN 0973–6972 Volume 10, Water Quality Monitoring System Based on IOT
5. Arbain N, Firdaus Nordin N, Mat Isa N, Saaidin S (2014) LAS: Web-based laboratory attendance system by integrating RFID-ARDUINO technology. In: 2014 2nd International Conference on Electrical, Electronics and System Engineering (ICEESE)

Multi-Function Digital Mirror Using Raspberry Pi with IOT

A. S. Sneha, J. Sonica, M. Pavithra, Syed Luqman, and S. Rohith

Abstract Intelligent mirrors, that continue the works nowadays and can take its place within the future technology, offer each mirror and pc motor-assisted info services to its users. There's never a finish to devices which will be created 'smarter' with the assistance of adequate technology. There are ample good show devices however mirrors offer associate interactive setting whereas displaying info. This paper presents the look and development of a digital mirror exploitation Raspberry pi with extra options. The contents are displayed on associate LED monitor that is basined during a picket frame and lined with a sheet of reflective two-approach mirror. The mirror provides basic amenities like weather of town, time and news details. All the computing is finished with the assistance of a raspberry pi. This might be achieved by the device because it provides valuable info on the show at a look, whereas additionally acting as a traditional mirror.

Keywords Smart Mirror · Raspberry Pi · LED monitor · Internet of Things

1 Introduction

In recent years most devices are connected to the net. The net has competed a very important role in connecting additional and more individuals across the planet. Devices began to become smarter and smarter, mobile phones became smartphones and most significantly net was connected to a range of devices and the construct came to be referred to as the 'Internet of Things'. Digital Mirror could be a mirror that permits touch-free user interaction with vital info displays within the kind of Associate in Nursing interactive system, created for easy access to basic vital info and enhancing utility. The fluorescent and incandescent lighting are the foremost common form of lighting in houses and official places, which are the foremost doubtless nature for the Digital Mirror.

A. S. Sneha (✉) · J. Sonica · M. Pavithra · S. Luqman · S. Rohith
Department of Electrical & Electronics Engineering, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_112

The problem lies to find some way to expeditiously check all factors that will have an effect on how someone prepares for the day whereas conjointly playacting all necessary tasks ahead of the mirror. Our goal is to form a product that fulfills one's would like of preparing whereas at the same time receiving the daily news, weather, time and alternative helpful info. The power to collect this info throughout one's morning preparation can lead to a lot of economical way of life.

The ability of the system to act with the external surroundings or the inner states is thanks to its embedded technology that conjointly plays a key role indecision creating [1]. This project concentrates on style and development of a sensible mirror that represents a sublime interface for glancing info for multiple individuals during a home surroundings powered with the interaction provided by Google Voice services. It provides a web-based interface to access knowledge feeds and alternative services. Its knowledge domain work comprising of domain information of computer science, Image process, and web of Things.

1.1 Background

In home automation domain, a wise mirror is usually used for displaying multimedia system knowledge, promoting healthy life-style, and dominant house appliances. The most attractiveness of this approach is that folks will access personalized data effortlessly whereas doing their daily activities. A staple article of furniture piece found in each bedchamber and toilet, the mirror has provided a method for effective personal grooming for thousands of years. Our team has brought it into the twenty first century with Digital Mirror [2]. Whereas the user is getting ready them for the day, they're going to be ready to look at their mirror and instantly retrieve vital bits of data resembling the current time, date, and climatic conditions. These straightforward bits of data are unremarkably asked for within the morning whereas preparing for the day or at nighttime before reaching to bed. Our Digital Mirror ensures that you just stay at the forefront of the rising smart-home revolution. You may make the most of each chance to seem at yourself within the mirror so as to be full of weather data and today's date. With a short cross-check the mirror, users are ready to check this time, peek at the forecast, and to not mention, consult the mirror for ancient personal look changes [3]. whereas the user is getting ready themselves for the day, they're going to be ready to ask their sensible mirror to examine what reminders that they had set for the up and coming back week, what the weather are like so that they will dress consequently, and set any calendar events they have to recollect all whereas having the ability to pay attention to their favorite songs via Spotify [4].

2 System Requirements

Our framework utilizes Raspberry Pi, interfacing parts much the same as intersection rectifier Monitor, receiver, speaker, HDMI link and an intelligent sheet. Raspbian OS, Python and Google Voice were acclimated coordinate our framework.

2.1 Hardware

2.1.1. **Raspberry pi:** One of the center board of the reasonable show and furthermore the mike interface. This focal gadget is that a tiny low single-board pc, that runs by Python programs inside the foundation [5] on A put in Raspbian OS. This presentations area basically based estimate, the most elevated 5 most inclining news features and furthermore the date and time, which includes the different web principally based administrations taking after climate and news path to extricate information from the net and show it to the client, inside such a gadgets. So as to accomplish this, the Raspberry Pi unit is associated with the net by means of a constituent Wi-Fi module connected inside the Raspberry Pi [6]. It conjointly gets contribution from the associated mike by means of AN outside USB card, that allows the client to perform changed capacities along with net pursuit, fitting of updates and in any event, participating in 2 strategy discussions with the voice aide, by using the Google's voice colleague API.

2.1.2 **Light Emitting Diode Monitor:** The intersection rectifier system is utilized inside the great Mirror in light of the fact that the essential show that interfaces with the Raspberry Pi unit by means of Associate in Nursing HDMI link. The intersection rectifier Monitor is utilized for showing imperative information to the client inside such a gadgets like current climate, universal news, time. This can be accomplished with the help of the foundation forms run on Raspberry Pi. The intersection rectifier show is roofed in a thin 2 methodology intelligent sheet that empowers foundation LED screen to strong the on screen show all over the sheet, though moreover allow it to go about as an intelligent mirror for the client.

2.1.3 **Microphone:** A 3.5 mm mouthpiece is associated with the Raspberry Pi unit through outside Universal Serial Bus sound card which permits client to include sound orders to the Google's Voice Assistant Application Program Interface which is introduced on the Raspberry Pi unit and perform different errands, for example, setting up of cautions and updates.

2.1.4 **Speaker:** A 3.5 mm jack supporting speaker, associated with the Raspberry Pi, is utilized to gives the sound yield.

2.1.5 **HDMI:** The HDMI link will be utilized to interface the Raspberry Pi to the Display Monitor so as to produce the video yield, for example the gadgets that are shown on the Monitor.

2.1.6 **Reflective Sheet***:* It gives the mirror viewpoint to the entire framework, permitting the client to see their appearance, while simultaneously empowering the LED Display Monitor at the back to cast gadgets through the sheet.

2.2 Software

2.2.1. **Raspbian OS:** Raspbian is the fundamental and essential programming for RPi gadgets, formally bolstered by Foundation [6]. Truth be told, it is a working framework, in view of Debian and upgraded for Raspberry Pi equipment.

2.2.2. **Google Assistant SDK:***:* This can be utilized as Voice Assistant.

3 Design

This mirror is developed with the usage of partner degree precious stone rectifier screen in light of the fact that the fundamental show, to show the data to the client. It's lined in an thin dual-way intelligent sheet, that goes about as an intelligent mirror to the client, conjointly allowing the gem rectifier screen at the back to show content through the thin intelligent sheet [7]. Raspberry Pi module is associated with the gem rectifier screen with the help of partner degree HDMI link that bolsters video yields through the screen. To utilize Google's voice collaborator API, voice orders are inputted through 3.5 mm Mic and furthermore the sound is yielded by means of a three 3.5 mm speaker.

Other miscellaneous parts embody a memory card for the Raspberry Pi, cables to allow power to the system. All this instrumentation is assemble within the frame, that added with of each aesthetic price and protection.

The housing of the good Mirror are designed during a method that's value effective however additionally displays the mirror and software package within the clearest way attainable. The housing are designed from wood with a face panel product of four meter cut boards as shown in Fig. 1 The wood are finished a burned end then stained for a pleasant presentation that individuals would possibly wish to have in their homes. it's vital that this can be conferred as a high quality show piece otherwise the general public would have very little interest in having it in their home. All sensors and buttons are hidden to gift the image of an everyday mirror.

Since this mirror isn't quite sort of a regular mirror, there ought to be housing designed round the diode Monitor which is able to extend five inches behind the face panel. The diode Monitor will mount to the rear of the face panel and can be control in victimization brackets and screws. The speakers from the diode Monitor are mounted to the edges of the mirror housing to permit music to be competing via the LED Monitor. The management buttons for the diode Monitor are mounted to

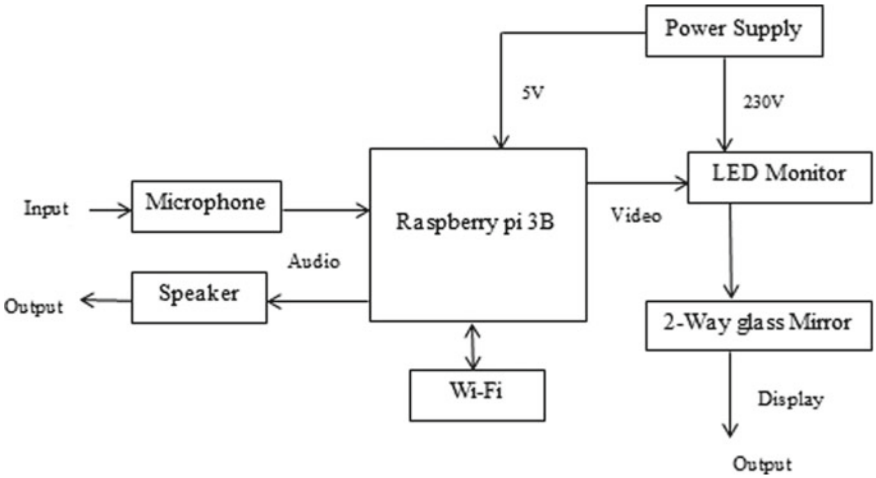


Fig. 1. Block diagram of interfacing device

the lowest of the mirror housing permitting manual control of the LED Monitor if necessary.

The method the mirror works is by, utilizing a 1 method mirror behind that is placed a monitor. this can permit lightweight to shine through the mirror where the diode Monitor is lit however it'll merely mirror sort of a mirror wherever there are black pixels on the show [7]. To accomplish this, the technique mirror can set into the face panel slightly and therefore the monitor are mounted directly behind it.

An opening for the Raspberry Pi are situated on one facet of the mirror housing to permit access to the USB ports and also the LAN port if required as shown in Fig. 3. If the mirror has internal speakers these will be enforced on the edges of the mirror housing to output sound suitably however to confirm a clean presentation from the front of the mirror.

Lastly, variety of tiny holes is created within the key locations on the face panel of the mirror to permit for placement of the electro-acoustic transducer, speakers to require user voice commands [8].

4 Implementation

STEP1:

Turn on the power supply for both Raspberry and LED monitor.

STEP2:

Turn on the hotspot and connect it to the Raspberry Pi.

Fig. 2. 3D Model of mirror housing design is shown in Fig. 2

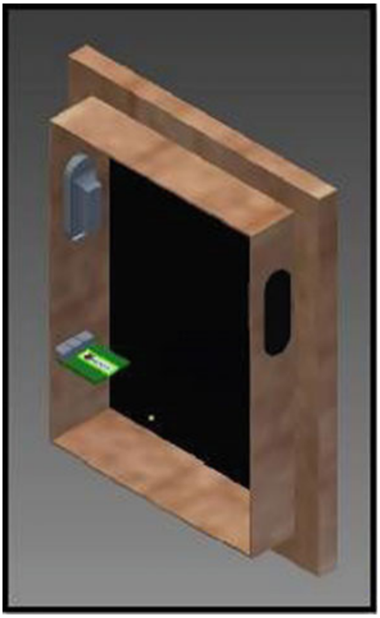


Fig. 3. 3D Model of the rear view of the mirror housing

Fig. 4 Final output is shown in Fig. 4

STEP3:

After observing the raspberry pi connect your device to raspberry through VNC Viewer it needs associate degree information processing address thus we've to enter the proper IP address of raspberry pi so it'll be connected [9].

5 Result

The final result of the digital mirror is as shown in the Fig. 5.

This displays weather, time, calendar, top most trending news and voice services for browsing, setting alarms and events is included.

Fig. 5 Digital mirror after hardware and software installations

6 Conclusion and Future Scope of Work

Technological integration into homes, therefore known as "Smart Home Technology" is turning into progressively widespread in our society nowadays. the first advantage of good home technology is to change our day to day lives in any means potential. Some edges embrace saving time, relieving stress, or maybe saving cash. These edges may be accomplished during a variety of how together with automation of tasks, improved access to media and knowledge, also as varied degrees of private comfort.

A Digital Mirror is supposed as a sensible home technology to permit convenient access to media and knowledge that individuals may well be curious about on a each day. The mirror can show custom-made info to the user to permit users to urge some basic daily information at a look. this may save users time in their busy morning routine which can, as a result, relieve stress and facilitate individuals get out the door on time within the morning. The mirror can show info users could access within the morning via their phone or alternative technology during a means that's promptly offered as shortly as they walk into the toilet. this may enable the user to set up for his or her day whereas at the same time finishing their daily routine ahead of the mirror.

Through the creation of this device, we have a tendency to were able to offer associate interactive expertise to the user of the Digital Mirror, displaying valuable every day info to the user by group action computing and communication technologies, in this manner protecting time, while that have practical experience in elective utilities of the gadget as a standard mirror or partner crisis activating gadget to expand the business case. in contrast with elective existing great mirror- item, in this project it is the one ready for offering intelligent crisis activating innovation though keeping the gadget modest for clients of differed monetary foundations. In the midst of the expansion in innovative headway, this gadget will pack the fundamental choices that a client would require, as they subsume great innovation in their daily practice.

References

1. Smither JW (2016) Maker culture chapter in innovating pedagogy 2013. The Open University. Retrieved 20 April 2016
2. Johri A (2018) Smart mirror: A time-saving and affordable assistant. In: 4th International Conference on Computing Communication and Automation (ICCCA)
3. Growth from knowledge. One-third of US consumers own two or more smart home devices, 2018
4. Lazar A, Koehler C, Tanenbaum J (2015) Why we use and abandon smart devices. In: 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York
5. Grayson JE (2000) Python and Tkinter Programming, Manning
6. Schmidt M (2014 Raspberry Pi: a quick-start guide, 2nd edn, Pragmatic Bookshelf
7. Herman SL, Garrard CG (2016) How can I get started with home automation?. Retrieved 20 April 2016
8. Choudhuri KBR (2017) Learn Arduino Prototyping in 10 days. Packt Publishing

9. Mixter K, Implementations for voice assistant on devices, U.S. Patent Application No. 15/592,137.

Forecasting Age Adjusted Rates of Lung Cancer in Mumbai by Fitting ARIMA Models

Manjula S. Dalabanjan and Pratibha Agrawal

Abstract There are several methods to forecast the values of any time series. The best model is one which has minimum error. If the given time series is stationary, we can use the autoregressive integrated moving average (ARIMA) model. If it is not stationary, it has to be made stationary by differencing the time series. ARIMA model was identified using the Bayesian Information Criteria. The Auto Covariance Function (ACF) and Partial Auto Covariance Function (PACF) were studied along with the stationarity of the time series. The future values were forecasted and the past values were predicted using recursive method.

Keywords Stationary · Time series · Age adjusted rates · Autocorrelation · Aurtoregressive models · Akaike information criterion · Bayesian information criterion

1 Introduction

Lung cancer is one of the dreaded diseases prevailing in the world. The incidence of lung cancer has increasing tendency in most of the metropolitan cities. The data of Age Adjusted Rates of prevalence of lung cancer in Bengaluru and Mumbai is available in the website of "National Cancer Registry Programme from the year 1982 to 2009" [1]. The data available is classified according to males and females.

2 Literature Survey

The actual time series data may not be linear. Most of time series have non-stationary behaviour. Thus, we cannot fit these time series adequately using autoregressive

M. S. Dalabanjan (✉)
DBIT, Bengaluru, India

P. Agrawal
AMCEC, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_113

moving average models. Several other models have been suggested by (Battaglia and Protopapas [2]). Among them, a simple class and particularly interesting models are multi-regime threshold models. Using alternative linear autoregressive equations, the series is generated [2].

ARIMA (*p, d, q*) is a general model where *p* is the order of autoregressive model, d is the order (degree) of differencing of time series and *q* is the number of lagged forecast errors in prediction equation called moving average order.

Liquid bulk is important for Indian revenue. Sankar and Poorvaraghavan [3] made a forecast of the liquid bulk export in Chennai by using different forecasting techniques. ARIMA *(0, 1, 1)* was selected as a best fit for the data. [3].

In time series analysis, we used the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion and the Bayesian Information Criterion (BIC) to select order of the ARIMA model.

Sankar and Ravikumar [4] made an effort to develop an ARIMA model for tourist arrival to Tamilnadu and applied the same to forecast the number of tourist arrivals to Tamilnadu in the future years.

Balasiddamuni et al. [5] proposed two modified estimation procedures for time series linear statistical models involving Moving Average MA (*1*) and MA (*q*) process errors. If we plot the trend of Age Adjusted Rates (AAR) of cancer over the years, there are oscillations in the graph. We can say that the oscillatory AAR of cancer are due to unknown risk factors affecting cancer.

C. Umasankar et al. [6] developed a linear statistical model with first order autoregressive scheme for the errors and estimated the parameters of the model by an iterative method of estimation using studentized residuals. Later, this model was used to obtain the feasible forecasts. C. Umasankar et al. [6] suggested a model compared to Box-Jenkins original fit to series D, where errors are propagated to find future values of the time series. Since the errors are unobservable, the fitting of the moving average models is more complicated than autoregressive models [6].

Heewon and Fumitake [7] used the lag weighted lasso (Least Absolute Shrinkage And Selection Ordering) model for forecasting. Recent studies discovered that the lasso estimator may be inefficient and the true model selection result could be inconsistent.

The chemical viscosity data is named as series D by (Box and Jenkins [8]). The series D was fitted with ARIMA (*1, 0, 0*) and ARIMA (*0, 1, 1*) in 1976. To fit the same series using ARIMA models Etebong [9] used the Bayesian Information Criterion.

An ARIMA (*1, 1, 1*) model was fitted. This model was tested using Ljung-Box goodness of fit test. The normalized Bayesian Information Criterion (BIC) was explored to confirm the adequacy of the model. The ARIMA (*1, 1, 1*) model to the chemical viscosity data was superior.

3 Methodology

3.1 Maximum Likelihood Estimation Method of ARIMA Model ([George et al. [8])

For modelling and forecasting of time series data, the ARIMA model was used. This model can be used only if the time series data is stationary. If it is not stationary, it has to be made stationary by applying transformations namely, difference method, logarithmic transformation or other transformations ARIMA (*p, d, q*), is a general model where *p* is the order of autoregressive model (no. of auto regressive terms), *d* is the order (degree) of differencing of the time.series and *q* is the number of lagged forecast errors in prediction equation called moving average order. ARIMA (*p, d, q*) models are ARMA (p, q) models with d^{th} -order difference transformation. Let the difference operator be $(1 - B)^d$ where *B* is called the backward shift operator and

$$BZ_t = Z_{t-1} \tag{1}$$

$$B^m Z_t = Z_{t-m} \tag{2}$$

$Z_t, t = 1, 2, 3, \ldots n$ is the time series data set.

In ARIMA (*p, d, q)* model, let the time series be stationary after taking d^{th} -order difference transformation. For stationary time series, ARMA (*p, q)* models we have the following equations in the form:

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \ldots \varphi_p Z_{t-p} + a_t + \theta_1 a_{t-1} + \ldots \theta_q a_{t-q} \tag{3}$$

That is

$$(1 - \varphi_1 B - \varphi_2 B^2 - \ldots \varphi_p B^p) Z_t = (1 - \theta_1 B - \theta_2 B^2 \ldots \theta_q B^q) a_t \tag{4}$$

Or

$$\varphi(B) Z_t = \theta(B) a_t \tag{5}$$

where $\varphi(B)$ and $\theta(B)$ are polynomials of degree *p* and *q* in B.

$$a_t = w_t - \varphi_1 w_{t-1} - \varphi_2 w_{t-2} - \ldots - \varphi_p w_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots + \theta_p a_{t-p} \tag{6}$$

3.2 Conditional Likelihood for an ARIMA Process

Assume that the $N = n + d$ original observations $Z_{-d+1}, ...Z_0, Z_1, Z_2 \ldots Z_n$ form a time series. Let us also assume that this time series is generated by an ARIMA model of order d. A series of differences w are generated.

The general problem of fitting the parameters φ and θ of the ARIMA model is equivalent to fitting to the w's. The stationary, invertible ARMA (*p, q*) model is

$$\varphi(B)\nabla^d Z_t = \theta_0 + \theta(B)a_t$$

On the choice of values of w's, the $a_1, a_2, a_3 \ldots a_n$ values are calculated. Thus the starting values of a's and w's can be calculated using the invertible ARMA (*p, q)* model given by Eq. (6). Given a particular set of data, the log-likelihood associated with the parameter values $(\varphi, \theta, \sigma_a)$ conditional Sum of squares is given by

$$S*(\varphi, \theta) = \sum_{t=1}^{n} a_t^2(\varphi, \theta/w's, a's)$$

The likelihood expression for ARIMA Model is $L*(\varphi, \theta, \sigma_a) = -n\ln(\sigma_a) - \frac{S*(\varphi,\theta)}{2\sigma_a^2}$.

The parameter values obtained by minimizing the conditional sum of squares function $S*(\varphi, \theta)$ will be called conditional least squares estimate.

3.3 Unconditional Likelihood for an ARIMA Process

The Unconditional log-likelihood for an ARIMA model is given by

$$L(\varphi, \theta, \sigma_a) = f(\varphi, \theta) - n\ln(\sigma_a) - \frac{S(\varphi, \theta)}{2\sigma_a^2}$$

where $f(\varphi, \theta)$ is a function of φ and θ. The unconditional sum of squares function is given by

$$S(\varphi, \theta) = \sum_{t=1}^{n} a_t^2$$

During calculation of the conditional sum of squares, a's are computed recursively by taking conditional expectations in (6). Computation of back forecasts start with forward recursion. For the MA (1) model for w_t, this method involves first computing the values $a_t(a_0 = 0)$, which we abbreviate as a_t^0, given by

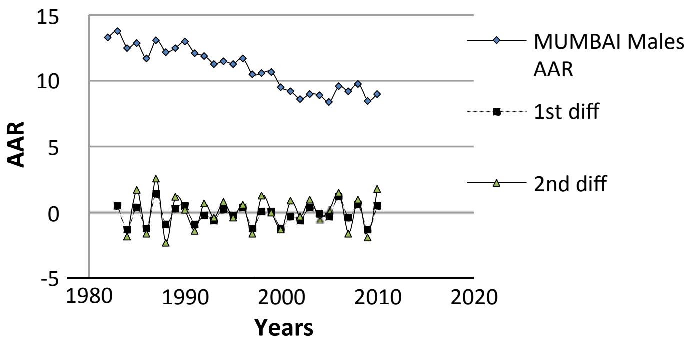


Fig. 1 The trend of lung cancer for males in Mumbai

$$a_t^0 = w_t + \theta a_{t-1}^0 \; t = 1, 2, 3, \ldots n$$

$$BIC(p, q) = \ln v * (p, q) + (p + q)[\ln(N)/N]$$

where p and q are the order of AR and MA processes respectively and N is the number of observations in the time series and $v*$ is the estimate of white noise variance σ^2. The graph of first order differences and second order differences appears to be stationary from Fig. 1. Taking initial value, $a_0^0 = 0$ a backward recursion is performed to obtain $u_t = a_t^0 + \theta u_{t+1}$, starting from $t = n$ down to $t = 0$, with $u_{n+1} = 0$ as the starting value. Finally, the exact back-forecast value of Z_t is calculated by taking

$$a_0 = \frac{-u_0(1 - \theta^2)}{(1 - \theta^{2(n+1)})}$$

And using this starting value, the $a_t's$ are computed from

$$a_t = w_t + \theta a_{t-1} t = 1, 2, 3, \ldots n$$

The exact sum of squares is

$$S(\theta) = \sum_{t=0}^{n} a_t^2$$

3.4 Model Identification

The time series of Age adjusted Rates (AAR) in lung cancer for males in Mumbai is plotted in Fig. 1. The series is not stationary. The graph of the first order differences of the time series seems to be stationary that varied over a time around a constant mean and variance. If the original series $\{Z_t\}$ is not stationary, we can look at the first order difference process.

$$X_t = \nabla Z_t = Z_t - Z_{t-1}$$

Or second order differences

$$X_t = \nabla^2 Z_t = \nabla(\nabla Z)_t$$

And so on.

The tool for identifying model is estimation of Autocorrelation Function (ACF). Suppose we have $(X_1, X_2 \ldots X_n)$ from a stationary series, we can estimate the mean

$$\overline{X} = \frac{\sum_1^n X_t}{n}$$

The auto covariance is given by

$$C_k = \hat{\gamma}_k = \frac{\sum_{t=k+1}^{n} (X_t - \overline{X})(X_{t-k} - \overline{X})}{n}$$

The check of stationarity is done with the help of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of differential data. Hence the ACF and PACF values of first differenced. Z_t values were computed and presented in the Table 6. From the graph of ACF and PACF, the first difference data is stationary. The different ARIMA models were discussed. The available data was differenced once (d = 1) (Tables 1 and 2). The autocorrelation is given by

$$r_k = \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}}$$

Table 1 ACF of the time series of AAR of lung cancer in Mumbai males

Lag	Autocorrelation	Std. error	Box-Ljung Statistic		
			Value	d.f	p-Value
1	−0.601	0.179	11.254	1	0.001
2	0.353	0.176	15.271	2	0.000
3	−0.236	0.173	17.146	3	0.001
4	0.071	0.169	17.322	4	0.002
5	0.075	0.165	17.527	5	0.004
6	−0.288	0.162	20.704	6	0.002
7	0.241	0.158	23.022	7	0.002
8	−0.155	0.154	24.033	8	0.002
9	0.228	0.150	26.341	9	0.002
10	−0.235	0.146	28.909	10	0.001
11	0.024	0.142	28.938	11	0.002
12	0.087	0.138	29.338	12	0.004
13	−0.132	0.134	30.309	13	0.004
14	0.109	0.129	31.016	14	0.006
15	−0.240	0.124	34.747	15	0.003
16	0.277	0.120	40.132	16	0.001

Table 2 PACF of the time series of AAR of Lung cancer in Mumbai Males (AARMM)

Lag	Partial autocorrelation	Std. error	Lag	Partial autocorrelation	Std. error
1	−0.601	0.189	9	0.159	0.189
2	−0.014	0.189	10	−0.089	0.189
3	−0.046	0.189	11	−0.264	0.189
4	−0.135	0.189	12	−0.026	0.189
5	0.114	0.189	13	−0.059	0.189
6	−0.290	0.189	14	−0.061	0.189
7	−0.139	0.189	15	−0.213	0.189
8	0.043	0.189	16	−0.094	0.189

3.5 Model Estimation

The estimated parameters of the models were represented in Tables 3 and 4. Hence the ARIMA (0, 1, 1) has the lowest normalized BIC value and a better R squared and RMSE values. This model can be said to be the best to Forecast the future AAR and predict the past values.The Bayesian Information Criteria is least for $p = 0$, $d = 1$ and $q = 1$ (Figs. 2 and 3).

Table 3 Measures of error and BIC Values of ARIMA (*p, d, q*)

ARIMA (*p, d, q*)	R-Squared	RMSE	*p*-value	BIC
0,1,0	0.79	0.76	0.002	−0.307
0,1,1	0.85	0.64	0.290	−0.515
0,1,2	0.86	0.64	0.251	−0.399
1,1,1	0.87	0.63	0.279	−0.461
1,1,2	0.87	0.64	0.233	−0.3
2,1,0	0.87	0.63	0.278	−0.416
2,1,1	0.88	0.62	0.329	−0.376
2,1,2	0.88	0.62	0.272	−0.229

Table 4 Estimated ARIMA Model fit Statistics

Fit statistic mean	Mean
Stationary R-squared	0.308
R-squared	0.855
RMSE	0.647
Normalized BIC	−0.515

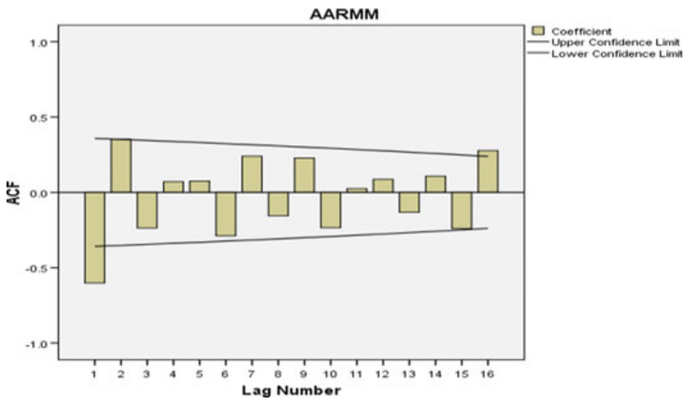


Fig. 2 Plots of ACF for AAR of lung cancer in Mumbai males

Thus the time series of Mumbai AAR can be identified as autoregressive integrated moving average (ARIMA) process of order (*0, 1, 1*) which had minimum BIC was

$$\nabla Z^t = (1 - \theta B)a^t - 1 \leq \theta \leq 1$$

$$a_t = w_t + \theta a_{t-1} \quad w_t = \nabla Z_t \quad E[w_t] = 0$$

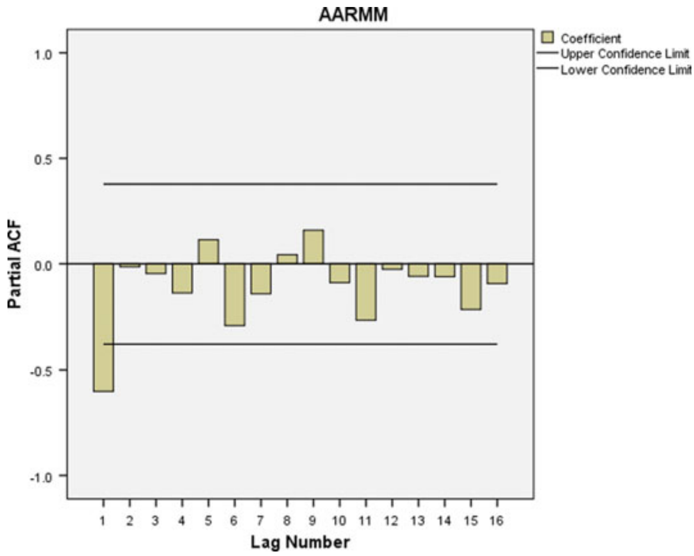


Fig. 3 Plots of PACF for AAR of lung cancer in Mumbai males

Recursive calculations of first few a's can be obtained for a particular value of θ. The sum of squares is minimum when $\theta = 0.4$. Proceeding in this way we find all the "a" values. Now the conditional sum of squares for an ARIMA (0, 1, 1) process is

$$S * (0.4) = \Sigma a_t^2 (a_0 = 0/\theta = 0.4) = 12.7900 \tag{7}$$

The unconditional sum of squares

$$S(0.4) = \sum_{t=0}^{28} [a_t^2] = 12.8215 \tag{8}$$

When we fitted ARIMA (*0, 1, 1*) model to the prevalence of lung cancer for males in Mumbai, the conditional residual sum of squares was nearer to the value of unconditional residual sum of squares (Tables 5 and 6).

3.6 Diagnostic Checking

Various autocorrelations were computed up to 16 lags and the same along with their significance tested using Box-Ljung statistic are provided in Table 7. As the result indicate, none of these autocorrelations were significantly different from zero at any reasonable level. This proved that ARIMA model is an appropriate model or

Table 5 Sum of squares for ARIMA(*0, 1, 1*) process for various values of **θ**

θ	$S*(\theta)$	θ	$S*(\theta)$
−0.5	24.9228987	0.1	14.65889291
−0.4	22.4526705	0.2	14.01161003
−0.3	20.2862665	0.3	13.66815137
−0.2	18.4236868	0.4	12.7900
−0.1	16.8649313	0.5	13.89270671
0	15.61		

forecasting Age Adjusted incidence Rate of lung cancer in males in Mumbai. Both the ACF and PACF are approximately normally distributed about their population values, and have standard deviation of about $1/\sqrt{n}$, where n is the length of the series.

A **'portmanteau' test** of white noise (due to Box and Pierce and Ljung and Box) can be based on the fact that

$$Q_m^{'} = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\gamma_k^2$$

approximately follows γ^2 distribution with m d. f.

The sensitivity of the test to departure from white noise depends on the choice of *m*.

Since, for this process, $\lambda = 1 - \theta$ the weights for the one-step-ahead forecast, we can write,

$$\hat{Z}_{t-1}(\lambda) = \lambda \sum_{j=1}^{\infty}(1-\lambda)^{j-1} Z_{t-j}$$

Thus for the IMA (0, 1, 1) model, the forecast for all future time is an exponentially weighted moving average of current and past $Z's$ (EWMA).

4 Results and Discussions

Manjula and Pratibha [10] have used the variate difference method to explain trend analysis of lung cancer in Bengaluru and Mumbai. [10] also forecasted the future AARs of Lung cancer in Bengaluru and Mumbai. In our study we have fitted an ARIMA (0, 1, 1) model and forecasted the future values of incident rates of lung cancers. ARIMA models of higher order may become good models to forecast the future AARs of time series of some types of cancer available in 2013 registries.

Table 6 Recursive values of a_t

t	Year	Z_t	$w_t = \nabla Z_t$	$a_t = w_t + 0.4a_{t-1}$	a_t^2
0	1982	13.3		0	0
1	1983	13.8	0.5	0.5	0.25
2	1984	12.5	−1.3	−1.1	1.21
3	1985	12.9	0.4	−0.04	0.0016
4	1986	11.7	−1.2	−1.216	1.47866
5	1987	13.1	1.4	0.9136	0.83466
6	1988	12.2	−0.9	−0.5345	0.28575
7	1989	12.5	0.3	0.0861	0.00743
8	1990	13.0	0.5	0.5344	0.28565
9	1991	12.1	−0.9	−0.68621	0.47088
10	1992	11.9	−0.2	−0.47448	0.22513
11	1993	11.3	−0.6	−0.78979	0.62377
12	1994	11.5	0.2	−0.11591	0.01343
13	1995	11.3	−0.2	−0.24636	0.06069
14	1996	11.7	0.4	0.30145	0.09087
15	1997	10.5	−1.2	−1.07941	1.16514
16	1998	10.6	0.1	−0.33176	0.11006
17	1999	10.7	0.1	−0.03271	0.00106
18	2000	9.5	−1.2	−1.21308	1.47156
19	2001	9.2	−0.3	−0.78523	0.61659
20	2002	8.6	−0.6	−0.91409	0.83556
21	2003	9.0	0.4	0.03436	0.00118
22	2004	8.9	−0.1	−0.08625	0.00743
23	2005	8.4	−0.3	−0.33450	0.11189
24	2006	9.6	1.2	1.06619	1.13678
25	2007	9.2	−0.4	0.026479	0.00070
26	2008	9.8	0.6	0.61059	0.37282
27	2009	8.5	−1.3	−1.05576	1.11463
28	2010	9.0	0.5	0.07769	0.00603

Table 7 Estimated ARIMA model

Model	Number of predictors	Model fit statistics	Ljung-box Q		
			Statistic	d. f	p-value
ARIMA (0, 1, 1)	1	0.308	19.693	17	0.290

5 Conclusions

The fitted ARIMA model gives the Forecasts of Age Adjusted incidence Rate of lung cancer in Mumbai Males. The forecasted AAR of lung cancer is provided in Table 8. Table 9, shows that there is decreasing tendency in the AAR of lung cancer in Mumbai. Thus we may infer that the forecasting is a good one because the tendency of the AAR from Fig. 1 is also decreasing in nature.

When we fitted ARIMA (*0, 1, 1*) model to the prevalence of lung cancer for males in Mumbai, the conditional residual sum of squares obtained in Eq. (7) was nearer to the value of unconditional residual sum of squares obtained in Eq. (8). We have taken $\theta = 0.4$ and obtained the residual sum of squares.

The column headings of Table 8 denote certain expressions which are defined below:

Col 1: t
Col 2: year
Col 3: Z_t
Col 4: $[w_t] = \nabla Z_t$
Col 5: $[a_t] = [w_t] + 0.4\left[a_{t-1}\right]$
Col 6: $0.4\left[a_{t-1}\right]$
Col 7: $0.4\left[e_{t+1}\right]$
Col 8: $e_t = w_t + (0.4 * e_{t+1})$
Col 9: $u_t = a_t^0 + \theta u_{t+1}$

From Table 8, the Age Adjusted Rate—back forecast value for the year 1981 for Mumbai males is 13.1355. The future forecasted values are shown in Table 9.

Table 8 Recursive calculations of a's using $\theta = 0.4$

Col 1	Col 2	Col 3	Col 4	Col 5	Col 6	Col 7	Col 8	Col 9
−1		**13.1355**		0		0.0650	0	
0	1982	13.3	0.1645	0.1645	0.2263	−0.0019	0.1625	0.19353
1	1983	13.8	0.5	0.5658	−0.4295	−0.5048	−0.00482	0.07258
2	1984	12.5	−1.3	−1.0737	−0.0118	0.0379	−1.26207	−1.23303
3	1985	12.9	0.4	−0.0295	−0.4847	−0.3052	0.09482	−0.39839
4	1986	11.7	−1.2	−1.2118	0.3661	0.4371	−0.76294	−0.92230
5	1987	13.1	1.4	0.9153	−0.2135	−0.3074	1.09262	0.72370
6	1988	12.2	−0.9	−0.5339	0.0345	0.1315	−0.76843	−0.47895
7	1989	12.5	0.3	0.0864	0.2138	0.0289	0.32891	0.13733
8	1990	13	0.5	0.5345	−0.2745	−0.4277	0.07229	0.12722
9	1991	12.1	−0.9	−0.6861	−0.1898	−0.1692	−1.06927	−1.01838
10	1992	11.9	−0.2	−0.4744	−0.3159	−0.2232	−0.42318	−0.83054
11	1993	11.3	−0.6	−0.7897	−0.0463	0.0420	−0.55796	−0.89018
12	1994	11.5	0.2	−0.1159	−0.0985	−0.0949	0.10508	−0.25098
13	1995	11.3	−0.2	−0.2463	0.1206	−0.0372	−0.23729	−0.33768
14	1996	11.7	0.4	0.3014	−0.4318	−0.4932	−0.0932	−0.22830
15	1997	10.5	−1.2	−1.0794	−0.1327	−0.0330	−1.23307	−1.3243
16	1998	10.6	0.1	−0.3317	−0.0131	−0.1826	−0.08267	−0.61241
17	1999	10.7	0.1	−0.0327	−0.4852	−0.5567	−0.45668	−0.70162
18	2000	9.5	−1.2	−1.2130	−0.3141	−0.1917	−1.39171	−1.67228
19	2001	9.2	-0.3	−0.7852	−0.3656	−0.1793	−0.47929	−1.14801
20	2002	8.6	−0.6	−0.9140	0.0137	0.1517	−0.44824	−0.90695
21	2003	9	0.4	0.0343	−0.0345	−0.0206	0.37937	0.01783
22	2004	8.9	−0.1	−0.0862	−0.1338	0.0484	−0.05155	−0.04131
23	2005	8.4	−0.3	−0.3345	0.4265	0.4211	0.12112	0.11236
24	2006	9.6	1.2	1.0661	0.0106	−0.1472	1.0528	1.11715
25	2007	9.2	−0.4	0.0264	0.2442	0.032	−0.368	0.12739
26	2008	9.8	0.6	0.6105	−0.4223	−0.52	0.08	0.25228
27	2009	8.5	−1.3	−1.0557	0.0311	0.16	−1.3	−0.89576
28	2010	9	0.5	0.0776		0	0.4	0.07769

The Age adjusted rates are available for the years 1982 to 2010. Bold value indicates The back-forecast value for year 1981 is 13.1355

Table 9 Forecast of age adjusted incidence rate of lung cancer in Mumbai males

t	Year	AAR
29	2011	8.9698
30	2012	8.9697
31	2013	8.96975
32	2014	8.96975

References

1. NCRP-Time trends in cancer Incidence Rates, 1982–2010. National Centre for disease Informatics and Research. National Cancer Registry Programme (Indian Council of Medical Research), Bangalore, 2013. www.ncrpindia.org
2. Battaglia F, Protopapas MK (2012) Multi-regime models for non linear non stationary time series. Comput Stat 27:319–341 (2012)
3. Jai Sankar T, Poorvaraaghavan J (2012) Forecasting export of liquid bulk in Chennai Port. In: Proceedings of the Second International Conference on Stochastic Modelling and Simulation, pp. 207–212 (2012)
4. Jai Sankar T, Ravikumar M (2012) Time series analysis of tourists arrivals in Tamilnadu. In: Proceedings of the Second International Conference on Stochastic Modelling and Simulation, pp. 339–343 (2012)
5. Balasiddamuni P, Sarojamma B, Ramakrishna Reddy P, Durga Prasad S, Naik JP (2012) Estimation of time series linear statistical model with moving average process errors. In: Proceedings of the Second International Conference on Stochastic Modelling and Simulation (2012)
6. Umasankar C, Vijaybhaskar Reddy M, Balasiddamuni P, Murali K, Yadavendra Babuand S, Babu SK (2012) An iterative statistical first order autoregressive forecasting model. In: Proceedings of the Second International Conference on Stochastic Modelling and Simulation, pp. 207–212 (2012)
7. Park H, Sakaori F (2013) Lag weighted Lasso for time series model. J Comput Stat 28:493–504
8. George E, Box P, Jenkins GM, Reinsel GC (1976) Time series analysis, forecasting and control, 3rd edn. Prentice-Hall International, Inc.
9. Clement EP, Using normalized Bayesian information criterion (BIC) to improve Box-Jenkins model building. Am J Math Stat 4(5):214–221
10. Dalabanjan MS, Agrawal P (2017) Trend analysis to forecast the prevalence of lung Cancer in Bengaluru GRENZE. Int J Eng Technol, 3(3):215–220 (2017)

COVID 19 Outbreak Analysis, Prediction & Forecasting

N. J. Anasuya, A. S. Chaithra, and Umme Athiya

Abstract COVID 19, alias scales up as Severe Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The virus which started showing up late November 2019 in Mainland China. The primary host of this spreading virus are bats (Chiroptera). Humans feeding on bats or the intermediate hosts a re the carriers of this virus. The outspread virus is so fast and vigorous that it happens through human to human transmission, thus it is handled by a genuine activity termed "Social Distancing".

Keywords COVID 19 · SARS-CoV-2 · Info graphics · Support vector machine · Bayesian ridge regression · Polynomial regression · Arima model · FB prophet model · Sigmoid model · Linear quadratic estimation · Forecasts · OLS regression results · AIC · BIC · Omnibus · Covariance · Skew · Kurtosis · Choropleth · Durbin–Watson · Jarque–Bera

1 Introduction

Coronaviruses originate from the subfamily called Orthocoronavirinae, a cluster of polymeric molecule Ribonucleic Acid (RNA) virus, predominantly a source of harm in vertebrates. They are majorly known to cause respiratory tract infections, in least cases is the occurrence of common cold. The enfolded micro-organism is a positive sense single- stranded RNA genome and a nucleic acid core encircled by capsid of spiral uniformity. The genome size of microorganism ranges from approximately 26 to 32 kb. They have attribute of clover leaf shaped spikes that reckon from their facet, thus an electron micrograph holds an image similar to the solar corona, thus named. The world's specialized agency responsible for health management issues WHO on March 11, characterized COVID 19 as pandemic.

N. J. Anasuya (✉) · A. S. Chaithra · U. Athiya
Don Bosco Institute of Technology, Bangalore, India
e-mail: anasuyaprakash@dbit.co.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_114

2 Proposed System

COVID 19, it is almost 7 months passed bearing the rising number of cases in the virus. The Government has taken various measures to break the pandemic chain spread. Despite the continuous efforts of Government in bringing the lockdown effect, the situation is still increasing to its peak. This is majorly caused by the people who lack maintaining the 'social distancing' policy, wearing 'face mask', & constantly having 'sanitized' (washing hands regularly).

With the above rising issues stated in the existing system, the work focuses on building a platform where in the people are educated by showing them with stuffs like visualizations, forecasting models, increase of cases, deaths & recoveries on daily basis. All these concepts involve the idea of Machine Learning algorithms, the time series forecasting models, & Web Scraping [1].

3 System Design

- Web Scraping, a process of uprooting of broad outstretch amounts of data from the websites & having them stored in the local file of your computer, or into other formats like database, json files, or .csv files.
- Info graphics is the amalgamation of "information" & "graphics", used in building higher levels of visualization & cognitive abilities of living beings to understand well.
- The analysis of thinking machine algorithms that boost spontaneously through experience. It is seen as a sub-element of artificial intelligence.
- Choropleth is a type of colorful map which shows the areas (locations) shaded for the time series statistical data or the geographical characteristics of various countries [1] (Fig. 1).

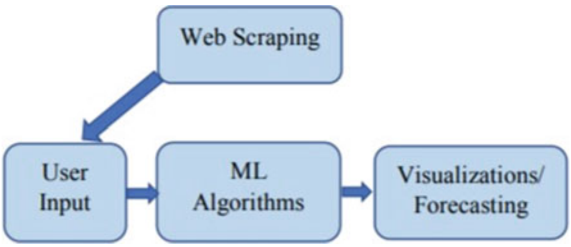


Fig. 1 Block diagram of the project

4 Algorithms

4.1 FB Prophet Model

FB Prophet Model is one the Machine Learning Model used in forecasting with the statistical data. This model is an additive model which basically fits the data on yearly, weekly, daily seasonality, & holiday effects. These are sinewy to missing data and capable of handling the outliers [2] (Figs. 2 and 3).

The working of prophet model depends on the following functions

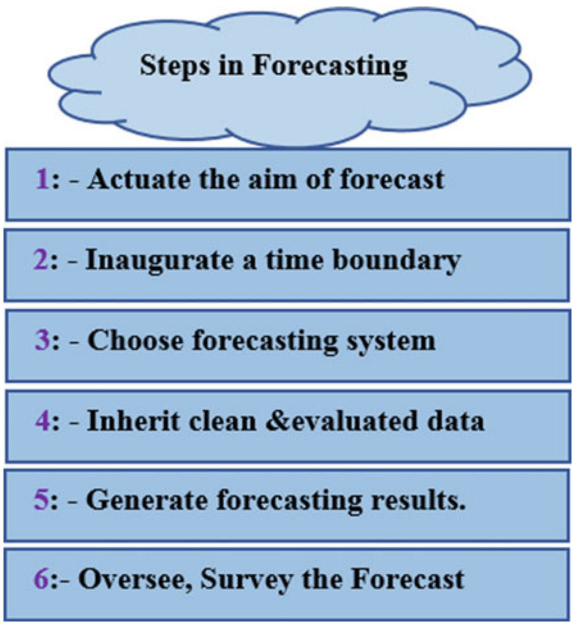


Fig. 2 Steps in forecasting

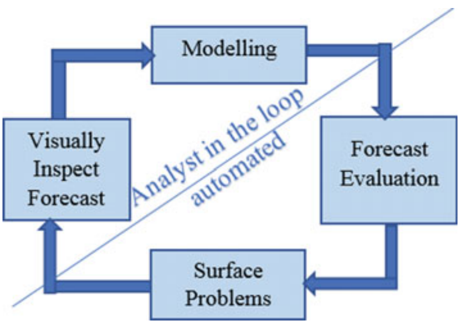


Fig. 3 Block diagram of FB model

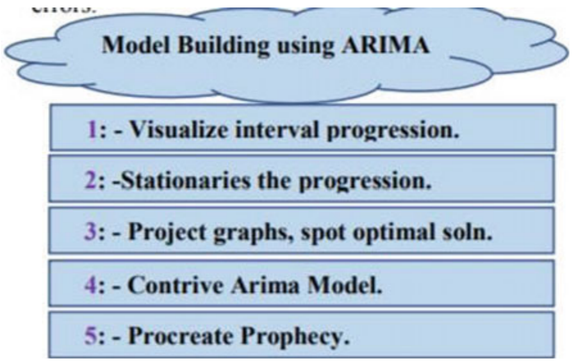


Fig. 4 Model Building using Arima

$$y(t) = g(t) + s(t) + h(t) + \in$$

where,

- ***g*(*t*)** stands for model trend, which means it handles the incessant increase or decrease in the data.
- ***s*(*t*)** stands for seasonality models which uses the ***Fourier Series*** concept to predict the change in the seasonal factors occurring during the year.
- ***h*(*t*)** stands for holidays models, which determine the effects of launch of products or event organizing which play a major role in the business development.
- **ϵ** represents the discordant fallacy term.

4.2 Arima Model

Arima, abbreviated as Auto Regressive Integrated Moving Average. It captures the distinct level of profane structures. Expansion of Acronym

(i) **AR: Auto Regression**—specifies the relation among some random observations & lagged observations.
(ii) **I: Integrated**—involves differentiating the raw observations into categories.
(iii) **MA: Moving Average**—it is dependent on two factors: Observations & the residual errors [3] (Fig. 4).

4.3 Support Vector Machines

Support Vector Machine algorithm is the one which can be employed in both regression as well as categorical (binary) circumstances. Support Vectors are the coordinates of unique observations. It is a delineation exemplar as extremity in arena, correlated to the disembodied grouping which are split by a comprehensible space that are spread far extent. SVM, an exemplar of Empirical Risk Minimization (ERM) algorithm employed for the hinge loss. These belong to an intuitive category of algorithms for analytical conjecture, and many of its unique attributes caused by the deportment of the hinge loss. Thus this visualization can provide future viewpoint into how and why SVMs work and allow us to better analyze the demographic effects [4].

4.4 Polynomial Regression

Polynomial Regression analysis is effective in finding out the correlation between an independent variable x and a dependent variable y which is collectively compounded into an nth degree polynomial in variable x. It is meant to best fit in a nonlinear model where the regression function would be represented as $F(y|x)$. The major goal of this analysis is to model the estimated value which belongs to a dependent variable y which is corresponds to independent valued variable x [5].

5 Results

The proposed system contains the live tracking of covid19 cases through web scraping, the choropleth graphs. The home page contains the live tracking of worldwide, India & Karnataka cases of active, cured & deaths from the web browser [6–8]. The proposed system also shows the preventive measures and symptoms which indicates precautionary measures to be taken by the users (Figs. 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, and 24).

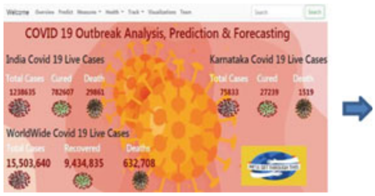


Fig. 5 Home page~analysis

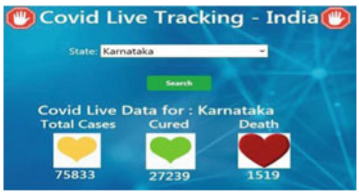


Fig. 6 Live tracking of cases in India

Fig. 7 Worldwide affected countries

Fig. 8 Top 12-countries affected

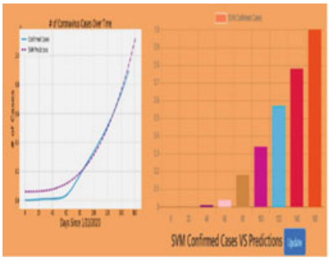


Fig. 9 SVM Prediction worldwide

Fig. 10 SVM forecasting worldwide

Dates	SVM Predicted Cases
23/07/2020	15334588.0
24/07/2020	15574049.0
25/07/2020	15816127.0
26/07/2020	16060836.0
27/07/2020	16308191.0
28/07/2020	16558206.0
29/07/2020	16810894.0
30/07/2020	17066271.0
31/07/2020	17324349.0
01/08/2020	17585145.0

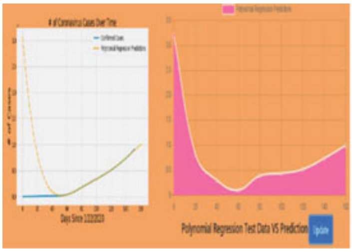


Fig. 11 PR prediction worldwide

Fig. 12 PR forecasting worldwide

Dates	Polynomial Predicted Cases
23/07/2020	14978564.0
24/07/2020	15171345.0
25/07/2020	15361989.0
26/07/2020	15550193.0
27/07/2020	15735641.0
28/07/2020	15918008.0
29/07/2020	16096955.0
30/07/2020	16272131.0
31/07/2020	16443174.0
01/08/2020	16609706.0

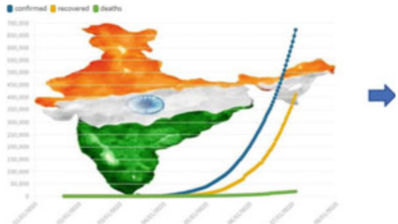


Fig. 13 COVID cases in India

Fig. 14 COVID cases state wise

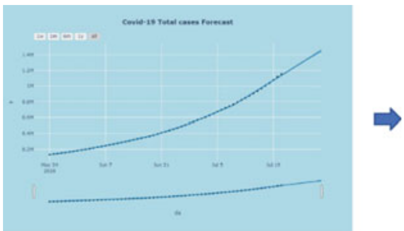


Fig. 15 FB model—India (COVID cases)

Fig. 16 FB model forecasting-India

Dates	FB Model Forecasting Values
24/07/2020	1.2312M
25/07/2020	1.2631M
26/07/2020	1.2967M
27/07/2020	1.3297M
28/07/2020	1.3612M
29/07/2020	1.3898M
30/07/2020	1.4213M
31/07/2020	1.4527M
01/08/2020	1.4892M

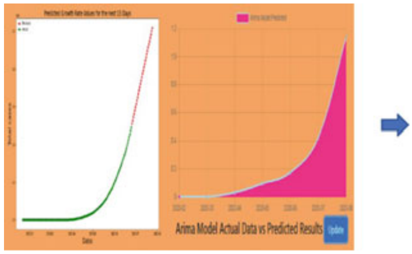


Fig. 17 Arima model—India (COVID cases)

Fig. 18 Arima model forecasting

Dates	ARIMA Model Forecasting
24/07/2020	1.2353M
25/07/2020	1.2672M
26/07/2020	1.2998M
27/07/2020	1.3302M
28/07/2020	1.3625M
29/07/2020	1.3921M
30/07/2020	1.4325M
31/07/2020	1.4581M
01/08/2020	1.4927M

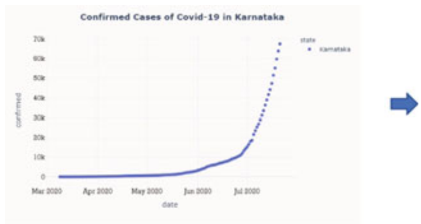


Fig. 19 Confirmed cases—Karnataka

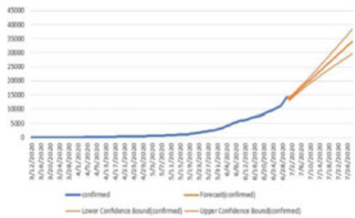


Fig. 20 Probability—confirmed

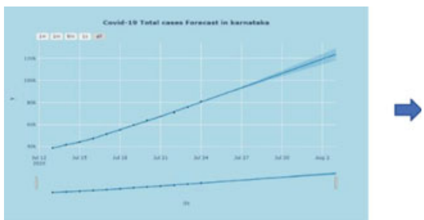


Fig. 21 FB Model—Karnataka (COVID cases)

Fig. 22 FB Model—Forecasting Karnataka

Dates	FB Model Forecast
24/7/2020	80863
25/7/2020	84753
26/7/2020	89096
27/7/2020	93488
28/7/2020	97781
29/7/2020	102124
30/7/2020	106467
31/7/2020	110810
01/8/2020	115152

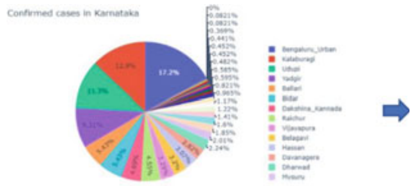


Fig. 23 Confirmed cases—Karnataka

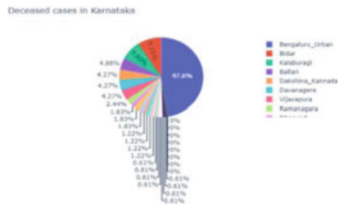


Fig. 24 Deceased cases—Karnataka

6 Conclusion

The proposed project aims to provide the users with knowledge & insights of the rise in the number of confirmed cases & the deaths happening. The people have to ta ke necessary precautionary measures like inhibiting the social distancing activity, avoiding unnecessary roaming, wearing face mask, keeping ourselves sanitized. The main aim is the proposed system provides with the forecasting graphs and its corresponding values through which necessary actions can be taken. The graphs that helps in predicting the forecast values indicate a sign of danger in India and Karnataka, the actual cases cross the line of mark from the predicted value. Thus stay at home and be safe…!

Acknowledgements The development of the proposed system was extremely supported by Don Bosco Institute of Technology. I present gratitude to my guides Dr. Anasuya N Jadagerimath, Professor, Department of Information Science & Engineering, DBIT & Mrs. Chaithra A S, Assistant Professor, Department of Information Science & Engineering, DBIT. They provided the idea in coming up with this system & extremity in resources.

References

1. https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/. This site provides the information on building a ARIMA Model.
2. https://www.karnataka.com/govt/covid19-hospitals-in-karnataka/. This contains the data of Karnataka state hospital facilities.
3. https://facebook.github.io/prophet/. This site contains the time series forecasting model information, installation steps, procedure~FB Prophet Model
4. https://en.wikipedia.org/wiki/Infographictext=Infographics%20(a%20clipped%20compound%20of,to%20see%20patterns%20and%20trends. Basic knowledge on Info graphics
5. https://en.wikipedia.org/wiki/Support_vector_machine. Support Vector Machine Algorithm
6. https://www.kaggle.com/sudalairajkumar/covid19-in-india?select=covid_19_india.csv
7. https://github.com/CSSEGISandData/COVID19/tree/master/csse_covid_19_data/csse_covid_19_time_series. This site contains the world wide dataset in time series form from January to recently update till date
8. https://www.kaggle.com/c/covid19-global-forecasting-week-4. This site contains the datasets which can be used in building a model or in forecasting

IoT Enabled School Bus Monitoring and Notification System

M. Selvam, Aishwarya R. Yadahalli, Monisha M. Dindi, and B. Nithin

Abstract In the past working model of the school bus system, there was a chance to categorize many methodologies. Nowadays there is an increase in crimes and accidents in the boarding and operation of school bus. Parents are worried about their kids on many reasons when they are going to schools like falling asleep in the school bus, children may miss the bus or may leave at the wrong place, too much delayed in pickup and drop, waiting long time in boarding and receiving, missing and tracking, no reasons on abnormality, unusual behaviour of children, boarding in other bus etc. To overcome these problems, use of GSM, GPS, Wi-Fi and RFID technologies for tracking and monitoring the school bus has proposed as the solution in this work. This has become an advantage for efficient tracking with low cost GPS and Wi-Fi Module or GSM which helps parents to keep track their children, sends an alert message to the parents about the present location of the vehicle and sends the emergency message alerts to the school on abnormal situations. Sensors play a vital role in this research for sensing the problems like fire, break down, accident, consumption of alcohol or metal detection.

Keywords Radio frequency identification · Global positioning system · Global system for mobile communication · Wi-Fi module

1 Introduction

One of the most common problems is traffic jam where vehicles of the parents carrying their children are stuck in traffic jam during school and college hours which gives a negative impact. This happens when parents are using their private vehicles instead of public transport or school bus for children. Parents who don't own private vehicles, they use public transports to carry their children to school.

Another scenario is that parents send their children to the school and stay tensed whole day because they don't get any updates about their children until they return from school. There is also a chance that students may miss the school bus or public

M. Selvam (✉) · A. R. Yadahalli · M. M. Dindi · B. Nithin
Don Bosco Institute of Technology, Banglore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_115

transport or passing the time out of the institution where parents do not get any updates regarding their children.

Generally parents will believe that school buses are safe, reliable, timing and accountable so that they can focus on their job or employment but they need the status on the abnormality. Global Positioning System (GPS) is an essential part in day to day life. It is used to track school bus location or nearby taxi. But there is no proper application that gives the exact location and real-time location of different vehicles [1]. In this research, a school bus tracking and notification model is introduced to get rid of all the problems that the children face during their school time journey.

This model helps parents to get exact location of the school bus via Google Maps. Essentially maps give a clear idea about the school bus location and also it shows an estimated time to reach the exact location through Short Message Service (SMS) alert or via the online application. It also has a feature called online attendance which helps the school to keep the count and account of every child daily which is based maps and Application Program Interface (API) [1].

This model can also be implemented in the college buses also. Due to the negligence of the driver which may lead to accidents and children may lose their life or might get hurt. In this model, the goal is to give safety transportation to children during school journey. Radio Frequency Identification (RFID) cards/Near Field Communication (NFC) tags give information about the entry and exit of every student to / from the school bus.

This system tracks the exact location of the school bus and sends the alert message to their parents [2]. Metal sensor is employed to detect any metals like gun or knife found or used inside the bus and Alcohol or Grove-Gas (MQ3) sensors are used for detecting the consumption of alcohol used by the driver and others in the school journey [3].

There are three units in this working model which are:

- Parent unit: In this unit, parents register into a particular application using their mobile number for more updates regarding their children.
- Bus unit: Various sensors are used in this monitoring and controlling like a smoke sensor, metal sensor, RFID reader, and Alcohol sensor. As soon as a student enters into the bus / exits from bus, the database is updated with the information of entry and exit with date and time.
- School unit: School unit consists of some applications where administrators can log in and add, delete, update and modify the student details, school bus information; in and out attendance of every student and many more information is stored with the help of GPS which is installed inside the school bus. In this unit, the school gets every message alert regarding every student on the security breech and emergency situations like fire, accident, rally, medical sickness/illness, deadlock on traffic, riot, etc.

2 Scope

The main aim of this research is to provide safety for students between the place and time of home and school. This solution intends to do the following.

- This system [3] detects alcohol content if the driver has consumed it.
- This system [1] can send message-alerts to the parents about the school bus location and abnormality
- Detects fire using a smoke sensor and sends a message to the school and nearby fire stations.
- Detects weapons using a metal sensor and sends a message to the school on the abnormality.
- Provides online attendance facility and stores data about every student with date and time.

This model helps in many ways to the school administration and parents. [1] Administration department gets the message about every student about on-boarding and off-boarding in the bus with the date and time and sends message alert to the parents about the abnormality if any.

3 Overview

The school bus monitoring system is implemented for tracking the exact location of the school bus at a specific time. The proposed system makes use of technology that combines a smart phone application with a microcontroller. GPS is implemented inside the school bus for tracking purposes and the Wi-Fi module acts as a Global System Mobile Communication (GSM) module for sending a message to the parents [1]. This model uses various sensors like alcohol sensor, smoke sensor, metal sensor, etc. for the safety of the students. In case of emergency, the status of the school bus is known to the school administration as well as parents.

4 Related Work

One of the reasons behind this research is to prevent accidents by using an alcohol sensor [3]. The research is formulated by integrating the sensor with the help of a microcontroller. There are different types of sensors to detect alcohol. In this research, MQ2 sensor is used which detects the presence of alcohol from the breath of human. The ignition system works based on BAC that is Blood Alcohol Content.

This model gives the easiest way to determine either the parents or any other person who can pick up their kids [3]. This model makes use of smart phone NFC functions and tags to identify an individual. Every student has a unique NFC tag ID which is stored in the smart phone of teacher and parent. When parent or any other person wants to pick up their kids from the school, the teacher scans the NFC tag of their parents or individual's ID that verifies the tag with the content table within no time. This procedure for the tags gives the correct mechanism to use while picking their kids from the school.

At present, parents are bothered about their school-going children because of the increase in missing cases and crimes. Students and parents/guardians/maids have to wait at their stop until the bus arrives sometime it may take a longer time than usual [3]. In order to overcome this, some communication system is used for student safety and parent awareness. But this does not provide sufficient services to the parents.

This research is capable enough for providing services through IoT. This system tracks the school bus with the help of GPS and sends the message to the parents via GSM. RFID system records details about in and out from the school bus. Prediction algorithm is attached for detecting the location of the school bus and arrival timing of the school bus via android application so that parents can get the information about boarding and arrival of their wards.

5 Existing System

Ying-Wen Bai et al. did the research on NFC function and NFC tags to determine the authorized individual which provides user interface to manage school class table [3]. NFC tag ID is used to construct school class table which is already stored in the teachers' phones. When authorized person wants to pick up their child, the respected teacher's phone scans the authorized person's NFC tag ID that verifies with the school class table of children. This method gives a reliable mechanism while picking up children from schools.

Ajit Jadhav et al. did the research on smart school bus system [2]. This paper describes about RFID tags and GPS module to track down the school bus location. Panic button used in the emergency situation which is connected to the Arduino board that creates the link between sensor and the Arduino board. Android application is used as a user interface which interfaces with the system.

Gowthamy J et al. dealt with school bus monitoring system which was developed to give extensive new features [1]. The school bus tracking gives safety transportation to children that consists of an accurate equipment and also provides online attendance, drunk and drive message.

R C Jisha et al. developed a mechanism with arrival time of school bus with the help of android application and tracking mechanism of school bus using RFID tags, GPS and GSM module for continuous alert about school bus location and predicted arrival time [4].

Anwaar Al-Lawati et al. presented about pick up / drop-off of the children to establish more safety to children from and to the home and school [5]. This consists of two units that is school unit and bus unit where school unit identifies which of the children is not boarded yet and bus unit identifies when a child leaves the school bus. This system provides promising school bus transportation.

Prashanth K P et al. proposed the idea and project to detect alcohol content if driver has consumed. There is a mechanism implemented to avoid / detect accident and send alert messages to the nearby hospitals and school management regarding accident and its location [6].

Sabira Khanam et al. did the research on RFID tag detection using antenna which is efficient and improved formulation to fulfill the requirements for RFID tag detection and resulted from mathematical and extensive simulation analysis. Demonstration shows significant enhancement over traditional approach from detecting tag and power allocation technique [7].

Micheal Drieberg et.al did the research on vehicle tracking using GSM and GPS. Sensor with GPS used to track location of the vehicle and GSM sensor to send alert message on the location of the vehicle [8].

6 Proposed System

The proposed system is designed to accommodate the following features and facilities.

- Parents will know basic information about the driver and the location of the school bus.
- This model contains the online attendance system with time and date which helps the school administration to keep count and account of children as soon as children enter the bus.
- This model sends the message regarding the location of pick up and drop of the students with time and date.
- RFID card is embedded in each of the student's bag or RFID cards are given as student's identity card. When students enter into the school bus student id is recognized by RFID reader of the system.
- RFID reader reads via radio waves and displaying the student name on LCD.
- It detects alcohol content in the bus if the driver or anybody has consumed and sends to the administration department.
- The metal sensor detects metals like gun, knife, iron rods and any other notified weapons and reports to the administration.
- Smoke sensor is employed for detecting fire and smoke in the bus which sends a message to the administration and nearby fire stations if there is any alarming situation.

As there are many disadvantages in the existing system where the proposed system will reduce and overcome the disadvantages of the existing system.

7 Problem Statement

Scope arises on looking into the striking factors studied and defines the promising implementation.

- This model provides many facilities like SMS alert, detection of alcohol consumption, alert on fire and accident, and missing children.
- This allows parents and management to check the status of the secure smart school bus by IoT
- This gives notification about inappropriate pickup and drop which is helpful for child security.

Hence the problem statement ensures the need of a new system to track and monitor the school bus all the facilities and features identified.

8 Methodology

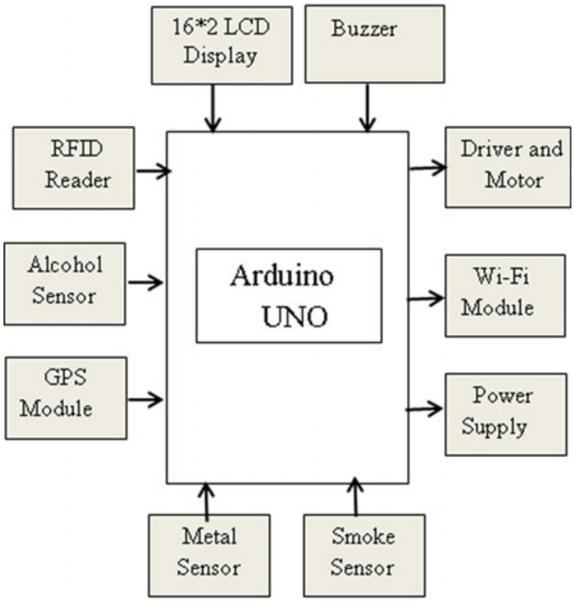
In the proposed system, there is a wireless module where the main board receives the signals from the sensor implemented and sends the signal to the wireless board. There are four main elements i.e. Arduino kit, Sensors, RFID, and android application. This system has the following components.

- Connect the sensors to the kit: At first, all the sensors are connected to the Arduino board. After connecting it start functioning and sensors will send their values to the boards and the system responses based on the value and responses Arduino reacts.
- Arduino UNO board: It is the main unit that acts as a bridge between the sensor and the mobile. Sensors are connected for further processes.
- Wireless module and application: Wireless module helps the user to get connected to the kit and processes the entire data which is transferred to the user using the wireless module. The android application used as a user interface that can connect the system. Using the android application the values sent via a module that can be seen through an android application.

8.1 Block Diagram

In Fig. 1, every sensor has its functionality where Arduino UNO is also known as ATmega328 which has 32 KB built-in memory and also it consists of 28pins.

Fig. 1 Block diagram



9 Overview of Technology

In our proposed system, the following components are used.

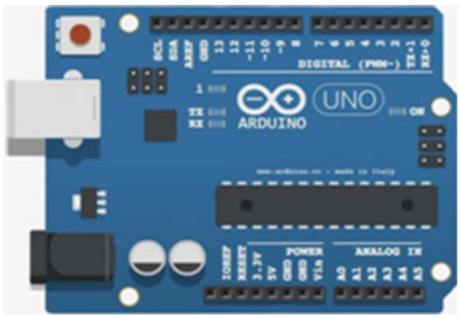
Arduino UNO

It is a microcontroller kit of ATmega328 and it has 28 pins totally where 14 pins are input pins and 14 pins are output pins like reset pin, etc. shown in Fig. 2. Arduino UNO is an inexpensive, open-source, cross-platform.

Alcohol Sensor

This sensor detects alcohol content in the breath which acts like a breathalyzer shown in Fig. 3. This can measure the concentration of 0.04 mg/L to 4 mg/L and also this can operate at the temperature of -10 to 50°C.

Fig. 2 Arduino UNO



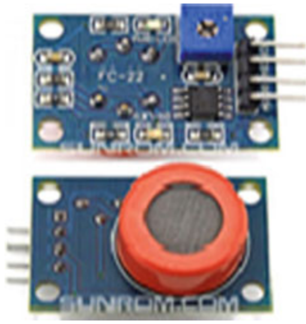


Fig. 3 Alcohol sensor or MQ3 sensor

LCD (Liquid Crystal Display)

LCD is easy to interface with microcontroller which has low cost and high flexibility shown in Fig. 4. It is a standard display compare to many other displays.

GPS (Global positioning System)

GPS is used for locating, navigating and determining time. It identifies the location at any point around the globe shown in Fig. 5. The values of longitude, latitude and height coordinates can be identified at any point on the earth.

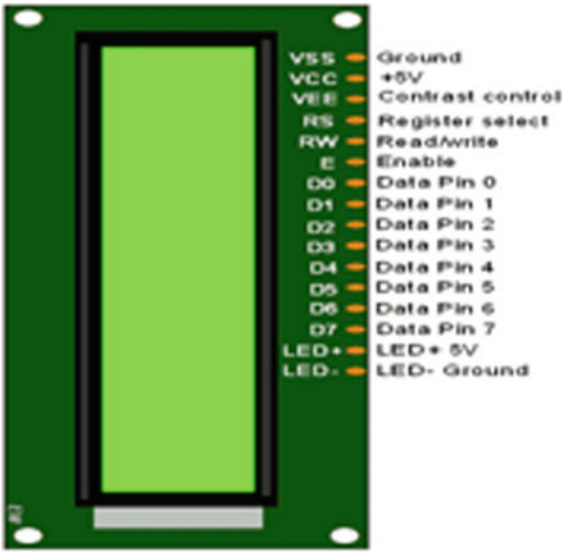


Fig. 4 Liquid crystal display

Fig. 5 GPS sensor

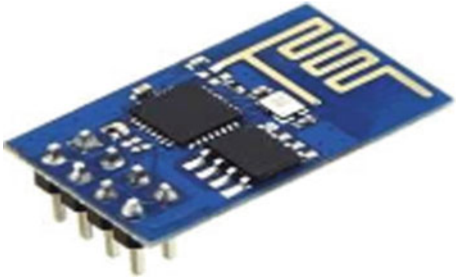


Fig. 6 Wifi module

Wi-Fi Module

There are three different supplies which are: 9 V/1Amp for GSM, 5 V/2Amp for ESP 8266, ESP 8266 wires on 3.3 V. This module converts 5 V to 3.3 V and the third one is for microcontroller, it can take supply from a USB port or 9 V battery shown in Fig. 6.

RFID Tags

RFID tags are encoded by smart labels captured by radio waves shown in Fig. 7. There are different types of RFID like Low frequency, High frequency, Active and Passive RFID. RFID code is used for tracking purposes.

RFID Reader

RFID Reader is made up of an electromagnetic field that automatically detects and tracks the tags attached to the smart cards. In the proposed system, RFID reader used to track the child location shown in Fig. 8

Arduino IDE

Arduino software (IDE) or Integrated Development Environment contains a message area, console, toolbar with buttons, etc. where this software is used to show the output of sensors.

Fig. 7 RFID tag

Fig. 8 RFID reader

C Programming Language

C Programming Language which is called as a middle-level language because it connects between machine-level operations with high-level language application. It has a small and fixed number of keywords.

10 Results

In this model, parents will get to know the basic information about the school bus driver, location, and school in and out information. This system sends the message alert to the respective parents regarding school bus location and arrival time of the school bus and reduces the risk of student waiting at pick up point.

RFID cards are implemented on the schoolbags or given it as ID cards to reduce confusion as soon as a student gets into the bus by scanning their RFID tags by the RFID Readers and displaying the student name in the LCD. This helps in online attendance system which stores the data along with date and time.

In case of the drivers have consumed alcohol, sensor detects and sends a message to the school. School administration gets the message regarding every student, unusual behaviour like alcohol consumption by anybody, abnormal occasion like fire, breakdown, accident and alarming situation with weapons if anything detected by metal sensor.

Implemented system, Display Unit and Simulation of messaging system are shown in Figs. 9, 10 and 11.

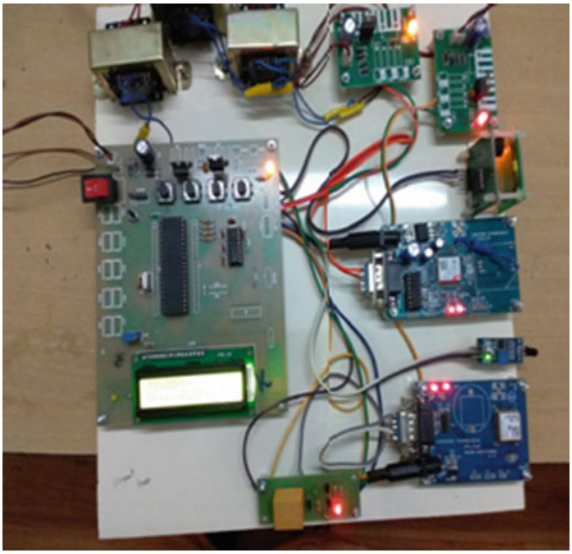


Fig. 9 Proposed system

Fig. 10 LCD module display

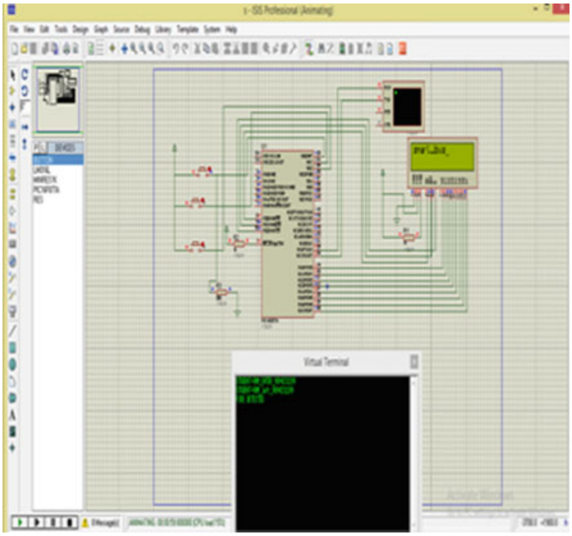


Fig. 11 Simulated Result

11 Conclusion

This system provides safe transportation for school children during their school journey and also parents to know about their children daily activities. Here GSM and GPS play an important role like tracking and monitoring school bus location and RFID provides information about the children in and out and students' attendance with date and time with the help of RFID tags.

12 Future Enhancement

Extension can be done in this model in the future is expected on the following areas.

- Implementation of biometric devices instead of RFID.
- Digitizing the images of the students who enter into the bus with the use of their RFID cards and Camera
- A provision to unlock the door only upon ID and identification by the capture of the image.

References

1. Gowthamy JAP, Patel J, Sohani T, Dutta A, Samal A, Sharma A (2018) School bus tracking and monitoring system, Department of CSE, J IAETSD, 4 April
2. Jadhav A, Ashish NN, Ranmode G, Gavali AB(2018) RFID based secure smart school bus system, Department of CSE, IAETSD J, March
3. Bai Y-W, Fu C-N, Yang J-H (2018) Using NFC tags and smartphones to design a reliable mechanism to pick a child up from school. In: IEEE, 29 March
4. Jisha RC, Jyothindranath A, Sajitha Kumary L (2017) IOT based school bus tracking and arrival time prediction. In: IEEE, 4 December
5. Al-Lawati A, Al-Jahdham S, Al-Belushi A, Awadalla M, Al-Adawi D, Al-Abri D (2015) RFID-basedSystem for School Children Transportation Safety Enhancement. In: 8th IEEE GCC Conference and Exhibition, Muscat, Oman
6. Prashanth KP, Padiyar K, Kumar N, Santhosh Kumar PHK (2014) Road accident avoiding system using Drunken Sensing Technique. IJERT 3
7. Khanam S, Mahbub M, Mandal A, Shamim Kaiser M, Mamum SA (2014) Improvement of RFID tag detection using smart antenna for tag based school monitoring system. In: International Conference on Electrical Engineering and Information and Communication Technology (ICEEICT)
8. Oat PH, Drieberg M, Cuong NC (2013) Development of vehicle Tracking System using GPS and GSM Modem. In: IEEE Conference on open System, December 2–4, Sarawak, Malaysia.

Impact of Artificial Intelligence (AI), Internet of Things (IoT) & STEM Social Enterprise Learning Based Applications in the Teaching and Learning Process of Engineering Education

K. Balaji, M. Selvam, and R. Rajeswari

Abstract This paper describes the benefits of Artificial Intelligence, Internet of Things and Social Business based Teaching methodology to foster new Innovations from the students studying Engineering education. The study shows the impact of promoting AI, IoT, Social Business in the young minds to provide a platform for the budding Engineers and the Science students to think and use the novel ideas to solve the problems in the society. The paper also discusses the case studies of successful students helped by the idea of Social Business to solve social problems using STEMSEL (Science Technology Engineering Mathematics Social Enterprise Learning) Microchip Programming methodology. The Authors use AI, IoT based Social Business Model and STEMSEL Programming technique to train the students in the area of pervasive applications. The applications and projects involving the Internet of Things have led to a numerous challenges in the research. This paper explains the use of Social Business Model and a new prototyping methodology that can achieve faster creation of working model for IoT and Pervasive application. This IDE is used to develop prototypes involving the learning of technologies like Bluetooth and Wi-Fi for pervasive infrastructures. It has shown that this IDE is very useful in the Engineering education to learn such advanced technologies in a simple and easy format. The use of such tool to create prototype and working models for the applications in Teaching and Learning process can definitely create a quick learning method. The case studies of students shown in the paper, prove that the integration of the Social Business in the Learning of Technology has created an interest and joy in learning which has the potential to change the life of poor.

Keywords Artificial intelligence · Internet of Things · Social business · Pervasive computing applications · STEMSEL microchip

K. Balaji (✉)
SSM College of Engineering, Komarapalayam, Tamilnadu, India

M. Selvam
Don Bosco Institute of Technology, Bangalore, Karnataka, India

R. Rajeswari
M.P.Nachimuthu M. Jaganathan Engineering College, Chennimalai, Tamilnadu, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_116

1 Introduction to Pervasive Computing

The Authors in India have tried the use of Artificial Intelligence (AI), Internet of Things (IoT), STEM & Social Business model to foster Innovations in creating new applications in IoT and Pervasive computing to create next generation techno-preneurs. In recent years a new class of applications have emerged that has the need to transfer high volume of data and control signals using different technologies available for communication. Such applications are found everywhere in this digital world and are referred to pervasive in nature. It is also called as Ubiquitous systems to indicate the concept "Existing Everywhere" [1]. Lot of innovations is seen in the field of pervasive application development [2]. Apart from pervasive application, Internet of Things is expected to be the technology of the future [3] that is capable of producing an impact in the life of every human. Such technologies are used in Medical, Home automation, Defense etc. The use of such technologies in these fields will definitely change the way of using the digital gadgets in traditional way. Such technologies can change the way of using things around us like a kitchen appliance or an advanced control system from a desktop way of usage to a highly mobility supported usage.

2 Literature Review

There are more number of Social Business opportunities and Research challenges expected within the area of pervasive application modeling and IoT. In [4], wide selection of IoT related applications within the field of Smart Agriculture is presented. The sensors and controllers for various purposes for agriculture are combined using different technologies within the IoT model. In [5] the IoT Application namely "Smart & Connected Health Care" is discussed. In this work, authors modeled a Acquisition system using Body area networks, Bluetooth & Zigbee to provide a transmission technology, a Mobile application as the data collector and hence the processing of the collected data which are sent to the cloud based storage.

3 Promoting AI, IoT Based Applications & Social Business

3.1 Smart Home

It is a new kind of constructing a home where the appliances used are monitored by sensors and controlled using a sophisticated algorithm. As mentioned in [6] "The total amount of funding for Smart Home Start-ups currently exceeds $2.5billion."

3.2 Wearable Gadgets

These are gadgets that can be used to wear on the body and can act as human activity tracker for real time data collection [7] using tiny electronic sensors.

3.3 Smart City

Smart city services include digitally controlled distribution of water, advanced management solutions for waste and recycling methods and the control of security in the full city using the intelligent algorithms and networking capabilities [8] built in the city.

The above mentioned applications widely use Microcontroller chips and intelligent algorithms for building Pervasive and communication intense applications. These are referred as IoT connected applications. There are other smart enabled applications in retail sector [9], Power Grid, Smart Agriculture and farming..

4 Social Business Opportunities in IOT

In the paper [6] the searches that are made in social media and the popular search engines are used to analyze the potential applications and opportunities for the prospective students. The survey results are based on searches carried by the users in popular websites like Twitter and the professional platform LinkedIn shown in Fig. 1.

The hit for these applications are used to provide rank. The top five ranks are secured by Smart Home, Wearable Gadgets, Smartly managed cities, Internet for Industries. The application in the rank one, namely the smart home is searched by sixty one thousand users over a month. This indicates the need of the huge volume of applications and hence the trained people to create and service such applications.

5 Challenges in IoT

There are some challenges in the field of IoT to solve the issues of security and compatible standard creation. The developers who can solve such issues can find better opportunities as the start-ups in this field are trying to create more sophisticated products [10]. At the same time, the students should have the right programming skill to grab this opportunity. Hence a variety of software and tools are used by training companies and Educational Institutions to provide a Hands-on experience. Although such training is provided, the industry is facing the problem of finding a skilled

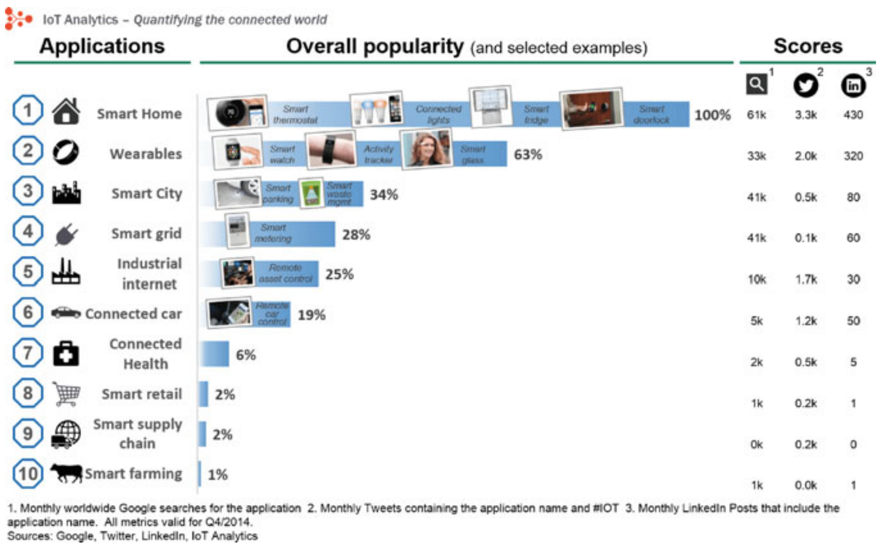


Fig. 1 Survey report on the opportunities. https://iot-analytics.com/10-internet-of-things-applications/ Source

programmer and the hardware designer for their projects to meet the rising demand. The survey conducted in [11] by CampatiaIT Industry Trends, a high demand for designers in IoT and AI is reported. It also stressed the need to develop the skill in programming to reduce the gap between the requirement and the available technical skills.

6 Triggering Social Business Case Studies in Technology Using STEMSEL & Runlinc

The authors found through various case studies from Engineering students that the addition of Social Business component through the teaching of AI & IoT in the minds of young students trigger a powerful thought to think and come out with a solution to the common man problems in the society. As the Social Business is introduced in the early stage of Technology Learning, the students start thinking about becoming Entrepreneur in a Social Business way as it is assumed to be the easiest way to become an Entrepreneur and also a way to be useful to the society and make service for the poor. The authors teach Social Business through STEMSEL Technology. STEMSEL stands for Science, Technology, Engineering, Mathematics, Social, Enterprise, Learning that utilizes the programming concepts for Teaching and Developing products for various Pervasive and IoT applications.

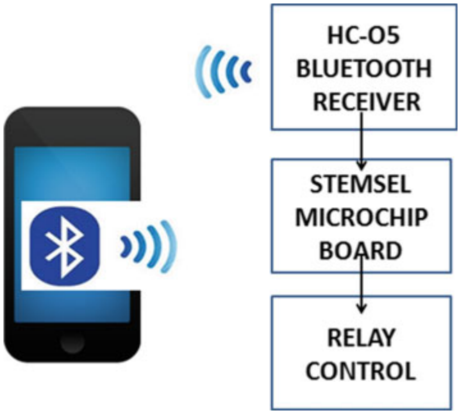


Fig. 2 Architecture of automation project

The STEMSEL methodology helps to teach & enable the innovations in a quick way. The Board uses a new Software tool for programming using an easy Graphical Drag and Drop elements for various processing related and control algorithms required for an Microchip Programming like Time Delay requirements, ON and OFF control, Decision control, Looping etc.

A Bluetooth based application created by a student for Home automation using STEMSEL is presented in the following example. The simplicity offered by the tool helped the student to program the application in a short time duration when a social business related application is intended.

The architecture developed for the Home automation project is shown in the Fig. 2. This is developed using the STEMSEL board. The main controller is modelled using the STEMSEL controller and it gets the control triggers from the user through a HC-05 Bluetooth transceiver. A mobile application is used by the user to transfer commands to control the devices.

The Fig. 2 Shows the Hardware Board for the creation of Bluetooth Based Automation system using STEMSEL technology. This prototype helps to control the home appliances in real time fashion.

7 Features of STEMSEL Board

The Home Automation project is built by the students after learning STEMSEL. It helped them to program an application quickly and apply the concept of STEM as shown Fig. 3.

The board gives an interfacing virtual facility which helps us to connect different peripherals and core chart software to create the program flow and transfer the program to chip. Figure 4 shows the technique of programming the STEMSEL board using a flowchart idea without using an assembly level language or a high

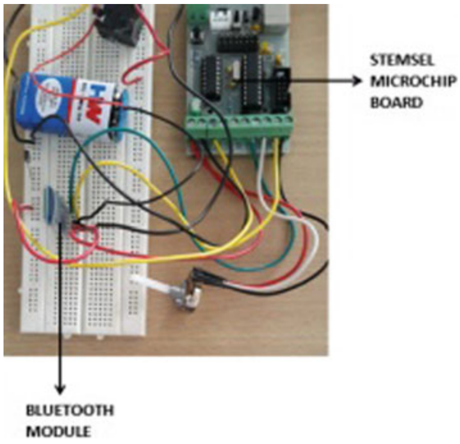


Fig. 3 Bluetooth project model using STEMSEL

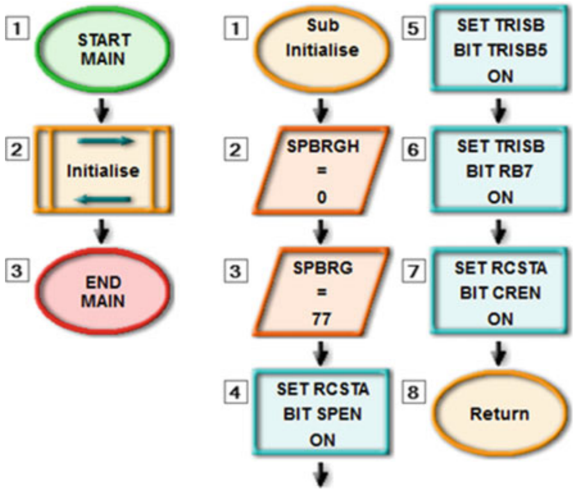


Fig. 4 Core chart sample to initialize the prototype

level language. This sample core chart is used to initialize the speed of the Bluetooth and specify the Port pins for transmission and reception as shown in Fig. 4.

The following case studies show the impact of teaching Social Business to the students at an early stage of learning. As social Business concepts are taught to technology learners, the students get an idea of solving problems in Poverty, Pollution and Unemployment. This has led to trigger innovative ideas in the minds of young students to start thinking on developing new prototypes to solve such problems. Such thoughts of Social Business and 3 Zeros created successful Technopreneurs to win Laurels in International Competition like Australia Royal Adelaide Show. By Teaching Social Business and 3 Zeros, SSM YUNUS STEMSEL Innovation centre has helped 12 Batches of students to learn Social Business in Technology by creating Project Prototypes & Working Models as shown in Fig. 5.

Fig. 5 Prototype winning best exhibit award

The Zero Net Pollution Monitor Project developed using the STEMSEL microchip board guided by the author received the Best Exhibit Award in the Australia Royal Adelaide Inventors Competition. This shows the importance of STEM and Social Enterprise based teaching in the Engineering Education.

8 Student Survey of Social Business and STEMSEL Learning

The Author has made a survey among the students to find out the impact of teaching social Business and Technology Learning and the differences of a normal way of learning in Engineering education. The study shows that the students were able to learn technologies, if it is applied to a real world scenario using social Business thought at an early stage. Even though the students are not asked to market or kick start a product development, the thought of Social Business helps them to learn fast and easy to teach.

8.1 Feedback Samples

The Responses given by the participants are available online in the following link for public viewing shown in Table 1.

https://docs.google.com/forms/d/1EfHJw7IWECq-8QNJB06G4L3cPSA-zMmqsDQ3dkvIDK4/edit#responses

The Authors after teaching the Microchip programming and Hardware interfacing using STEMSEL, conducted a student survey. The Learners belong to different age groups. The survey is used to Identify Social Business Opportunities, Solving the problems, Programming methodology, Time needed to Develop, Time needed to Debug, Time needed to Learn etc. Figure 6 shows the kind of the teaching methodology based on STEM and its application in the learning process of Engineering.

Table. 1. Feedback samples

Give your feedback about STEMSEL workshop	Email address of the participant	Overall response
Good	deepaahavi1111@gmail.com	Positive
Good	p.swathipalani96@gmail.com	Positive
Good good	karthi94id@gmail.com	Positive
Essential and Excellent Workshop	sugilingam@gmail.com	Positive
It is very useful	thangavel272@gmail.com	Positive
It is a good opportunity to learn latest technology	chamuarjun2005@gmail.com	Positive
Yes	p.maheshwaran41@gmail.com	Positive
Yes	krishkalai64@gmail.com	Positive
Essential for students to promote their skill and knowledge	shartamil@gmail.com	Positive
This type of social net work is very useful to students	sabarinathan0505@gmail.com	Positive
From the STEMSEL workshop i got more ideas about projects and the methods to implement it by using STEMSEL kit itself	r.dhayalni@gmail.com	Positive
From the STEMSEL workshop i got more ideas about the projects and how it is implemented in STEMSEL kit	nandhinivellaichamy@gmail.com	Positive
It is very useful.it helps to improve my knowledge	kavithaappusamy972@gmail.com	Positive
We execute our innovative ideas	sssuganthi98@gmail.com	Positive
No i am not worked in STEMSEL	smani0408@gmail.com	No idea
No idea	selvamani2502@gmail.com	No idea

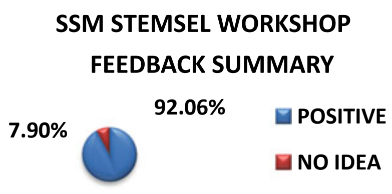


Fig. 6 Survey results showing the impact of STEM learning

Figure 7 shows the outcome of the survey pertained to the easiness in learning programming techniques. The survey results show that ninety percentage of the students were happy to use such easy technology. In this way the method used by the authors can create a positive impact in the skill development in the areas of Artificial Intelligence, STEM, and IoT Programming.

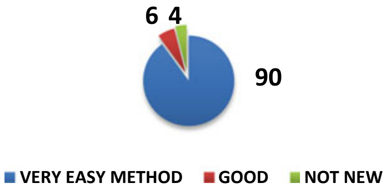


Fig. 7 Survey outcome of social business model to learn technology

9 Conclusion

This work presents a new training methodology in the Engineering Education implemented by the authors using STEMSEL & Social Business at early stage of learning in the field of Microchip programming for AI, IoT and Pervasive Communication intensive environments. The case studies and the prototypes developed in the paper focus a sample application based on Bluetooth wireless technology which were generated based on the Social Business and 3 Net Zeros. The Social Business Teaching and Learning Technology method can also help to create projects using Wi-Fi, AI, RFID etc. The method helps to remove the skill gaps that reduce the pace of IoT product development. The survey outcome shows the simplicity of the Social Business method that assists the Training companies and Institutions to create more innovations in the near future. This definitely creates talent pool in the fields of Smart Agriculture, Medical applications, and Modern digital Lifestyle. This method will help students in Engineering to be inspired to create their own STEM Inventions for Social Good.

References

1. https://www.techopedia.com. Pervasive computing
2. Guruduth B, James B, Eugene G, Jonathan M, Jeremy S, Deborra Z (2000) Challenges: an application model for pervasive computing. In: 6th ACM/IEEE International Conference on Mobile Networking and Computing
3. Steegen A (2015) Technology innovation in an IoT Era. In: 2015 Symposium on VLSI Circuits, Kyoto
4. Gondchawar N, Kawitkar RS (2016) IoT based Smart Agriculture. IJARCCE 5(6)
5. Hassanalieragh M, Page A, Soyata T, Sharma G, Aktas M, Mateos G, Kantarci B, Andreescu S (2015) Health monitoring and management using internet-of-things (IoT) sensing with cloud-based processing: opportunities and challenges. In: IEEE International Conference on Services Computing
6. https://iot-analytics.com/10-internet-of-things-applications/
7. Tasneem Usha R, Sazid Sejuti F, Islam S (2019) Smart Monitoring Service through Self Sufficient Healthcare Gadget for Elderly," 2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS). Rourkela, India
8. Kirimtat A, Krejcar O, Kertesz A, Tasgetiren M (2020) F, "Future Trends and Current State of Smart City Concepts: A Survey,." IEEE Access 8:86448–86467

9. Trappey AJC, Trappey C, Wang WC, Hsieh HI (2018) Patent analysis of Key technologies for smart retailing and their projected economic impact. In: 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design (CSCWD), Nanjing
10. Mohsen A (2017) The Internet of things: limitless opportunities for Business and society. J Strategic Innovation Sust
11. https://www.comptia.org/resources/it-industry-trends-analysis-2017

Sentiment Analysis of Twitter Data Using Naïve Bayes Classifier

S. R. Shankara Gowda, B. R. Archana, Praajna Shettigar, and Kislay Kumar Satyarthi

Abstract Nowadays social networking is creating a rush of users' data. Since Microblogging sites like Facebook, Twitter, etc. provides an easyand short way to express and share their view of millions of users daily. With increase of use of the microblogs people express their opinions on variety of topics, companies, products etc. These helps a lot for the company to propagate their business by analyzing people's opinion. The intention of this paper is to develop a model that analyzes the sentiment behind the tweets gathered from the Twitter. The challenge is the data that are there in Twitter are extremely unformed which makes difficult to study. In our model, the sentiment analysis process is as follows: (1) Collect tweets directly using Twitter API, (2) Data Cleaning, (3) Training the data into our model i.e. Naïve Bayes, (4) Extracted tweets are classified by its sentiment i.e. its polarity (positive, negative or neutral). In our work we have used a supervised machine learning algorithm i.e. Naïve Bayes Classifier.

Keywords Sentiment analysis · Opinion · Microblogging · Tweets · Naïve Bayes

1 Introduction

These days large amounts of data are within reach through various blogs. Most of the researchers and practitioners focus on sentiment analysis which is one of the booming applications in today's business sphere.

In these years, there has been an enormous growth in microblogging platforms like Twitter. Twitter is an inventive microblog started in 2006 which has more than 560 million users. Tweets are the messages created by the users. Companies and other organizations seek many ways to analyze these tweets to know the polarity (positive, negative or neutral) of the product, company etc.

Sentiment analysis is identifying and categorizing the views expressed by the people on particular topic. We classify the tweets either positive, negative or neutral.

S. R. S. Gowda (✉) · B. R. Archana · P. Shettigar · K. K. Satyarthi
Don Bosco Institute of Technology, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_117

Sentiment Analysis is a great benefit in today's business world. The companies or organizations can divert their business plan based on the sentiment analysis results.

In this paper, we extract twitter data through TwitterAPI—tweepy using hashtags (e.g. #kfc, #burgerking etc.). Then we identify the sentiment behind the tweets whether they are conveying positive, negative or neutral sentiment.

There are various approaches like lexicon analysis, machine learning approaches, hybrid method etc. Tweets has the limitation of only 140-character length. This creates a restriction whenever tweets include hyperlinks. Slangs and misspellings words in tweets is one of the challenges that has to be overcome.

We propose a work which approaches twitter data and build a machine learning model to classify its polarity.

2 Related Work

Sentiment Analysis has created a great change in business era as well as in research field. Sentiment Analysis for microblogging like Twitter has its own challenges.

Sentiment Analysis comes under Natural Language Processing at many steps of speckle. Beginning from document level classification, it has been done for sentence level and then to phrase level.

Microblogging platforms like Twitter, we get enormous data where people share their opinions about almost "all" the topics. Refining these data and extracting only the required data or tweets is actually a hurdle.

There are many other methods to classify tweets. Different approaches include Statistical method, lexicon, bag of words, machine learning approach, hybrid method etc. [1]. In our work we try to implement machine learning approach. Machine learning approach for sentiment analysis usually uses supervised machine learning algorithms. We have also deployed a supervised machine learning technique i.e., Naïve Bayes classifier to classify the tweets. Many other supervised machine learning techniques are also there such as SVM (Support Vector Machine), Maximum Entropy etc.

We can observe this with three different models: unigram model, a feature-based model and kernel-based model. They inferred that when the unigrams were combined with parts-of-speech gives the better result. And the tree kernel method topped the other two approaches.

ML (Machine Learning) and NLP (Natural Language Processing) algorithms provides more accurate results and has made sentiment analysis automated. These automated approaches require highly precise data.

For training a machine learning modelit uses the data i.e. unigrams or bigrams and they are trained to a supervised machine learning model (here Naïve Bayes Classifier) like SVM (Support Vector Machine), Maximum Entropy etc. [2].

The Paper "Twitter Sentiment Classification using Distant Supervision" 2009, provides a new method of automatically classifying tweets [3]. The paper says that data pre-processing must be highly accurate in order to get appropriate result. They also used a new idea of using emoticons that are there in tweets.

There are many areas where they use sentiment analysis. But sentiment analysis for microblogs like Twitter is a new and recent trend.

3 Algorithm Used-Naïve Bayes Classifier

A machine learning model that is used to classify or discriminate different values based on particular attributes or features is known as a classifier. Naïve Bayes is one among the machine learning classifiers. It is a probabilistic classifier used to classify the objects.

Naïve Bayes Classification algorithms are based on the theorem called "Bayes' Theorem". This is not a single algorithm in Naïve Bayes, it has many algorithms. All the algorithms in Naïve Bayes follow same principle, i.e. all feature pairs that will be classified is not dependent to one another.

An important assumption is that all feature makes an equal and independent contribution for the result. However, this assumption made are normally correct in real-time circumstances. But, assumption of independence typically works good in practice.

Naïve Bayes algorithms are easy and fast to deploy. Spam filtering, sentiment analysis, recommendation systems etc. are some of the common applications of naïve bayes algorithm.

3.1 Bayes Theorem

Bayes' Theorem is used to detect the probability of an occurring event where the probability of an already occurred event is specified. Mathematically Bayes' theorem is mathematically expressed as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

where *A* and *B* are events (1).

- Here, we are attempting to detect the probability of event A, where assuming that event B is true. Event B is also called as evidence.
- Evidence is value of an attribute of an unknown event.
- *P*(*A*) is the prior probability or priori of *A*, which means probability of event *A* before knowing probability of event *B*.
- *P*(*A*|*B*) is the posterior probability of event *B*, which means probability of event after knowing the evidence.

3.2 Kinds of Naïve Bayes Classifier

1. *Multinomial Naïve Bayes*: This type of Naïve Bayes classifier is used for problems like document classification, sentiment classification etc. This classifier uses word frequency present in documents as predictors or features.
2. *Bernoulli Naïve Bayes*: This type is almost same as multinomial Naïve Bayes classifier but here it uses Boolean variables as predictors.
3. *Gaussian Naïve Bayes*: Whenever the values of the predictors are not discrete, i.e. continuous value. These values are formed from gaussian distribution.

4 Existing System

Existing systems of sentiment analysis are statistical methods, knowledge-based techniques, hybrid approaches etc. Bag-of-words model is another approach where it converts words in documents into vectors where each word is assigned with score. Another approach is RT2M (Real Time Twitter Trend Mining System) where it operates on large stream of datasets of Twitter through query, similarity calculation of two users etc. This method stores data in a database where we require queries to access them.

Though existing system is popular, its time consuming and accuracy is low. When it's the matter of large datasets, these existing systems may generate inappropriate results.

5 Proposed Work

Machine learning is one of the main techniques that is gaining focus of the researches. The main reason for growth of machine learning approach for sentiment analysis is its adaptability and accuracy.

Sentiment analysis is mostly based on the supervised machine learning technique. In this paper we use Naïve Bayes Classifier which is also a supervised machine learning algorithm which is simple and convenient to analyze.

Sentiment analysis involves following main steps: (1) Gathering Data (2) Data Pre-processing (3) Training the Data (4) Classification and plotting the results.

In training the data, a set of data is considered as training set. Naïve Bayes model is created and trained based on this training dataset. Then the tweets are classified.

Following steps briefs steps involved in this work [3]:

(1) *First Step*: *Gathering Data*—We collect Data that has to be analyzed from the Twitter. The first step is to register in the twitter application developer's portal and get the authorization.

We need: (1) consumer_key: This is Twitter Consumer key also called as Client ID. This key is used to identify the clients. (2) consumer_secret: This is the password key of client to authorize with the authentication server. (3) access_token: This is the key given to user when the client authentication is successful. This defines the permissions that are given to client like what are the data that he/she can access and cannot access. (4) access_secret: This is the key similar to consumer_secret key. This will be along with access_token key and acts as password.
For this we need Tweepy library.

(2) *Second Step: Data Pre-processing*—In this step, the collected data is processed into required format. Data pre-processing is also known as data cleaning. Here the required key words are extracted. All uppercase and lowercase are made into a common case. Removing the unwanted texts like hyperlinks etc. Deleting unnecessary white spaces, tabsetc.
Below mentioned are the few things that has to be removed in this step:

   (a) Usernames: Usernames are mentioned unnecessarily. Usually these do not help us much to achieve sentiment classification.
   (b) URLS: URLS are not at all required for sentiment classification.
   (c) Repeated Letters: Tweets usually contains repeated letters which shows stress on particular feeling / emotion.
   (d) For example: "good" sometimes stressed as "gooood' or sometimes even "goooooooood". All three have the same meaning. Yet it is difficult to differentiate.
   (e) Hashtags: Hashtag words must be read as same as the normal words which do not have hashtags.
   (f) Punctuations and Additional Spaces must be removed.

(3) *Third Step: Training the Data*—A set of data is fed to the classifier for training. The algorithm will learn based on this data.
We use train_test_split to split our dataset into training and testing data. We have given test_size as 0.33 which means 33% of data is splitted as testing remaining as training data. Then we fit these data into our Naïve Bayes model.

(4) *Fourth Step: Classification*—This is the main part of whole technique. Naïve Bayes classifier classifies the extracted tweets into its particular sentiment.
Below is the code snippet for classifying tweets into positive, negative or neutral.

```
for text in list_text:
obj = TextBlob(text)
if obj.sentiment.polarity < −0.1:
negative _sentiment += 1
elifobj.sentiment.polarity>0.1:
positive_sentiment + =1
else:
neutral_sentiment+ = 1
```

(5) *Fifth Step: Result*-In our work, result is plotted in a pie-chart and the total number of positive, negative and neutral tweets are shown. Result is discussed further.

The main hurdle in this technique is availability of required amount of training data. At the same time, it overcomes the restriction of other existing systems that are discussed.

6 System Design

System Design is a conceptual model that defines the structure, behaviour and more view of a system.

The description for Fig. 1. is as follows, the [4] Input text (keyword) is raw tweets that are extracted using a Python library "tweepy". In Tweets Retrieval, we further filter out the tweets to be labelled so that we have the greatest amount of variation in tweets without loss of generality. We also remove non-English tweets, similar tweets, retweets etc.

Data pre-processing consists of three steps:

(1) ***Tokenization*****:** Tokenization is breaking the sentences or phrases into single meaningful elements. These elements are called as tokens.
(2) ***Normalization*****:** The elements are converted into a common case i.e. either all uppercase or all lower case. Abbreviations are replaced with its actual meaning.
(3) ***Parts-of-speech (POS)*****:** Here the words or tokens are tagged to its grammatical parts-of-speech that it belongs to (like noun, verb, adjective etc.).

Next is the classification algorithm, we have used Naïve Bayes Classifier to classify the tweets. After this we get the classified tweets i.e. total number of positive, negative and neutral tweets present in our collected tweets. These classified tweets are then represented in a pie-chart which shows the percentage of positive, negative and neutral tweets.

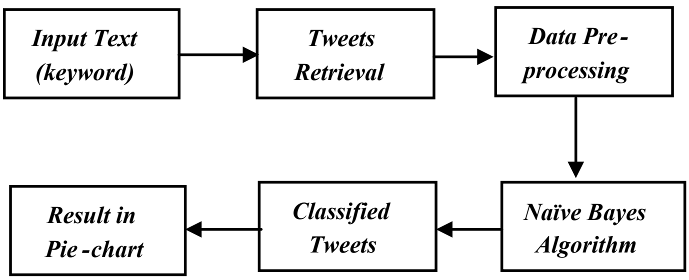


Fig. 1 Flow chart of sentiment analysis

7 Result

Figure 2 is the classification result of our proposed model. This shows that 73.5% tweets are neutral, 20.0% tweets are positive and only 6.5% tweets are negative for the restaurant "KFC".

Our model overcomes the problems of existing system especially knowledge-based system. Proposed work completely optimises the time consumed compared to previous models. Also, in this model it only considers the main words/adjectives like happy, sad, love, hate, annoying, beautiful etc. which increases the efficiency and decreases the time taken to classify.

Table.1 says that Naïve Bayes algorithm has precision of 63% which is quite high result compared with others.

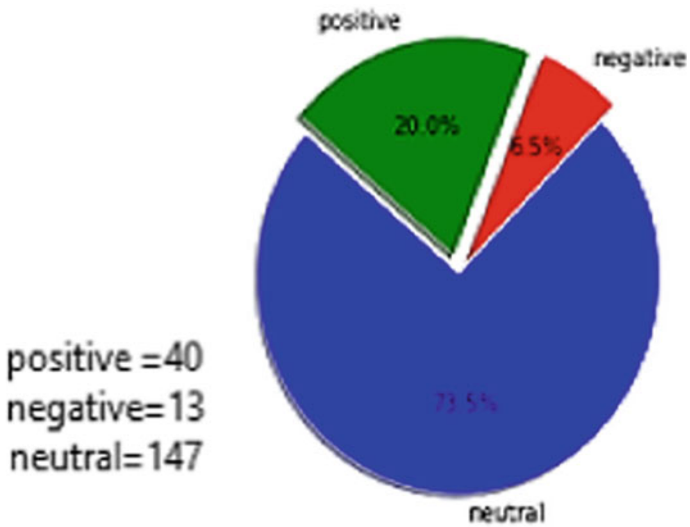


Fig. 2 Result of our proposed model

Table. 1. Accuracy result when compared with other models [4]

Algorithm	Accuracy			
	Prescision (%)	Recall (%)	Fscore (%)	Cross validate (%)
Naïve Bayes	63	56	51	41
SVM	50	33	40	56
Random forest	33	11	16	58
Bagging	50	33	40	43

8 Conclusion and Future Enhancement

As discussed before, Machine Learning techniques provide more accurate results than other techniques that are mentioned in existing system and literature survey. The proposed classification technique in our work is very simple and easy to understand. It also provides more accuracy for even small amount of data.

Sentiment Analysis in the microblogging field is still evolving and its away from complete. So, there are many things that has to be improved and consider as future work.

Our model is only developed for unigram. Further we can implement the same for bigram or trigram. But bigram and trigram require more amount data set. Also, while classifying more work has to be carried out for negations and on parts of speech.

References

1. Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau R (2009) Sentiment analysis of Twitter data, EACL
2. Zhang L, Ghosh R, Dekhil M, Hsu M, Liu B (2011) Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis, 2011 HPL
3. Krithika M, Woonna S, Giri P (2016) Sentiment analysis of Twitter data. IJIET 6(4). ISSN 2319–1058
4. El_ Rahman SA, AlOtaibi FA, AlShehri WA (2019) Sentiment analysis of twitter data, IEEE

A Robust Layered Security Approach Using Fogcomputing

B. E. Manjunathswamy, T. Pavithra, Niveditha, K. R. Pooja Shree, and M. Manasa

Abstract Cloud computing is mainly used for storage service. The security issues such as privacy breach, malicious modification and loss of data make it more vulnerable whereas fog computing is a recent research tool which is trending to bring cloud computing services to network edges. Fog computing is been used as a storage employment in multiple clouds. To make the cloud storage more reliable, in this paper we have used AES and Base64 algorithm, where AES algorithm is faster and can encrypt data in blocks that is also suitable for image encryption. Base64 algorithm converts the binary encrypted data to a string. Many studies have proved the AES encryption is most secure method for data encryption. Also, we have implemented a special feature of splitting data and storing it on distinct cloud severs. This makes the proposed scheme more reliable for storage of data.

Keywords Fogcomputing · AES algorithm · Base64 algorithm · Splitting

1 Introduction

Fog computing is an improving technology which includes all the benefits that the cloud computing technology provided and also includes new things that the cloud failed to achieve. It is very important to protect the data from unauthorized access, fog computing plays a major role in securing data by providing many tools and techniques to protect the data.

In the recent years, many different industries have seen a good progress in the system improvement which offers huge bandwidth to the end users where the users share the information, it can be of text, image, audio…etc. When these contents are transmitted security issues takes place. And, it is necessary to protect all this sensitive information from the illegal users. To protect these data encryption and decryption takes place. Encryption is divided mainly into two types Symmetric and Asymmetric, where Asymmetric uses private and public key to conduct encryption and decryption process. If a user uses a symmetric key algorithm, he can use only

B. E. Manjunathswamy (✉) · T. Pavithra · Niveditha · K. R. P. Shree · M. Manasa
Department of CSE, Don Bosco Institute of Technology, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_118

one key and if the attacker can guess that particular key using brute force approach, then he can easily find out the original message from the cipher message using the key. So, a more secured algorithm has to be used, which can construct and use a key which is hard to guess.

In this proposed scheme, we have used AES algorithm. Earlier schemes have been proposed using, XOR combination algorithm which is used to conceal data but this algorithm is too slow and may also result in loss of important data to the servers and will fail to protect data against destruction, unauthorized access and modification.

Considering these flaws, in the current scheme, we have used AES algorithm which is fast and designed to encrypt and decrypt data in block which is also suitable for image encryption. AES is abbreviated as Advanced Encryption Standard and also known as Rijndael. This algorithm supports 128 bits of fixed length block size and variable length key size of 128, 192 and 256 bits.

The main objective of the paper is multimedia content takes a huge space. When too many files are encrypted it takes huge time. In order to get the suitable encryption and decryption technique which takes less time. Secondly, security issues where the data that has been stored in the cloud should secured from cyber-attacks. Second algorithm that is been used is Base64 which converts binary data to text format that could be transmitted into environments where it can handle text safely. It is sometimes referred to PEM [Privacy enhanced Electronic Mall. The encrypted data will then be split into three different blocks. Those blocks are prepared for storage purpose.

2 Related Work

D. Zissis and D. Lekkas [1] explains the solution for cryptography based on public key infrastructure, which operates basically with Single Sign On (SSO) and Light Weight Directory Access Protocol (LDAP), which make sure the authentication, integrity and confidentiality of the data that's been involved. So, this paper tells us about the horizontal level of service, and the available entities, security mesh is maintained.

Garima Gupta, P.R Laxmi and Shubhanjali Sharma [2] explained about few of the security issues in cloud in different aspects which includes insider attacks, outsider attacks, loss of control, loss of data, multitenancy, network security, elasticity and availability. It also contains different security scheme and methods available for cloud security. It delivers the idea on various security issues and tools to the professionals and researchers.

M Preetha et al. [3] introduce three-layer privacy preserving cloud storage scheme using fog computing. By making use of this scheme, we can partition the whole data into three parts of data. Each data is encrypted and stored in the three-layer cloud. By using this method, we can easily overcome with the location awareness and we can protect the data. The lost data from the cloud could be recovered by introducing bucket where lost data can be recovered and complete set of data can be obtained.

Md MartuzaAhamad and Md Ibrahim Abdullah [4] tells about the encryption techniques. Where AES is one of the best and it is also time efficient than others. If we compare both symmetric and asymmetric cryptography, symmetric takes lesser time than asymmetric technique.

M A Manazir Ahsan, Ihsan Ali and Muhammad Imran [5] proposed a fog-centric secure cloud storage scheme where the data will be protected against unauthorized access. The proposed scheme apprises a new technique to prevent illegitimate access, which is XOR combination to conceal data. Block Management outsources the outcome of XOR combination to secure better recoverability in data loss and to prevent malicious retrieval of data. Hash algorithm has been proposed in order to facilitate modification, detection with higher probability.

T. Wang et al. [6] says about fog-based solutions for the better secure cloud storage and to protect the data against cyber threat. Based on Tian et al.'s proposed scheme the author proposes a secure cloud storage device.

Somchai, Wen Dong [7] says security of data information has become a very big problem that need to be solved. For ensuring security data encryption and decryption is required. Considering these flaws, they proposed the principle of encryption and decryption using Base64 algorithm.

3 System Model

The model proposed is three-layer architecture. It comprises of an end user, fog server and cloud servers. As per the designed system model the end user is any user who wants to use cloud severs for data storage.

The data is uploaded by the end user for storage. To maintain confidentiality of the data that has to be stored on cloud we perform few encryptions and split the data to be stored. These processes are performed in the fog layer, a backend server. The split data is stored on different cloud servers to provide a robust storage.

End User: The person who wants to store or retrieve data in a secured cloud server is the end user. The user may store text, audio, image or a video file. It is ensured that the user is authenticated and do not allow unauthorized person (Fig. 1).

Fog Layer: Fog is a decentralized layer of network environment that is closely related to the cloud server. Here we implement fog on the backend sever, where the encryption and decryption process take place. Both encryption and decryption are done using AES and Base64 algorithms. The encrypted data is then split into three different blocks. This block is the prepared for storage. The same process is then carried in retrieval.

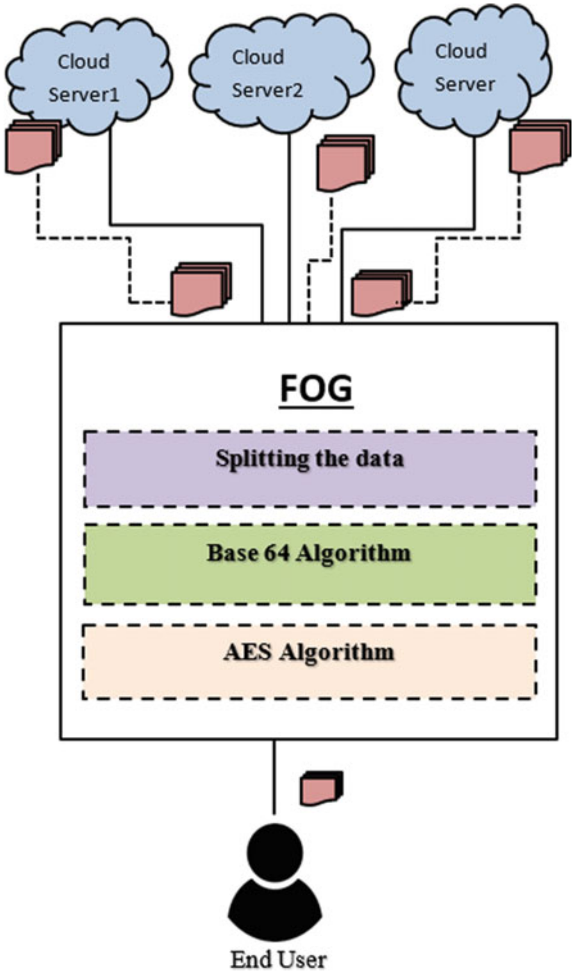
Cloud Server: The cloud server is hosted for data storage.

Also, system model includes encryption, decryption and splitting scheme. Security in the proposed system is provided using two encryption algorithms. They are AES and Base-64. Also, along with encryption of data the file is split into three blocks to prevent data tampering by any server end attack.

AES Algorithm: Advanced Encryption Standard is most widely used symmetric encryption algorithm. It is considered vulnerable against key search attack.

The features of AES are as follows:

Fig. 1 System model



- Symmetric key symmetric block cipher.
- 128/192/256-bit keys Stronger.
- Less process-time.

Working

AES is a 'substitution permutation networks' using iterative method. The working involves series of linked operations done by shuffling and replacing input bits. In AES, computations are done in bytes. The data given to the algorithm is converted into bytes and are arranged in a square matrix. For 128 bits of plaintext a 4x4 matrix is formed. The number of rounds varies with length of key. For each round different key is generated using the original AES key.

Encryption

- The data is divided into blocks: The plaintext to be encrypted is divided into blocks. If the plaintext is in bits its converted into byes to form a matrix. For instance, 4x4 matrix of 16 bytes is formed for 128 bits of plaintext.
- Key Expansion: A new key is required for each round of encryption; in this step these keys are generated using the initial key. These keys a derived using a simple random key generation method, one of the fastest ways to generate keys.
- Add Round key: this is the first round of the initial key. The initial key is added to block message using XOR cipher which is an additive method. The addition is done by converting the message and key to binary.
- Substitute bytes: In this step each byte is substituted with new byte according to the predominant table which is available.
- Shift rows: In this step the rows are left shifted by one.

  – Row one will not be altered.
  – Row two is left shifted by one position.
  – Row three is left shifted by two positions.
  – Row four is left shifted by three positions.
  – By shifting the initial matrix with respect to each row, resultant matrix of 4x4 of 16 bytes is obtained.

- Mix columns: Here the resultant matrix of the step shift rows is taken as input. Each column of four bytes is feed as input to a defined mathematical function to produce a new column of four bytes, which will replace the original column. The same process is repeated with all column to produce a new 4x4 matrix. Note: In the last round of encryption this step will not be performed (Figs. 2 and 3).
- Add round Key: In this round, the key generated in key expansion and the resultant matrix of Mix Column is XORed. The obtained 4x4 square matrix of 16 bytes is converted into bits and is XORed with the round key. The encrypted text will be obtained in the last round of encryption. Else, the resulting bits will be fed to another similar round.

Decryption

The decryption is performed in reverse order of encryption with inverse methods. The steps are in the following order:

- Key Expansion: The algorithm for 10 rounds of round key is used such that each byte is moved once in the matrix.
- Inverse Add round key: In this step XOR operation of cipher text and intermediate expanded key is performed.
- Inverse Mix columns: Along with shift rows step, the mathematical function is reversed.
- Inverse Shift rows: The rows are right shifted by one.
- Inverse Byte substitution: This step the rows of matrix are replaced by the entry of Inverse S-BOX.

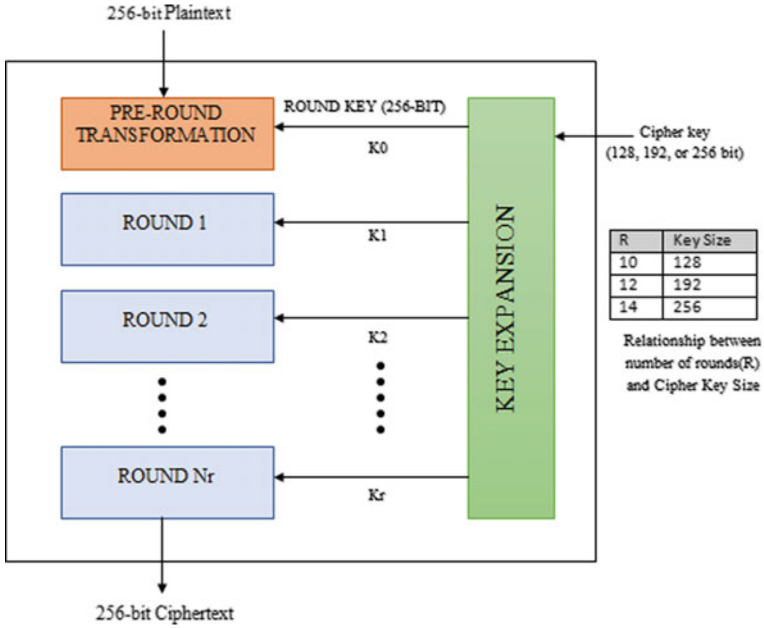


Fig. 2 AES scheme diagram

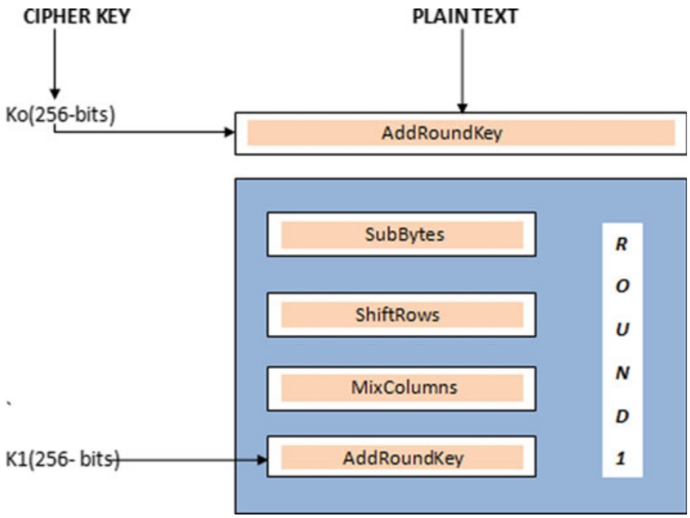


Fig. 3 Rounds of AES

Base-64:

The data obtained from the AES encryption is then encrypted with Base64 algorithm as it the one the best methods to encode bytes. Base64 encodes the original binary databy dividing it into tokens of three bytes. A byte consists of eight bits, so **Base64** takes 24bits in total. These 3 bytes are then converted into four printable characters from the ASCII standard.

Decoding of Base64 is done by converting the six-bit bytes into eight-bit bytes. The data is split into 8 characters and the decoded. Each byte is considered as ASCII can we can refer to the table for decoding the data (Fig. 4).

Splitting Data:

The data is split with the following equations:

- a=n/4
- block1=str [0: a]
- block2=str [a: 2a]
- block3=str [2a: n] where n is length of the string obtained after Base64, str is the string. We split it into three blocks for convenient storage.

The above steps use the length of the cipher text to split the data. The first block of data is from 0th byte to n/4th byte. Second block consists of n/4th byte to n/2nd byte. The last block is n/2nd byte to nth byte. This is an important step in the proposed

Value	Char	Value	Char	Value	Char	Value	Char
0	A	16	Q	32	g	48	w
1	B	17	R	33	h	49	x
2	C	18	S	34	i	50	y
3	D	19	T	35	j	51	z
4	E	20	U	36	k	52	0
5	F	21	V	37	l	53	1
6	G	22	W	38	m	54	2
7	H	23	X	39	n	55	3
8	I	24	Y	40	o	56	4
9	J	25	Z	41	p	57	5
10	K	26	a	42	q	58	6
11	L	27	b	43	r	59	7
12	M	28	c	44	s	60	8
13	N	29	d	45	t	61	9
14	O	30	e	46	u	62	+
15	P	31	f	47	v	63	/

Fig. 4 Base64 index table

scheme. The data is split such that the integrity of data is not affected. Also the data confidentiality is maintained as, retrieving data becomes harder when only a part of the data is available.

4 System Design

As per the proposed system model the end user is one who wants to store or retrieve his data in secure cloud storage. The user before storing his data has to prove him as a legitimate user. Hence the user has to have a verified email-id and password. With this the user can sign-up and he can use this email-id and password to have a secure sign-in each time signs-in. There many ways in which an attacker can guess the email-id and password to breach this security level. Hence there is other level of authentication for the user where he has to send a request for an OTP to his registered number. Then the server will send him the One-Time password which will be entered on the login-page. This OTP is verified and if correct the user is allowed to view his cloud storage. Else the user will be blocked from using or viewing the data uploaded (Fig. 5).

After the user is authenticated, he can upload or download file (audio, video, text or image). The user selects his file from his system and upload it. Then, file undergoes encryption process. The file is encrypted twice, first time using AES encryption and second time using theBase64 encryption. In the AES encryption the file is divided

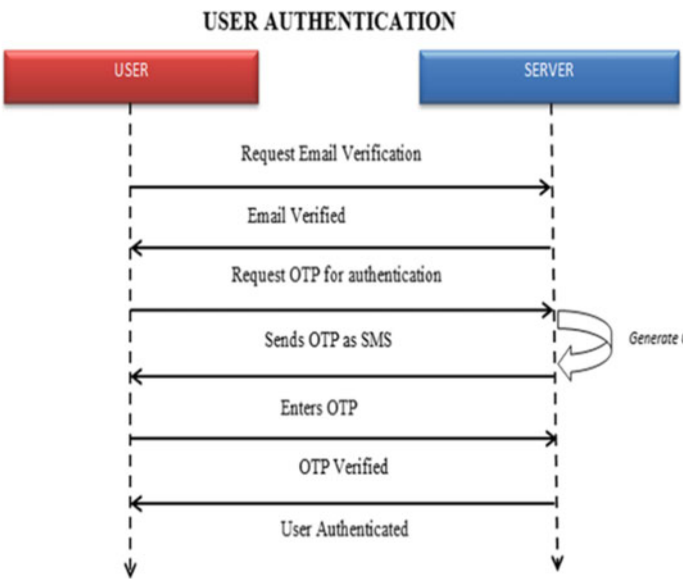


Fig. 5 User authentication scheme

into data blocks of size 128 bits. Then a key is chosen and we come up with many round keys for each round. The key is variable for each file and each user. This key is XORed with the data blocks; each byte is substituted with the data from Predominant table. After the byte substitution, the rows are shifted by one to left. Then the columns run through a mathematical function. The last step is to XOR the output of column function and the first key obtained.

After encryption with AES encryption the data is encrypted with Base64. Base64 is chosen as it is convenient to encode bytes with Base64 than any other encryption method. After encryption the data is split into blocks. The splitting process is done with the above-mentioned method of slicing.

For the second decryption i.e. AES decryption is done by inverse functions of encryption process. The XOR operation of cipher text and intermediate expanded key is performed. The rows are right shifted by one. The rows of matrix are replaced by the entry of Inverse S-BOX.

The user can now download and view his file. If at the server end data was changed or deleted, he could recover the original data if and only if the data backup was completed on backend. Also, if the data was tampered in the cloud server the user is notified through a mail that his data has been tampered.

5 Proposed Scheme

Data Storage and Retrival:

Data Storage: This step showed in Fig. 6 stores the data uploaded by the user to cloud. The user proves to be authenticated and log-in to the upload page. The user selects to upload any file text, video, audio or image. The uploaded file is loaded onto the fog for encryption and splitting. The data is initially encrypted using the AES Algorithm. The output of AES is then converted into a string using Base64 encryption. The encrypted data is split into 3 blocks of data using splicing. This data is now stored in three different Cloud Servers.

Data Retrieval: This step majorly includes decryption and concatenation of data. The data is requested by the user by clicking on the download option. As the data has been requested the fog layer retrieves the data from the cloud where it was initially suited. The retrieved data is a random block of encrypted data. The blocks are concatenated into a single block of cipher. This cipher first undergoes Base64 decryption. The Base 64 decrypted data is then feed as an input to AES decryption. The data then is decrypted using AES decryption algorithm. With the end of this step the data is completely retrieved from cloud and is now available and downloaded by the user.

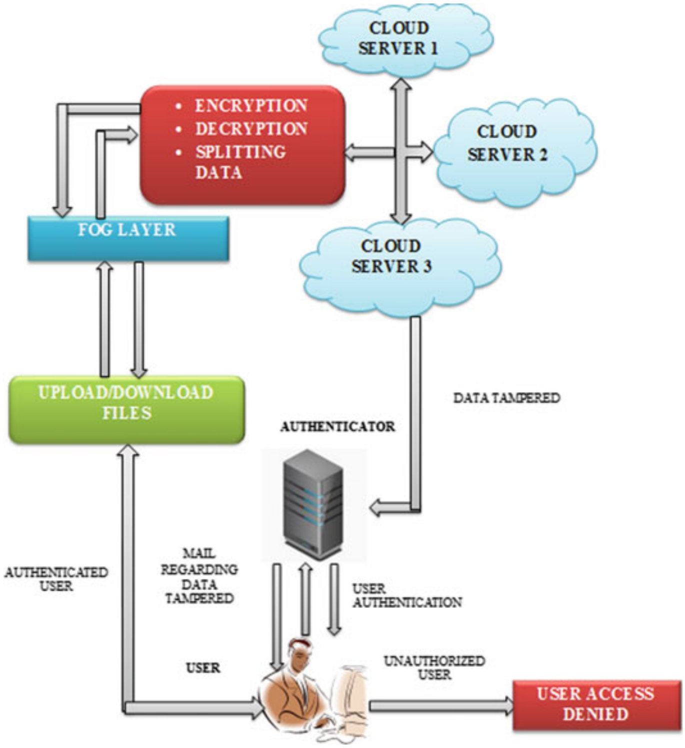


Fig. 6 System architecture

6 Result

In this proposed scheme we have designed a strict model for user authentication. As the user before storing his data has to prove him as a legitimate user so that there would be no problem from the user. This is also to make sure that the user is not malicious and does not try to access other user's data. Also, we make sure that others will not be able to access the user from front end.

The data is encrypted twice using AES and BASE64 algorithm where AES algorithm is an efficient and strong encryption method. AES algorithm is the most secured protocol and as it uses longer key sizes i.e. 128, 192 and 256 bits for encryption, it is more resistant towards hacking and no can hack personal information. Base64 algorithm converts the binary encrypted data to a string. If any malicious user tries to retrieve the data breaching the cloud server security, the data available to him would be in encrypted form. As the data is encrypted twice in our encryption method, there are very few possibilities for the malicious user to decrypt the data. Also, the method of using password to generate key makes it impossible to guess key to decrypt the data (Fig. 7).

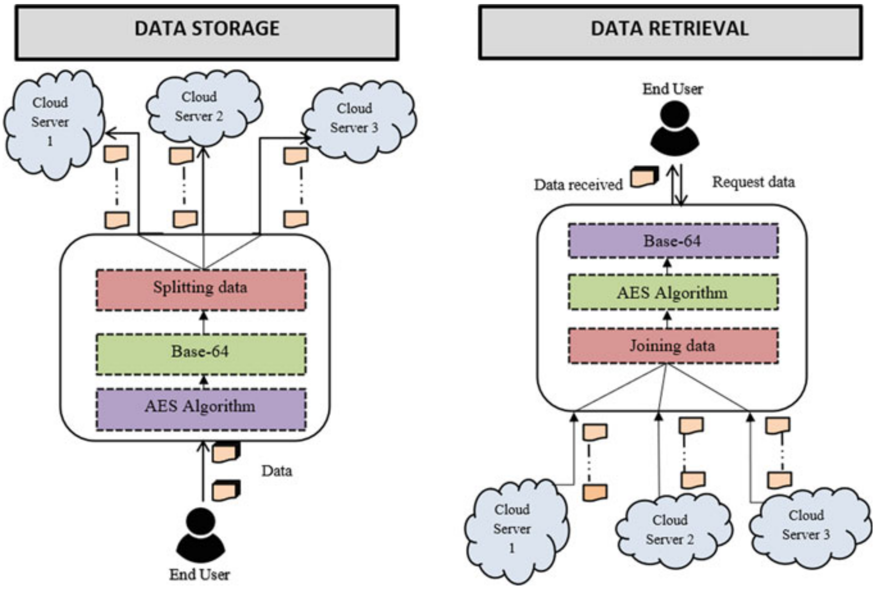


Fig. 7 Data storage and retrieval

If the attacker manages to decrypt the data, he will not be able to retrieve the whole data, as we have used the splitting method to split data and store it in three different locations. If the malicious user could decrypt and retrieve data from cloud server, he will retrieve just a part of the data. To retrieve the whole data this malicious user needs to find three different pieces of data. He will not be able find all locations of storage. This makes the proposed system more secure to store data on cloud server.

7 Conclusion

The user authentication is treated as one of major characteristics. This makes the proposed scheme strict model, such that not all people gain access. The data is still vulnerable to many attacks and malfunction which permanently damages the data. So, we have proposed this system in order to provide a secure solution for robust cloud storage to prevent cyber threats. The encryption methods AES and Base-64, makes the model more secure. Also, as the special feature is to store data in cloud by splitting into blocks makes it hard to retrieve data. The splitting of data could be done in many ways, but we have chosen a simple and efficient method. This makes the scheme efficient and fast.

The proposed system helps in preventing data loss. In this scheme, we maintain confidentiality of data by securing data from unauthorized access and from data loss and the hackers cannot access the sensitive information that has been stored in the server.

References

1. Zissis D, Lekkas D Addressing cloud computing security issues. Future Generation computer systems.
2. Gupta G, Laxmi PR, Sharma S A survey on cloud security issues and techniques
3. Preetha M, Pravinkumar G, Sabari Nathan S Three-layer privacy preserving cloud storage scheme based on computational intelligence in Fogcomputing
4. Multimedia M, Ahamad M, Abdullah I Comparison of encryption algorithms for Multimedia
5. Manazir Ahsan MA, Ali I, Imran M A fog-centric secure cloud storage scheme
6. Wang T, Zhou J, Chen X, Wang G, Liu A, Liu Y (2018) A three-layer privacy preserving cloud storage scheme based on computational intelligence in fog computing. In: IEEE Transactions on Emerging Topics in Computational Intelligence
7. Somchai WD Research on Base64

Deep Learning Based Malware Detection for IoT Devices

N. Naveen, Mohammed Asim Safwan, T. G. Manoj Nayaka, and N. Nischal

Abstract Internet of Things (IoT) in military environment for the most part comprises of a different scope of Internet-associated gadgets and hubs (for example clinical gadgets to wearable battle outfits), which are an important objective for digital crooks, especially state-supported or country state on-screen characters. A typical assault vector is the utilization of malware. In this paper, we present a profound learning- based strategy to recognize the Internet of Battlefield Things (IoBT) malware by means of the gadget's Operational Code (OpCode) arrangement. We transmute OpCodes into a vector space and apply a profound Eigen space learning way to deal with group malevolent and benign application. We additionally exhibit the strength of our proposed approach in malware identification and its manageability against garbage code inclusion assaults. Finally, we make accessible our malware test on GitHub which ideally will profit future exploration endeavors (for example for assessment of proposed malware recognition draws near).

Keywords IoT · ML · Deep learning · IoBT · OpCode

1 Introduction

In general Internet of Things (IoT) architecture consists of connected devices and a wide range of Internet networks. Where all devices are connected together, some of the smart objects are sensor and actuators, it is embedded into the systems, so the system will inevitably transfer and stores the precise data and it will fetch the data whenever it is required to process it. Some of the important applications of IoT systems are benefitted to the customers with respect to the domains are, smart-city application, Agricultural-application, smart-transportation system, Electrical-management system are usually benefitted on peoples towards the improvement of civilians. The use of IoT in the field of military defense is called IoBT (Internet-of-Battle-field-Things) [1].

N. Naveen (✉) · M. A. Safwan · T. G. Manoj Nayaka · N. Nischal
Department of CSE, Don Bosco Institute of Technology, Bangalore, India
e-mail: naveensetty@dbit.co.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_119

Environments running IoT have a lot of privacy and security concerns. However, IoT and IoBT share a lot of cybersecurity issues. The confidentiality of these IoBT stations (military and defense) also increases the risk of IoBT devices and stations becoming targets for cybercriminals. The IoT facilities are directing and to be sponsored by people from other countries as well as adversaries. Malicious code detection and prevention of malicious code attacks on IoT facilities is an active research area [2].

A constraint that delays the development of powerful malware detection. Generally, IoBT has many devices, hardware as well as custom OS to operate the system efficiently. Most IoT device malware attacks exploit low-level vulnerabilities in damaged devices.

2 Related Literature

Static and dynamic analysis are broadly categorized methods for detecting malware. Static detection includes byte arrays, identification of sequence in op-code, N-gram analysis and graph navigation of control flow. These methods are checked by statistics and detect the program if it is healthy or malwared. In dynamic malware detection, the program is to be classified as congenital or malware, when the programm is execute, it is going to be preset or controlled in the environment, and the execution path is the requested permission. Study such behavior.

David proposed a framework called Deep sign that automatically detects malware based on signature generation methods. When Deep sign calls the API, it will generate a dataset in the sandbox based on activity logs such as web search, number and details of ports accessed, registry entries, etc., and convert the log of these data into a binary vector. To do. An accuracy of 98.6 was achieved using a deep trust network for their classification [3].

In another study, Pascanu proposed a method of modeling execution using natural language processing. We then used a neural network to extract the appropriate functionality and predict API calls based on a record of past events. A true positive rate of 98.3% and a positive rate accuracy of 0.1.1% were achieved [4].

Demme, investigated the utility of building malware detectors from IoT node hardware using performance counters and K Nearest Neighbor, Decision Tree and Random Forest classifiers with learning capabilities. The accuracy achieved was between 25 and 100% [5].

Alam applied an RFC sorter to a dataset of smartphones connected to the Internet to recognize malicious code. We ran APK (Android Application Package) on the Android-emulator, other various parameters, like memory details, networking, classification permissions, and studied and evaluated approaches [6].

The experimental result shows, classifier has the optimal about 40 tree's, with root mean square of 0.0171 was achieved.

Mahesh et al. In live streaming, it recognizes the faces using the LBPH Recognition algorithm. If the camera detected face matches with the trained datasets, then it will print the name of the person or else the face doesn't match with the faces in datasets then the algorithm takes the snap of the person and sends it to the owner through email, using MQTT [7].

3 Experimental Details

The above project is design and development of system Architecture is based on 3 modules, like (Fig. 1).

- Feature-selection based on NGram
- Data classification using SVM
- Deep learning using eigenspace

3.1 Feature-Selection Based on NGram

A data set of benign and malware samples was created for an ARM-based IoT application. We collected all these malware samples from March 2015 to December 2016 using the VirusTotal [8] threat intelligence platform. We collected all high-quality software from various certified IoT application stores, like Pi Store (Fig. 2).

IoT and IoBT applications are most likely to consist of a series of operation codes (operation codes of IoT devices), which contain instructions and addresses that need to be executed in the processing unit of the device. We use Objdump to disassemble

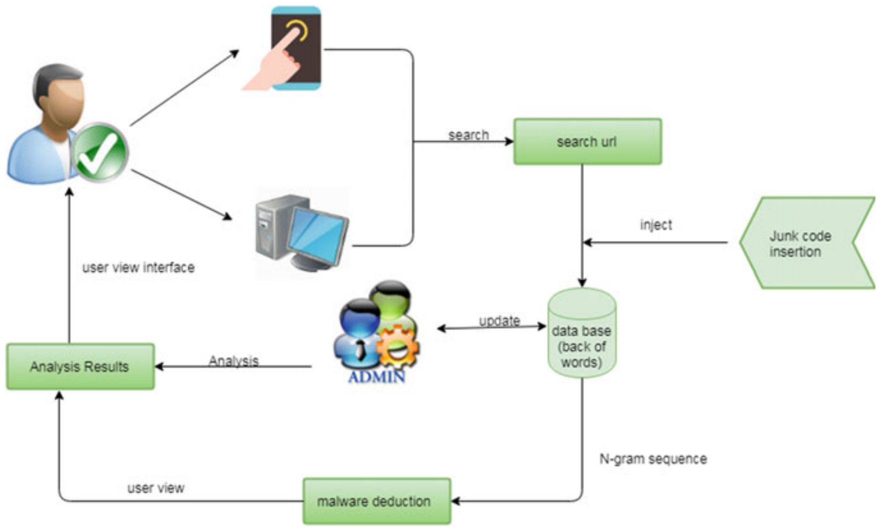


Fig. 1 System architecture

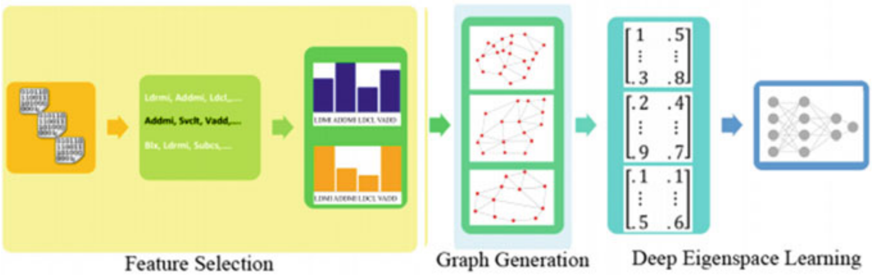
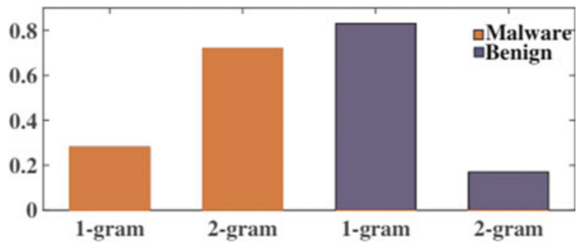


Fig. 2 Phases of implementation

the sample and extract the opcode. We created an NGram OpCode sequence to classify the malware based on its teardown code. The number of features that are important to length is CN, where C represents the size of the instruction set. From this we can infer that a large increase in N will cause the feature to explode. Additionally, reducing the function size will increase the strength and efficiency of detection, because invalid functions will reduce the performance of machine learning methods. Based on this, there is a tendency to apply feature selection algorithms to narrow the attribute set first, avoiding feature selection to find the best features. Information Gain (IG) is a widely used information retrieval technique for function selection. The features in the ranking are selected according to the amount of information available in the ranking problem. IG uses statistical tools to identify global features and does not consider selected class information. In some scenarios, such as unbalanced data sets, the global feature selection method ignores class-specific secondary features that reduce the efficiency of the system (Fig. 3).

In our research, we extracted 4500 1 gram and 600,000 2 gram different opcode sequences, and calculated their category information gain (CIG). CIG was proposed to overcome the incompleteness of global feature selection, aiming to identify more useful features based on the available class information. As shown in Fig. 1, most malware features are 2-gram sequences, and most 1-gram sequences are composed of harmless application features. This knowledge is very important for detecting malware in IoT and IoBT devices.

Fig. 3 1-gram and 2-gram feature distribution for benign and malware samples



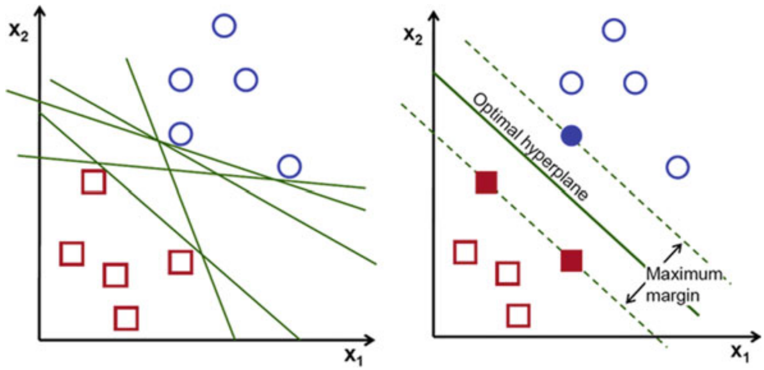


Fig. 4 Possible hyperplanes

3.2 Classification Using Support Vector Machine

The goal of the SVM algorithm is to classify the data points, where SVM algorithm is to finding a hyperplane, here N-dimensional space is used to separate the data points, where N stands for the features [8]. In general, many possibilities are there while choosing a hyperplane from isolate classes. But our goal is to find the hyperplane, with the maximum margin in data points, i.e. the distance between two data-points.

For the data-points classification, hyperplane is required, it provides the boundary level decisions, where, data-points may fall on the either side of the hyperplanes, that can be featured to other classes. In general, the size of the hyperplane is totally depending on the number of features used by the system, if the size of the input functions is 2, then hyperplane is a simple line. If the size of the input functions is 3, then hyperplane is a two dimensional plane, If the size of the input functions is exceeding to 3, it is very hard to imagine. The support-vector is a data point, it affects the location as well as the orientation of the hyperplane (Figs. 4 and 5).

3.3 Eigen Spatial Deep Learning

Graph is one of the common types in ML, even though it is very important for representing relationships between data types. Taking data in graph form for data mining and ML, algorithms consider different stages by transfer the graph into space vector. Eigenvectors and eigenvalues are two characteristic elements in a spectrum of graph, where it transforming the adjacency matrix of graph into space vector, out of which v, λ, A are signifying the eigenvector, eigenvalues, adjacency matrix of the graph, respectively.

$$A_v = \lambda_v \tag{1}$$

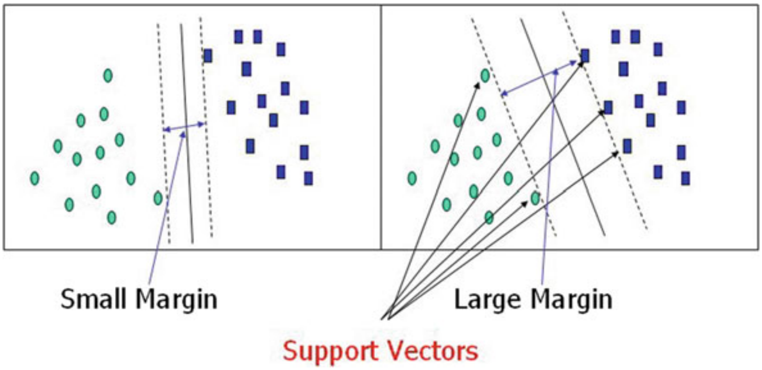


Fig. 5 Support vectors

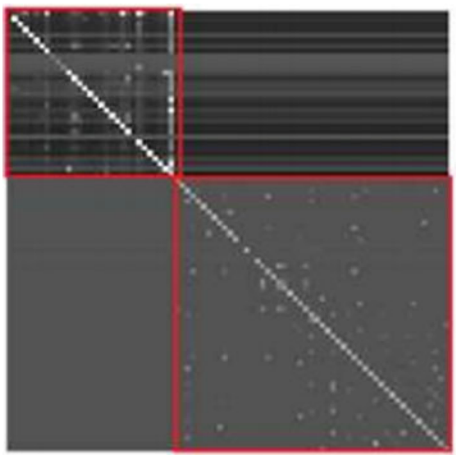
A control flow graph (CFG) represents the executable working code structure is plotted inside a graph data set representing the accumulation of all samples to gain sufficient knowledge of the structure.

Figure 6 contains 2 main diagonal-blocks, it is marked by red color in borders. These blocks show the two main distributions present in a particular sample. Based on the spectrum of graph theory, the eigenvalues of a matrix will have distinct eigenvalues.

Figure 7 shows that a gap exists between λ_2 and λ_k ($k > 2$). So the first two eigenvectors of the sample matrix contain detailed information's on matrices, and compared with rest of the eigenvectors, so the system can be represented as a whole matrix.

Also learning steps to improve the performance of v_1 and v_2, as well as the eigenvalues λ_1 and λ_2 of the malware and good ware samples, how to detect these

Fig. 6 An overview of samples—cumulation affinity matrix



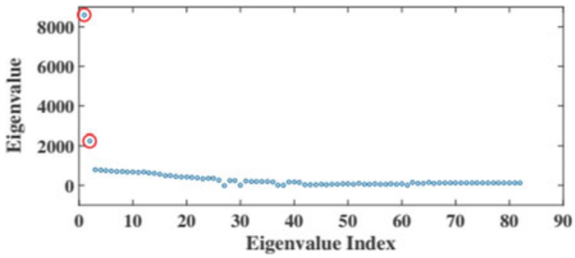


Fig. 7 Sample's of cumulation eigenvalues

sample labels. Figure 6 show the transformation between the distribution of malware and good ware eigenvalue ($\lambda_{1,2}$) data, Fig. 7 show that it is acceptable to use λ_1 and λ_2 as the features of the classification task.

4 Result Analysis

For the algorithm evaluation, accuracy and performances are the two different criterion and very important for the evaluation and identify of malware in a products/classes. The proposed approach reaches 99.86% of a height increased by 99.86%, while the proposed approach reached 98.59% and 95.91 shares respectively, are existing approaches. The detection rate is significant reference and Proposition System reaches 98.37% accuracy, compared to 81.55% and 77.70 reports of existing approaches. In our project, we use JUNK OPCODE to avoid scanning, using the standard Like Lomess Based at Features.

5 Conclusion

IoT devices are becoming more important in the future. Even though we can create absolutely be a senseless malware detection solution. We must always have the coherence of efforts to defend Google system. It is important that Cyber defender retains cyber-attackers. In this project, we have put in place Internet for malware software detection (IOBT) detection of malware, based on the selection of the Selection of the OPCODE sequence according to the classification. For the verification purpose graph is developed, the graph comprises of various features, where the system is created the sample as well as a unique apprenticeship approach was used for the classification of malicious software.

References

1. Kott AS, West BJ (2016) The internet of battle things. Computer 49:70–75
2. Gardiner J, Nagaraja S (2016) On the security of machine learning in malware c&c detection: a survey. ACM Comput Surv 49(3)
3. David OE, Etanyahu S (2015) Deepsign: deep learning for automatic malware signature generation and classification. In: International joint conference on neutral networks (IJCNN). IEEE, pp 1–8
4. Pascanu R, Stokes JW, Sanossian H, Marinescu M, Thomas A (2015) Malware classification with recurrent networks. In: 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 1916–1920
5. Demme J, Maycock M, Schmitz J, Tang A, Waksman A, Sethumadhavan S, Stolfo S (2013) On the feasibility of online malware detection with performance counters. ACM SIGARCH Computer Architecture News, vol 41(3). ACM, pp 559–570
6. Alam MS, Vuong ST (2013) Random forest classification for detecting android malware. In: 2013 IEEE international conference on green computing and communications and IEEE Internet of Things and IEEE cyber, physical and social computing, pp 663–669
7. Mahesh K, Ashok Kumar PS, Naveen Raj HN, Naik PP, Apoorv MN (2019) IoT based smart surveillance security system using Raspberry Pi. IJARIIT 5(3):1356–1358
8. https://www.virustotal.com/

Single Picture Super-Resolution Using Generative Adversarial Network

G. S. Gowramma, R. Kishor Kumar, S. Manish Kumar, and D. E. Monusha

Abstract Picture Super-resolution is a by and large analysed issue in computer vision, where the objective is to change over a low-resolution picture to a high-resolution picture. As of now, deep learning methods such as convolution neural systems and generative adversarial networks are being utilized to perform super-resolution with results competitive to the best in class. In this paper, a generative adversarial network, SRGAN, is proposed for super-resolution with a perceptual loss work comprising of an adversarial loss, mean squared loss, and content loss. The target of our usage is to get familiar with an end-to-end mapping between the low and high-resolution pictures furthermore, enhance the up-scaled picture for quantitative measurements as well as absolute quality. We at that point think about our outcomes with the present cutting-edge strategies in super-resolution, lead proof of idea division study to show that super-resolved pictures can be utilized as a compelling pre-processing step before division and approve the findings measurably.

Keywords Deep learning · Generative adversarial networks · Super-resolution · Python · PSNR · SSIM · MSE

1 Introduction

In this paper, we investigate the usage of Generative Adversarial Networks for performing single picture super-resolution on different cases, for example, CCTV (Closed Circuit TV) night footage estimations and SAR (Synthetic Aperture Radar) pictures. In most exceptional imaging applications, high-resolution pictures are liked and every now and again required to whole assignments. Picture super-resolution (SR) is a comprehensively propose issue in computer vision, where the point is to make a high-resolution picture from the low-resolution picture. Super-resolution calculation intends to make nuances better than the examining system of a given imaging gadget by expanding the quantity of pixels per unit division in an image. SR is prevalently known to be a reverse methodology, though a low-goal picture

G. S. Gowramma (✉) · R. Kishor Kumar · S. Manish Kumar · D. E. Monusha
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_120

(typically defiled by commotion, movement obscure, associating, optical mutilation, and so on. A high-resolution picture is restored [1, 2]. Super-resolution has its approach in different fields such as Satellite imaging (e.g. remote sensing) where several images of a single region are obtainable, in security and surveillance where it may be need to extend a specific point of interest in a scene (such as zooming the face of a criminal or the numbers on a license plate), in computer vision where it can upgrade the performance of pattern recognition text picture examination, bio-metric identification, fingerprint image enhancement, and so on.

The issue of creating a high-resolution picture (HR) from a low-resolution picture (LR) is an undetermined inverse problem that doesn't have a one of a solution. This is made worse by the way that a wide range of different solution exist for some random low-resolution pixel. While capturing a digital picture, there is a much loss of spatial resolution brought about by optical distortions; movement obscures because of constrained shade speed, noise that happens inside the sensor or because of transmission bringing about noteworthy contrasts between the original scene and the captured scene. Thus, aside from scaling the low-resolution picture, the Super-resolution algorithm likewise needs to represent these components. Standard errors in gaining a picture aside, CCTV footage pictures are increasingly powerless to experiencing issues like photograph fading.

This paper will investigate whether Deep Learning (Generative adversarial networks) can be used to achieve better results for Super-resolution in the SAR and CCTV images (compared to conventional methods).

The ultimate objectives of this project are to

1. Propose and Implement a broad learning procedure for a solitary picture super-resolution (SR) that straightforwardly learns a start to finish arranging between the low/high- resolution pictures and takes a low-resolution picture as the info and yield the high-resolution picture.
2. Evaluate the created HR pictures by the technique with the original LR pictures, differentiate the got output with the state-of-the-art methods and consider the reasonableness of Deep Learning for accomplishing SR in pictures.

2 Background

2.1 Observation Model

Image Resolution—The term resolution in picture preparing relates to the measure of data contained in an image that can be utilized to judge the quality of the picture and picture obtaining/handling devices. Resolution can be grouped into a several classes, for example, Pixel or Spatial Resolution, Temporal resolution, Spectral Resolution, and Radiometric resolution. For this task, we will manage with spatial resolution, and the term resolution utilized from now on will suggest spatial resolution. Spatial Resolution is the quantity of pixels that are utilized to build the picture and is estimated by some pixel columns (width) × the quantity of pixel rows (height) state for, e.g., 800 × 600.

Pixels—They are the smallest addressable pieces of a picture. Each image can be considered as a framework comprising of a few pixel values. Each pixel stores a worth corresponding to the light intensity at a specific region, and for an 8-piece grayscale image, the pixel can take esteems from 0 to 255.

Low-Resolution—A Low-Resolution image suggests that the pixel thickness of the image is small in this manner giving less subtleties.

High-Resolution—A High-Resolution images suggests that the pixel thickness of the image is high prompting more subtleties.

Super-Resolution—SR is developing an HR picture from a single/multiple LR picture. Super-Resolution methodologies can be ordered into two groupings subject to the quantity of pictures included—(a) Multi-frame super-resolution (b) Single image super-resolution.

Multi frame Super-Resolution—This strategy uses various LR images to remake a HR images. These numerous images can emerge out of different cameras at independent location catching a scene or a several photos of a similar scene. These various info LR images more or less contain a similar data, in any case, the data of interest is the sub pixel moves that happen because of the development of object, scene shifts, movement in imaging frameworks (e.g., satellites).

Single Image Super-Resolution (SISR)—In SISR, the super resolving algorithm is applied to just one input image. Since in most cases there is no underlying ground truth, the critical issue is to make a worthy image. Most of the SISR algorithm utilize some learning algorithm to fantasize the missing subtleties of the output HR images using the connection among LR and HR images from a preparation database.

The SR reconstruction issues can be formulated regarding an observation model [1] as shown in Fig. 1 which compares the HR image with the input LR images (Figs. 2 and 3).

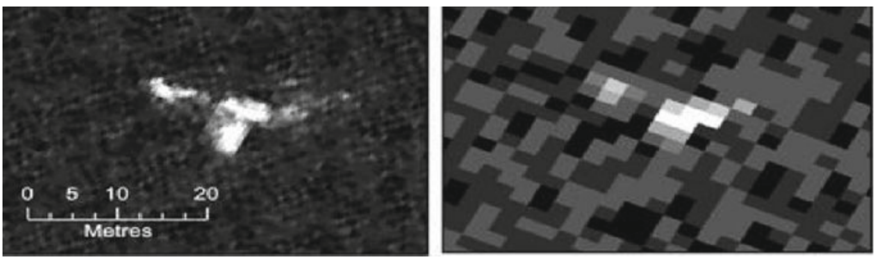


Fig. 1 **a** As can be seen in the above figures, L has more spatial resolution than R

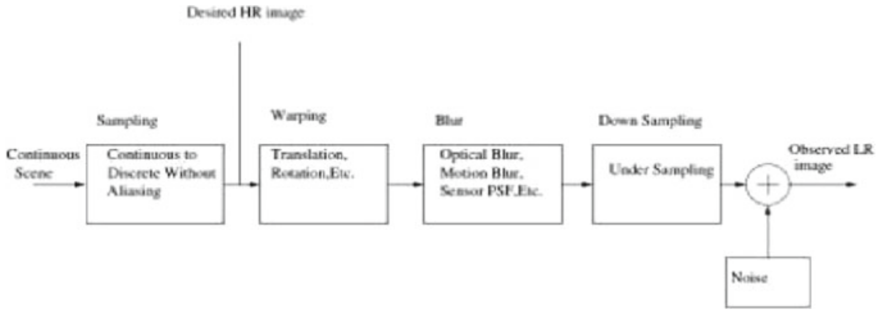


Fig. 2 Observation model

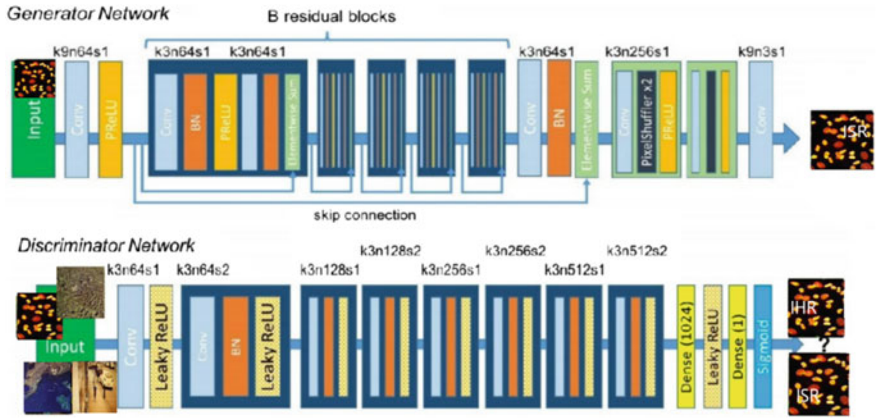


Fig. 3 Architecture of generator and discriminator network with corresponding kernel size (k), number of feature maps (n) and stride (s)

3 Proposed Method

We discuss the exploration strategies in this thesis utilizing the portal gave by Hakånsson [6] as the reference.

Quantitative and Qualitative techniques—Quantitative strategies are worried about quantitatively estimating factors and supporting/disproving hypothesis by methods for thorough experimentation. Qualitative techniques utilize deciphering assessments, practices and so on to formulate a hypothesis and is not explicitly quantifiable by means of quantitative metrics.

This theory utilizes the triangulation technique which is a both blend of quantitative and subjective strategies, in spite of the fact that we lean heavily towards quantitative strategies to draw deductions about the created image quality. Quantitative research strategies utilize trials and testing of explicit factors to validate/invalidate a theory which can be estimated utilizing quantifications and not ambiguous terms. We

Table 1 Software packages used in the paper

Library	Version
Tensorflow	2.0.0
Keras	2.0.0
OpenCV	4.0.0

propose and evaluate a quantifiable hypothesis that, our proposed model, SRGAN prompts better super-resolved images contrasted with Bicubic introduction and SRGAN in regards to a measurement, i.e., the peak signal to noise proportion.

Applied research strategy is utilized in this work since this project involves expanding upon the current exploration completed in super-resolution, information from this real world is utilized straightforwardly as images which are SAR images, CCTV footage analyses to perform super-resolution. Research approaches provide the processes for conducting research which later on helps in concluding the study by proving what is true or false. We follow the deductive research approach to verify/falsify the hypothesis that our proposed model (SRGAN) scores better than the state-of-the-art methods and is useful in other applications. This hypothesis is validated quantitatively by the measuring the PSNR. Research Design—Research design give a direction for performing the research which comprises of organizing, arranging, structuring and leading the research. Exploratory research methodology is utilized in this proposal but absence of extensive training data. Quality Assurance—The data analysis results must be approved and checked which is accurately what the quality confirmation strategy helps. We utilize severe quality affirmation measures to ensure that the software packages and frameworks were utilized accurately without bugs which may have affected the result of the experiments. Python programming language (variant 3.6) was utilized to code the whole project and the framework with their package are recorded in Table 1. To guarantee the reproducibility of our analyses.

3.1 Mathematical Formulations

Given a LR image (ILR), the essential goal in single image SR is to estimate a SR image (ISR) which should to be conceivable to the comparing ground truth High-resolution image (IHR). In particular, our generative adversarial network takes the low-res image 'x' as input and predicts its relating high-resolution image 'y'. The objective of the generator is to gain the function G in y' = G(x), where y' is the gauge of the ground truth high-resolution image y. Let wxhxc be the width, height, and the quantity of channels of the low-resolution image. If 'r' is the scaling factor, the particular width, height, and the quantity of channels of the SR image (ISR) are given by rxw, rxh, rxc. The LR image is acquired by down scaling the ground truth HR image (IHR) by factor 'r'. The generator network is prepared as a feed forward convolutional neural network parametrized by GθG, where θG = {W1:

Lr; b1; L} means the loads and biases for a 'Lr' layer deep neural network and upgraded utilizing a custom loss function intended for super-resolution. For a lot of HR preparing pictures.

IHR n, n = 1, …, N with corresponding ILR n, n = 1, …, N

$$\hat{\theta}_G = \mathrm{argmin}_{\theta_G} \frac{1}{N} \sum_{n=1}^{N} \left(G_{\theta G}\left(I_n^{LR}\right), I_n^{HR}\right) \tag{1}$$

We present a discriminator function $D_{\theta D}$ which is improved nearby the generator network $G_{\theta G}$ in a substitute manner, to discover the solution of the adversarial min–max issue as describe by Goodfellow et al. in their milestone paper [5]

$$\min_{\theta_G} \max_{\theta_D} E_{I^HR} \sim p_{\mathrm{train}}\left(I^{HR}\right)\left[\log D_{\theta_D}\left(I^{HR}\right)\right] + E_{I^{LR} \sim p_{\mathrm{train}}\left(I^{LR}\right)}\left[\log D_{\theta_D}\left(I^{LR}\right)\right] \tag{2}$$

4 Loss Function

4.1 Perceptual Loss

The proposed perceptual loss function in this paper will be the loss function upgraded to produce perceptually conceivable looking and high PSNR SR images. Building up and improving the loss capacities utilized by Ledig et al. [4], the recommended perceptual loss forms the crux of our generator network. Conventional Deep learning-based SR approaches depend on the pixel-wise MSE as the loss function to produce SR images. However, such techniques regularly make the generated images stray away from their ground truth characteristic complex bringing about excessively smooth pictures, regardless of having a high psnr as Ref. in [3]. The latest best in class SR arrangement totally ignores the pixel-wise MSE, rather, depends on utilizing the weighted sum of the content loss (include portrayal of VGG) and adversarial loss as the final perceptual loss. In any case, based on the results and by empirical findings after carrying out different tests we find that utilizing MSE, perceptual loss (as in SRGAN) and adversarial simultaneously together leads to better SR images not only in terms of the PSNR metric but also in terms of visual quality. However, we propose the accompanying perceptual loss function as the weighted combination of pixel-wise MSE, content loss and adversarial loss

$$l^{SR} = \alpha \times MSE + \beta \times l_X^{SR} + 10^{-5} \times l_{Gen}^{SR} \tag{3}$$

where,

l^{SR}	Total loss (Perceptual loss)
MSE	Pixel wise mean squared error between generated SR and ground truth HR.
$l^{SR}X$	Content loss (Feature Reconstruction loss from the mini VGG layers).
$l^{SR}Gen$	Adversarial Loss.
α	weight coefficient for MSE.
β	weight coefficient for content loss.

4.2 Pixel Wise Mean Squared Error

The pixel wise Mean squared error is calculated as

$$l_{MSE}^{SR} = \frac{1}{r^2WH}\sum_{x=1}^{rW}\sum_{y=1}^{rH}\left(I_{x,y}^{HR} - G_{\theta_G}\left(I^{LR}\right)_{x,y}\right)^2 \tag{4}$$

where W, H, r are the width, height and scaling factor respectively. $G_{\theta G}(I^{LR})x$, y is the SR image generated by the GAN and IHR x, y is the ground truth HR image. MSE is the most widely recognized loss function utilized in numerous Deep Learning based state of the art solutions for SR. Utilizing this function prompts superior PSNR values for generated images however lacking in high-frequency content results in perceptually implausible and excessively smooth textured results. We duplicate Eq. (4) by a MSE coefficient (α) to alter the weight given to MSE and add it to the content loss (likewise increased by a content loss coefficient, β) and the adversarial loss determined above to give the total loss.

4.3 Content Loss

Building upon the thoughts of Gatys et al. (Style transfer), Johnson et al. (Perceptual loss) and taking motivation from Chritian et al. (SRGAN) we propose a content loss that is characterized on the element activation of a miniVGG19 trained on images to optimize for the perceptual quality of the SR images.

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j}Hi,\,j}\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}\left(\phi_{i,j}\left(I^{HR}\right)_{x,y}\right)\phi_{i,j}\left(G_{\theta_G}\left(I^{LR}\right)_{x,y}\right)^2 \tag{5}$$

where $\Phi_{i,j}$—feature maps obtained from the j-th convolution before the i-th max-pooling layer in the miniVGG19 network. $W_{i,j}$ and $W_{i,j}$—dimensions of the respective feature maps stored in the miniVGG19 network. VGG loss is the Euclidean distance between the feature representations of the super-resolved image ($G_{\theta G}(I^{LR})$) and the ground truth HR image (IHR).

4.4 Adversarial Loss

The generative parts of the generator network are added to the perceptual loss involving the weighted sum of MSE and Content loss as portrayed above to push the generated SR images in their normal complex. The generator loss is given in condition below

$$l_{Gen}^{SR} = \sum_{n=1}^{N} -\log D_{\Theta_D}\left(G_{\Theta_G(I^{LR}))}\right) \tag{6}$$

where, $D_{\theta D}(G_{\theta G}(I^{LR})$ represents the probability of the reconstructed image given by $(G_{\theta G}(I^{LR}))$ is a real image.

5 Experiments and Results

We perform the experiments in the following request First we execute SRGAN and explore various configurations for the content loss to watch the impacts of various layers of the miniVGG on the image quality and PSNR. Then we move on to study the impact of loss functions on the performance measures (PSNR and visual quality) that SRGAN is bound to give better outcomes on colorized images.

5.1 SRGAN

Model Details—In this model, we give the content loss a somewhat higher weight (0.6) than MSE (0.4). In order to leverage the lower and higher-level feature representations of the VGG, we utilize all the layers to optimize the content loss. The analysers utilized for the generator and discriminator are ADAM and Stochastic gradient descent with learning rates of 1e−3 and 1e−4 respectively.

Table 2 PSNR, MSE, SSIM values for SRGAN and Bicubic for portrait HR Images

	Bicubic interpolation	SRGAN
PSNR	31.903751888	35.26725395
MSE	125.84223611	58.00644095
SSIM	0.94275597	0.9599276

5.2 Evaluation

The input low resolution and ground truth HR images. The comparison with bicubic interpolation and SRGAN. There is a remarkable increment in the PSNR of SRGAN contrasted with low-resolution (>0.80). SRGAN not exclusively does well PSNR wise, but the visual quality of the SR images generated by it are stunning and by a long shot the best among all the experiments we completed. Pixel level changes are considered, the contour lines are pleasantly drawn and above all the textures are reconstructed in remarkab.

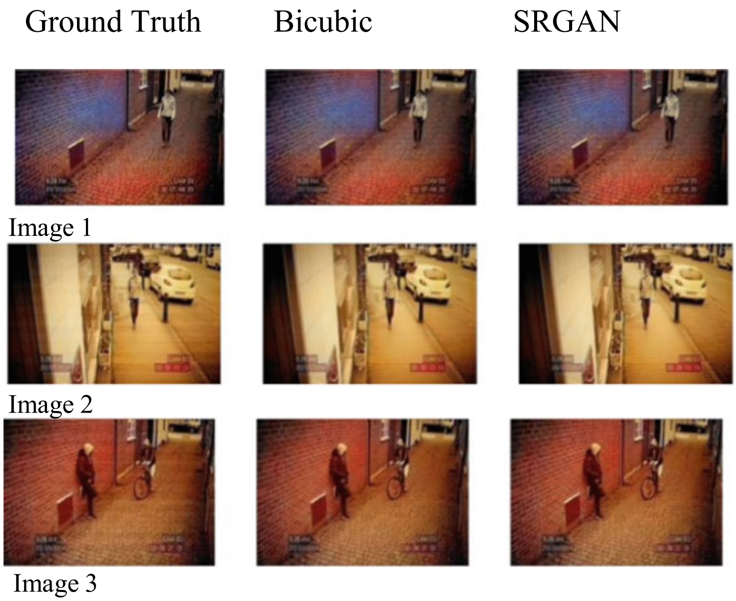
5.3 Results

1. Portrait Image (Table 2):

2. CCTV Footage (Table 3):

Table 3 PSNR, MSE, SSIM values for SRGAN and Bicubic for CCTV HR Images

Image 1	Bicubic interpolation	SRGAN
PSNR(avg.)	23.9217648	24.7746019
MSE	790.724389	649.7427238
SSIM	0.5689400	0.6425353



6 Conclusion

In this work, we introduced a first Generative adversarial network SRGAN for performing super-resolution on numerous kinds of images such as portrait, microscopic and SAR images. We feature the inefficiency of utilizing MSE alone for creating super-settled images that are not affable to the human visual framework. We introduce a more balanced perceptual loss, which expands on the qualities of MSE, content loss and adversarial loss, bringing about outwardly satisfying image near the common complex of real image which don't experience the ill effects of poor PSNR since pixel level changes have been suitably considered. We conclude that Generative adversarial networks optimized for carefully designed loss functions capacities can prompt aesthetically satisfying SR images and actualizing such inventive loss capacities equipped towards the application area holds a significant key towards producing increasingly reasonable pictures.

References

1. Park SC, Park MK, Kang MG (2003) Super-resolution image reconstruction: a technical overview. Sig Process Mag IEEE 20(3):21–36. https://doi.org/10.1109/msp.2003.1203207
2. Chopade PB, Patil PM (2015) Article: single and multi-frame image super resolution and its performance analysis: a comprehensive survey. Int J Comput Appl 111(15):29–34

3. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ (eds) Advances in neural information processing systems, vol 25. Curran Associates, Inc., pp 1097–1105
4. Ledig C, Theis L, Huszar F, Caballero J, Aitken AP, Tejani A, Totz J, Wang Z, Shi W (2016) Photo-realistic single image super resolution using a generative adversarial network. CoRR, abs/1609.04802
5. Goodfellow IJ (2017) NIPS 2016 tutorial: generative adversarial networks. CoRR, abs/1701.00160
6. Håkansson A (2013) Portal of research methods and methodologies for research projects and degree projects. In: Proceedings of the international conference on frontiers in education: computer science and computer engineering (FECS)

Automated Attendance Marking System Using Computer Vision and Deep Learning

R. Kishor Kumar, C. Ganesh Prasad, Reshma Upadhyaya, and Rahul A. Aithal

Abstract The advancement in the history of computer vision utilizing deep learning approaches especially convolutional neural networks have accomplished to solve difficult problems in face recognition field. Face recognition-based approach is one amongst the important identification methods which can be used as a possible substitution for conventional system of marking attendance manually, especially if a huge classroom of students is addressed for an hour session. In this proposed paper, an automated attendance system is implemented utilizing algorithms such as Histogram of Oriented Gradients (HOG) and Face landmark Estimation for face detection and finally, the deep convolutional neural network is trained for face recognition. In this system, a database containing the student's images is created. A camera installed in the classroom captures the image of all the student in the classroom. This image is processed using machine learning algorithms to detect faces and to mark the attendance automatically in an excel sheet.

Keywords Face detection · Face recognition · Deep learning · Machine learning · Attendance system · Python · HOG · Face landmark estimation · SVM · Triplet learning

1 Introduction

Face recognition is a paramount field in computer vision and analysis, which has become a very popular research paradigm from past decades. Facial recognition is a biometric technique that mathematically maps a person's facial features and stores the processed data as a faceprint. The general approach of a face recognition system is by comparing the specific facial features from an input image with a stored image in the database. This approach is widely adopted due to its contactless and seamless integration process, despite lacking accuracy compared to other biometric recognition systems such as fingerprint and iris recognition. It is prevalent in areas such as law enforcement, forensic investigations, healthcare. One such application

R. Kishor Kumar (✉) · C. G. Prasad · R. Upadhyaya · R. A. Aithal
Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_121

where it can be implemented is in attendance tracking. The principal impetus for developing an attendance tracking system is to automate and reduce the time involved in the traditional way of marking the student's attendance.

The predicament in computer vision is the task of determining the faces and to deal with issues such as the brightness problem, the posture problem, scale inconsistency, inferior quality image procurement and partially occluded faces. The face recognition algorithms must be robust against contingencies in the above-mentioned parameters. The objective of the project is to design and implement a system that is less sensitive to brightness and rotation invariant along with being robust enough to be feasible for practical applications.

2 Related Work

Ashish Choudhary et al., have proposed "Automatic Attendance System Using Face Recognition" [1] which implements face detection using Viola-Jones algorithm and face recognition using PCA algorithm. A database is created containing the images of all the students in the institution. In this system. two images of the classroom are captured that is one at the beginning and one at the ending of each class to avoid errors. Integral Images from Viola-Jones algorithm computes the value of each pixel, AdaBoost Algorithm extracts the feature of the face image and additional cascade degenerates the decision tree. Product Component Analysis (PCA) approach is used to recognises the faces, from a collection of face images, called as Eigen-pictures to form a set of basic features of every face. The attendance is marked by linearly combining the Eigen pictures which reconstruct the original training set.

Anushka Waingankar et al., have proposed "Face Recognition based Attendance Management System using Machine Learning" [2] which uses Histogram of Oriented Gradients (HOG) and Support Vector Machine (SVM) classifier for face recognition. In this system, a predefined database consisting of all the student's name and images is constructed. A mobile app is used for streaming the video of the whole class for marking attendance. HOG algorithm is applied to the input image obtained by converting the video to an image of the classroom. SVM classifier outputs the student name to mark the attendance in a spreadsheet by recognising the faces from the input image and converts the spreadsheet into a PDF file.

Sujoy Patole et al., have developed "Automatic attendance system based on face recognition" [3] which is divided into three modules- Face detection module using Viola-Jones algorithm; Face recognition module using Principal Component Analysis (PCA) algorithm along with Eigenface algorithm and Euclidean distance classifier which extracts the facial features from the input face images; Attendance marking module is implemented using a GUI which also considers the changes in the faces of the students in the future.

Priyanka Wagh et al., have developed "Attendance System based on Face Recognition using Eigen Face and PCA Algorithms" [4] which uses Viola-Jones algorithm along with illumination invariant principle and Principle component analysis (PCA) approach to overcome problems of head posing in different directions and the difference in light intensity. A database consisting of student's faces is formed to crosscheck the detected faces. This system pre-processes the input image taken from the camera in the classroom using skin clarification technique in which all the pixels related to the skin is converted into white and the rest as black. Viola-Jones algorithm detects the faces and Eigen Face approach verifies the faces on a basis of region of interest after which the attendance is marked.

Ju-Yuan Hsiao et al., have proposed "Hybrid face recognition system based on multiple facial features" [5] which implements Adaboost algorithm for face detection by separating the faces in the image from their background. The algorithm thus removes the background from the face image to detect the facial features without any errors. Face recognition is implemented using Principle Component Analysis (PCA) which compares the features in the input face image and the feature database which is created by integrating the features extracted from the Adaboost algorithm. Support Vector Machine (SVM) along with Euclidean distance classifier provides with the optimized results.

3 Proposed Architecture Design

Figure 1 indicates the entire block diagram of the methodology of the proposed system which consists of 3 main parts:

(A) Collection of datasets.
(B) Face detection and recognition system (training the neural network).
(C) Attendance marking system.

3.1 Collecting the Datasets

The name, USN and other basic details of all the students is registered at the time of admission which will be helpful while marking the attendance. A dataset is created by capturing several pictures of each student during the registration process which is used to train the neural network.

3.2 Face Detection and Recognition System (Training the Neural Network)

The methodology for implementing the automated face detection and recognition is as follows:

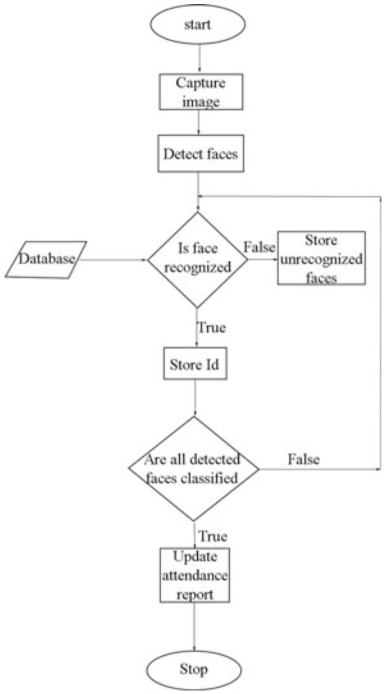


Fig. 1 Design of the proposed system

1. **Image Acquisition**: The first stage of training the network is by capturing the image of the classroom of students with a high definition camera installed inside the classroom.
2. **Image Pre-processing**: Histogram Normalization and median filtering techniques are used to remove the brightness and noise present in the input image. The image contrast is enhanced in the spatial domain by using the histogram normalization technique which removes the brightness or darkness in the captured image. This technique produces a greyscale image from the RGB image, after which the noise from the normalized image is removed using median filtering.
3. **Face Detection Using Histogram of Oriented Gradients (HOG) Algorithm**: The contrast issues in the input image is resolved using a feature descriptor called Histogram of oriented gradients (HOG) by converting the image into a black and white image to detect the faces, as color data is not required for detection. In HOG, a single feature vector is used to represent the complete image and a sliding detection window is used to traverse across the image. A HOG descriptor is computed at every position of the sliding detection window which compares each pixel with its surrounding pixels those which are directly touching. An arrow is then drawn over that pixel indicating the flow of light from brighter region to darker region in the image. The arrow that is to be computed for each pixel is determined by comparing the pixels as shown in

Fig. 2. Irrespective of the brightness in the image, these arrows which represent the significant attributes of the face image are known as gradients

Figure 3 indicates the detection window computing the gradients for the eye in the face image. Hence for each feature of the face image, the process of computing the gradients is repeated several times so that each pixel is replaced by an arrow that signifies the flow of light towards the darker region over the entire image. These gradients form a HOG descriptor for the entire face which indicates all the features of the face which has no brightness or contrast issues as shown in Fig. 4.

4. **Face Recognition Algorithms**: The algorithms which are used to implement face recognition are Face Landmark Estimation algorithm for feature extraction, Triplet Learning algorithm for the encoding of faces in the image, and finally Support Vector Machine (SVM) classifier.

In the Face Landmark estimation algorithm, the features are extracted by labelling and identifying the key facial attributes in the image as shown in Fig. 5. This algorithm solves the problem of a face posing in different directions in each input image. The machine learning algorithm is accustomed to detect 68 specific points (called landmarks) on the face image. Facial landmark detectors essentially try to label and localize the features that are present on a person's face such as the chin, eyes,

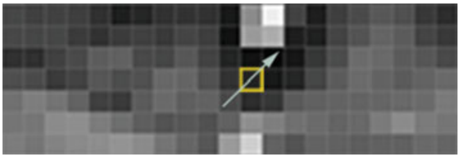


Fig. 2 Representation of gradients

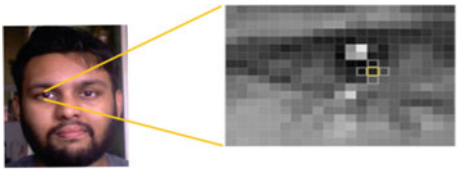


Fig. 3 Detection window computing gradients for an eye in the face image

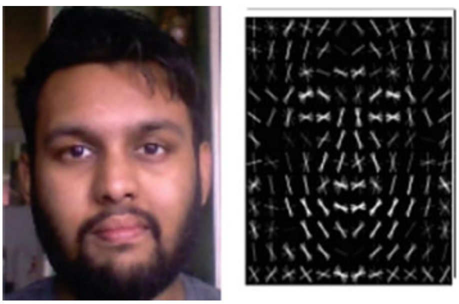


Fig. 4 HOG descriptor for the entire face image representing all the features of the face

Fig. 5 Identification of the key 68 facial attributes

jaw, nose, mouth, etc. Subsequently, the facial features are delineated by 68 (x, y)-coordinates using the facial landmark detector. Then, the image is warped so that the eyes and mouth are aligned to the centre as effectively as possible. Hence, irrespective of how the face is posed or turned in the image, the features of the face are scaled and sheared to the centre so that these features are approximately in the same orientation as that of the sample image in the dataset.

Using triplet learning the faces in the image are encoded by training the deep convolution neural network to produce 128 measured values of every face which are known as embeddings. The network is trained by taking three face images simultaneously, that is, two face images of a known person whom we need to detect after training, and the third face image of a completely different person. The measurements are generated by the algorithm for all three images that are loaded to the network. The neural network is moderately altered so that the generated values for the first two images are slightly nearer while the measurements for the second and third images are a bit far apart as shown in Fig. 6. Hence the algorithm finds the closest measurements for any number of images of the same known face.

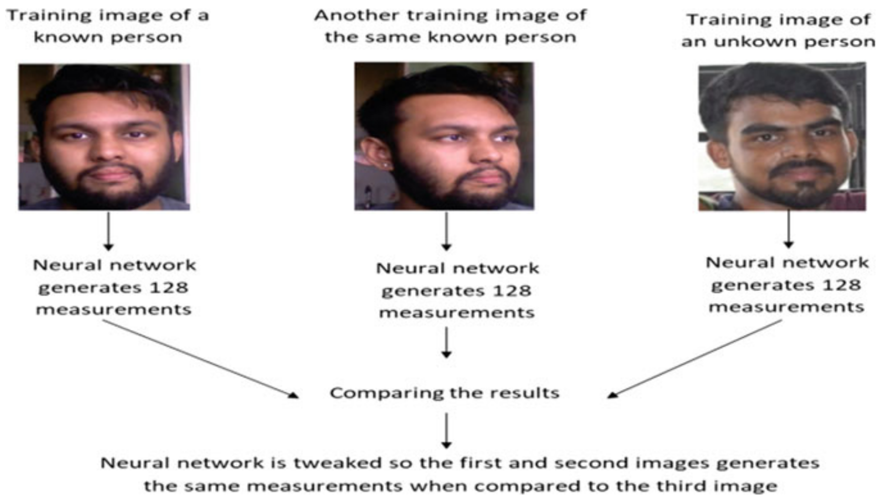


Fig. 6 A single step of triplet training

Fig. 7 Student registration

The final step in the face recognition system is user identity anticipation. To find the person in our test image who has the closest measurements from our database of sample images of known people, the machine is trained using a basic machine learning classification algorithm called linear SVM classifier. A Support Vector Machine (SVM) is a classifier which predicates the result by discriminating the objects (in our case images) by using a separating hyperplane. The SVM classifier is acclimated to acquire the measurements from an input sample picture and outputs a trained student's name which is of the closest match to the input sample picture.

3.3 Attendance Marking System

The students who are recognized by the SVM classifier are marked as present and the attendance is tabulated in a spreadsheet for viewing.

4 Implementation and Results

See Figs. 7, 8, 9, 10, 11 and 12.

5 Conclusion and Future Work

Deploying a system which uses machine learning algorithms along with deep learning and computer vision to automatically mark the attendance of the students in institutions helps in saving time, avoiding proxies and also in uniformed maintenance of the records, unlike the traditional attendance marking system. Here, the Deep Convolutional Neural Network is trained on the acquired image dataset. Histogram of

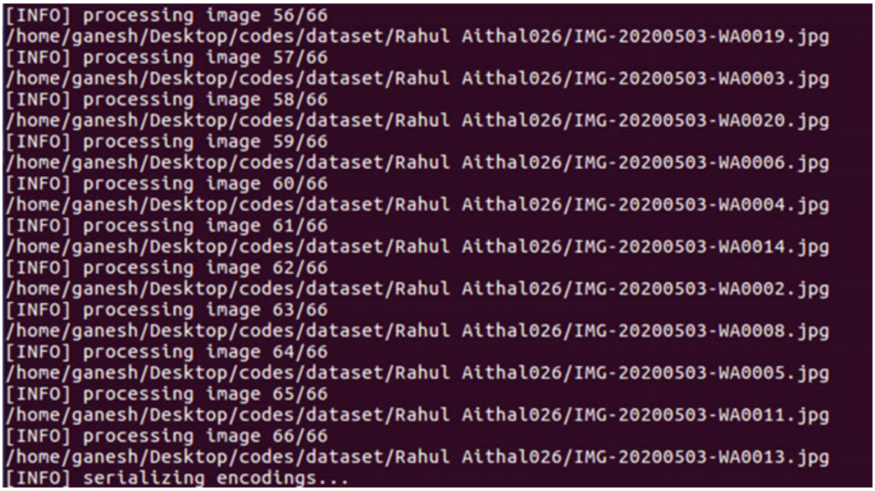


Fig. 8 Training the neural network

Fig. 9 Sample image used for marking the attendance

Fig. 10 The faces in the image are detected and cropped based on oriented gradients

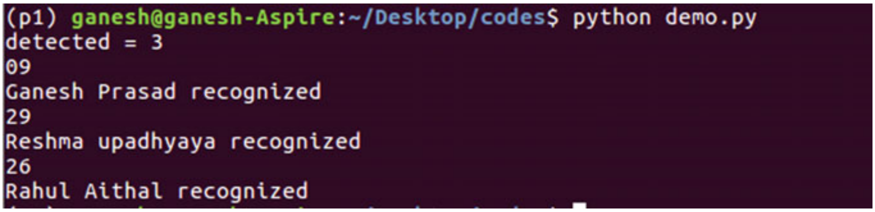


Fig. 11 Student face detection and recognition

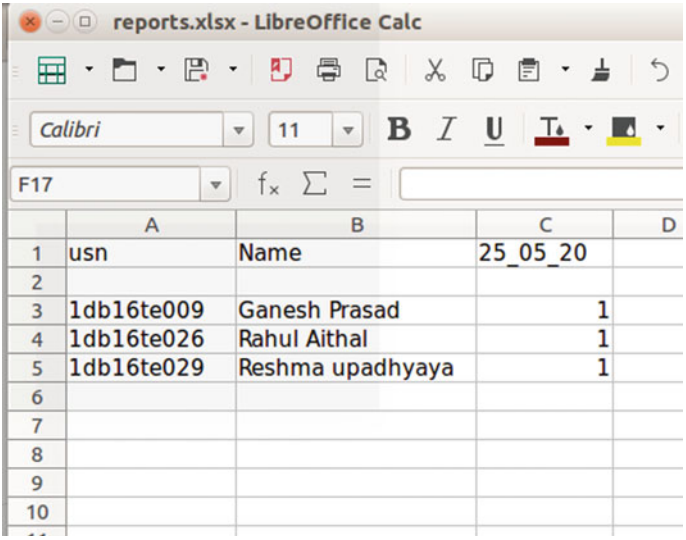


Fig. 12 Final attendance report

Gradients (HOG) approach, Support Vector Machine (SVM) and Face landmark estimation algorithms have been used for face detection, along with features of OpenCV (Computer vision) have been employed for face recognition. The proposed system is able to achieve an overall accuracy of 99% in Face detection and 96% in Face recognition.

Due to certain limitations like long training period and difficulty in recognizing identical twins, this system can be further improved in terms of computational speed which requires the use of improved processors and Graphics Processing Units (GPU). Also, by using cameras with higher resolution, better quality images can be captured which improves the overall performance of the system. With the help of technical security systems, the proposed system can be made more secure, thereby taking care of the privacy issues. The attendance system can also be integrated with a communication software to automate the process of sharing the attendance records with parents.

References

1. Choudhary A, Tripathi A, Bajaj A, Rathi M, Nandini BM (2016) Automatic attendance system using face recognition. Int J Mod Trends Eng Res 03(04):2393–8161
2. Waingankar A, Upadhyay A, Shah R, Pooniwala N, Kasambe P (2018) Face recognition based attendance management system using machine learning. Int Res J Eng Technol 05(06):2395–0072
3. Patole S, Vispute Y (2017) Automatic attendance system based on face recognition. Int J Innov Res Sci Eng Technol 6(8):2347–2410
4. Wagh P, Thakare R, Chaudhari J, Patil S (2015) Attendance system based on face recognition using eigen face and PCA algorithms. In: International conference on green computing and internet of things, Noida, pp 303–308
5. Hsiao J-Y, Chuang S-J, Chen P-Y (2016) A hybrid face recognition system based on multiple facial features. Int J Comput Appl 38(1):1–8. https://doi.org/10.1080/1206212X.2016.1188553
6. Okokpujie K, Noma-Osaghae E, John S, Grace K, Okokpujie I (2017) A face recognition attendance system with GSM notification. In 2017 IEEE 3rd international conference on electro-technology for national development (NIGERCON), Owerri, pp 239–244
7. Elmahmudi A, Ugail H (2019) Deep face recognition using imperfect facial data Centre for Visual Computing, Faculty of Engineering and Informatics, University of Bradford, Science Direct. Future Gen Comput Syst 99:213–225
8. Esmaili SS, Maghooli K, Nasrabadi AM, A new model for face detection in cluttered backgrounds using saliency map and C2 texture features. Int J Comput Appl 1925–7074. ISSN: 1206-212X (Print) (Online),
9. Guo G, Wang H, Yan Y, Zheng J, Li B (2019) A fast face detection method via convolutional neural network. ScienceDirect, m5G
10. Surekha B, Nazare KJ, Viswanadha Raju S, Dey N, Attendance recording system using partial face recognition algorithm, SpringerLink. Intelligent techniques in signal processing for multimedia security, pp 293–319

Smart Self-defense Gadget for Women's Safety Using IoT

A. B. Bhavya, S. Niranjan, A. H. Nithin, V. S. Sandhya, and B. Sharadhi

Abstract The intention of this work is to enhance a gadget for the safety and protection of women and girls. Women all over the world are facing and even subjected to physical harassment. It has an ability to help a female with technologies that are embedded in a compact device. Now we will be developing a gadget that communicates through the message consisting of location of the user and additionally sends the video, Photograph of the attacker via an E-mail to family members and to an emergency help line numbers. In this gadget we have used three push buttons to outline various sorts of accident the victim is facing. If any of the push buttons are pressed, then Raspberry Pi acquires it and sends a text message to the registered phone number. Here we are using Logi-tech Camera.

Keywords Protection · Raspberry Pi · Push buttons

1 Introduction

In today's world, women protection has become a major difficulty as they can't step out of their residence at any given time due to the physical harassment and concern of violence. Even within the twenty century whenever technology is quickly developing and new devices had been developed then again nonetheless many women and girls face issues. There are numerous applications to reduce the danger of assault by defending the system. Most of the gadgets used the GSM module for verbal exchange and a GPS module for tracking [1, 2].

Even some other units used Wi-Fi module, Bluetooth module to send their message to an emergency contacts. They can send it in an Audio format using the voice recognition technique, Video structure through using Micro Camera with low power. We create these gadgets by using Raspberry Pi model B+ involving IOT technology [3, 4, 7, 8].

A. B. Bhavya (✉) · S. Niranjan · A. H. Nithin · V. S. Sandhya · B. Sharadhi
Dept. of ECE, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_122

2 System Designing

(A) **Block Diagram**

The typical block graph of this venture is shown in this Figure 1. In this project, we have used Raspberry Pi model B+, 3 push buttons, GPS, GSM modem, 16 $\times$ 2 LCD display, shock generator, Voice recognition kit. Raspberry Pi model B+ is the main component of the system. We have used three push buttons for detecting the type of accident occurs to the user. When a consumer gets Eve-teased in the street if she presses the first button a message with her GPS coordinates will be sent to the particular phone number.

In the equal way the second button is for snatching and the third button is for kidnapping. Moreover, these messages will continue to ship in accordance to her location change. The camera will seize the picture of the attacker and it additionally makes the video for evidence. Shock generator is used to generate Shock in terms of electric power pulses. Voice recognition kit is used to pre store the keywords [5].

(B) **Circuit Diagram**

In Fig. 2 Switch 1, 2 & 3 are connected to pin no's 26, 19, 16 of Raspberry pi. The heart beat sensor is connected to pin no. 12 of Rpi and pin 3 to buzzer, pin 2 to relay, in turn connected to the shock generator. 5V of Rpi is connected to Vcc of ADC. And pin 9 is connected to Dout, pin 10 to Din, pin 11 to clk, CH7 of ADC is connected to temperature sensor [6].

(C) **Working Principles**

1st switch on the power supply then by pressing the key1 indicates eve teasing so that location will be sent to the registered contacts. By pressing key 2 and 3 indicating snatching and kidnapping respectively. Then the victim's photo will be sent to the registered mail id and shock generator generates shock.

1. Raspberry pi
2. Mcp3008
3. Temperature sensor
4. Heart beat sensor
5. Buzzer

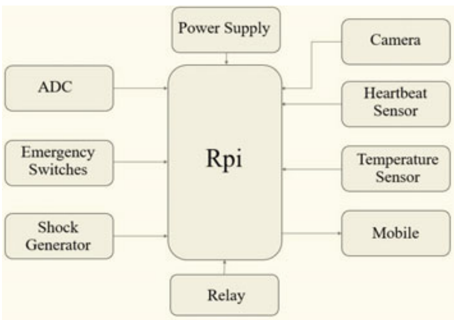


Fig. 1 Block diagram

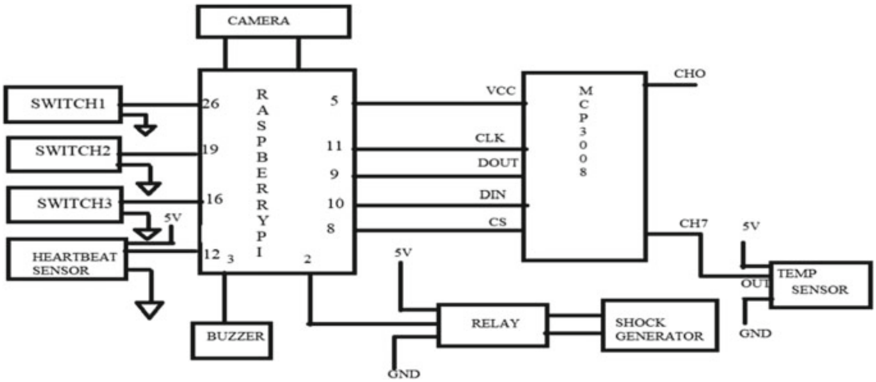


Fig. 2 Circuit connection of the model

6. Shock generator
7. Camera
8. Relay
9. Switches.

Raspberry Pi model B+ is the main driver of the system. We have used three push buttons for detecting the type of accident that happens to the user. When a person gets Eve-teased in the street, if she presses the first button a message with her GPS coordinates will be sent to the particular phone number.

The flow chart of the gadget as depicted in Fig. 4. It is divided into one sub method named SMS procedural, as shown in Fig. 3. In the beginning, the system will exhibit the interactor message. This will provide a fundamental appreciation of the theme of our project.

When a device is turned on successfully, will exhibit on display that a GPS has successfully power up and it reaches for GPS Satellite. When a person presses a push button, message will be sent to the registered contacts consisting of teasing, kidnapping and snatching. The Sufferer is at a sure location. Then the system will start to track the Customer's location. When an area changes, continuous message along with location will be sent to the contact. If the women is rescued or if she feels safe to turn off the device, she can reset the gadget, then the device turns on to the next accident detection. So that the continuous region track will be useful specially when the women will be kidnapped.

3 Results and Discussion

Victim is Snatched/Eve teased/Kidnapped at the particular location. The Attacker's photo is sent to the registered Mail id. The heart rate and temperature range of the victim is displayed. When a victim presses the first button Eve teasing sends the

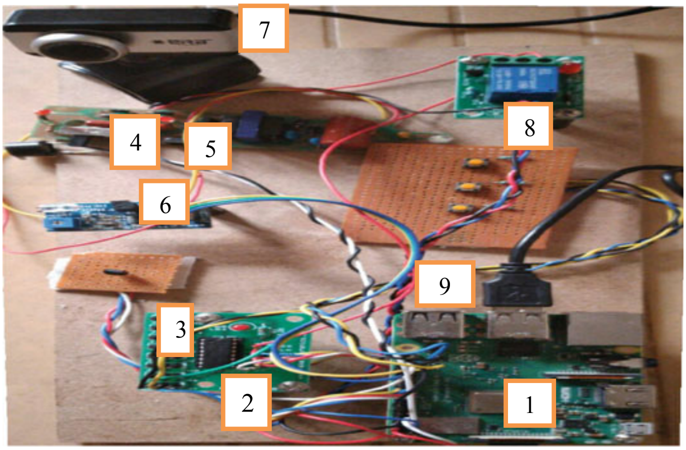
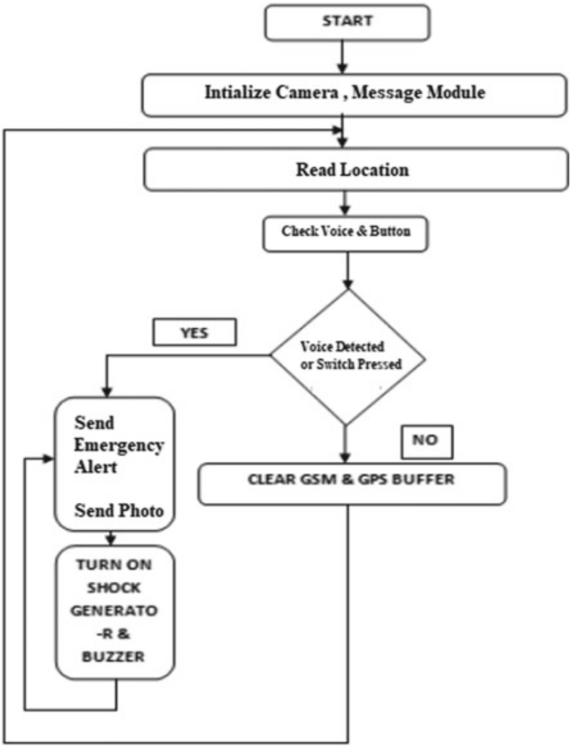


Fig. 3 Protocol implementation model

Fig. 4 Flow chart



message regarding the Eve teasing with location information of the victim to the registered contacts in the Twilio mobile application and to the Emergency helpline.

In the same manner if the victim presses other two buttons protocol sends the message regarding the snatching and kidnapping incidents to the preregistered contacts in Twilio mobile application.

In addition to this it also sends the image of the attacker through the Email to the registered mail id in the Twilio application. This will give a proof of evidence for the police investigation. In this model we have also used a shock generator to protect herself from the attacker. We have used a buzzer which acts as an alarm so that the people around her get to know that the victim is in trouble. Temperature sensor and Heart sensors are also incorporated to monitor the victim's health condition.

In Fig. 5 images is created by means of OpenCV evolving Haar cascade algorithm. The main Purpose of RGB is for the sensing, representation, display of images in electronic system. Gray scale image is the shaded image, black & white image measures the intensity of light at each pixel. Preprocessing means the improvement of image that removes the unwanted distortions and enhances the image features. Segmentation means the partitioning of a digital image into multiple segments. It

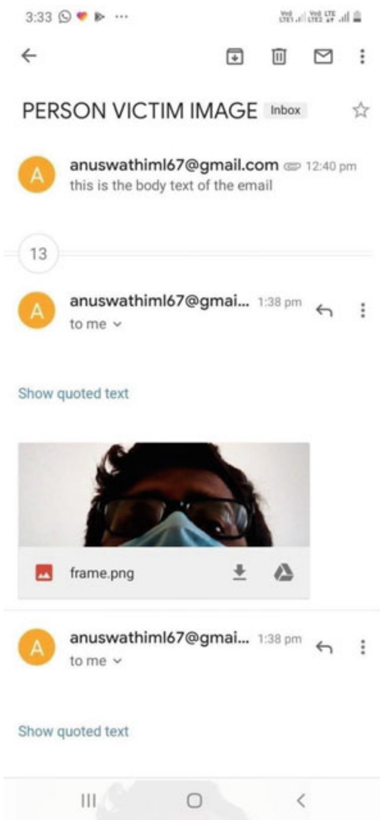


Fig. 5 Victim's image in E-mail

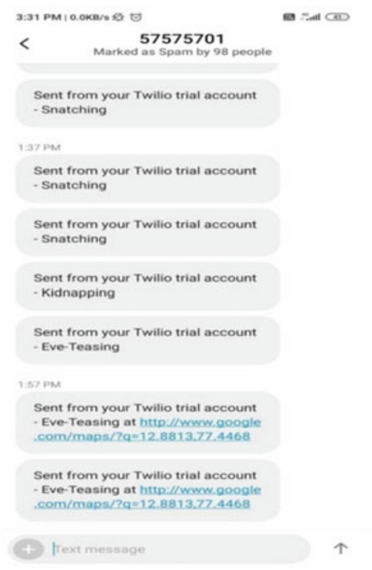


Fig. 6 Type of accident and victim's location

makes the image easier to analyze. Feature extraction OS used to detect shape, texture, color of an image.

In Fig. 6 shown the resultant messages are sent to the register contacts which are stored in the database. By seeing these messages we can see the victim situation and can save the life.

4 Conclusion and Future Scope

The main aim of this project is to protect women from the harassment and theft. The proposed model deals with expository issues faced by women in the recent past and will help to solve them through using safety devices. Since this model is a combination of Embedded system and Mobile applications, we will ensure that this project would enhance the chances of rescuing woman from the danger situation.

References

1. Sharma V, Tomar Y, Vydeki D (2019) Smart shoe for women safety. In: IEEE 10th ICAST
2. Raseduzzaman Ruman M., Badhon JK, Saha S (2019) Safety assistant and harassment prevention for women. In: 5th International conference on advances in electrical engineering
3. Tejonidhi MR, Aishwarya, Chaitra KS, Dayana MK, Nagamma (2019) IOT based smart security gadget for women's safety. In: 1st international conference on advance in information technology

4. Sindhu K, Subhashini R, Gowri S, Vimali JS (2018) A women safety portable hidden camera detector and jammer. In: 3rd ICCE
5. Sen T, Dutta A, Singh S, Kumar VN (2019) Pro-techt—implementation of an IOT based 3-way women safety device. In: 3rd international conference on electronics, communication and aerospace technology
6. Zikriya M, Parameshwar MG (2018) Smart gadget for women safety using IOT. IJERT 6(13)
7. Toney G, Jabeen F, Puneeth S (2015) Design and implementation of safety armband for children using ARMV7. In: 2015 international conference on power and advanced control engineering (ICPACE), pp 300–303
8. Viswanath N, Pakyala NV, Muneeswari G (2016) Smart foot device for women safety. In: 2016 IEEE region 10 symposium (TENSYMP), pp 130–131

Smart E-commerce Hub for Real Estate Web Application

Abhishek Kumar, Hardik Koul, Hemant Sharma, Tanmay Kumar, and P. S. Ashok Kumar

Abstract Real estate is known as one of the most important sectors of the economy, hence Machine Learning Techniques outstrips a real estate proficient when it approaches to conduct a demographic market research, accompanying the environmental and financial probes such as price predictions and loan approvals. The intent behind this paper is to find and apply the best model to predict the prices and loan approvals to develop a smart E-commerce application where users can search and inquire about a property, post their property Ads and can use the smart feature to predict the eligibility for their loan approvals using Machine Learning Model. The initial thing that we have taken into account is the dataset of the apartment prices and different representative cases of loan applications. The dataset was analyzed, pre-processed and calibrated for the real exploration. Thus, our paper will also have a focus on preprocessing of the raw dataset. Further after pre-processing the raw data, we will review the use of Neural Network and Random Forest and the outcome that it generates.

Keywords Real estate · Smart web application · AJAX · Neural network · Random forest · Data pre-processing

1 Introduction

In recent decades, every time the property searching is done, the focus is mainly on the location with adequate basic amenities available from designated location such as unveiling the distance of nearby bus stops, railway station, hospitals etc. The web application spares the user to choose any of the given location and get information appropriately along with the smart features to predict the loan approvals.

The main aim of this project is to design and develop Smart web application that can be used to keep track of all the properties in the country for the purpose of sell, share and rent along with all the relevant data related to the users in a

A. Kumar · H. Koul · H. Sharma · T. Kumar · P. S. Ashok Kumar (✉)
Department of Computer Science and Engineering, Don Bosco Institute of Technology, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_123

centralized database. The availability of the deployed website makes the system more user friendly and makes it more credible. Once after the successful registration user receives an email for the same though.

The project has been developed based on Django which is a high-level Python Web framework motivating the rapid development and practically clean design. The main motive here is to allow developers rather than of imposing the same solutions again and again, to focus on the sectors of the particular application that are new and show uniqueness to their project. The project uses PostgreSql as its database engine. The Dynamic and Smart Web application uses different machine learning models such as Random Forest Regressor to predict the prices and Neural Network Classifier to predict the loan approvals based on the user's data [1, 2].

2 Literature Survey

Technology has always been indicator of smartness. Real estate needs to improve its adoption of disruptive technologies to move from traditional to smart real estate [3].

The user can search the property by entering the keywords in the search tab after which it shows the property listings along with the listings nearby to those keywords. It is required for the user to register on the web application to avail all the features [4].

For developing the real estate web application consists of the various issues, Primarily, the search time should be minimal. Secondarily, the web application should render the services that is required by the both buyer and seller. Lastly, the web-based application must have a friendly and understandable user interface [5].

3 Proposed System

As we know that, the Data security and Data accessing is very slow in the existing system and also the existing system mostly fails to carry out ably when there is an adjustment in the environment of operation. The proposed system is a Smart E-commerce application which is majorly related to online system, which provides the centralized database.

The Web enabled application uses Neural Network to predict for the loan approvals based on the data given by the users. The system uses Tensor Flow, a python module which acts as a higher-level API to build the model. The algorithm applied here uses a multi-layer feed-forward neural network with back propagation learning algorithm to build the proposed model.

The application also uses Random Forest Regressor to predict the apartment prices which is useful for the accurate predictions based on the model. This application helps user to find the property listings based on their requirements and sends notification to the particular Realtors for the same. The Smart E-commerce application also uses the AJAX technique to run the system dynamically and smoothly.

4 Implementation of Web Framework

(a) Django

Django is mainly a free but also an open source web application framework written in Python. A framework may be called as a collection of modules that make development easier. They are basically coupled together, and permits you to build applications or websites from any source that is existing, instead building it from scratch. It's a group of python libraries enabling you to efficiently create quality web application very quickly which is suitable for both front end and backend.

4.1 Front End Implementation

- HTML and CSS
  A **HTML** (the Hyper Text Markup Language) and CSS (Cascading Style Sheets) are two of the core technologies for building web contents. HTML mainly provides the anatomy of the web page, whereas CSS provides the layout, or for a range of devices, together with the graphics and scripting. HTML and CSS are used to build online web page and Web Applications.

- JavaScript

More specifically JavaScript is a dynamic computer programming language, whose lightweight property makes it the most ordinarily used part of web contents, of which implementations allow users to interact or to connect with the client side scripts to develop a dynamic web page. It's an interpreted programming language with object-oriented capabilities/concepts.

- AJAX

AJAX stands for Asynchronous JavaScript and XML. It is an advanced technique for building beneficial, rapid, and more interactive web applications with the assistance of XML, HTML, CSS, and Java Script. In tandem with AJAX, once you hit or submit your request, JavaScript will make a call for participation to the server, unravel the results, and updates the same screen. Within the purest sense, the user would never know that anything was even mediated to the server.

4.2 Back End Implementation

- Python

Python is generally a high-level and interpreted object-oriented scripting programming language. It is intended to be highly readable. It has very less number of syntactical constructions than other programming languages.

5 Machine Learning Models

- Neural Network

Neural networks is an arising AI technology that imitates the human brain on the computing devices. It uses the concept of neurons similar to that of a human brain. These techniques are supported on the parallel and distributed processing design. The proposed machine learning model recalls artificial neural network as an efficient tool for the applications to support loan decisions within the commercial banks for evaluating credit applications.

- Random Forest

Random forest algorithm is more specifically used for the prediction of apartment prices apart from the other predictive algorithms as it delivers better accuracy in the prediction. The algorithm generally deals with the task of classification problems. Taking into account all likely consequences it takes the decisions, based on the average of all the decisions made by each trees.

6 System Design and Flow Charts

The data are taken for analysis and predictive modeling from an online platform particularly known as Kaggle. It allows the users to analyze their datasets and helping them out to prepare models and to grind with various data science engineers to find solutions for real-life challenges using data science. The dataset used in this proposed application has been downloaded from Kaggle, where the data is present in raw and uncleansed format. The initial step is to convert the whole raw data into processed and cleaned data which is done by using preprocessing and feature extraction using sk-learn library and is followed by a classification process. The training data set is exploited to train the model whereas the test data is exploited to predict the accuracy of the model.

Once the model is trained then it is pickled and applied into the Smart web application to work accordingly. The Model is then used to predict for the loan

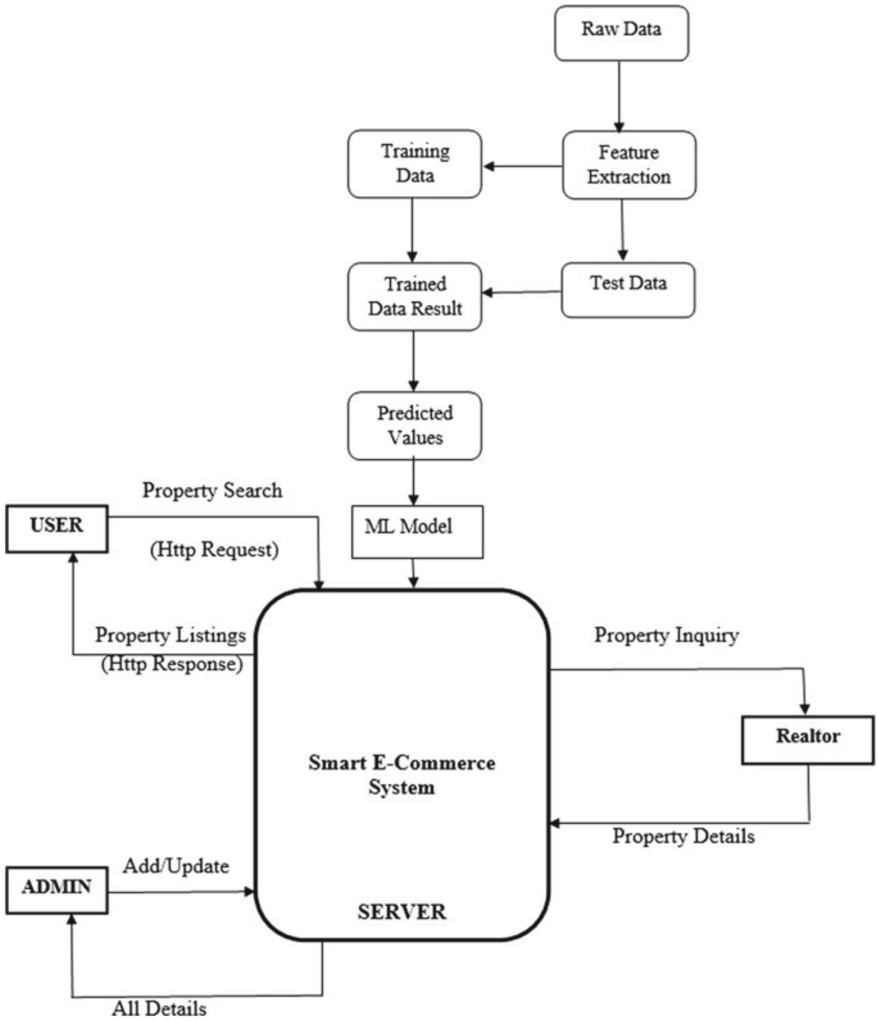


Fig. 1 System architecture

approvals and for price predictions as per user's HTTP request which intern return the result with a HTTP response (Fig. 1).

7 Experimental Analysis and Results

The dataset file consists of raw data based on which models are going to be analyzed. The first dataset that we took is of Loan status and the another is of Apartment prices.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Stat
2	LP001002	Male	No	0	Graduate	No	5849	0		360	1	Urban	Y
3	LP001003	Male	Yes	1	Graduate	No	4583	1508	128	360	1	Rural	N
4	LP001005	Male	Yes	0	Graduate	Yes	3000	0	66	360	1	Urban	Y
5	LP001006	Male	Yes	0	Not Graduate	No	2583	2358	120	360	1	Urban	Y
6	LP001008	Male	No	0	Graduate	No	6000	0	141	360	1	Urban	Y
7	LP001011	Male	Yes	2	Graduate	Yes	5417	4196	267	360	1	Urban	Y
8	LP001013	Male	Yes	0	Not Graduate	No	2333	1516	95	360	1	Urban	Y
9	LP001014	Male	Yes	3	Graduate	No	3036	2504	158	360	0	Semiurban	N
10	LP001018	Male	Yes	2	Graduate	No	4006	1526	168	360	1	Urban	Y
11	LP001020	Male	Yes	1	Graduate	No	12841	10968	349	360	1	Semiurban	N
12	LP001024	Male	Yes	2	Graduate	No	3200	700	70	360	1	Urban	Y
13	LP001027	Male	Yes	2	Graduate		2500	1840	109	360	1	Urban	Y

Fig. 2 Raw data of bank loan

We found that Neural Network best suits for the Loan Approval Model whereas Random Forest best suits in predicting the accuracy of apartment prices.

There are 13 columns or 13 attributes that describes the Loan Status of a person. Some of them are Loan_Id, Gender, Married, Dependents, Education, Self_Employed, Applicant_Income, CoapplicantIncome, LoanAmount, Loan_Amount_Term, Credit_History, Property_Area, Loan_Status. We dropped off the first column i.e. Loan_Id which doesn't play crucial role in predicting the status (Fig. 2).

After that we used One Hot Encoder method to encode the Categorical values or features. This encodes the categorical feature as a one-hot numeric array. This creates a binary column for each category and return a sparse matrix or dense array (Fig. 3).

The method was applied on the columns Gender, Married, Education, Self-employed and Property Area which contains categorical data (Fig. 4).

The training stopped as it gives us the best accuracy for the above dataset. The accuracy classified is between 86 and 90% was obtained using Neural Network Algorithm as compared to SVM model which yielded 81% of accuracy for the same.

Figure 5 shows Confusion Matrix states that: When it was Predicted as NO it was Actually correct 52 times. When it was Predicted as YES it was Actually correct 49 times. When it was Predicted as NO it was Actually YES 19 times. When it was Predicted as YES it was Actually NO correct 13 times.

There are 9 columns or 9 attributes that describes the different apartments data. Some of them are area_type, availability, location, size, society, total_sqft, bath, balcony, price. After splitting training and testing data we have used Label Encoder and after changing the categorical features into numeric array the data is as shown in Fig. 6.

After that we moved to Preprocess the columns and we found that there are many missing values in different columns such as 'bath' and 'balcony'.

Gender_Female	Gender_Male	Married_No	Married_Yes	Education_Graduate	Education_Not Graduate	Self_Employed_No	Self_Employed_Yes	Pr
0	1	0	1	1	0	1	0	
0	1	0	1	1	0	0	1	
0	1	0	1	0	1	1	0	
0	1	1	0	1	0	1	0	
0	1	0	1	1	0	0	1	
0	1	0	1	0	1	1	0	
0	1	0	1	1	0	1	0	
0	1	0	1	1	0	1	0	
0	1	0	1	1	0	1	0	
0	1	0	1	1	0	1	0	

Fig. 3 One hot encoded data

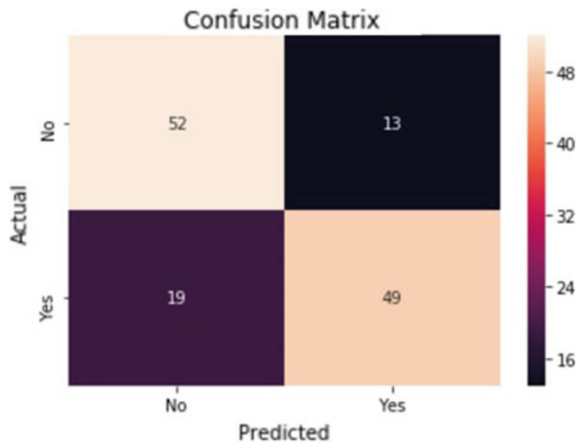
Fig. 4 Optimal set of parameter

Parameters	Values
Number of Training Data	615
Number of Testing Data	369
Number of Hidden Neurons	200 (Condensed to 1)
Number of Hidden Layer	4
Learning Rate	0.2
Epochs	100
Momentum	0.9
Threshold	0.05

In the next step we dropped off the price column from test data so that we can use the testing data on trained data result to predict the prices. The model was trained using Random Forest Regressor which is best suited to predict the apartment prices in a particular area.

Once we get the trained data result of both the algorithms we pickled the model using python's pickle module. Python objects are serialized and de-serialized using this pickle module. The object of any type in Python can be pickled and can be saved on disk for further use. First the object is "serialized" before writing it on the file to convert it into a stream of characters. The character stream contains all the related information to remodel the object in another python script.

Fig. 5 Confusion matrix



	area_type	availability	location	size	society	total_sqft	bath	balcony	price
0	0	2	430	14	487	1056.0	2.0	1.0	39.07
1	2	0	325	20	2569	2600.0	5.0	3.0	120.00
2	1	0	1220	17	1528	1440.0	2.0	3.0	62.00
3	0	0	778	17	2303	1521.0	3.0	1.0	95.00
4	0	0	736	14	1528	1200.0	2.0	1.0	51.00

Fig. 6 Encoded training data using label encoder

The pickled model is then loaded into the web application in the Django framework to iterate over the data given by the user and to predict the values (Figs. 7 and 8).

Fig. 7 Property search in smart webapp

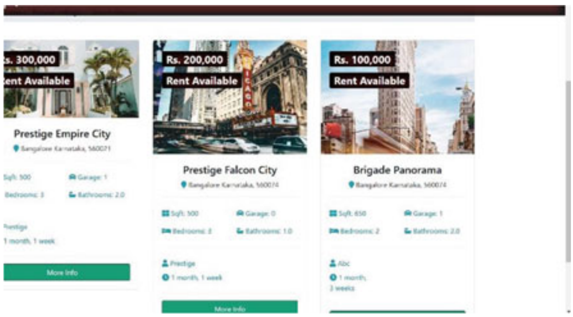
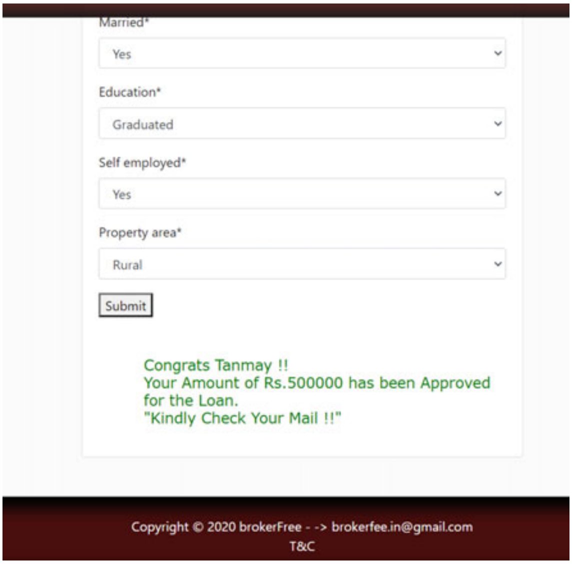


Fig. 8 Loan status off a user



8 Conclusion

Smart E-commerce application is user-friendly and flexible which will eventually save the time of the user in searching property listings and get right business solutions. Between the measured accuracy of different algorithms, we found that the most relevant algorithm for predicting the Loan Approvals based on certain data points from the labelled data is the Neural Network Algorithm. The algorithm might come up as a great asset for the associated banks and users, reason being it is trained on a large heap of pre-processed data and has been validated using the validation dataset, whereas the project uses Random Forest Regressor to predict the accurate prices of apartments, which in turn will be a great asset to the Real estate owners to set the apartment price accurately to get right business solutions.

The project illustrates the use of machine learning model for the prediction of the loan approvals and the price predictions with greater accuracy in contrast to the machine learning models implemented earlier.

References

1. Supriya P, Pavani M, Saisushma N, Kumari NV, Vikas K (2019) Loan prediction by using machine learning models. IJET
2. Joylin Zeffora A, Shobarani R (2019) Statistical analysis of random forest on real estate prediction. IJITEE (2019)

3. Necula S-C, Pavaloaia V-D, Strimbei C, Dospinescu O (2018) Enhancement of E-commerce websites with semantic web technologies. MDPI
4. Mahajan J, Mahajan G, Bhuvad V, Chaudhari R (2016) Real estate management system. IJARCCE
5. Xu W, Lu G, Jade LQXY (2008) The new e-business models in the modern enterprise applications—based on the example of Guangxi Sugar Network Social Survey. Coast Enterp Sci Technol (09)

Efficient Look Based Pin Validation Using Gaze Based Pin Entry

B. E. Manjunathswamy, R. Sai Arpitha, J. Yashwanth, E. Sushma, and S. Yogesh

Abstract The major goal of this work is to propose a Real-time application for Password Authentication using Eye Tracking, with a Gaze-Based pin entering employing an intelligent camera for eye tracking and detection to identify the password pin. PIN's, or Personal Identification Numbers, are primarily used for end-user security and verification. When a security pin is physically entered for password authentication, this raises the risk of PIN information being stolen via shoulder surfing and thermal tracking. Another highly secure method of authentication is a Gaze-Based PIN entering that does not require physical touch. This one-of-a-kind method of identification is based on a channelized disrupted image frame that tracks the eye centre on a regular basis.

Keywords Password · Authentication · Secure-authentication · Gaze-based authentication

1 Introduction

Alpha Numeric Password or Personal ID Number (PIN) is a coded message used for many transactions in electronic banking for several applications, and a popular user verification law. Authentication of Passwords using PIN requires end users Input the PIN manually which may be vulnerable to Techniques used to capture passwords. Current strategies for public safety are Generally weak, since most of them are easy to bypass or fool.

One of the security requirements for general terminal authentication systems is to be easy, fast and secure as people face authentication mechanisms every day and must authenticate themselves using conventional knowledge-based approaches like passwords. But these techniques do not seem to be safe because they're viewed by malicious observers who use surveillance techniques like shoulder-surfing observation user while typing the password through the keyboard) to capture user authentication data. Also, there are security problems due to poor interactions between systems

B. E. Manjunathswamy (✉) · R. S. Arpitha · J. Yashwanth · E. Sushma · S. Yogesh
Department of CSE, Don Bosco Institute of Technology, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_124

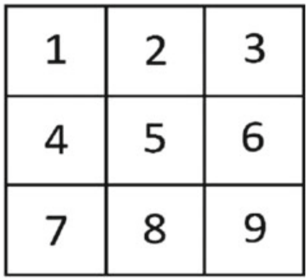


Fig. 1 EYE gazing and tracking using digital keypad

and users. As a result, the scientists proposed eye tracking systems, where users can enter the password by looking at the suitable sign in the appropriate order and thus the user is invulnerable to shoulder surfing [3]. The aim of this paper is to analyse techniques or solutions to deal with eye movement tracking in security systems.

The purpose of this device is to identify the embodied PIN by real time eye recognition and Intelligent tracking system using a smart camera. Using camera, the NI Vision Builder and LabVIEW software are used to identify the eye and record the location of the eye centre [1]. A smart camera lets you process, acquire and collect data.

Non-physical based authentication incorporates a protection surface into physical PIN impressions, and the vulnerability of authentication mechanism is assumed to be truncated. This project provides the user with a virtual on-screen keyboard and permits the user to authenticate by looking at the symbols in their password order. By using a keyboard on-screen, we preserve the flexibility of conventional password schemes that can still be used in protected settings such as home or office (Fig. 1).

2 Related Work

Many eye global positioning frameworks are recommended for writing, and are adapted for speaking practically. Zhu and Ji depict a PC vision system focused on powerful dynamic IR enlightenment for constant look and skilful practical presentation. Prototypical Look global positioning systems need a static head to function praiseworthily and involve a minor adjustment. The look tracker gives correct look estimations without alteration and also with head changes in the system depicted by Zhu and Ji. A look-adjustment action is performed that identifies and distinguishes the preparation from student boundaries to screen encourages the use of summed up neural relapse (GRNNs) systems [2]. The developers use student flash commitment, look adjustment and aim to carry out the process of eye tracking. No wearable contractions are fundamental in our application. Following the eye is practised using eye-discovery counts performed on the brilliant NI camera which has short preparation times suitable for Real Time applications [6]. Following this, the condition of the chief matches in the complex picture diagrams is ignored. We use Lab VIEW to execute real-time eye tracking and recognition image processing algorithms that include pattern matching, edge detection, and eye centre position computing and recording for tracking purposes [7].

3 System Design

(A) System Architecture

The overall design of the device is represented in the above figure. The figure shows the algorithms which the smart camera supports developed and implemented in Vision Builder AI 3.6 (Fig. 2).

The general engineering of the system is defined in the figure mentioned above. In Vision Builder AI 3.6, the figure shows calculations produced and executed by shrewd camera. The NI shrewd camera is connected directly to the standard PC through an Ethernet link in a nearby unit. In this link, the information and control signal is sent back phase in complete duplex structure. The PC is equipped with a Pentium 4 2.26 GHz processor, a 64 MB on board design card and 1.25 GB RAM [2]. The clever camera uses the LabVIEW work area interface to communicate.

(B) Proposed System

Our system contains the three layers which are:

1. Face reorganization
2. Eye- blink verification and
3. OTP (Fig. 3).

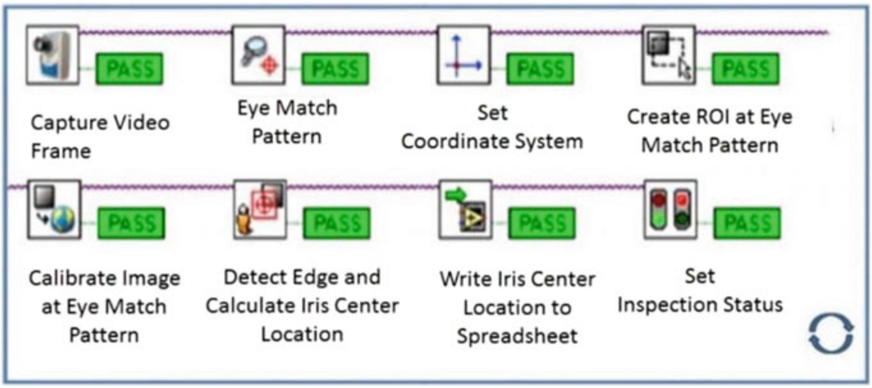


Fig. 2 Eye detection and tracking architecture

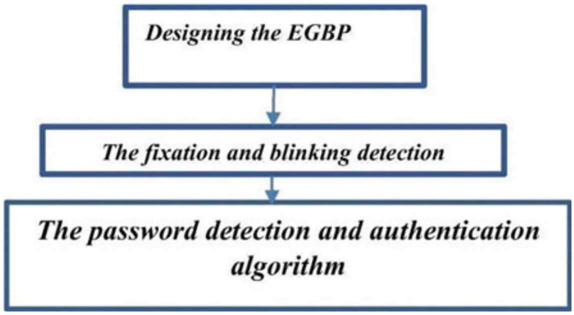


Fig. 3 System design

(A) **Registration System**

Registration system takes up individual details such as name, client id, email id and mobile number. Further we also train the system with Images. This is how we feed our information into the system. These details will be stored in MySQL database. It will also make sure whether the data satisfied certain conditions and throws up a message on successful registration.

(B) **Login System**

Login system authenticates the user and this authentication takes place in two steps:

1. Face and Iris Matching (Template Matching)

Matching templates is the technique of identifying the small sections of an image that match a template image. It goes through the template from top left to bottom right of the picture, and compares with the template for the best match [4]. It checks whether the prototype matches the picture from the real time frame.

2. Gaze Based OTP Entry to Successfully Login

Generated OTP from the registered mobile number is entered on the keypad provided by login system using gaze based blink technique.

4 Implementation

At first the face identification method must locate a face within the picture and accentuate the ranges. There are sort of algorithm calculations for this software. To make decision on the choices of contours within the picture and their contrast with the contour of faces.

1. Eye Detection
   Eye detection can be done by capturing frames from camera at regular intervals of the time. Difference is estimated from the consecutive frames [5].

2. Eye Tracking
   Tracking algorithm tracks the user face's eye position and checks for user's blink.
3. Activation and Blink Match to Crack Password.
   Each number on keypad will be provided with an activation for a second and this cycle repeats. Based on user's eye blink when particular number is activated, the number will be taken as password. Here the password is OTP received by the user on his registered mobile number. If gaze based entry matches with the OTP received.
   On the digital numeric keyboard, a cursor will keep moving because the person blinks, therefore the attention system will generate the pin supported sequence collected by the quantity of blinks. To watch the attentive eye movements, we are using web camera (Fig. 4).

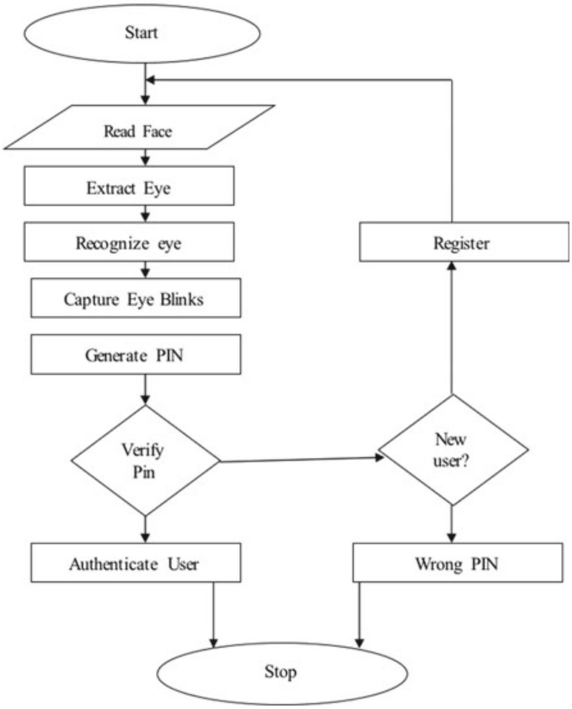


Fig. 4 Methodology flowchart

5 Real Time Eye-Detection System

(A) An Intelligent Camera

All this tracking system is an essential aspect. An Intelligent camera has a twin processor which enables the camera to perfectly configure in real time with applications such as optical character recognition, pattern matching, and data array code. The complete setup generates a very high-resolution Gray-scale images which is critical for the real-time eye tracking algorithm implementation.

(B) PIN Entry

It is a gaze based pin entry authentication which is completely a Non-physical contact PIN entry, as a result it leaves no footprints behind and thus providing a safer option for password entry. It is mainly based on tracking eye centre periodically and sequential image frames using eye location. This method involves a unique processing where the end user enters the password by visually examining the (PIN) pad. For a moment, the end user gets across each and every digit of pin and plots their respective password pin sequentially in the following PIN order. The smart camera encapsulates the sequential picture frames of the face when the user looks at the digits on the keypad, Using implemented image processing algorithms, calculates the eye centre and stores the Cartesian coordinates representing the eye centre in the on-board spreadsheet register, along with the related frame number. When PIN entry is accomplished the eye tracking process is completed.

(C) Eye Detection Algorithm

The model requires proper training at first, which means the algorithm must be trained using proper data sets. This method can be achieved by capturing and saving a single frame of the end user's eye in the form of an eye replica. It is accompanied by eye centre monitoring and identifying the face of the end users. Using the saved prototype, eye matching is performed, and the real-time image frame is scanned for best fit using normalised cross correlation. The transformation of the reference frame is done primarily to standardise the recorded eye position coordinates to allow the capture of eyes at various angles on the basis of head tilt. Using oval or circular matching, edge detection comes into the image when detecting their eye theirs.

6 Result and Analysis

Web based application which contains the three layered authentication to produce the safety to avoid the illegal access of our accounts. This application obviates shoulder surfing and thermal tracking. The steadiness or attentiveness of the end user's gaze will have an impact on the detected PIN's, which must be accounted for.

Finally, the identification of the PIN only succeeds after full eye monitoring, eye centre computations and recording. Future work includes incorporating the Pin identification algorithm for an all-in-one time password recognition method into a real-time environment. When studying these methods, on a significant note, To include them in mobile devices and other camera-based systems, the gaze-based password entry is expanded (Figs. 5, 6, 7, 8 and 9).

Figs. 5, 6, 7, 8 and 9 shows entire system working, where fig seven and eight shows the gaze based PIN entry with the accuracy rate of 80% and delay time 3–4

Fig. 5 Registration page

Fig. 6 Login page

Fig. 7 Virtual keyboard

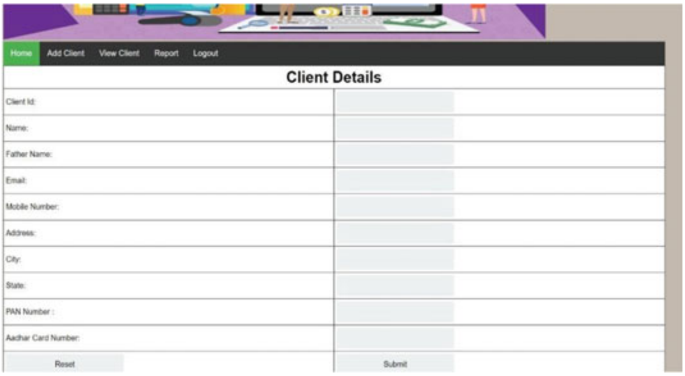


Fig. 8 Virtual keyboard for PIN entry

Fig. 9 Home page

min. The model are going to be trained from real time capturing of images. It captures 5 images per second. The performance will be enhanced by using highly efficient camera with >2 MP.

7 Conclusion

This Non-physical contact PIN entry used for authentication increases a layer of security to physical contact PIN entries by the end user. As a result it is proved to reduce the vulnerability faced by the ancient PIN entry techniques. An intelligent camera based eye blinking system is implemented into a replacement for eyelid blink method for password authentication. The whole model has been successfully tested with a ten digit digital keypad which appears on the screen and can also be extended for a character and digit combination password. Computation depends on the system speed. Further this may be enhanced with good computational systems and smart cameras. The steadiness or attentiveness of the end user's gaze will have an impact on the detected PIN's, which must be accounted for. Finally, the PIN identification gets successful only after eye-tracking, eye centre computations and recording are completely formulated.

(a) **Future Scope**
While our project provides promising results, it can still be upgraded within the coming future for more safer authentication. We will fortify a password by withdrawing supplemental bits from the gaze path that the utilizer follows while entering the password. Even different gaze patterns to be accepted by the system so that the assailer cannot mimic the user's gaze path.

References

1. Zhang T (2012) Solving large scale linear prediction problems using stochastic gradient descent algorithms. In: Proceedings of the twenty-first international conference on machine learning
2. Downey AB, How to think like a computer scientist
3. Sawale GJ, Gupta SR, Use of artificial neural network in data mining for weather forecasting. Int J Comput Sci Appl
4. Lashkari AH, Farmand S, Zakaria OB, Saleh R, Shoulder surfing attack in graphical password authentication
5. Chollet F, Deep learning with Python
6. Mehrubeoglu M, Nguyen V (2018) Real time eye tracking for password authentication. In: 2018 IEEE international conference on consumer electronics (ICCE)
7. Mehrubeoglu M, Pham LM, Le HT, Muddu R, Ryu D (2011) Real time eye tracking using smart camera In: 2011 IEEE applied imagery pattern recognition workshop (AIPR)

An Embedded Technology to Auto Monitor the Boat Height and Bridge Condition Along with Their Crash Avoidance

U. Mamatha, K. Kavitha, S. Kavya, Nagaratna, and N. P. Pramodini

Abstract The transportation of India is developing rapidly day by day. The different kinds of bridges, boats and the speed of navigation have been increasing rapidly, and the accidents of boat and bridge are increasing. At present the bridge inspection is done manually. Implementing the correction process for any crack in the bridge is very difficult. Travelling of heavy loads which are prohibited on the when compared to the given threshold are the main reasons for occurrence of crack in the bridge. The aim of the proposed project is to provide a safety navigation system and guide the navigation using wireless networks. We are focussing on implementation of system having sensors, Zigbee technique which informs bridge condition to nearby officers.

Keywords Sensors · Zigbee technology · MATLAB

1 Introduction

The transportation of India is increasing day by day and different types of bridges are being built and in river crossing areas many kinds of bridges are being built. Because of this increase in bridges there is increase in bridge-boat collisions and accidents are increasing rapidly. To avoid these accidents and to maintain bridge condition properly our proposed system is very useful. The most common reason for bridge collapse is crack in the element of the bridge and level of sand at the base of bridge. Monitoring the health condition of the bridge is an increasing concern for the benefit of all living beings.

The main challenge is to ensure that the health condition of the civil infrastructure bridge is able to stand with the cumulative weight of all the vehicles which travel on the bridge and the capacity to bare the speed of flow of water [1]. The role of bridge is very important in the nation's economic & infrastructure development for the conveniences to the people, transportation and for connecting & communicating two areas.

U. Mamatha (✉) · K. Kavitha · S. Kavya · Nagaratna · N. P. Pramodini
Department of Electronics and Communication Engineering, Don Bosco Institute of Technology, Kumbalgod-74, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_125

For the security of the Bridge point of view, purpose to monitor the health status of the bridge, especially monitoring the sand level at the bottom of bridge pillar, any crack detection in the bridge is the main topic in any research.

Safety is a major issue after the big incident happened such as Earthquake, Flood and Tsunami for examining purpose to determine the damages in what extent [5]. Monitoring and detecting the minor crack on the bridge will give the bridge more efficiency and the bridges will be in good condition for long time. So detecting the crack in early stage is very essential.

Crack detection during experimental testing may require researchers to mark crack on the specimens, whereas researchers can take photographs of the specimens from a safe distance and take suitable precaution. This system comprises of sensor technology and Zigbee technology.

2 Literature Survey

Authors [4] focussed on development of an IOT based bridge safety monitoring system. It is developed using ZigBee technology. In this system monitoring devices are installed at different parts of the bridge, communicating devices are connected to the monitoring devices and the information is collected, and calculation is done in the monitoring device.

Authors [2] IOT based bridge safety monitoring system is developed using the WSN technology. The sensor technology is used to monitor the conditions of the bridges properly and to have more life span and efficiency of the bridges. The data from the sensors are useful for monitoring.

Authors [3] bridge health monitoring system implements systems having sensors, GSM technique which informs bridge condition to the nearby officers. This project system uses wireless network on real time basis for bridge health monitoring purpose. It can transmit data for several minutes.

Authors [6] bridge monitoring and alert generation system using IOT. This project proposes a system which consists of a weight sensor, water level point contact sensor, Wi-Fi module and aurdino microcontroller. We measure load of the vehicle entering the bridge, we also take a count of level of water and pressure it has. If the value is more than threshold then alert message is sent through buzzer. If it is necessary, then the admin assigns the task to the employees for maintenance.

3 System Architecture

Our proposed methodology ensures safety across boat-bridge area. We use sensor and Zigbee technology in our project. Different sensors are used to monitor bridge condition and Zigbee is used to communicate between boat and bridge. ARM LPC

2148 is used as controller which takes inputs from sensors and monitors the bridge by providing all the sensor information to the nearby offices.

Load cell is used as weighing device. It measures the weight of vehicle entering the bridge, if the weight of the vehicle is heavy then gate closes. Figure 1 shows the block diagram of bridge module. Stepper motor is used as gate. There is limit for number of vehicles that enter the bridge. IR sensor is used to count the number of vehicles that is entering and exiting the bridge. If the limit is reached then gate closes.

Flex sensor is used to detect cracks in the bridges. Cracks are detected by the change in the resistances of the flex sensor. Vibration sensor detects the vibrations that occurs near the surroundings and informs to the nearby office. Figure 2 shows the block diagram of boat module. Water level sensor measures the level of water and the pressure sensor checks the pressure of water to avoid the collision between boat and bridge.

MATLAB is used to estimate the height of the boat. ARMLPC2148 takes the inputs from the sensors and MATLAB and communicates with technology.

Proposed System mainly deals with Height of prediction of boat, water level monitoring, pressure monitoring below the bridges. LCD and buzzer are used to inform the boat about crossing the bridge. It deals with comparison between bridge height and the boat height and Air gap and boat height is compared and the boat

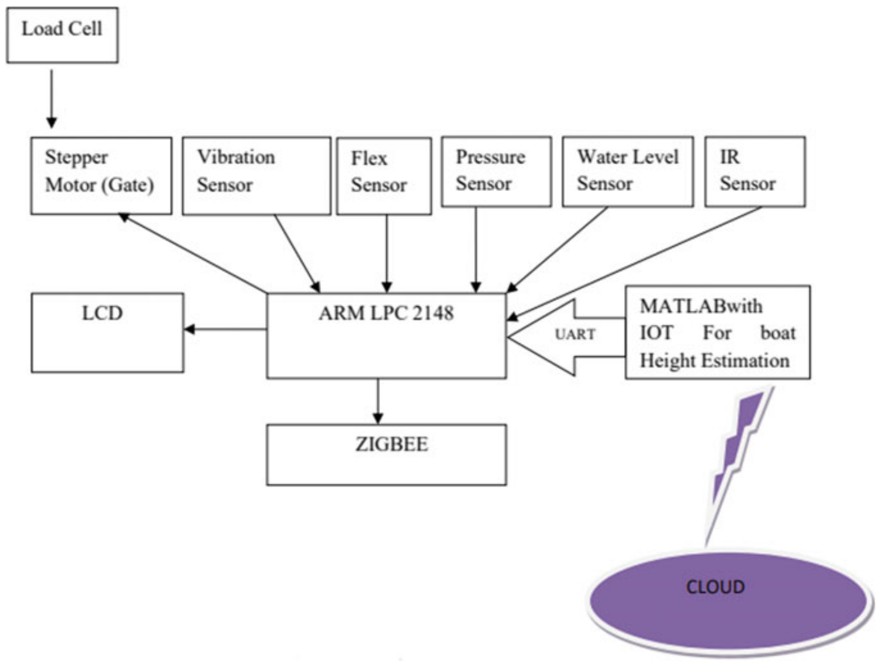
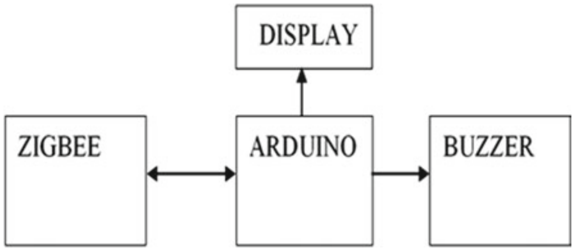


Fig. 1 Bridge model

Fig. 2 Boat model



owner is intimated using LCD predicts weather boat can pass under the bridge or not in order to avoid collisions.

The information's of the water level and pressure sensor are taken and the estimated boat height is collected from the Matlab. Now the processor compares all the inputs and decides whether boat can cross the bridge or not Arduino is used to connect LCD, buzzer and zigbee. The height of the bridge and level of water flowing under the bridge is measured. Zigbee is used to communicate between boat and bridge. The difference between these two values are calculated which gives the air gap between the bridge and the water. Height of the boat is measured using Image processing (MATLAB).

4 Experimental Setup

Figure 3 shows the hardware setup showing the Bridge and Boat module. This project intends to provide the continuous safety bridge monitoring system by using IOT technique, ZigBee and monitoring sensors. This system has the ability to monitor the structure environment to transmit the sensor data through the wireless communication and sends alerts to the concerned authority in the real time scenario for the actual output.

5 Results

Figure 4 shows the output of Load cell. Load cell is placed at the entrance of the bridge. It measures the weight of every vehicle entering the bridge. The threshold maintained in our project is 940 gm. If the weight of vehicle exceeds 940 gm then the gate gets closed.

Figure 5 is the output of IR sensor and is displayed in LCD. We use two IR sensors in our project. Both sensors are placed at the entrance and exit of the bridge respectively. The threshold maintained in our project is 5. If the number of entry of vehicles reaches more than 5 then the gate gets closed.

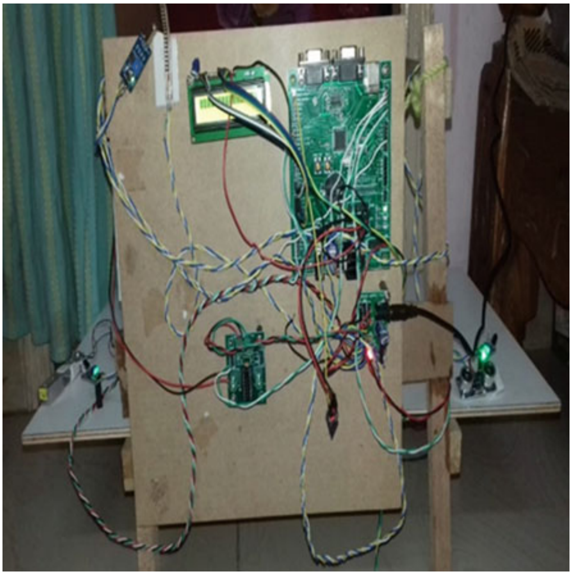


Fig. 3 Experimental setup

Fig. 4 Output of load cell

Fig. 5 Output of IR sensor

Figure 6 shows the output of vibration and flex sensor and also boat height estimation output. If any vibrations occur in the bridge sensor detects it at the initial stage. It shows alert message to the vehicles displaying it on LCD. It also sends the information to nearby offices. Flex sensor detects the crack present on the bridge.

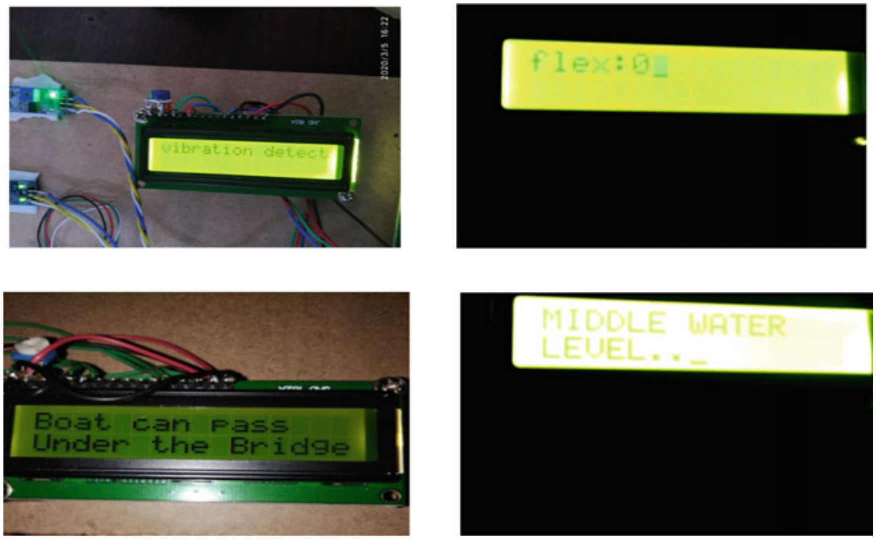


Fig. 6 Outputs of flex, vibration and boat height estimation

We measure the resistance of the flex sensor. When the resistance of the flex sensor increases then it means bridge is experiencing more pressure and there is crack in the bridge. The threshold maintained for flex sensor is 600. If it reaches more than 600 then alert message is displayed on LCD to inform vehicles.

Boat height is estimated using MATLAB image processing using Rgb color model. We use 3 cases where position of the water level sensor is changed. We check for two methods live boat and stream boat. Boat height is sent to the Arm and the inputs from water and pressure sensor is also sent to arm. The Arm calculates whether boat can cross the bridge and sends the information to the boat using Zigbee.

6 Conclusion

The proposed methodology is used in bridge health monitoring system. It is used to avoid cracks on the bridge, maintains bridge conditions. It keeps continuous track on the pressure level, water level and humidity level and provides the information to the nearby offices through Zigbee. As a monitoring system it provides safety transportation for bridge waters. It reduced the accidents by negligence and damage due to environmental conditions like flood, tsunami etc.

References

1. Sazonov E, Li H, Curry D, Pillay P (2009) Self-powered sensors for monitoring of highway bridges. Sens J IEEE 9:1422–1429
2. IoT based bridge safety monitoring system. Int J Res Appl Sci Eng Technol (IJRASET) 7:26–2331 (2019)
3. Bridge health monitoring systems. Int J Res Appl Sci Eng Technol (IJRAS-ET) 6:327–331 (2018)
4. Lee J-L, Tyan Y-Y, Wen M-H, Wu Y-W (2017) IEEE Development of an IOT-based bridge safety monitoring system. In: Proceedings of 2017 IEEE international conference on applied system innovation
5. Jha AK (2016) Bridge monitoring system. Int J Innov Stud Sci Eng Technol (IJISSET) 2:2044–2047
6. Bridge monitoring and alert generation system using IOT. Int J Adv Res Ideas Innov Technol (IJA-RIIT) 3(2017)

Ultrasonic Blind Stick Using ATmega328P

Sulochana I. Akkalkot, G. H. Nandini, Y. Krishna, L. Sowmya, and K. G. Ravikumar

Abstract Ultrasonic sensor assisted walking stick is developed as an assistive technology for visually impaired people. To avoid obstacles and potholes for the visually impaired people HC-SR04 module, ultrasonic sensor is used. The sensor would warn the visually impaired people about the obstacles and the pothole. This implementation is framed using ATmega328P microcontroller. Visually impaired persons can use this for obstacles and pot holes that is within a distance of 2–400 cm or 1–13 ft.

Keywords Microcontroller ATmega328P · Ultrasonic sensor · HC-SR04 · Smart · Blind stick

1 Introduction

In the year 2019 World Health Organization (WHO) published a report that around 39 billion people are blind. Blind people face lot of challenges to do their daily activities but chief among them is navigation. An average person needs about 150 min of exercise a week. Blind people can move easily around their house without any help because they know the position of things in the house [1]. Commercial places can be whole another challenge. Even with regulation to build ramps and tactile tiles in these commercial spaces, they are seldom enforced. This adds to the problem for blind people who might want to visit the place [2].

People with visual disability generally tend to keep their houses unchanged so that they have a mental map of the layout with them to navigates around without bumping into things and breaking them or worse cause harm to oneself. But when friends or relative visit them they might inadvertently change the location of things, which becomes another challenge for them [3].

As a solution to issues mentioned above a project is proposed to develop a smart portable assistance unit (blind stick) for the blind to navigation in public places. The

S. I. Akkalkot (✉) · G. H. Nandini · Y. Krishna · L. Sowmya · K. G. Ravikumar
Department of Electrical and Electronics Engineering, Don Bosco Institute of Technology, Bangalore 560074, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_126

uniqueness of this project is spatial awareness the stick offers to the blind with vacillating sound [4]. This is achieved by a using buzzer attached to the microcontroller that intelligently activates when blind person reached near an obstacle. Finally, the stick is embedded with GPS tracker to help the blind find their stick in case they misplaced it [5].

The blind stick architecture under consideration is presented in Fig. 1. There are four sensors primarily used. All of them interface to the microcontroller. The microcontroller chosen for this purpose is ATmega328P for the ease of interfacing, programming, and availability. The four sensors are ultrasound sensor, light sensor, moisture sensor and IR sensor.

When an average blind person is sweeping his stick from side to side it allows the ultrasound sensor to capture the relative distance of the object to his walking speed. If the person does not navigate around it in time, it sounds a buzzer. Larger the object louder the sound intensity but not to the point of deafening. IR sensors are used alongside ultrasound to increase reliability and redundancy.

To account for all road conditions and weather moisture sensor and light sensor are affixed to the stick. The moisture sensor detects whenever there is water or not. On rainy days where there is high chance of encountering a puddle or water pits, the stick can alert the person though hectic feedback.

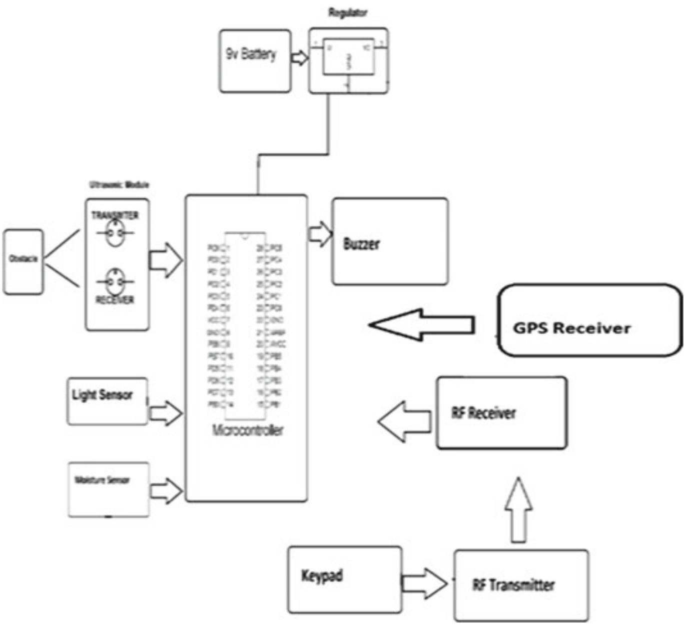


Fig. 1 Blind stick architecture

2 Hardware

2.1 Ultrasonic Sensor

The ultrasonic sensors module HC-SR04 emit sound with wavelength of 1.9 cm or less (>20 kHz), which is inaudible to human ears. The sound waves interfere with the obstacle and bounces back to detectors. The ultrasonic sensor records the incoming wave from the objects/obstacle. The microcontroller makes a record start time the moment the sound is send and stop time once the sound is received at the sensor. The distance is calculated as below:

$$\frac{1\,\mu\text{S}}{58} = 17.24 \tag{1}$$

Upon calculating the distance of the obstacle certain conditions are checked in the program to judge relative safe distance. The microcontroller operates the buzzer accordingly.

2.2 Buzzer

The buzzer is small sound device. Since it is low power it can be powered by dc voltage of 5–9 V.

2.3 Light Sensor

The light sensor used is ORP12 Cadmium Sulphide photoconductive cell. It is the most widely available sensor. It works on the principle of light dependent resistor, sensor's resistance decreased when it is exposed to sunlight.

2.4 Microcontroller

ATmega328P is an 8-bit microcontroller with 23 programmable I/O lines it is well suited for our application. It consumes a mere 1.5 mA at 3 V which makes it ideal for use in portable and light weight battery powered application.

3 Software

3.1 Algorithm

Figure 2 depicts the algorithm of the proposed work.

MCU Sends a high pulse to trigger pin of ultrasound sensor and waits for high pulse response from the object in the sensor's ECHO pin.

(i) The MCU waits till the bits are acquired and comes out of the loop. Otherwise it stays in the loop.

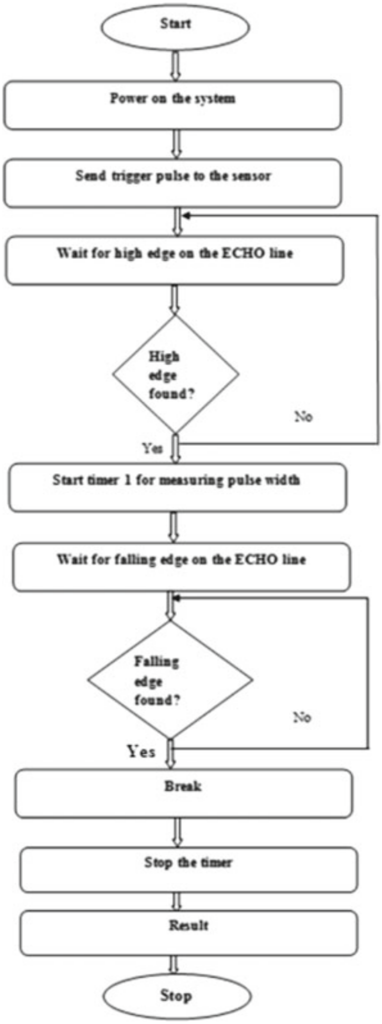


Fig. 2 Algorithm of proposed system

(ii) TIMER1 is used to measure the distance of the object that is proportional to the delay measured in the timer.
(iii) Timer1 is set with a pre scaler value of 2. Thus, the counting frequency is half that of the CPU frequency. For a 5 MHz frequency (crystal frequency, 20 MHz), the frequency of the timer is 2.5 MHz and period is 0.4 µs.
(iv) The I/O line connected to the ECHO pin on the sensor is continually polled. On receiving allow online the MCU breaks the loop.
(v) The pulse duration is used to calculate the distance using the formula as given in (1).

3.2 Code

```
#include <SoftwareSerial.h>
#include <TinyGPS.h>
SoftwareSerialmineSerial(9, 10);   // (Rx, Tx)
intldrpin = A0;                  // ldr
intbuzzpin = 2;
constinthydrogenmeter = A1;  // analog pin A1
constinttriggerPin = 3;            //ultrasound
transmitter
constintechoPin = 4;               //ultrasound sensor
receiver
constintbuttonPin = 11;
constint trigPin2 = 5; // ultra 2
constint echoPin2 = 6;
intbuttonState = 0;
void setup ()
{
Serial.begin(9600);    //ldr
pinMode (ldrpin, INPUT);
pinMode (buzzpin, OUTPUT);
pinMode(buttonPin, INPUT);
Serial.begin(9600);
pinMode(hydrogenmeter,INPUT);
pinMode(triggerPin, OUTPUT);       // triggerPinis set as
Output
pinMode(echoPin, INPUT);           // echoPinis set as
Input
Serial.begin(9600);                // Serial
communication is started
```

```
pinMode(triggerPin2, OUTPUT);        // triggerPin2is set
as Output
pinMode(echoPin2, INPUT);            // echoPin2is set as
Input
Serial.begin(9600);                  // serial
communication is set to 9600hz
mineSerial.begin(9600);              // Setting the baud
rate of GSM Module
Serial.begin(9600);                  // baud rate of Serial
Monitor is set to 9600hz
delay(100);
}
void loop ()
{
while(1)
{
if (buttonState == HIGH) {
Serial.println("gsm button");        // turn LED on:
msend();
}
intldrval = analogRead(ldrpin);
Serial.println();
Serial.print("ldr_val:");
Serial.println(ldrval);
delay(1000);
if (ldrval>= 1020)
{
tone(buzzpin,2000,500);
}
int value = analogRead(hydrogenmeter);    // Read analog
value
value = constrain(value,400,1023);
value = map(value,400,1023,100,0);// Map value : 400 will
be 100 and 1023 will be 0
Serial.print("Soil humidity: ");
Serial.print(value);
```

4 Conclusions

This work presents a reliable and portable embedded blind stick system. It considers cost of developing an economical device affordable to all spectrum of income individual by using off the shelf readily available parts. Real world use case is demonstrated for ease of use and practical feasibility for all the weather and road conditions. Future scope for this project is the use of LIDAR sensor for special sensing and wireless earphone which can communicate with the individual thought the use of machine learning and AI.

References

1. Sohini D (2018) Ultrasonic sensor based smart blind stick. In: International conference on current trends towards converging technologies (ICCTCT)
2. Sharma H, Tripathi M (2018) Embedded assistive stick for visually impaired persons. In: IEEE—43488 9th ICCCNT 2018 July 10–12
3. Liu Y-Q, Gao Z-K, Shao Z-J, Liu G-Y (2017) Intelligent ultrasonic detection of walking sticks for the blind. In: ECAI 2017—International conference—Electronics, computers and artificial intelligence, 9th edn
4. Sen A, Sen K, Das J, Ultrasonic blind stick for completely blind people to avoid any kind of obstacles. In: IEEE conference
5. Tekade A, Shekar M et al, Ultrasonic blind stick with GPS tracking system. Int J Eng Sci Comput

A Study on Virtualized Sensor Cloud Infrastructure

S. J. Akhila and N. J. Anasuya

Abstract In recent days, the most leading services among cloud services are sensors as a service, which provides on-demand sensor services to the users. With this increased usage of sensor applications, it is not possible to use the deployed sensors to meet their requirements, hence this can be provided using virtualization of cloud known as sensor cloud. It is an assimilation of the WSN and cloud-computing. This combination is advantageous for both WSN suppliers and CSP. In the sensor-cloud environment, multiple physical sensors are formed into a virtual-sensor that provides on-demand services to the users. Even though sensors have many applications, there are several research issues which include scalability, power managing, scheduling resources, mobility administration, QoS management, and security. This paper initiates a survey on sensor-cloud Infrastructure.

Keywords WSN · Sensor-Cloud (SC) · Cloud-Computing (CC) · Virtual sensor · Virtualization

1 Introduction

Comprising of spatially allocated independent sensors, wireless sensor networks (WSNs) can sense the environmental situations which include temperature of Surrounding, humidity and pressure of nature, etc. [1]. For providing useful and on-request network access to a shared pool of configurable computing, cloud-computing is most used computing model.

Sensor-cloud is a group of WSN and cloud. It delivers a stage for storing and managing information using cloud computing elements. It operates in way that; all the sensor points gather the information then transmits to cloud for storing and processing, which permits users to retrieve the information from anyplace and at any point of moment instantaneously. SC provides features for the users like, data

S. J. Akhila
KLE Society's Degree College, Nagarbhavi, Bangalore, Karnataka, India

N. J. Anasuya (✉)
Don Bosco Institute of Technology, Bangalore, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_127

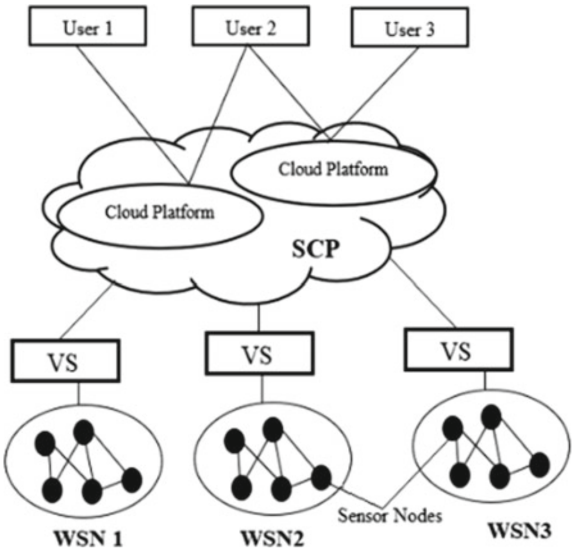


Fig. 1 Sensor cloud architecture

processing, visualization, analyzing, collecting, and sharing the sensed information [2, 3]. The following explains the structure of Sensor-Cloud (Fig. 1).

Some of the factors that sensor-cloud includes are

- User: End-consumer or client who uses services from sensor cloud.
- Cloud Platform: Cloud Servers which provide services to clients.
- VS: VS means virtual-sensor is logical sensor group that receives information from sensor nodes.
- WSN: A group of spatially allocated independent sensors.

The remaining paper is formed as, second part gives the introduction to sensor-cloud environment, third part gives details on sensor-cloud, and Fourth part concluded with future studying objectives.

2 Background Details on Sensor-Cloud

Some of the preliminaries of sensor cloud are as follows,

(A) **Cloud Computing (CC)**

CC is an emerging and growing technology in IT industry that includes features like virtualization, power processing, storage management, sharing resources, distributed networks and connectivity. In this rapidly developing world Services of cloud are used in various application [4]. Cloud Computing is a method of providing computed resources to consumers through internet. The four types of cloud services are shown in following diagram (Fig. 2).

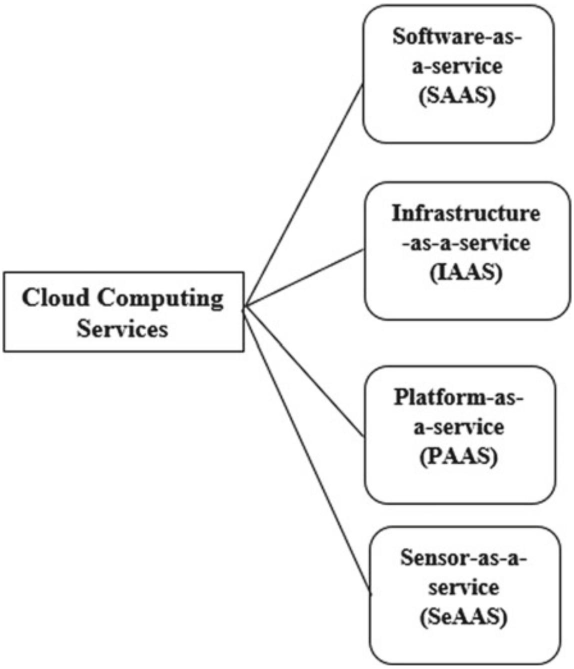


Fig. 2 Types of cloud services

(B) **Wireless Sensor Network (WSN)**
WSN can be defined as collection of wireless scattered sensor nodes. These scattered sensor nodes commonly come up with minimal processing energy and storage availability. WSN can be divided as i.e., structured WSN and unstructured WSN. In Unstructured WSN, the sensor nodes can be deployed in an adhoc system whereas in structured WSN proper planning to done before deploying [3] (Fig. 3).

(C) **Characteristics of Sensor-Cloud**
Sensor-cloud provides an open source, smart and inter-operable wireless sensor network. When request for services occurs, the demanded services

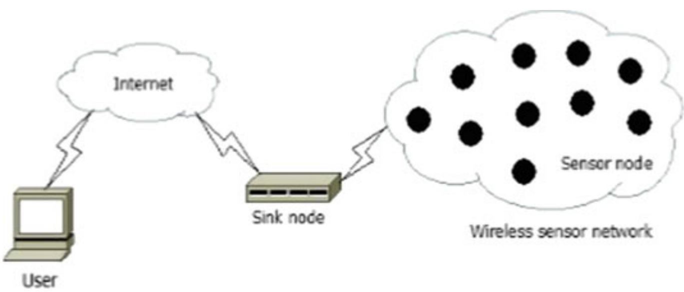


Fig. 3 Wireless sensor network

are automatically directed to WSN, these WSN gathers the data, forms the Virtual sensor group and send responses to the users, and the allocated sensors become free after the completion of request. Some of the characteristics of sensor cloud are

(a) Analysis: Gathered sensed information is beneficial for various types of examination or research purposes.
(b) Scalability: Sensor-Cloud provides scalability, means the groups or the owner can insert any sensor related services from cloud-providers without help of hardware.
(c) Visualization: The SC will provide visualization Application Program Interface which envision the sensor information from various network.
(d) Multi-occupancy: To enable the demand of users request, the sensor-cloud simply combine services of various service suppliers into one. This assimilation permits access-gain to many cloud (data centers) maintained anywhere from the world.
(e) Automation: This feature increases the data broadcasting time.
(f) Resource Allocation: Sharing resources for various of applications allows resource allocation.
(g) Fast response time: Sensor cloud provides quick responses that mean as and when user wants sensor services.

(D) **Sensor-Cloud Benefits**

(a) Transport: In transport field SC is widely used for traffic signal control, emergency warning, routing, fee or toll gathering, number plate identification etc.
(b) Military-Filed: The WSN are helpful in military field, tackle's and missiles, forces investigation, battle area supervision, battle-destruction calculation, target findings, information related to nuclear, chemical related attacks and enemy findings based on biological factors.
(c) Weather Examining: Monitoring the information related to weather contains upcoming prediction, data gathering, estimating and weather forecasting.
(d) Health Field: WSN are immensely helpful in the field of health. Some hospitals use sensor-nodes to examine the patient's information so that, medicine dosage will be tracked thus controlled, and the particulars of patients and doctors can be kept in cloud for further use.
(e) Automation: Automation played an important role in scheduling the Sensor-Cloud environment, automation services enhance the delivery time.
(f) QoS Provision: Quality of Service is one the most significant aspect of the cloud computing atmosphere, using sensor cloud better service will be provided by scheduling a particular data center (Fig. 4).

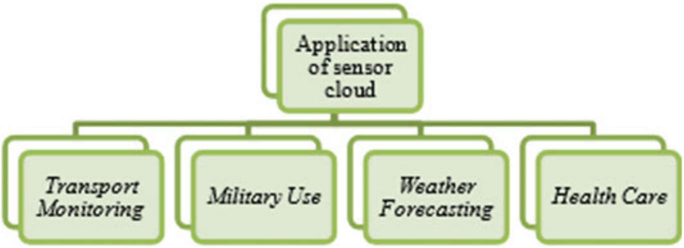


Fig. 4 Application of sensor cloud

(E) **Cloud Simulators**

It is not possible in real world for cloud users to conduct experiments on cloud, in that cases cloud simulators plays an important role. Some of the advantages of cloud simulators are

- Simulators provides dynamic and flexible configuration for executing cloud programs.
- Simulators are simple-to-use.
- Simulators are the very cost-efficient solution for real world implementation.
- Simulators also enable scientists to create cloud environments with their own planned execution, security and other provisioning algorithms.

One of the important cloud simulators which is useful for researchers is Aneaka cloud simulator developed by Manjrasoft Pvt ltd. designed by Dr. Rajkumar Buyya, which includes the concepts of cloud sim, a simulator tool developed by cloudlabs, Melbourne University Australia with the guidance of Dr. Rajkumar Buyya and his group which uses java language. The following figure shows layered cloudsim architecture which is achieved by expanding the features of gridsim simulator that helps in testing and developing new algorithms [15] (Figs. 5 and 6).

(F) **Issues and challenges**

Some of the challenges in sensor cloud are

(i) QoS: QoS is the measure of overall performance/efficiency of network which is a big challenge in the sensor cloud.
(ii) Energy efficiency: WSN and cloud-computing both have the same disadvantages. Power or Energy of sensors vanishes because of processing and the pre-defined storage capacity. Usage of Memory and efficiency of sensor nodes are enhanced via sensor-cloud structure. It is like a central layer which focuses on information which will be useful in long term maintenance.
(iii) Bandwidth constraint: The defined ideal and effective bandwidth depends on allocation methods, due to increased number of device resources users cannot manage bandwidth allocations.

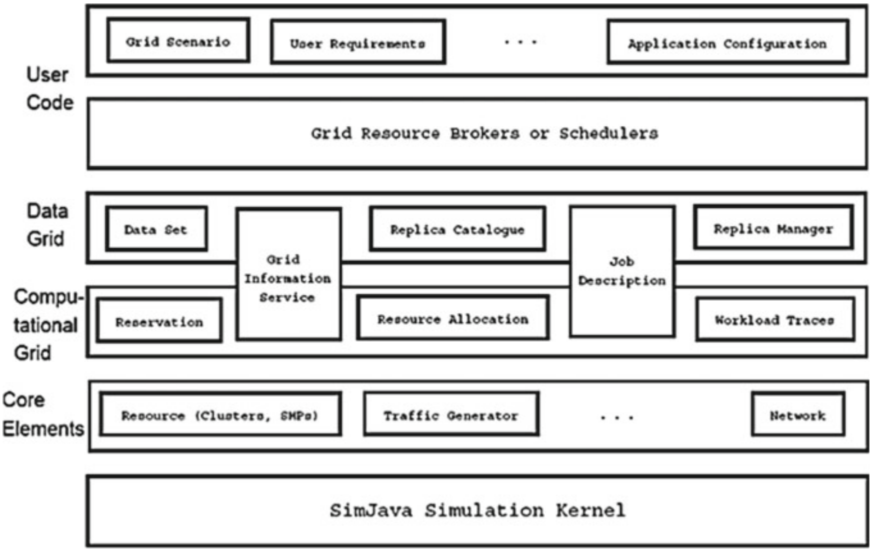


Fig. 5 Structure of cloud simulator

```
Starting CloudSimExample1...
Initialising...
Starting CloudSim version 3.0
Datacenter_0 is starting...
Broker is starting...
Entities started.
0.0: Broker: Cloud Resource List received with 1 resource(s)
0.0: Broker: Trying to Create VM #0 in Datacenter_0
0.1: Broker: VM #0 has been created in Datacenter #2, Host #0
0.1: Broker: Sending cloudlet 0 to VM #0
400.1: Broker: Cloudlet 0 received
400.1: Broker: All Cloudlets executed. Finishing...
400.1: Broker: Destroying VM #0
Broker is shutting down...
Simulation: No more future events
CloudInformationService: Notify all CloudSim entities for shutting down.
Datacenter_0 is shutting down...
Broker is shutting down...
Simulation completed.
Simulation completed.

========== OUTPUT ==========
Cloudlet ID    STATUS    Data center ID    VM ID    Time    Start Time    Finish Time
    0        SUCCESS       2           0        400       0.1        400.1
*****Datacenter: Datacenter_0*****
User id         Debt
3               35.6
**********************************
CloudSimExample1 finished!
```

Fig. 6 Output for the execution of cloud sim example 1; (A simple example showing how to create a datacenter with one host and run one cloudlet on it)

(iv) Security and confidentiality: Since the communication occurs via internet data need to secure. Data integrity must be protected by providing access to approved or authorized transaction. Users must know that his/her information should be encoded or encrypted before sending in o cloud.

(v) Interface standardization: Now a days, web interfaces performance as intermediary between cloud and sensor-cloud users. Even though, web interface is not suitable for smartphone and devices in mobility which causes overhead.

3 Literature Review

Literature Review gives a brief view on various sensor cloud parameters. Chatterjee et al. [7] explains the scheduling of data centers in sensor cloud infrastructure. Misra and Chakraborty [8] explains dynamic scheme for mapping virtual sensor to sensor-points around sensor-cloud environment. El Rachkidi et al. [9] defines an algorithm that focus on optimize the distribution of VS and VMs in sensor-cloud structure. Sheng et al. [10] focuses on latest developments in industrial wireless sensor network for well-organized administration in IoT. Zhu et al. [11] developed a model for energy efficient green IoT. Sheng et al. [12] invented new method that minimize energy utilization of application. Rahman [13] propose a method that reduces the bandwidth utilization and response time. Khan et al. [14] explains how information will transmitted in Sensor Cloud in energy efficient manner. Dwivedi [3] provides a brief study on sensor cloud. More et al. [15] explains the simulation tool called cloud-sim which is responsible for executing various cloud-based research programs.

Table 1 shows the comparative analysis of related work.

4 Conclusion and Future Directions

This paper relies on the background study of sensor cloud and presents various research workings which are taken in this area. Many structures and architectures were created in sensor-cloud field. By this study, our future direction focuses on parameters effecting the QoS in sensor cloud.

Table 1 Comparative study of related work

Authors	Impact	Techniques
Chatterjee et al. [7]	Data center scheduling for quality of service management	Optimal decision rule
Misra and Chakraborty [8]	QoS-Aware dispersed dynamic mapping of VS	QADMAP scheme
El Rachkidi et al. [9]	Resources optimization and efficient distribution	Algorithm reduces the number of deployed virtual machine
Sheng et al. [10]	Efficient management in IoT	Energy efficient
Zhu et al. [11]	Green IoT	Energy efficient
Sheng et al. [12]	Energy efficient cooperative computing in mobile wireless sensor network	Minimize energy consumption
Rahman et al. [13]	Efficient sensor cloud communication	Lossless data compression algorithm
Khan et al. [14]	A review on energy efficient data transmission in sensor cloud	Survey
Dwivedi [3]	A review on sensor cloud	Survey

References

1. Zhu C, Shu L, Hara T, Wang L, Nishio S, Yang LT (2014) A survey on communication and a data management issues in mobile sensor networks. Wirel Commun Mobile Comput 14(1):19–36
2. Dwivedi RK, Saran M, Kumar R (2019) A survey on security over sensor-cloud. In: 2019 9th IEEE International conference on cloud computing, data science & engineering (Confluence), Noida, India, pp 31–37
3. Dwivedi RK, Kumar R (2018) Sensor cloud: integrating wireless sensor networks with cloud computing. In: 2018 5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON)
4. Moghaddam FF, Ahmadi M, Sarvari S, Eslami M, Golkar A (2015) Cloud computing challenges and opportunities: a survey. In: 2015 1st international conference on telematics and future generation networks (TAFGEN)
5. Singh MK, Amin SI, Imam SA, Sachan VK, Choudhary A (2018) A survey of wireless sensor network and its types. In: International conference on advances in computing, communication control and networking (ICACCCN 2018)
6. Khan MF, Dwivedi RK, Kumar R (2019) Energy efficient data transmission in sensor cloud: a review. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI)
7. Chatterjee S, Misra S, Khan SU (2015) Optimal data center scheduling for quality of service management in sensor-cloud. IEEE Trans Cloud Comput
8. Misra S, Chakraborty A, QoS-aware dispersed dynamic mapping of virtual sensors in sensor-cloud. IEEE Trans Serv Comput
9. El Rachkidi E, Agoulmine N, Chendeb N, Belaid D (2017) Resources optimization and efficient distribution of shared virtual sensors in sensor-cloud. In: 2017 IEEE international conference on communications (ICC)
10. Sheng Z, Mahapatra C, Zhu C, Leung VCM, Recent advances in industrial wireless sensor networks toward efficient management in IoT, vol 3. IEEE Access
11. Zhu C, Leung VCM, Shu L, Ngai EC-H, Green Internet of Things for smart world, vol 2015. IEEE Access

12. Sheng Z, Mahapatra C, Leung VCM, Chen M, Sahu PK, Energy efficient cooperative computing in mobile wireless sensor networks. IEEE Trans Cloud Comput 6:114–126
13. Rahman MT (2017) Efficient sensor cloud communication using data classification and compression. Int J Inf Technol, 9–17
14. Khan MF, Dwivedi RK, Kumar R, Energy efficient data transmission in sensor cloud: a review. In: 3rd international conference on trends in electronics and informatics (ICOEI)
15. More NS, Ingle RB (2019) Study of Aneaka and Cloudsim simulators. Int J Emerg Trends Eng Dev 1(9). January 2019

Digital Solution to Campus Life: Campus Mate

G. S. Gowramma, Ramanand Sirvi, G. Manish Kumar, P. B. Harshith, and Manjunath D. Murdi

Abstract In everyday campus life both students and staff struggle to keep up with the schedule and tasks assigned to them which are to be completed within the given timeframe. Students often find it difficult to maintain a fine balance between activities of academic nature and co-curricular activities, be it Hackathon, activities of Cultural nature etc., (Tung et al. in Global conference on technical workshops in knowledge networking and technology [4]) Staff often face problems in keeping track of all the student details collected over some time, and often these records are lost. The idea is to create a solution that provides easy access to information collected in an institution by organizing and maintaining the information clean and secure. The other goal is to make campus life easier for active members with any role in an institution. It is very difficult for the students to navigate through the tasks, syllabus, and scoring aspects related to academics. It is also not easy to keep up-to-date with the activities that happen on the campus. To remove these restrictions and to save time for both students and institution staff, a tool must rise which can handle all these tasks efficiently. The tool or application in context should be capable of automating most of this or at least should have an easy to use UI which would make it very simple to do these tasks. Campus Mate solves these problems and provides a custom interface for each user role. E.g.: Student would see a personalized homepage showing his marks, attendance, assignments, etc., whereas staff would see a dashboard on the homepage showing a graph of student's marks, their active participation in the academic activities, etc., and Principal would have a view where he could view statistics and important higher-level details on the homepage.

Keyword Institution · College · LMS · Management software

1 Introduction

In everyday campus life, both students and staff struggle to keep up with the schedule and tasks assigned to them which are to be completed within the given timeframe.

G. S. Gowramma (✉) · R. Sirvi · G. Manish Kumar · P. B. Harshith · M. D. Murdi
Department of Information Science, Don Bosco Institute of Technology, Bengaluru 560074, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_128

Students often find it difficult to maintain a good mix between education and curricular activities sports, be it Sports, Cultural, etc., Staff often face problems in keeping track of all the student details collected over some time, and often these records are lost. The idea is to create a solution that provides easy access to information collected in an institution by organizing and maintaining the information clean and secure. The other goal is to make campus life easier for active members with any role in an institution. Campus Mate is a platform that provides a One-stop digital solution which makes the whole campus life easier by eliminating all the manual work and making it completely digital hence reduces time and paperwork making the whole process environment friendly. A platform made for colleges that provide complete digital solutions concerning day to day activities at the institutions. Creating a platform using modern frameworks and technologies to digitalize the entire paperwork process carried out in an institution and to provide an app for students and staff to save their time in many activities.

It is very difficult for the students to navigate through the tasks, syllabus, and scoring aspects related to academics. It is also not easy to keep up-to-date with the activities that happen on the campus. To remove these restrictions and to save time for both students and institution staff, a tool must rise which can handle all these tasks efficiently. The tool or application in context should be capable of automating most of this or at least should have an easy to use UI which would make it very simple to do these tasks. Campus Mate solves these problems and provides a custom interface for each user role. For e.g.: Student would see a personalized homepage showing his marks, attendance, assignments, etc., whereas staff would see a dashboard on the homepage showing a graph of student's marks, their active participation in the academic activities, etc., and Principal would have a view where he could view statistics and important higher-level details on the homepage.

As we witness in and around the campus of the college, we notice all kinds of people in the institution including students, teaching and non-teaching staff face various kinds of problems and hurdles in any task they do at any time of the day; be it paperwork, everyday routine or even students catching up with the academic progress at a proper pace and managing all their curriculum-related work [2]. We aim to develop an application ecosystem that would enable each and everyone in the institution (the principal, teaching faculty, HODs, students, and non-teaching staff including the library) to manage their tasks efficiently by making use of an application platform built using digital technologies. With this platform, we aim to provide smart digital solutions for problems faced in the institutions.

1.1 Key Technologies

PostgreSQL (Database)

This open-source database is taking the world by storm, both on the ground and up there in the cloud. It gives you an advanced relational database for the low, low price of free. Supports a large number of SQL standards.

ExpressJS (Server)

Express is a fast and an opinion minimalist web framework for nodejs. You can write JavaScript in your front-end and the back-end and you can create a fully functional server out of it.

ReactJS (Front-end)

It is a JavaScript library for building a fast and interactive user interface. It was developed by Facebook in 2011.

NodeJS (JS runtime)

It is a runtime environment where you can run JavaScript on a standalone machine. Using NodeJS you build standalone applications.

React Native

A framework for building native apps using JavaScript created by Facebook in 2015. Increased developer productivity and faster time to market. It has an amazing feature called hot reload where the moment you save the changes to see results on the simulator.

Typescript

Typescript is a language when you compile, it will compile to a JavaScript language. Typescript creates an equivalent JavaScript file when executed.

2 System Design

2.1 Architect Diagram

Machine design is the paradigm that describes a system's configuration, actions, and specific views. It consists of an integrated element and established sub-systems that will operate together to execute the system as a whole. The students must sign on with the college administration program. When the computer logs in such an admin area, the admin the students and professors must be allowed to log in to the program, only the admin will provide full access to the system. He will give the students circular and also the faculties about the forthcoming activities that will occur at

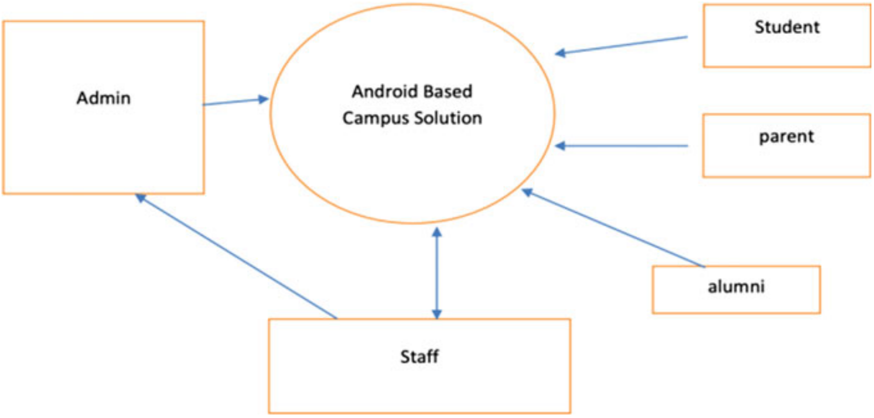


Fig. 1 App architecture

the campus. Administrators should set the system schedule, and even, generates the professorship study as well as the graduates. If a member of the faculty logs into the program, his/her task is to control the participation and internal appraisal points of the professorship study as well as the graduates. If a member of the faculty logs into the program, his/her task is to control the participation and internal appraisal points and to give a memo to his graduates about his/her topic updates, such as class exams, laboratory exams, etc. The teachers would be allowed to produce the student's study dependent on their success in their class. Once a student logs into the program, he can only access his specific profiles and can review their educational excellence and participation. The reports of a student like internal assessment marks, attendance, assignment submissions, etc., will be automatically generated in their profile. All this information is stored in the database (Fig. 1).

2.2 Data Flow Diagram

A data flow chart is a graphical visualization of the process of data management visualization. In a Data model, data objects migrate into an internal mechanism from an external entity or internal information origin to an inner data source or a live data sink. DFD does not include details on the pacing of the method or whether the project should be operating sequentially in sequence (Fig. 2).

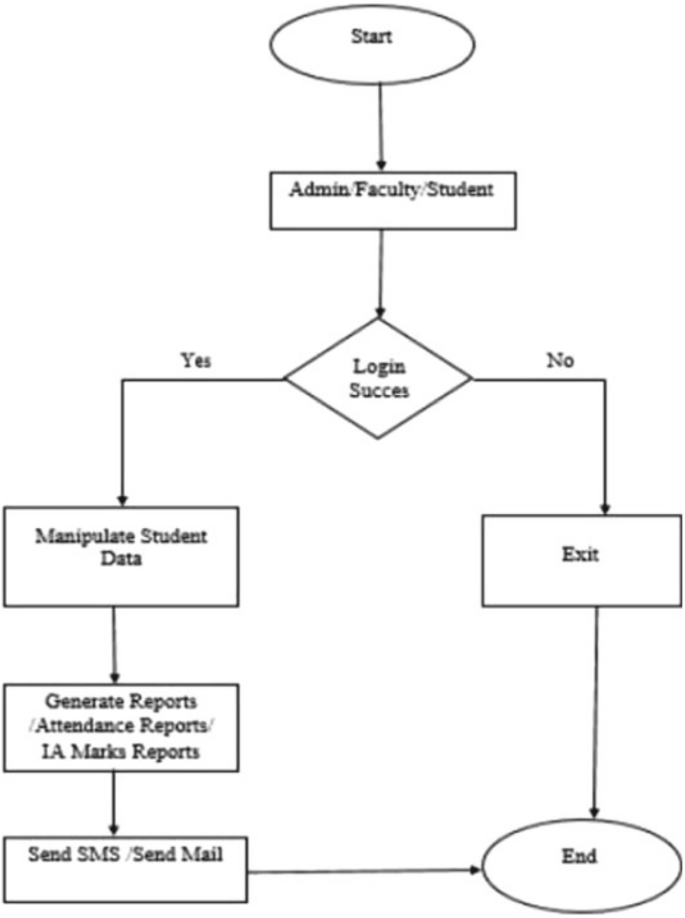


Fig. 2 Data flow diagram of application

2.3 Database Schema

A relational database schema is a framework that is defined in a structured language provided by Database Management System. The word “schema” applies to the data structure as a drawing-up of the process of assembling the database [3]. The database schema is formally defined as a Collection of formulae (sentences) named database-imposed quality restrictions. The combined restrictions guarantee all portions of the model are consistent. In the same language, all the restrictions are always expressible. Through the understanding of the language of the database, a database can be known as a framework. The relational schema generated is transformed into an actual mapping, the structure of the database. It illustrates how the Server models the real-world entities.

A database schema specifies, based on the database administrator's knowledge of possible applications, the facts that can enter the database or those of interest to the possible end-users [1]. The concept of this "theory" coincides strongly with the database, which can be regarded as a mathematical at any moment object. A scheme includes formulas describing the validity constraints unique to an application and the category of database represented in the same language of the database the schema describes the tables, rows, associations, lists, indices, sets, processes, roles, queues, causes, forms, sequences and materialized views in the relational database.

In general, the database stores the schema in the dictionary of results. While the text database vocabulary describes a schema, the word is often referred to a GUI description of the configuration of the database. In other terms, a schema is the database layout, which determines the items in the database.

There is also a cause for security maintenance, as database contains all the factual data about institution and its participants [5].

3 Structural Design of the System

3.1 API Architecture

There were several design considerations and concerns as to how the server architecture be and what layers it should include.

Figure 3 shows different layers used for server architecture. The base architecture used for development on the server is MVC. This architecture is tweaked for its use on an enterprise scale. Keeping those requirements in mind, there are four main layers involved in the final architecture.

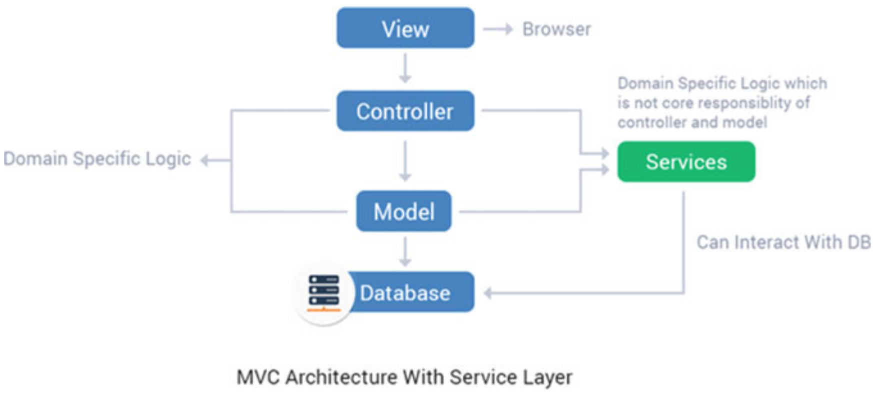


Fig. 3 MVC architecture of the application

Business or Service Layer: This is probably the most important layer included in the architecture. The responsibility of this layer is to separate business concerns or logic from the other layers. This is particularly useful if the application is to scale on many platforms as this layer doesn't need any modification what-so- ever on any platform.

Controllers Layer: The responsibility of this layer is to take a request object, response object, and other headers **from** the request and give an appropriate response. It is important to make sure that even if an error occurs it's handled gracefully and appropriate response is sent back.

Router Layer: The responsibility of this layer is to get the route string, Http verb, and query parameters. This **layer** also applies middleware and attaches the appropriate info to the request header and then forwards it to the controller to handle the rest of the operations.

Models Layer: The responsibility of this layer is to handle data schema, storage and transactions. Services interact with this layer to perform operations on data required by the business. The most used pattern for data translation from database to native language is "Repository Pattern". Repository pattern encourages translating business requirements into class based data models.

Leveraging this architecture in our case: The problem can be divided into these layers and service layer can be implemented as the base layer containing all the business classes and methods. The controllers layer can be used to write the controllers and gateways for the API solution.

4 Modules

4.1 Team Building Module

Problem: Students often find it difficult to find other students in the campus with similar skills or interests.

Solution: Students would fill out their profiles with their skills and interests. Students with similar skills/interests can be found on Campus Mate and they can team up to achieve something that they are interested in.

4.2 Assignment Module

Problem: Students find it difficult to search their assignment questions in WhatsApp or mobile phones.

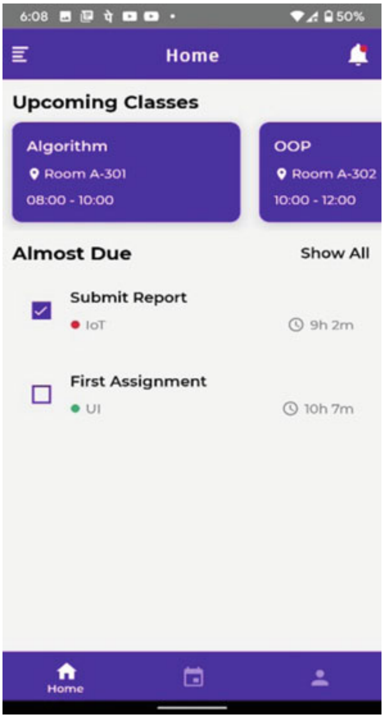


Fig. 4 Assignment module

Solution: To overcome this kind of problem this module is used so that the assignment given by the teachers is organized in order. Hence students can view the assignment questions whenever they need it (Fig. 4).

4.3 Attendance Module

Problem: Students find it difficult to know their eligibility criteria for attendance in academics.

Solution: This module helps to keep the track record of attendance of each student individually in the academics and displays in their profile (Fig. 5).

4.4 Grade Module

Problem: Most of the time students don't even know their grades properly in their academic year.

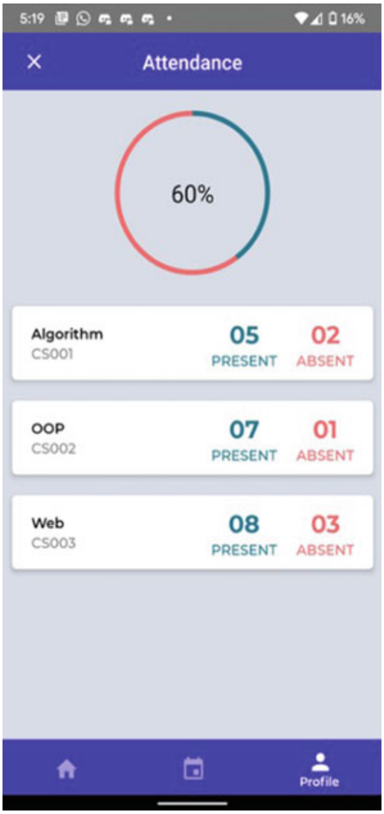


Fig. 5 Attendance module

Solution: Grade module will update the student internal marks of all the subjects scored in all the internals test and provides the average of best of 2 internals marks to each student.

4.5 Notes Module

Problem: Students find difficulties in searching for proper notes required according to their syllabus.

Solution: The purpose of this module is to provide the proper notes and question banks for the students.

5 Conclusion and Future Enhancement

Education management solution is particularly helpful in an organization, college, or university. This new program will not require documentation. Surveillance may be performed from everywhere. Such a project specifically minimizes the required human effort. The institutes manage this task, ensuring that there is no information loss and data is protected. As it is an internet-based program, anybody can access the framework anytime at any time so without the lag, it is very simple to encourage the necessary details. Because the college can administer this program if they require any improvements to a program they will do it without the initial cost, so the framework would be better than managed by the college itself (Fig. 6).

Even with the automation done in the proposed system, there is a huge scope for automating general to specific tasks for both faculties and students. Machine Learning can be used in major concepts of the project to give analytics of all students and provide good predictions and suggestions related to Academics and Extracurricular activities. The platform can be extended to include the schools and other Institutions making it easy for them to access the content and stay in touch with the activities in the said Institution.

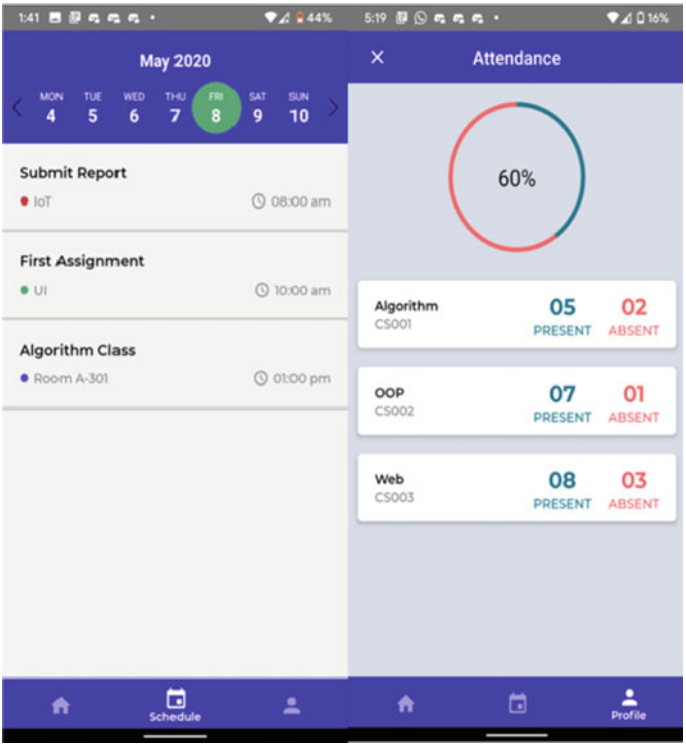


Fig. 6 Assignment and attendance report

References

1. Gandhi L, Silva C, Gualotuna T (2017) Mobile application development process—a practical experience. In: 2017 12th Iberian Conference in Information Systems and Technologies (CISTI)
2. Cruz L, Abreu R, Performance-based guidelines for energy-efficient mobile applications. In: The fourth international conference on wireless computer technology and applications. IEEE/ACM
3. Vantova Z, Paralic J, Gaspar V, Mobile application for creating presence lists. In: IEEE international symposium on applied artificial intelligence and informatics (SAMI)
4. Tung T-Y, Lin L, Lee DT (2012) Pandora messaging: an enhanced self-message-destructing secure instant messaging architecture for mobile devices. In: Global conference on technical workshops in knowledge networking and technology
5. Oh T, Stackpole B, Cummins E, Gonzalez C, Ramchandran R, Lim S, Best security practices for android. Blackberry and iOS. IEEE Laboratory New Innovations for Smartphones and the Internet

Implementation of an Effective Hybrid Partial Transmit Sequence Model for Peak to Average Power Ratio in MIMO OFDM System

R. Shivaji, K. R. Nataraj, S. Mallikarjunaswamy, and K. R. Rekha

Abstract In this paper, we have proposed an effective hybrid partial transmit sequence (EHPTS), which addresses the problem of peak-to-average power ratio (PAPR) in a reconfigurable high-speed multi-input, multi-output orthogonal frequency division multiplexing (RHS-MIMO-OFDM) technology and efficiency minimizes PAPR in orthogonal frequency division multiplexing (OFDM).But the PTS technique needs to analysis all the probability of phase rotation factors which results in increases of mathematical model computation complexity grows exponentially with the increase of various subcarrier and sub-blocks. To reduce the analysis complexity, we propose a method mainly operated on to select the optimized phase factors to reduce the high PAPR. The proposed RHS-MIMO-OFDM system incorporated the novel local analysis operations that are designed and implemented to explore the optimum phase factors. The results of the simulation show that the proposed EHPTS method are reduced PAPR effects on OFDM system when it compared to the conventional PTS techniques (U = 4).

Keywords Multi input multi output · Single input single output · Peak to average ratio · Orthogonal frequency-division multiplexing

R. Shivaji (✉)
Visvesvaraya Technological University, Belgaum, Karnataka 590018, India

K. R. Nataraj
Don Bosco Institute of Technology, Bangalore, Karnataka 560074, India
e-mail: director.research@dbit.co.in

S. Mallikarjunaswamy
JSS Academy of Technical Education, Bangalore, Karnataka 560060, India
e-mail: mallikarjunaswamys@jssateb.ac.in

K. R. Rekha
SJB Institute of Technology, Bangalore, Karnataka 560060, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_129

1 Introduction

The advancement in the field of wireless digital communication has resulted in more adoption of highly efficient and more reliable advanced wireless systems as shown in Fig. 1. Various techniques are used for communication in wireless medium; one such technique, which is gaining popularity in recent years is OFDM which is based on multi-carrier modulation scheme that is used in the wide range of systems such has DSL, WLAN's, broadcasting digital media etc. Due to its highly efficient and effortless managing of inter symbol interface in channels that extremely scatters the transmitted signals, OFDM is used in applications that requires high data rate [1, 2].

OFDM is based on block modulation approach where a block consisting of M data symbols are sent parallelly over M subcarriers. The block of data is transmitted M times faster as in comparison with a single carrier system [3]. A modulator for OFDM is constructed on a block of M data symbols using inverse Discrete Fourier Transform (IDFT) trailed by analog-to- digital converter (ADC), a G sampled guard interval or an cyclic prefix (CP) who's length is equal to or greater than the channels length is used to lead the coefficients of M IDFT block [4, 5]. Each antenna's in wireless medium gets the OFDM signals from source end Inverse Fast Fourier Transform (IFFT) and the destination end uses the fast Fourier Transform (FFT) to process a signal at the fastest rate. The disadvantage of using OFDM is that it creates distortion in the signal due to its high PAPR, if the transmitter in the wireless communication system includes components like power amplifier, which is a non-linear. However, there are many ways like clipping and filtering, cyclic coding, selective mapping to reduce the PAPR factor in the signal [6, 7]. Due to its high tolerance property to multipath signals, the MIMO-OFDM arrangement is used efficiently when the transmitters and the receivers in the wireless medium use multiple antennas. However, the drawback of this arrangement is that the signal peak ratio is m times higher than the optimized average power, where m corresponds to number of MIMO-OFDM carriers. In an environment where the signals are scattered in multiple routes, a MIMO system

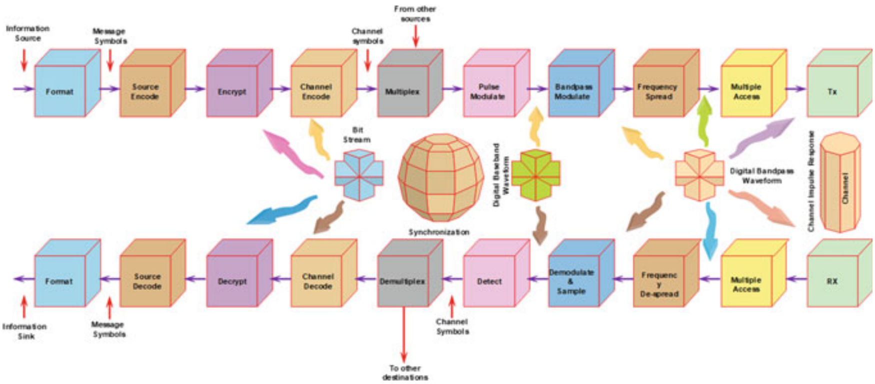


Fig. 1 Basic block diagram of wireless digital communication system

makes use of antenna diversity scheme to enhance the quality and reliability of the wireless communication system by making use of two or more antennas that are spatially separated [8]. To reduce fading of signal and also to improve the capacity gain, MIMO system can be implemented in numerous ways. The MIMO techniques are mainly categorized into three types, the first kind of technique comprises of space-time trellis code (STTC), space–time block codes (STBC) and delay diversity process which targets to maximize the antenna diversity to increase the efficiency of the power [9]. After fast Fourier Transform, the received xth sub-carrier's MIMO-OFDM symbol and the yth OFDM symbols of the ith receiving antenna is written as:

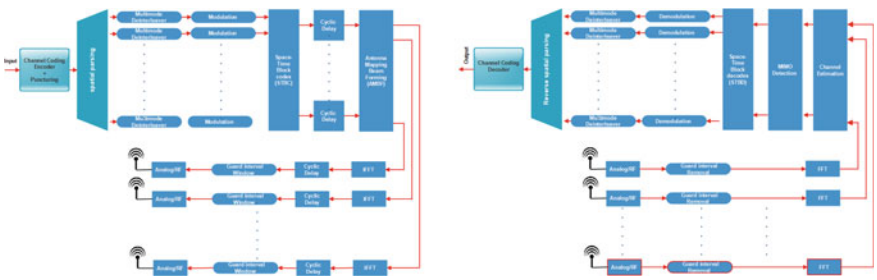
$$L_i[x, y] = \sum_{y}^{x} Q_{i,j}[x, y] N_j[x, y] + O_i[x, y]$$
$$i = 1, 2, \ldots, Y \tag{1}$$

The data symbol transmitted on xth carrier and yth OFDM symbol is denoted as $N_i[x, y]$. The contribution of additive noise at ith receive antenna is denoted as $O_i[x, y]$ and the frequency domain channel co-efficient is denoted as $Q_{i,j}[x, y]$. The data $(n_1, n_2, n_3, \ldots, n_X,)$ are simultaneously transmitted independently across multiple antennas using the same frequency characteristics bandwidth [10, 11].

Constraints of MIMO decoder antenna at receiver side is $X \geq Y$ .The data symbols that are received is expressed as follows.

$$\begin{gathered} l_1 = m_{11}n_1 + m_{12}n_2 + m_{13}n_3 + \cdots + m_{1X}n_X \\ l_2 = m_{21}n_1 + m_{22}n_2 + m_{23}n_3 + \cdots + m_{2X}n_X \\ \vdots \\ l_2 = m_{X1}n_1 + m_{X2}n_2 + m_{X3}n_3 + \cdots + m_{XX}n_X \end{gathered} \tag{2}$$

The implementation of MIMO-OFDM system consisting of X-transmitters and Y-receivers is shown in Fig. 2. The distortions caused by inter modulation increases



(a) Transmitter (b) Receiver

Fig. 2 The fundamental block diagram of MIMO-FDMA. **a** Transmitter and **b** receiver with the increase in the large peaks thereby increasing the error rate in the data transferred from transmitting antenna to receiving antenna [12]. The contents in this paper are organized as follows. Next section briefs about signals including SISO-OFDM, MIMO-OFDM and the effect of PAPR on these systems; later new EHPTS technique is introduced. The results of the simulation are shown in Sect. 3 and at the last we conclude the study and work done in Sect. 4.

2 PAPR Effects on Characteristics of MIMO-OFDM and SISO-OFDM Systems

2.1 PAPR Effect on Characteristics of Conventional SISO-OFDM System

In 'T' OFDM signal period and with Δf = 1 = T subcarrier separation, a block consisting of N Modulation symbols and N orthogonal sub-carriers generates OFDM symbol and the blocks OFDM signal xn can be written as follows

$$n_p = \sum_{x=0}^{x-1} N_x e^{j2\pi f_x / T^k} \tag{3}$$

The PAPR can be expressed as follows

$$PAPR = 10 \log_{10} \frac{P_{peak}}{P_{avg}} \tag{4}$$

where the OFDM symbol's peak power is Ppeak and the average power is Pavg. PAPR0 is defined as follows.

$$PAPR = \frac{\max\limits_{0 \le s \le T} |n_s|^2}{E\left|n_s^2\right|} = \frac{\max\limits_{0 \le s \le T} |n_s|^2}{\frac{1}{T}\int_0^T |n_s|^2 dt} \tag{5}$$

where the value expected is represented as $E[.]$. Then, the probability that OFDM symbol's PAPR surpasses a specified threshold $PAPR_0$ is specified by complementary cumulative distribution function (CCDF) can be written as

$$CCDF = P_l(PAPR > PAPR_0) \tag{6}$$

OFDM signal's CCDF for conventional OFDM considering there is no technique for PAPR reduction [13, 14] can be written as

$$CCDF = P_l(PAPR > PAPR_0) = 1 - (1 - e^{-PAPR_0})^X \tag{7}$$

2.2 PAPR Effect on Characteristics of MIMO-OFDM System

In distinct time, multicarrier system can be implemented using IFFT as modulator and a FFT as demodulator. A MIMO-OFDM system with Y transmitter which uses X subscribers, is considered, a data block which is an OFDM symbol $\{N_x, x = 0, 1, \ldots, X-1\}$ consisting of X symbols are sent in parallel over X subcarriers from a set $\{OF_x, x = 0, 1, \ldots, X-1\}$ that are orthogonal. i.e.,$OF_x = x\Delta OF$ where $\Delta OF = 1 = XT$ and T is the time period [15, 16]. The baseband OFDM signal subsequent block's is $n_{y,s}$ can be written as

$$n_{y,s} = \frac{1}{\sqrt{X}} \sum_{x=0}^{X-1} N_x e^{j2\pi f_x s}, \quad s = 0, 1, \ldots, X-1 \tag{8}$$

where N_x denotes the OFDM signal that is transmitted from yth transmitting antenna over the x subcarriers. Then the definition of PAPR can be written as

$$PAPR = \frac{\max\limits_{0 \le s \le XT} |n_{y,s}|^2}{E\left|n_{y,\ s}^2\right|} = \frac{\max\limits_{0 \le s \le XT} |n_{y,s}|^2}{\frac{1}{XT}\int_0^T |n_{y,s}|^2 dt} \tag{9}$$

The probability that OFDM symbol's PAPR that is randomly generated over Y transmitters that exceeds the threshold $PAPR_0$ in MIMO-OFDM systems is written as follows

$$P_l[PAPR > PAPR_0] = 1 - OF(PAPR_0)^{YX} \tag{10}$$

$$P_l[PAPR > PAPR_0] = 1 - (1 - e^{-PAPR_0})^{YX} \tag{11}$$

Next, each sequence of signal with lowest PAPR transmitted computes the PAPR. Figure 4 shows the block diagram of MIMO-OFDM system using RSPTS method.

2.3 EHPTS Algorithm

The conventional approach to reduce PAPR is to implement the available algorithms on each transmitter's separately. The MIMO-OFDM system uses proposed EHPTS methods that breaks the block of data into multiple sub-blocks, the sub-carriers use phase factor rotations to weigh each sub-block in order to nullify PAPR effect on system [17, 18]. Inverse phase factor rotations are used at the destination end to get the original data. Figure 3 shows the PTS technique block diagram. The downside of EHPTS method is that it requires side information (SI) at the receiving antennas. Increasing the amount of side information, which drastically reduces the throughput, can increase its performance.

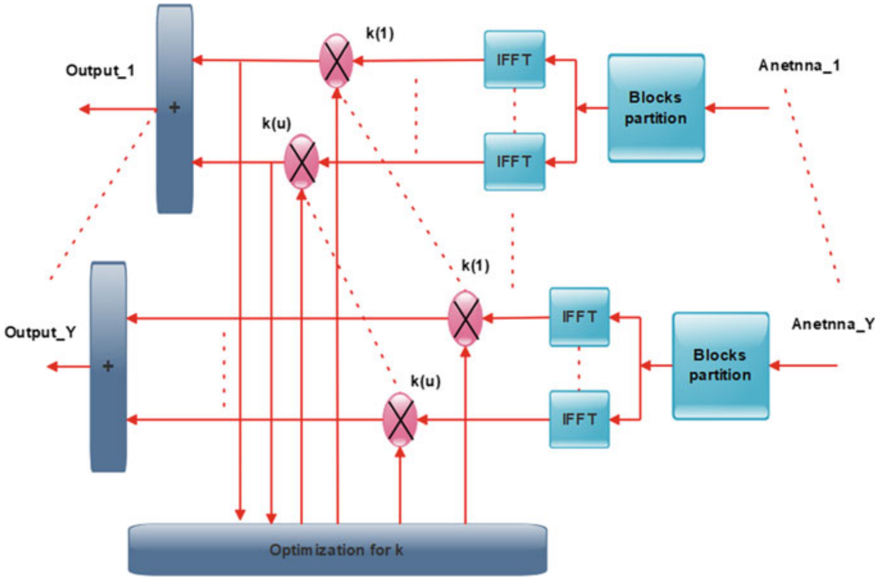


Fig. 3 PTS technique fundamental signal flow architecture

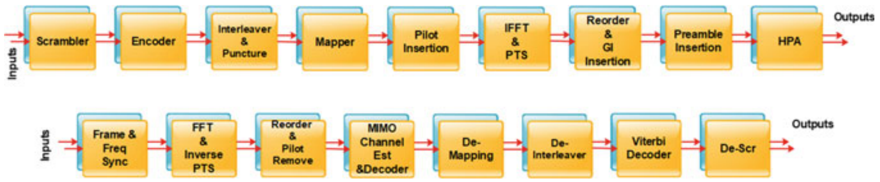


Fig. 4 Proposed high speed MIMO-OFDM systems by using EHPTS method

At transmitter end, vector is used to define an input block

$$N = [N_1, N_2, \ldots, N_X]^T \tag{12}$$

where X represents the number of sub-carriers in the OFDM signal sequence, U disjoints symbol subsequences are obtained after partitioning the vector N.

$$N = \sum_{V=1}^{V} X_V, X_V, v = 1, 2, 3, \ldots V \tag{13}$$

The combined sequence obtained after using phase rotations on equal sized sub-blocks and considering the combination of V sub-blocks partial transmit sequence-weighted sum [19, 20] can be written as

$$N = \sum_{V=1}^{V} k_V . N_v \tag{14}$$

The weighted factor that includes each sub-blocks phase rotation is denoted by k_v, Where $v = 1, 2, \ldots, v$. The conventional systems Chooses $k_v \in \{\pm 1, \pm i\}(O = 4)$, Where O indicates the number of phase factors allowed. The value of b_1 can be set to 1 without compromising the performance. The IFFT, which is a time domain vector, can be written as follows.

$$n_v = IFFT\left\{\sum_{V=1}^{V} k_V . N_v\right\} = \sum_{v=1}^{V} k_v . IFFT\{N_v\} \tag{15}$$

Here, optimal set of phase factors can be discovered by overlooking into O^{V-1}, which is a set of phase factor.

The search complexity is denoted by V (the number of sub-blocks complexity search enhancing exponentially), which increases exponentially. Data block sequences with reduced PAPR is attained.

2.3.1 Statistical Model of Nonlinear Operational High-Power Amplifier (OHPA)

A memory-less device is used to demonstrate a nonlinear OHPA. The input signal's complex envelope to the OHPA in [21] and [22] can be denoted as

$$n(t) = \gamma(t).e^{j\varphi(t)} \tag{16}$$

The following equation defines the output signal's complex

$$z(n(t) = OF\big[\gamma(t)\big].e^{\left(j\varphi(t)+\varphi(\gamma(t)\right])} \tag{17}$$

where $OF\big[\gamma(t)\big]$ *and* $\varphi\big[\gamma(t)\big]$ denotes the AM/AM and $\varphi\big[\gamma(t)\big]$ denotes AM/PM nonlinear amplifier conversion characteristics. In memory-less devices, to evaluate a system solid-state power amplifier (SSPA) is used. The $OF\big[\gamma(t)\big]$ and $\varphi\big[\gamma(t)\big]$ for SSPA can be defined as follows

$$OF\big[\gamma(t)\big] = \frac{\gamma}{\left[1 + \left(\frac{\gamma}{B_0}\right)^{2P}\right]^{\frac{1}{2P}}} \tag{18}$$

$$\varphi\big[\gamma(t)\big] = 0 \tag{19}$$

where B_0 represents the maximum peak amplitude output signal and from linear to limiting of saturation, the parameter that controls the transition smoothness is

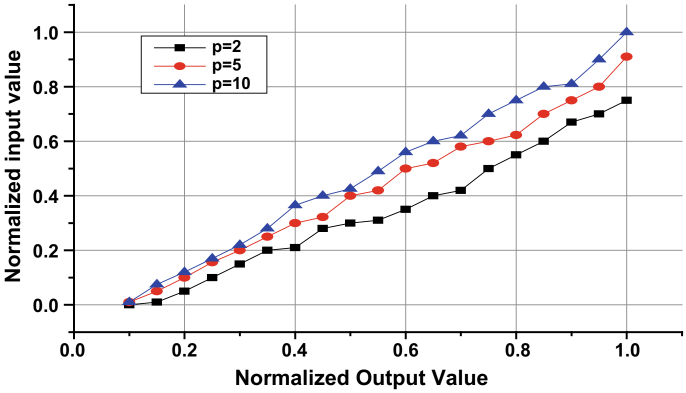


Fig. 5 Performance analyses of AM/AM transformation for a SSAP system

denoted by P. The input saturation to the OHPA model is defined by the value of A_2 and Fig. 5 depicts the AM/AM conversion characteristics for an operational solid-state power amplifier model. The 'backoff' factor is used to recognize the amplifiers operating point thereby adopting the input backoff parameter.

$$BOI = 10\log_{10}\left\{\frac{B^2}{E[|n|^2]}\right\} \tag{20}$$

3 Simulation Results

The simulation analysis shows the performance of PAPR in the algorithm that is proposed. The CCDF is used to evaluate the performance of proposed PAPR using EHPTS. The following table gives the details the included simulation parameters (Table 1).

Table 1 Performance analysis of proposed EHPTS by considering following simulation parameters

S. No	Parameters	Range
1	FFT length	3.6 μs
2	Signal bandwidth	36 MHz
3	No. of data subcarriers	68,128
4	Channel model	TGn channel model D
5	FFT size	64,128
6	Modulation type	32 QAM
8	Guard interval duration	0.6 μs
9	Coding rate	½

Figures 6 and 7 show the performance of MIMO-OFDM system with respect to 64 and 128 subcarriers.it is observed that when the value of CCDF = 10^{-4}, the proposed PAPR with EHPTS method that uses 8 sub-blocks and conventional PTS method that uses 4 sub-blocks are improved by 5.3 and 3.6 dB compared to existing MIMO-OFDM signal.

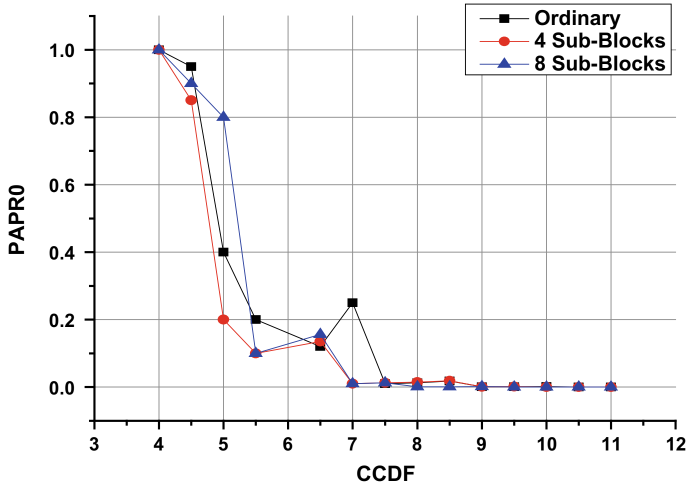


Fig. 6 PAPR performance analysis of MIMO-OFDM system using EHPTS in case 64

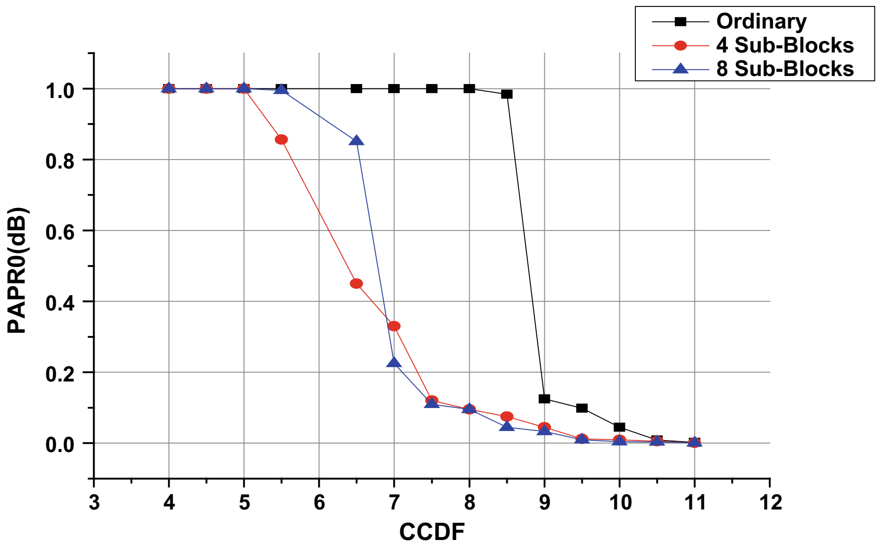


Fig. 7 PAPR performance analysis of MIMO-OFDM system using EHPTS in case 128

4 Conclusion

This study includes description and simulation of EHPTS technique for PAPR reduction in an OFDM signal. For the PTS approach, the rearrangement of the optimal combination of phase rotation factors involved the application of an effective hybrid partial transmit sequence model improvement in the performance of algorithm was obtained in the proposed EHPTS approach by embedding a local sub-carrier allocation algorithm in each generation in the MIMO-OFDM model each phase factor sequence is represented by each chromosome and the phase factor is represented by the number in genes. The computational complexity of PTS scheme is efficiently decreased by the proposed EHPTS method. For wireless communication system, that use MIMO-OFDM technology, the performance is observed to be better in reduction of PAPR factor with the proposed PAPR reduction method. The evaluation performance is carried out using simulation parameter specified by IEEE 802.11n. As we increase the number of sub blocks, the proposed method has better performance in PAPR reduction.

References

1. Shen J, Wang J (2019) Single LED-based 46-m underwater wireless optical communication enabled by a multi-pixel photon counter with digital output. Opt Commun 438:78–82 (Elsevier Publications)
2. Kakati D, Arya SC (2018) A full-duplex optical fiber/wireless coherent communication system with digital signal processing at the receiver. Optik-Int J Light Electron Opt 171:190–199
3. Bicakci K, Bagci IE (2012) Communication/computation trade-offs for prolonging network lifetime in wireless sensor networks: The case of digital signatures. Inf Sci 188:44–63 (Elsevier Publications)
4. Ro J-H, Seung-Jun Y (2020) An adaptive QR-based energy efficient signal detection scheme in MIMO-OFDM systems Comput Commun 149:225–231 (Elsevier Publications)
5. Taheri T, Mohamad M (2020) Joint spectral-spatial precoders in MIMO-OFDM transmitters. Signal Process 26:565–598 (Elsevier Publications)
6. Taşpınar N, Yildirim M (2015) A novel parallel artificial bee colony algorithm and its PAPR reduction performance using SLM scheme in OFDM and MIMO-OFDM systems. IEEE Commun Lett 9:1830–1833
7. Nejakar SM, Benakop PG, Sharanabasappa RR. Orthogonal frequency division multiplexing modulation scheme for 4G/5G cellular network. Eur J Adv Eng Technol 2(3):46–50
8. Heo S-J, Seoul National University Seoul (2007) A modified SLM scheme with low complexity for PAPR reduction of OFDM systems. IEEE Trans Broad Casting 53:804–808
9. Saadia R, Khan NM (2020) Single carrier-frequency division multiple access radar: waveform design and analysis. IEEE Access 8:35742–35751
10. Tong WU, Xin-chun QU (2007) Resource optimization in distributed cooperative multi-relay system based on STBWFDM. J. China Univ Posts Telecommun 14:1–6
11. Singh D, Joshi HD (2019) Error probability analysis of STBC-OFDM systems with CFO and imperfect CSI over generalized fading channels. Int J Electron Commun 52:156–163
12. Kildal P-S, Vosoogh A (2015) Fundamental directivity limitations of dense array antennas: a numerical study using Hannan's embedded element efficiency. IEEE Antennas Wirel Propag Lett 15:766–800

13. Bo TANG, Kaiyu QIN (2020) A hybrid approach to reduce the PAPR of OFDM signals using clipping and companding. IEEE Access 8:18984–18994
14. Sam DD, Jayadevappa D DTCWT-MBOFDM with reconfigurable down converter for smart grid communication. Int J Grid Distrib Comput 10(10):21–32
15. Jinga J, Ni W (2014) A low complexity channel estimation algorithm for massive MIMO system. Int J Grid Distrib Comput 7(4):81–92
16. Santhiya V (2017) Optimization of path selection in MIMO. Int J Recent Innov Trends Comput Commun 5(6):1020–1025
17. Krishna Kishore K (2020) Analysis of PAPR in PTS-OFDM and modified UFMC using firefly algorithms. Int J Adv Sci Technol 29(1):469–476
18. Vijayakanthan K, Hemachandran K (2020) High throughput and mixed radix N-point parallel pipelined FFT VLSI architectures for advanced wireless communication. Int J Grid Distrib Comput 13(1):400–411
19. Pethunachiyar GA, Sankaragomathi B (2020) OFDM based energy detection algorithm for spectrum sensing in cognitive radio networks. Int J Adv Sci Technol 29(03):7296–7306
20. Jayasankar1 T, Sujatha M (2020) Unconstrained global optimization base partial transmit sequence for OFDM PAPR reduction. Int J Adv Sci Technol 29(03):7346–7354
21. Valli C, Mathuram S (2020) VLSI implementation of OFDM transceiver using customized IFFT/FFT algorithm. Int J Adv Sci Technol 29(7s):794–801
22. Anil Kumar R, Prasad KS (2020) Comparative analysis of OFDM, FBMC, UFMC & GFDM for 5G wireless communications. Int J Adv Sc Technol 29(5):2097–2108
23. Somasekhar B, Sai Kishore CH (2020) Inter carrier interference mitigation in space frequency block code MIMO-OFDM system. Int J Adv Sci Technol 29(9s):3305–3314

Wireless Sensor Networks: A Methodical Analysis

M. S. Sinduja, K. R. Rekha, and Raghavendra Manjegowda

Abstract Wireless Sensor Networks are gaining significance over time as technology is advancing. With more usage of Wireless Sensor Network (WSN), few implementations require data communication be with minimum latency. Throughput is considered a significant parameter in some implementations where the delay is given less significance. Every application has requirement based on what parameters are important. Understanding the structure of network and the ability to pick the best routing protocol is significant and it should be suitable for the use case. In this work, we present an overview of existing routing methods in WSN. Firstly, we diagram the design considerations for protocols in WSN following this we provide an overview of distinctive techniques used in routing.

Keywords Wireless sensor network · Architecture · Routing protocols · Sensor nodes (motes)

1 Introduction

Because of advancement in technology, the low-cost sensors are gaining popularity. Processing the information surrounding the nodes in network uncovers a few properties about objects found or the event that occurred in the area of the sensor network [1]. Organizing such sensors in large amount will help in numerous applications that needs unattended activities.

Wireless Sensor System (WSN) contains thousands of motes. Motes have the capacity to transfer data between each other or legitimately to an externally placed Base Stations (BS). Accuracy achieved in bigger topographical regions is directly

M. S. Sinduja (✉)
Department of ECE, SJBIT Research Centre, Affiliated to VTU, Bangalore, India

K. R. Rekha
ECE Department, SJBIT, Bangalore, India

R. Manjegowda
Nvidia Corporation, Santa Clara, USA

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_130

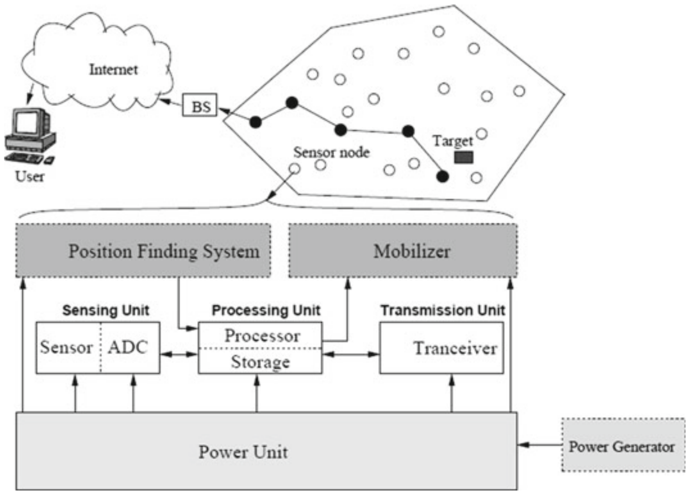


Fig. 1 Components of sensor

proportional to number of sensors taken into consideration [1, 2]. The sensor node (motes) components is as shown in Fig. 1.

The sensor field will have sensor nodes that are scattered. High-quality data about the environment will be provided by the motes as they coordinate among other sensor nodes [3]. These scattered nodes in the sensor field is capable in collecting and communicating the data to different nodes otherwise directly to a BS that is placed externally.

BS will connect sensor nodes to the internet or communicating infrastructure that already exists. Users can view the information that is reported. If the network area is lesser in which BS and motes will be closer so that they can exchange data with single-hop communication which is also termed as direct communication. In most of the WSN, the coverage network will be larger and therefore these criteria may require more nodes to be deployed in the network. This network takes multi-hop communication which is indirect communication due to the large distance between the BS and the father nodes [4].

In indirect communication, the nodes will not only transfer the data packets but also serve as a route between other motes towards the BS. The way of establishing the perfect route from src to dst is termed as routing. It is the fundamental requirement that has to be fulfilled by the network layers [3].

2 Design Problems of WSN

Despite numerous uses of WSNs, these systems have various requirements, for example, low energy consumption, less processing power, and limited data transfer capacity or bandwidth of the remote associates joining the motes.

One of the objectives of WSNs is to do information exchange while enhancing the network lifetime which can be accomplished by effectual energy management strategies [2, 5].

While structuring new routing protocols, these requirements needs to be satisfied.

2.1 Node Arrangement

In WSNs Node arrangement is dependent on the application and it can be physical or random deployment. Physical arrangement means sensors will be placed physically, where information is communicated via prearranged ways. Whereas, in the randomized deployment of nodes, the sensor hubs are randomly scattered, creating ad hoc or specified routing framework [2]. As a result, clustering is mandatory to increase the efficiency of energy and for networking. Inter sensor correspondence is usually inside minimum broadcast range because of energy and data transmission limitations (bandwidth limits) [6].

Thus, it is expected that a data transfer direction will comprise of numerous hops in the wireless network.

2.2 Transmission Models and Data Transmission

There are four methods of information transmission relying upon the applications in WSN to be specific as continuous type, event-driven, query-driven and lastly, hybrid type (a mixture of all three approaches). A node initiates the data transmission just when the sink makes a query or when the events happen in an event-driven or query-driven model. In continuous type, Information conveyed is intermittent. Routing protocol performance is combination of transmission medium and size of the network.

Thus, good medium of transmission improves the performance of the wireless network [5].

2.3 Association Heterogeneities

In numerous studies, Deployed nodes in the network was accepted as similar means it has identical power, computation, capacity [7, 8]. Nevertheless, upon request nodes may have other capacity.

2.4 Network Adaptability

In Sensing Zone, the quantity of nodes deployed can be between hundred to thousand. The designed protocol should be capable to adapt to such conditions. Furthermore, the designed protocol must adapt and respond to any event. Most of the nodes could be stale until any incident happens [1, 7].

2.5 Delay

A few applications require immediate response or low latency, for example, temperature sensor or caution checking and so on. In this way, designed protocol must produce lesser delay which means- expected time in sending the detected information should be smallest [5, 6].

2.6 Quality of Services (QoS)

Moment the data is detected, it has to reach the destination within no time otherwise it is impractical. As a result, data distribution with constrained expectancy is another constraint [7].

Nevertheless, various applications define energy conservation, which is related to the lifespan of a network, as sensibly more critically than the quality of information.

2.7 Complexities of Routing Protocol

The network performance is impacted by the routing protocol complexity. The explanation is that we have limited hardware capabilities and we furthermore face energy limitation in WSN.

2.8 Energy Efficiency

WSNs are generally battery operated. Energy deficiency is one of the significant problems in sensor network particularly in applications like a war zone, battlefields, and so on [2, 9].

When battery falls below the threshold level, performance of the network is impacted. One of the primary constraints to focus on while designing a network is Energy constraint. In a network, there can be hundreds or thousands of motes. Energy consumption is the limitation due to less available power in the nodes. Thus, designed protocol must have good energy efficiency.

3 WSNs Routing Challenges

Designing a protocol for routing is challenging due of many traits [5]. Some of the challenges are as follows:

- It is practically impossible to allocate "one-fit-all" scheme for a large amount of sensor hubs. Thus, WSNs are not capable of utilizing traditional protocols that are IPbased.
- Communication of sensed information from various source to a particular BS is necessary. But this does not hold good for conventional network.
- In most of the case there can be repeated information from the sensed data. Since most nodes will create the exact information during detecting. Hence, such redundancy should be avoided so that the protocol can efficiently make use of energy and bandwidth available.
- Also, WSN is confined to transmission of energy, bandwidth, storage, capacity. Because of such dissimilarities, many protocols is been formed to address such routing difficulties in WSN.

4 Routing Protocols

In this section, classification of routing protocols for WSNs is presented. The routing protocol specifies how motes will interact with one another along with communication of data in network. Numerous approaches are available to characterize the protocol in WSN as shown in Fig. 2.

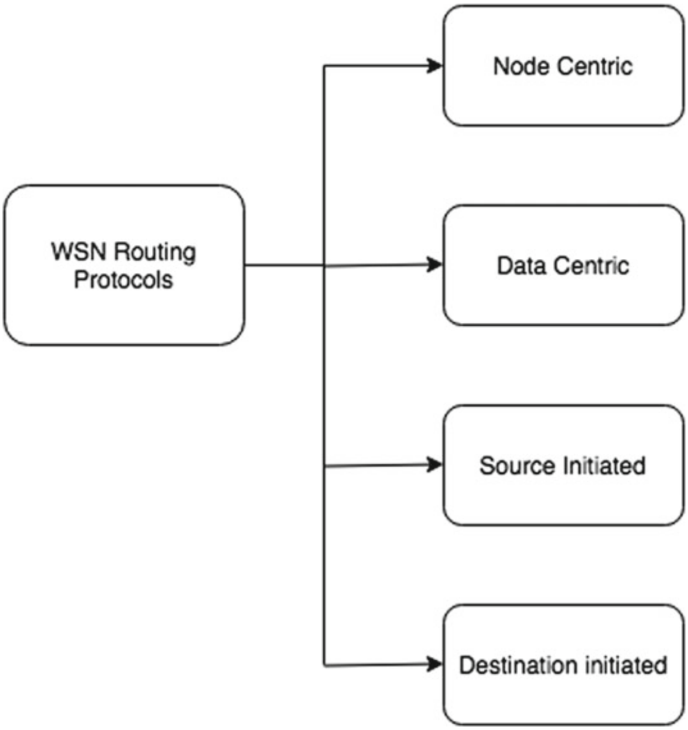


Fig. 2 Classification of routing protocol

4.1 Node Centric

In this routing, destination will be given an address which is a numerical identifier. Hence, motes can forward information to its destination by means of its ID. An example of LEACH is quoted below.

Low Energy Adaptive Clustering Hierarchy (LEACH) In this protocol, clusters are formed using sensor nodes in the network with equal energy among all the nodes [8, 10]. One of the sensor nodes will be characterized as cluster head (CH) and this CH will be the routing point for other nodes as shown in Fig. 3.

The CHs are picked before the whole communication starts and fixed all through the network lifetime like conventional clustering algorithms, it is not desired to have CH fail due to any problems within it as there are more chances that battery would die early [11].

Among the node's CH is chosen using randomization. Hence, choosing CH among nodes for temporary timeline helps the protocol to have increased lifetime. CH is chosen based on the probability characterized by the routing procedure and reports the same to remaining nodes.

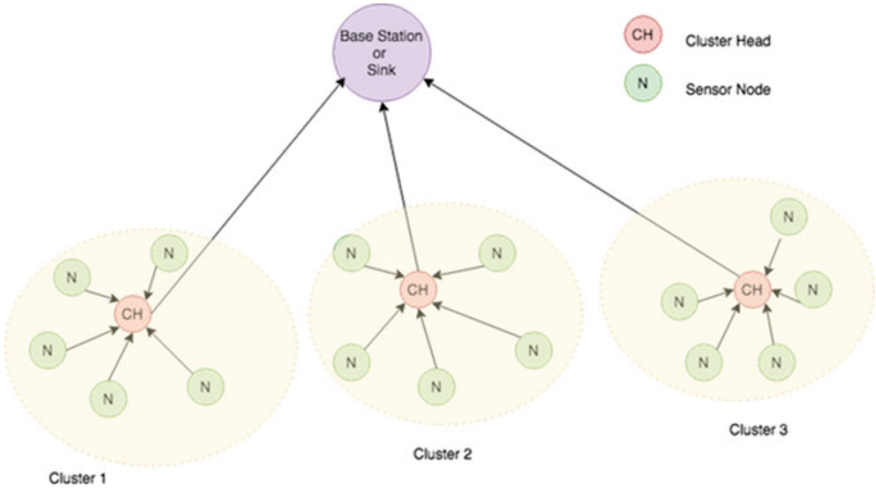


Fig. 3 The LEACH protocol in WSN

4.2 Data Centric

Detected information or data is unquestionably more important than the actual node itself. Consequently, prime focus for data-centric routing method is communicating the data indicated through specific parameters other than gathering information [12].

This protocol is utilized to control the redundancy of information. Sensors nodes will not have an identification number which is used to identify them uniquely; henceforth information is communicated to every node within a redundant manner.

In this protocol, the sink requests data by sending the query, so that the closest motes communicate the data after query. All the intermediate motes collaborate the data received and then transfers the data to the destination which will help in saving the loss of energy during repeated information transmission. An example of SPIN is explained below.

Sensor Protocols for Information via Negotiation (SPIN) Key element of SPIN is that it utilizes an advertisement mechanism. In this, the Meta information is traded among sensors. Nodes upon getting new data sends an ADV packet to advertise to all of the neighbors. The neighbor nodes that have an interest in data will recover the data by communicating request (REQ) messages for the node that sent the advertised packet. When REQ is received, the node that was advertising will send the data to that node. The model is illustrated in Fig. 4.

The benefits of the SPIN are that nodes need to advertise only its single-hop neighbors. It avoids the Resource Blindness and no repetitive data being passed subsequently improving energy efficiency.

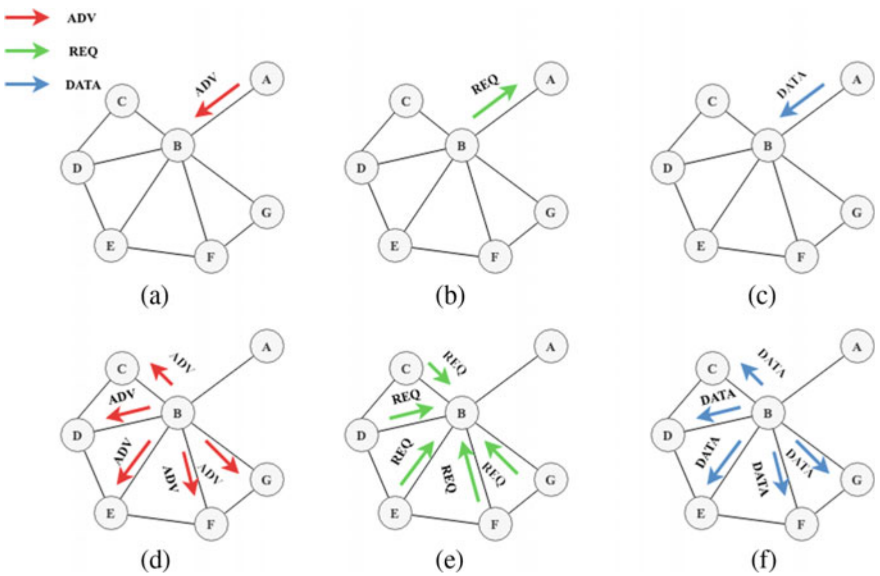


Fig. 4 SPIN protocol in WSN **a** A sends ADV message to B. **b** B reciprocates to A by sending REQ to A. **c** B receives data from A. **d** Node B then sends an advertising (ADV) message to its immediate neighbors. **e** B gets REQ messages from its neighbor nodes. **f** B then communicates the DATA message to its neighbor nodes

The disadvantage is that it doesn't ensure data delivery, for example, if the src node and dst node is too far and if transitional nodes do not participate in data conveyance as they are not interested which leads to data not reaching the destination.

4.3 Dst-Initiated (Destination-Initiated)

When path formation begins from dst nodes then such protocol is known as Dst-initiated. Examples: LEACH, Directed Diffusion (DD) [13].

Directed Diffusion (DD) Data Centric methodology is used for collecting and circulating information. Energy saving is improved through this protocol and has a direct relation with the network lifetime. Addressing is not required in this protocol as all the communication between a node to node uses directed diffusion [6].

At the point when a node known as the sink node needs data about a specific attribute, it sends interest messages to its entire neighbors. When interest messages are flooding through the wireless network, they are added to every nodes' interest cache.

Every interest record in the cache might have at least one inclination that relate to the neighboring nodes that transfer the interest.

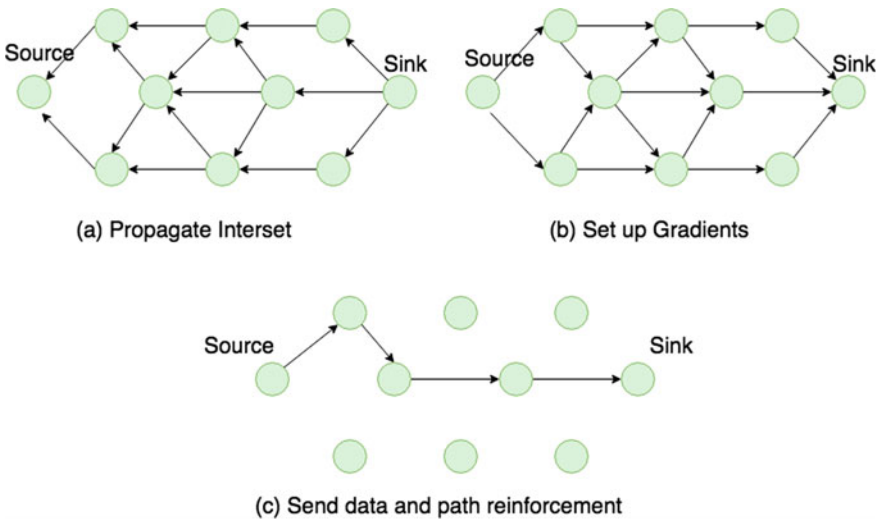


Fig. 5 Phases of DD protocol

The attributes stored are–desired data rate, interest duration, and timestamp. At the point when a node forms information that matches the cached interest, it sends the information back to the source along with the attributes. Instinctively, through gradients the information is communicated to the destination as shown in Fig. 5. Sink or base station may choose the shortest way i.e., the one with the quickest response, by sending interest with a higher information rate along that way.

4.4 Src-Initiated (Source-Initiated)

SPIN is best examples for an src-initiated routing protocol, As explained above how SPIN work the source node have data to share it uses advertisement technique and sends out an ADV message to all its neighbors. Routes are discovered from src to dst node.

5 Routing Protocols Categories

Two methods are utilized to send information in the sensor network- flooding and gossiping protocol. These protocol does not make any use of topology and algorithm. The data is communicated to all the neighbors until the data reaches the targeted destination, this is termed as flooding [14].

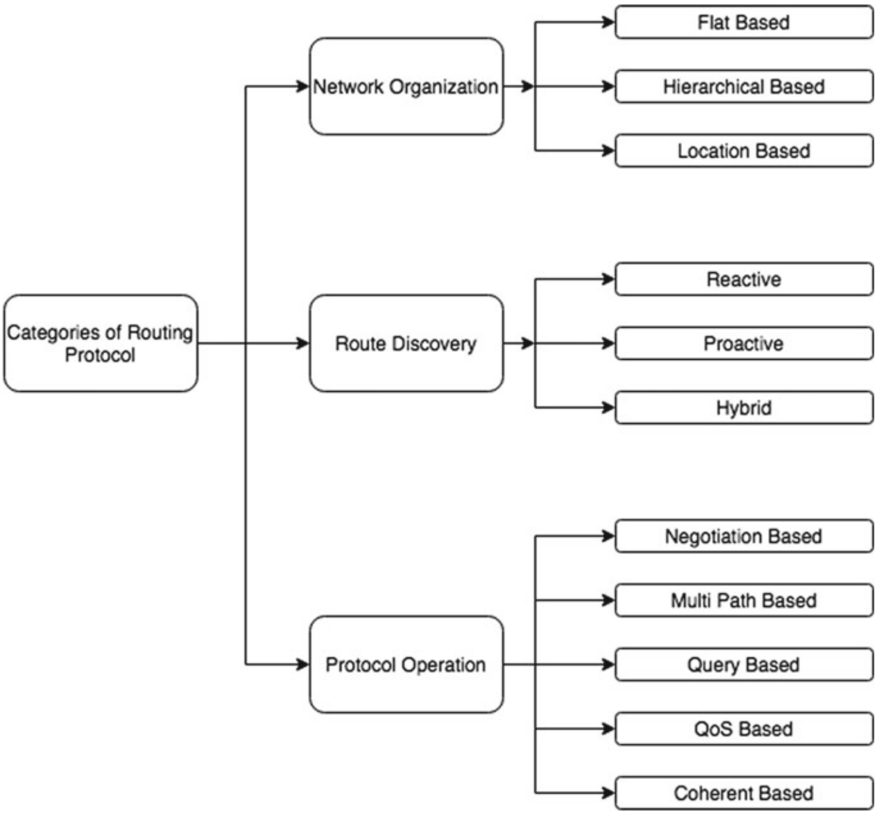


Fig. 6 Categories of routing protocols

Its advantage is that its implementation is easy. Some limitations include Implosion-Problem, Overlap-Problem, and Resource-Blindness.

In the Gossiping protocol, data is transferred to its random neighbor, the received packet is transferred to its randomly selected neighbor and this process continues. This mechanism may induce the propagation delay of data between the nodes. Nevertheless, this avoids implosion-problem. The broad Categories of routing protocols is as shown in Fig. 6.

5.1 Network Organization

These protocols depend on the organization of the network [5].

Flat Topology This topology is principally for networks that are homogeneous, and it treats each node similarly, all the nodes will have the same attributes and same functionality.

Hierarchical Based Routing Generally, networks that are heterogeneous make use of this protocol when nodes are advanced and powerful than different nodes, yet not generally this is the situation, Occasionally, Cluster is formed by grouping the nodes in the hierarchical protocol.

From each cluster, CH is chosen and it communicates to the base station after aggregating all the data from all nodes. The clustering mechanism is more efficient in terms of energy.

Location-Based/Geo-centric Routing This makes use of localization protocols to detect the current location of nodes. The location detected will improve the routing process and also helps the sensors to provide extra service.

5.2 Route Discovery

This category of the protocol is chosen based on the route discovery process.

Reactive Protocols As the name says, Reactive routing protocol will not maintain the network topology. The networks are active only when any nodes want to communicate with other nodes. Henceforth, the routes are created with the initiation of a query [6].

Proactive Protocols Another terminology for proactive protocol is table-driven routing. This protocol forms the routes prior to using them (even when the traffic is zero) and the routing table is preserved the entire network through sending the information of the network to every node present in that network.

Hybrid Routing Protocols This Protocol has the benefits of both reactive and proactive routing protocols.

5.3 Routing Protocol Based on Operation

Multi-path Routing Protocol A multi-path protocol is chosen over a single path to enhance network performance. The non-critical failure of a protocol is estimated based on the probability that another route exists from src to dst during the main route fails [12, 15]. By keeping up numerous routes from the src and the dst at the cost of increased energy utilization and generation of traffic, the lifetime can be improved. These substitute paths are kept alive by-passing intermittent signals. Henceforth, the quality of the network can be improved by keeping up the substitute routes to the destination [16].

Negotiation Based Routing Protocols This protocol utilizes significant level information descriptors so as to avoid or eliminate repetitive information transmissions. The negotiation process happens between the node and communicates the data with its neighbors within the available resources [15, 16].

Query-Based Routing Protocol The protocols fall under the classification Dst-initiated. When there is a query initiated by the destination node only then the nodes will respond [5, 16].

QoS Based Routing Protocol The protocols fall under this category provide good Quality of Service. The path will be established between source to sink and will have optimum metrics providing good QoS. Metrics may include energy, delay, data delivery, and throughput along with using the network resource effectively.

Coherent Based Routing Protocol Coherent routing protocol enables the information communicated to aggregators after the least processing like copy suppression, time-stamping, and so forth. The nodes that handle processing are known as aggregators [5, 15].

The coherent protocol is chosen when efficient energy routing has to be achieved.

6 Conclusion

Technological advancement of sensors has made WSN popular. This work presents a comprehensive study of how the protocols classification is done and categorized based upon certain features. Constrained attributes of the network should be considered for the development networks. Reduction in energy resources, adaptability, and versatility emerges as the fundamental limitations concerning the routing protocols in WSN. Routing protocols have a huge impact to create lesser interference along with effectual communication among src and dst nodes. Reliability and performance of a system generally rely upon the choice of good routing protocol. organization of networks, discovery of routes, and protocol-based operations are the classes of routing protocol is WSN.

References

1. Al-Karaki JN, Kamal AE (2004) Routing techniques in wireless sensor networks: a survey. IEEE Wirel Commun 11(6):6–28. https://doi.org/10.1109/MWC.2004.1368893
2. Mishra S (2012) Wireless sensor networks design issues
3. Matin MA Islam MM (2012) Overview of wireless sensor network. In: Wireless sensor networks—Technology and protocols. InTech
4. Busi Reddy V, Venkataraman S, Negi A (2017) Communication and data trust for wireless sensor networks using D–S theory. In: IEEE Sens J 17(12):3921–3929. https://doi.org/10.1109/JSEN.2017.2699561

5. Shabbir N, Hassan SR (2017) Routing protocols for wireless sensor networks (WSNs). In: Wireless sensor networks—Insights and innovations. InTech
6. Ehsan S, Hamdaoui B (2012) A survey on energy-efficient routing techniques with QoS assurances for wireless multimedia sensor networks. In: IEEE communications surveys and tutorials, vol 14, no 2, pp 265–278. https://doi.org/10.1109/SURV.2011.020211.00058
7. Randhawa S (2014) WSN routing challenges: a methodical analysis. Int J Adv Comput Sci Appl 3:780–785
8. Villalba LJ, Orozco AL, Cabrera AT, Abbas CJ (2009) Routing protocols in wireless sensor networks. Sensors (Basel, Switzerland) 9(11):8399–8421. https://doi.org/10.3390/s91108399
9. Yin G, Yang G, Yang W, Zhang B, Jin W (2008) An energy-efficient routing algorithm for wireless. In: International conference on internet computing in science and engineering (ICICSE'08). IEEE, China
10. Usman MJ, Xing Z, Chiroma H, Gital AYU, Abubakar AI, Usman AM, Herawan T (2014) Modified low energy adaptive clustering hierarchy protocol for efficient energy consumption in wireless sensor networks. Int Rev Comput Softw (IRECOS) 9(11):1904–1915
11. Tao L, Qing-Xin Z, Luqiao Z (2010) An improvement for LEACH algorithm in wireless sensor network. In: Proceedings of the 5th international conference on industrial electronics and application. IEEE Explore, pp 1811–1814
12. Khan AA (2016) A survey of routing protocol in wireless sensor networks [thesis]. GCU, Lahore. Available from http://library.gcu.edu.pk/theses.htm
13. Bakr BA, Lilien L (2011) LEACH-SM: a protocol for extending wireless sensor network lifetime by management of spare nodes. In: International symposium on modelling and optimization in mobile, ad hoc and wireless networks (WiOpt). IEEE, New Jersey, Princeton
14. Hassan SR (2014). Performance analysis of ZigBee based wireless sensor networks [thesis]. GCU, Lahore Available from http://library.gcu.edu.pk/theses.htm
15. Sabor N, Sasaki S, Abo-Zahhad M, Ahmed SM (2017) A comprehensive survey on hierarchical-based routing protocols for mobile wireless sensor networks: review, taxonomy, and future directions. Wirel Commun Mobile Comput 2017(Article ID 2818542):23. https://doi.org/10.1155/2017/2818542
16. Shabbir N, Nawaz R, Iqbal MN, Zafar J (2015) Routing protocols for a small scale WLAN based wireless sensor networks. In: 9th International conference on sensing technologies. IEEE, New Zealand

Development of Robust and Real Time Web Based IVF Success Rate Prediction Using Machine Learning

G. S. Gowramma, Shantharam Nayak, K. Rakshitha, R. Varsha, and T. Jayashree

Abstract Research into the impact of in Vitro Fertilization (IVF) enables customer systems is still in its infancy. Our method represents a tool to support decision making for patient analysis with fertility problems whether the patient will conceive or not. An in-depth formative evaluation of the essential attributes which are necessary for the IVF treatment are achieved by data collection (data set), data pre-processing, a model formulation which includes applying of the three supervised ML techniques namely, Multinomial Naive Bayes, Random Forest, Decision Tree which predicts the accuracy and then followed by uploading the patient records into the cloud. There were 28 attributes used to build the project. The results obtained by using these three algorithms are 62.41% for multinomial Naive Bayes, 69.14% for the random forest, and 41.31% for decision tree. IVF patients seek support from the real-time web based decision support system by using machine learning and AWS cloud services to know the accuracy, retrieving, and uploading the patient records to the on demand cloud service platform. More research is needed to inform the real-time web based disease decision support system design that effectively tailors support for patient's wellbeing.

Keywords IVF · Data collection · Data pre-processing · Machine learning · Multinomial Naive Bayes · Random forest · Decision tree · Cloud computing

1 Introduction

In Vitro Fertilization is an infertility procedure. According to the World health organization, infertility is characterized as one year of regular, unprotected intercourse that did not occur during pregnancy. There are numerous assisted reproductive techniques for conceiving a child, one of them being IVF. Infertility is suspected when a woman cannot get pregnant without the use of birth control within one year after having

G. S. Gowramma · S. Nayak · K. Rakshitha · R. Varsha · T. Jayashree (✉)
Don Bosco Institute of Technology, Bengaluru, India

G. S. Gowramma
e-mail: gowribasu.dbit@gmail.com

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_131

sexual intercourse. Several of the causes of infertility in women are impaired or blocked Fallopian tubes, endometriosis or hormonal imbalance. Low sperm content or low sperm count may cause infertility in men.

IVF is one of the possible methods for making an infertile pair more likely to become pregnant. Its use depends on whether it's infertile. If there is a blockage in the woman's endometriosis or fallopian tube, or sperm quality or low sperm count in the males, IVF may be a choice. The other potential therapies for such disorders are surgery for endometriosis or blocked channels, which can be done before IVF.

This project mainly focuses on the IVF patients, gathering the information of the patient records from the cloud and predicts whether the patient can conceive or not conceive by using ML techniques such as multinomial Naive Bayes, Random Forest, and Decision Tree. A web based system is developed and deployed in the ART centers and we can directly upload or fill the data on a cycle basis. Each ART center is also able to review the status of its data and can access global and individual statistical reports of the past years. This information is stored in the cloud allowing statisticians to extract data to produce yearly reports. Finally, the system predicts the appropriate results by analyzing the patient reports.

2 Literature Survey

See Table 1.

3 Methodology

The purpose here is to find the success factors of IVF treatment. To have high-quality results before machine learning algorithms can be used, data cleaning is required. When the data is incomplete, it indicates that some values are missing, or they are corrupt and inaccurate data is present in the database. This may cause some problems for research. Therefore, the initially collected data must go through the data selection and removal stages. Once the process is completed, the dataset is uploaded into the AWS cloud and hence the data can be retrieved from the cloud for future predictions.

3.1 Data Collection

Data collection is necessary for identifying and collecting the data needed by the Human Fertilization and Embryology Authority (HFEA) for the identification of infertility in women. This dataset has 28 attributes. Among 28 attributes, 18 attributes have been used for prediction based on the Recursive Feature Selection algorithm.

Table 1 Literature survey

Citation, Year	Methodology	Attributes used	Accuracy (%)
Kaufmann et al. [1], 1997	Artificial neural network	14 variables	59
Jurisica et al. [2], 1998	TA3IVF(classification)		71.2
Saith et al. [3], 1998	Artificial intelligence	53 features (embryo grade, cell number, follicle size and follicular fluid volume)	61
Trimarchi et al. [4], 2003	C5.0 decision tree data mining algorithm	100 features	75
Davis et al. [5], 2005	CMIFN and MCAR algorithms		88
Guh et al. [6], 2011	Hybrid intelligence method combining genetic algorithm and decision learning techniques	70 features	73.2
Durairaj and Meena [7], 2011	RST and ANN	Spermatological data set	
Girija and Shashidhara [8], 2012	Decision tree algorithm	9 features	56
Gil et al. [9], 2012	MLP, SVM, and decision tree	34 features	86
Durairaj and Kumar [10], 2013	Attribute selection and multilayer perceptron network, weka	27 features reduced to 4 features	73
Durairaj and Thamilselvan [11], 2013	ANN predictive model, Neuro solution	8 features (endometriosis, follicles in ovaries, sperm concentration, tubal factors, infertility duration, BMI, stress levels and embryos transferred)	73
Girela et al. [12], 2013	Multilayer perceptron network	Questionnaire from 123 young healthy volunteers	90
Uyar et al. [13], 2014	Receiver operating characteristic (ROC) analysis, Naïve Bayes model	IVF clinics have EHR. After ICSI the data set had 2453 embryos. The embryo had 18 clinical characteristics and a class label, + 1 or − 1	80.4

(continued)

Table 1 (continued)

Citation, Year	Methodology	Attributes used	Accuracy (%)
Gowramma et al. [14], 2015	A review of applications in health care sector using data mining and data mining techniques in IVF prediction	To collect large real time IVF/ICSI datasets to develop an efficient algorithm to predict the IVF outcome successfully	
Gowramma et al. [15], 2017	Multi-layered perceptron with attribute selection algorithm	24 features	87.7
Gowramma et al. [16], 2019	In this paper a review of various intrinsic and extrinsic factors which contribute toward the success rate of IVF treatment are used		
Gowramma et al. [17], 2020	The review of various data mining applications such as statistical, traditional, computational, and hybrid approaches used in the prediction of IVF success rates	Applying ANN and GA can serve the purpose of enhancing the accuracy of IVF success rate prediction	

The attributes used and the 18 attributes that are selected using RFE model are Highlighted in the above Table2.

3.2 Data Preprocessing and Feature Selection

Some of the missing instances are handled by preprocessing using mean median computations. For the feature selection process the RFE algorithm has been used. Recursive feature elimination (RFE) is used to pick features by taking feature sets smaller and smaller recursively consideration. RFE is popular because it is easy to configure and easy to use and also because it is very effective at selecting those columns (features) in the training dataset that are most relevant in predicting the target variable.

Table 2 Attributes of IVF dataset

Patient_Age	Ovulatory_Disorder	Partner_Sperm_Immunological_Factor
Date_started_patient_trying_for_pregnancy	**Patient_Unexplained**	**Stimulation_used**
Total_no_of_the_past_IVF_cycles	**Male_aspect**	**Specific_treatment_type**
Total_no_of_IVF_pregnancies	Cervical_Factor	**Egg_Source**
Total_no_of_live_births- a_through_IVF_definition	**Endometriosis**	Sperm_Form
Type_of_Infertility-Female_First	Partner_Sperm_Condensing	**Fresh_Eggs_Collected**
Type_of_Female_Secondary_infertility	Female_Factor	**Embryos_Transferred**
Type_of_Infertility-Male_First	Partner_Sperm_Morphology	Stage_Day_of_Embryo_Transfer
Type_of_Male_Secondary_ Infertility	Partner_Sperm_Motility	**Occurrence_of_Live_Birth**
Tubal_disease		

18 attributes are made bold out of 28 attributes that are selected using the RFE model

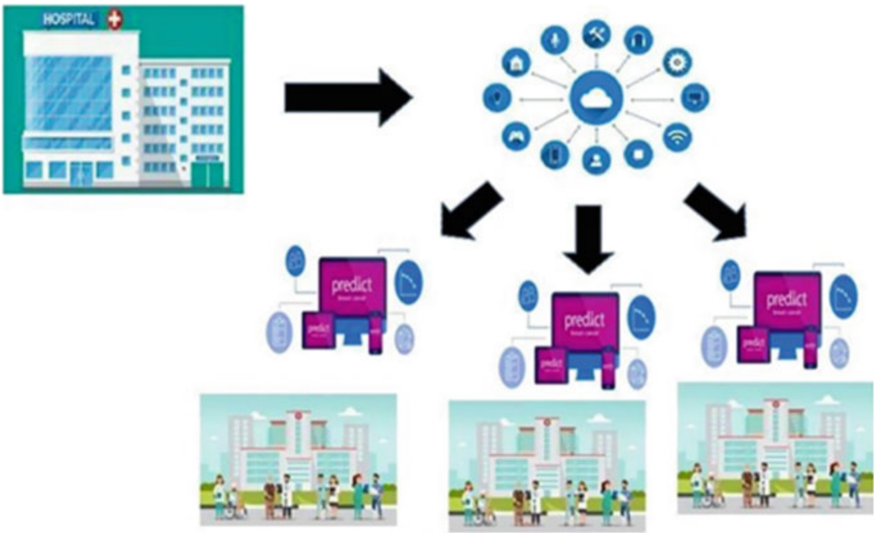


Fig. 1 Represents the architecture of the proposed system

3.3 Usage of Cloud

In this project, we have used AWS cloud as it offers a reliable, scalable, and pays as you go basis. A web-based system (Fig. 1) is developed and deployed in which hospitals having different branches can directly upload the records and retrieve the records on a cycle basis. These records can be easily accessible to the doctors for further prediction. Therefore, the patients can avoid unnecessary testing and can get a clear image of whether the patient can successfully conceive or not through IVF treatment.

3.4 Proposed Algorithms

3.4.1 Multinomial Naive Bayes

The multinomial classifier Naive Bayes is used for discrete feature classification. Normally the multinomial distribution requires integer feature counts. However, in practice, fractional counts may also work.

The samples in this model represent the frequencies with which a multinomial (p_1, …, p_n) has produced such events where p_i is the probability that event i occur. A function vector $x = (x_1 \ldots, x_n)$ is then a histogram, with x_i counting the number of instances that i was observed in a particular case. It is the event model for the classification of documents, with events describing a term occurs in a single text.

The probability of observing a histogram x is given by

$$p(x|c_k) = \frac{(\sum_i x_i)!}{\prod_i x_i!} \prod_i pk_i^{x_i} \tag{1}$$

3.4.2 Decision Trees

The Decision Tree algorithm is one of the algorithms for supervised learning. To solve regression and classification problems, using the decision tree algorithm. Decision Tree aims to build a training model that can be used bylearning basic decision rules derived from training data to predict the value of the target variables. Prediction in Decision Trees starts at the root of the tree. We compare the root attribute values with the attribute of that record. We obey the branch corresponding to that value based on the relationship and move to the next node.

We can understand the working of the ID3 algorithm from the following steps:

1. From the original set S it begins as a root node.
2. It tierson every iteration of the algorithm over every unused attribute of set S and calculates the information gain (IG) and Entropy (H) of that attribute.
3. Then it selects the attribute with the largest gain of Information or smallest Entropy.
4. The selected attribute thensplitsthe set S to generate a subset of data.
5. The algorithm continues to iterate on each subset, taking into account only attributes never previously picked.

3.4.3 Random Forest

Random forest is one of the algorithms for supervised learning. The "forest" that the random forest creates is a collection of decision trees normally equipped with the process of "bagging". The concept for the method of bagging is that a mixture of learning models wouldmaximize the overall result we get. The concept of the Random forest is tocreate multiple decision trees and combine them to get a prediction that is more reliable and precise. The advantage of random forests is that they can be used for both regression and classification issues.

The workings of the algorithm Random Forest can be interpreted using the following steps

1. First, select some of the random samples from a given dataset to start.
2. The algorithm will then create a decision tree for each sample after the collection.
3. Thealgorithm thengets the results of the prediction from each decision tree.
4. Voting for every predicted result will be done in this 3rd step.
5. At last, the outcome of the most voted prediction will be chosen as the outcome of the prediction.

4 Result and Discussions

After the analysis the results obtained before and after feature selection are observed to be as follows (Table 3).

The results of the application of various classifications in Table 2 shows algorithms on the dataset, which consists of accuracy before and after feature selection. Accuracy of various algorithms is as indicated in Fig. 2. Random Forest algorithm outperforms than Decision Tree, Multinomial NB algorithm.

The ROC curve analysis against true positive rate and false positive rate for before and after feature selection is as shown below Fig. 3. According to the ROC curve of Fig. 3 the Area Under Curve of Random Forest algorithm before preprocessing is 0.685 and after preprocessing is 0.690. Hence we can conclude that the Random forest algorithm outperforms than Decision Tree and Multinomial NB (Table 4).

Table 3 Accuracy of algorithms before and after feature selection

Algorithm used	Before feature selection	After feature selection
Random forest	69.12	69.14
Decision tree	41.24	41.31
Multinomial NB	58.73	62.41

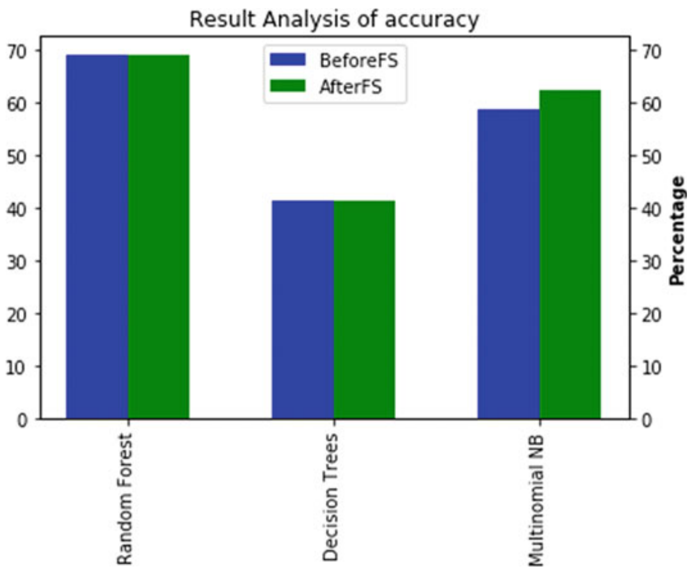


Fig. 2 Result analysis of various algorithm

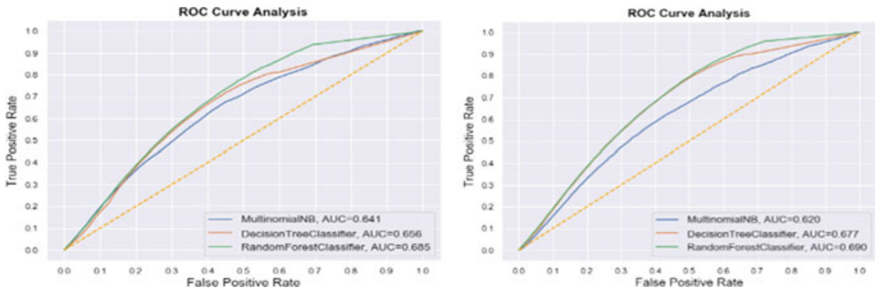


Fig. 3 ROC curve analysis before and after feature selection

Table 4 AUC of algorithms before and after feature selection

Algorithm used	AUC before feature selection	AUC after feature selection
Random forest	0.685	0.690
Decision tree	0.656	0.677
Multinomial NB	0.641	0.620

5 Conclusion and Future Enhancement

The assisted reproductive technology used for gestational surrogacy and infertility treatment is called InVitro fertilization. The web based system is a blessing for the couples who long for a child. Among the three algorithms used, the random forest algorithm gave the best accuracy of 69.19% and also according to the ROC curve the AUC of Random Forest algorithm before preprocessing is 0.685 and after preprocessing is 0.690. Hence we can conclude that the Random forest algorithm outperforms than Decision Tree and Multinomial NB both in terms of Accuracy and AUC. The patients can get a clear picture of whether they can conceive or not through IVF treatment which can be achieved by using machine learning algorithms.

In the future, other algorithms can be used for greater accuracy with more instances and attributes which are found as an impact of association between IVF and Assisted Reproductive Technologies. Nowadays, many algorithms can detect the possibility of the disease in early stages i.e., before even getting it. This would be very efficient to take future steps before it is too late. Hence these kinds of systems can be developed.

References

1. Kaufmann SJ, Eastaugh JL, Snowden S, Smye SW, Sharma V (1997) The application of neural networks inpredicting the outcome of in-vitro fertilization. Hum Reprod 12(7):1454–1457
2. Jurisica I, Mylopoulos J, Glasgow J, Shapiro H, Casper RF (1998) Case-based reasoning in IVF: Prediction andknowledge mining. Artif Intell Med 12(1):1–24

3. Saith R, Srinivasan A, Michie D, Sargent I (1998) Relationships between the developmental potential of human in-vitro fertilization embryos and features describing the embryo, oocyte and follicle. Hum Reprod Update 4:121–134
4. Trimarchi JR, Goodside J, Passmore L, Silberstein T, Hamel L, Gonzalez L. Comparing data mining and logisticregression for predicting IVF outcome. Fertil Steril 80:100–100
5. Davis J, Illingworth P, Salam A (2005) Applications of data mining techniques in assisted reproductive technology. In: 16th Australasian conference on information systems, Sydney
6. Guh RS, Wu TCJ, Weng SP (2011) Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. Expert Syst 4437–4449
7. Durairaj M, Meena K (2008) Application of artificial neural network for predicting fertilization potential of frozen spermatozoa of cattle and buffalo. Int J Comput Sci Syst Anal 1–10
8. Girija DK, Shashidhara MS (2012) Classification of women health disease (fibroid) using Decision Tree algorithm. Int J Comput
9. Gil D, Girela JL, De Juan J, Jose Gomez-Torres M, Johnsson M (2012) Predicting seminal quality with artificial intelligence methods. In: Expert systems with applications. Elsevier
10. Durairaj M, Kumar RN (2013) Data mining application on IVF data for the selection of Influential parameters on fertility. Int J Eng Adv Technol (IJEAT) 2(6):262–266. Appl Eng Sci 2(3):205–209
11. Durairaj M, Thamilselvan P (2013) Applications of artificial neural network for IVF data analysis and prediction. J Eng Comput Appl Sci 2(9):11–15
12. Girela JL, Gil D, Johnsson M, Gomez-Torres MJ, De Juan J (2013) Semen Parameters can be predicted from environmental factors and lifestyle using artificial intelligence methods. Biol Reprod 88(4):1–8
13. Uyar A, Bener A, Ciray HN (2014) Predictive modeling of implantation outcome in an in vitro fertilization setting: an application of machine learning methods. Med Decis Mak. Retrieved from Research Gate athttp://www.researchgate.net/publication/262536801.
14. Gowramma GS, Mahesh TR, Patil SB (2015) A study of in vitro fertilization/intracytoplasmic sperm injection prediction in data mining: a review. In: Int J Appl Eng Res. 10(86). ISSN 0973-4562. Available www.ripublication.com/ijaer.htm
15. Gowramma GS, Mahesh TR, Gowda G (2017) An automatic system for IVF data classification by utilizing multilayer perceptron algorithm. In: ICCTEST-2017, vol 2, pp 667–672. ISBN 978-81-931119-5-6. https://doi.org/10.21647/icctest/2017/49043
16. Gowramma GS, Nayak S, Cholli N (2019) Intrinsic and extrinsic factors predicting the cumulative outcome of IVF/ICSI treatment. IJITEE 9(2S):269–273. ISSN 2278-3075. https://doi.org/10.35940/ijitee.b1007.1292s19
17. Gowramma GS, Nayak S, Cholli N (2020) Data mining models applied in prediction of IVF success rates: an overview. Lecture Notes in networks and systems. ISSN 2367-3370, ISSN 2367-3389 (electronic). https://doi.org/10.1007/978-981-15-5309-7

Eye Blink to Voice Communication Aid for Paralyzed Patients

G. S. Gowramma, S. S. Arumugam, K. P. Pranav, V. Soundharya Lakshmi, and V. Mithun Gowda

Abstract Paralyzed is also known as Locked-in syndrome happens when an individual is a quadriplegic and the patient cannot speak or do a facial movement (www.cdn.shopify.com). The affected patient cannot communicate but he will be aware of his surrounding; the patient can hear and see has a normal person. The patient has thinking capacity and reasoning ability he can only use eye movement to communicate with others. Hence, the key purpose of this paper is to propose a real-time interactive system for a paralyzed patient using eye blink an eye motion detection algorithm by assisting them to communicate to caretakers to their needs through blink which is converted to voice. For example, clinical help, S.O.S, essential utility. This is done using a video outline that is handled by Open Computer Vision libraries for Image Processing which is an open-source programming method which is aimed at real-time computer vision.

Keywords Eye blink detection · Eye motion detection · Human-PC communication · Paralyzed · Open computer vision

1 Introduction

Loss of movement results from the spinal line injury or brain or by sicknesses, for instance, unique sclerosis or amyotrophic sidelong sclerosis. In the extreme loss of motion patients, correspondence limits are restricted. For sure, in some cases, patients can control their eye movement and blink. The respectable point behind this errand is to study and grab regular attention improvement and apply a glint revelation to make a fluid interface for the patient. With the help of HCI, an elective channel can be built up without talking and hand improvements, which will construct the personal satisfaction for a stifled patient [1].

The various advances utilized for executing the correspondence between paralyzed patients and the individuals joining in and thinking about them are mouth

G. S. Gowramma (✉) · S. S. Arumugam · K. P. Pranav · V. Soundharya Lakshmi · V. Mithun Gowda
Don Bosco Institute of Technology, Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_132

impelled joysticks, inhale puffing straws, tongue development examination, switch mounted close to the client's head, and so forth. These frameworks are exorbitant to execute, increment weight on the patients, and need gifted work to arrangement and keep up the framework for appropriate working.

This paper has a framework that utilizes a webcam on a PC which receipts contribution to the type of film outlines and breaks down eye boundaries outline by outline. The proposed framework at that point deciphers them as control orders. There is no need for clinical hardware to communicate with the patient. With the assistance of the Open Computer Vision library video outlines that are taken by webcam can be investigated and prepared to utilize different picture handling methods and produce wanted yield.

The facial milestone identifier actualized inside dlib produces 68 focuses on the face. These 68 point-mappings were acquired via preparing a shape indicator on the marked iBUG 300-W dataset. Utilizing these focuses, the area of the eye can be identified and student movement and flicker location can be executed. An incapacitated patient can explore through the proposed framework and turn on or off the framework as indicated by his will with no outer impedance. Distinctive flicker signals can be modified like intentional twofold squints or triple squints as indicated by the patient accommodation. The recognized eye flickers are of two sorts short squint (under 200 ms) and long squint (more noteworthy than 200 ms) which can assist us with differentiating among wilful and automatic eye squints. Utilizing these short squints and long flickers a blink gesture interface can be utilized in the framework for better exactness of the framework.

2 Proposed Methodologies

In this paper the system includes two algorithms namely the Eye Motion Detection and Eye Blink Detection. The balanced figuring from the outset perceives the face utilizing facial accomplishment finder acknowledged inside dlib and perceive the eye zone and draw a bounding box around the restricted eye co-ordinates. By at that point, the eye an area is scaled and changed over into grayscale. While playing out these checks, the idea that is made that the patient's head is steady. The PC screens console the headway of the eye towards right or left or perceive squints. The shape lines are then perceived by strategies for a mass around the eye. The output from these process of figuring can be used as far as possible for the basic PC interface.

2.1 Eye Blink Algorithm

The eye-flicker location calculation comprises of four key steps: Face Detection, Eye-area extraction, Eye-squint identification, and Eye-squint characterization. In the first place, the computation recognizes the face using Open Source Computer Vision

Fig. 1 Facial landmarks detected using dlib

Library and dlib libraries, remove the eye region in the face. Using this hopping area, it perceives the number of gleams close by their time length and organizes the squint reliant on the predefined classifier. The revelation of the eye region should be conceivable using a facial achievement locater completed inside dlib. The eye squint is recognized using eye limitation after which thresholding is used to find the whites of the eye and choose whether that region of eyes disappears a white (Fig. 1).

2.1.1 Face Detection

Face detection determines the face location and size of a human face in the given digital image [2]. Before recognizing eye flashes, the huge development will perceive the characters of the patient. Face limitation is begun at the basic stage to perceive the outward appearances of the patient. This methodology should work effectively in fluctuating light conditions, head positions, facial features, haircuts or glasses, etc. The face can be perceived with the assistance of a facial accomplishment finder executed inside dlib constantly. There are different blueprints advanced for face affirmation which wire part-based techniques utilizing eyes, mouth, and nose-based exposure. The following framework is the setup sorting out technique subject to the pre-recorded plan of the face. The third technique perceives the face utilizing neural structures, that prepares the framework utilizing haar falls and facial accomplishment finder executed inside dlib to recognize facial highlights like eye, nose, ear, and so on. The dlib library is utilized to plot 68 bases on the face to see a specific facial section. Utilizing these records, the basic zone of the face can be taken out and a jumping box can be enrolled around the eyes.

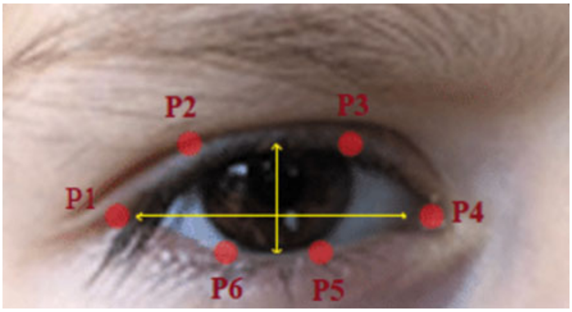


Fig. 2 Eye localization

2.1.2 Eye Localization

In this estimation, the eye region is orchestrated from live film plots. After the apparent look, the piece of the eye is orchestrated with the assistance of facial accomplishment highlights. From the eliminated eye an area, it is additionally managed for eye squint affirmation. The eye is an area constraint that is done on the introduction of the framework. The found eye district from the apparent face is also utilized for eye following and squint territory. As appeared in the underneath figure (Fig. 2).

2.1.3 Thresholding

In thresholding, we discover whites of the eyes in the taken-out eye an area. With the assistance of the time library, it's very well, maybe settled if the white area of the eye dissipates for a while which is shown as a glimmer. Asquint divulgence can be settled with the assistance for EAR (Eye Aspect Extent) which is given as shown in the formula below.

$$\mathrm{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

3 Eye Motion Detection Algorithm

In this algorithm, red marking is determined to recognize the right and left eye. The right half of the image area must contain the left eye and the opposite way around. Using this, the computation crops out two separate sub-images containing the benefit eye and the left eye. By direct expansion, the structure produces eye photographs of a fixed size. The size relies upon the scale at which the face was seen. It is ordinary that shows limitation head is steady, and there is no improvement in the circumstance of the eye. A small movement of the head may result in a bumble or mistaken eye

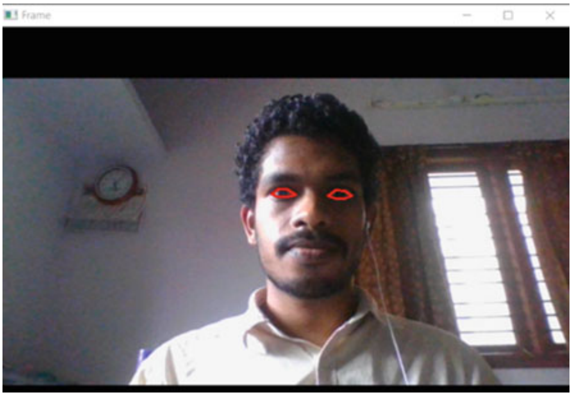


Fig. 3 Locating both eye on face

development. Accordingly, in this stage, the development of the eye is offset after the tended to eye picture is scaled. In this adjustment, the framework attempts to separate developments from wilful and automatic eye developments. Accordingly, utilizing this calculation, a definite eye development can be settled and can be used as a noteworthy limit for the graphical UI (Fig. 3).

4 Graphical User Interface and Proposed System

In this paper, the suggested framework is to develop a unique and intelligent GUI which is represented by developing a straightforward idea, which can be easily adopted by a patient [3] The GUI is instinctual that any patient can be trained how to use the system easily and efficiently. This GUI can be utilized by all age people from kids to adults. The project interface works with the assistance of two factors, eye movement discovery and eye squint location.

The eye movement discovery calculation can be used to pick a decision on the screen with the help of the left or right eye development. The eye squint calculation can be used to turn on/off with purposeful eye flickers. It can moreover be used to fix the picked other option, which will gigantically push the patient to deliberately control the system as demonstrated by his will. The virtual console is partitioned into 12 keys. The choices are Medical Assistance, Entertainment, Basic Utilities, and Emergency. In a virtual console, there are 12 keys in which it contains all the fundamental needs like a fan on, fan off, light on, light off, food, tv, music, medicine, snacks, juice, S.O.S. By selecting the keys in the virtual console, the system outputs the need of patients through voice format which is pre-recorded at the same time text message sent to the caretaker of the patient by using Twilio as shown in Fig. 5.

Figure 4 shows the virtual console which contains 12 options, the virtual console is pointing at 6 which means the patient is requesting water. The system outputs the audio, as well as the text message sent to the caretaker of the patient.

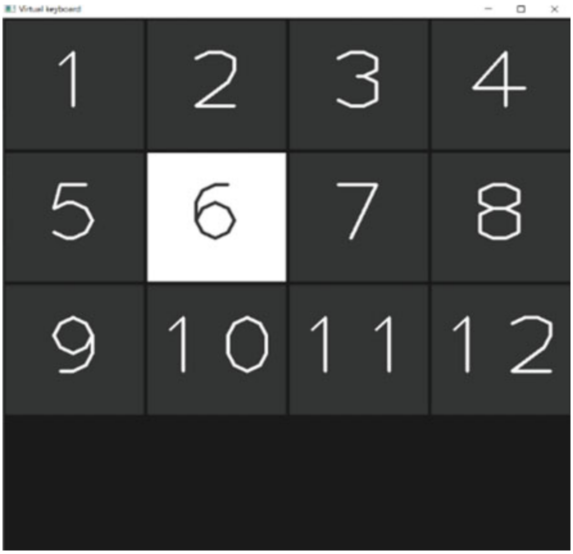


Fig. 4 Virtual console pointing at 6

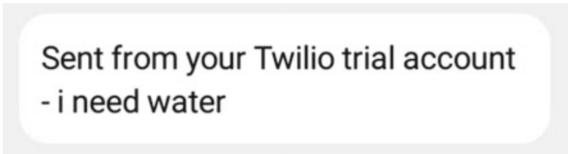


Fig. 5 Text message sent to the caretaker of the patient

From Fig. 6, the virtual console is pointing at 7 which means the patient is requesting a washroom. System outputs the audio, as well as the text message sent to the caretaker of the patient as shown in Fig. 7.

From the Fig. 8, the virtual console is pointing at 9 which means the patient is requesting medicine. The system outputs the audio, as well as the text message sent to the caretaker of the patient as shown in Fig. 9.

From the above Fig. 10, the virtual console is pointing at 12 which means the patient is feeling uncomfortable or he may have some health issues so he is calling for a doctor and the message is directly sent to Doctor saying as “Emergency”.

5 Result and Discussion

The output of the project is developed and tested on the Windows 10 Operating System, with 8 GB RAM, i5 seventh GEN, and 1 TB HDD inbuilt webcam and speaker. The virtual console is partitioned into 12 keys a graphical user interface among the display screen with 300 mm and the best detachment should be under 600 mm depending on the webcam and with sufficient lighting was provided. The

Fig. 6 Virtual console pointing at 7

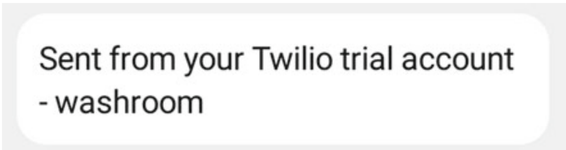


Fig. 7 Text message sent to the caretaker of the patient

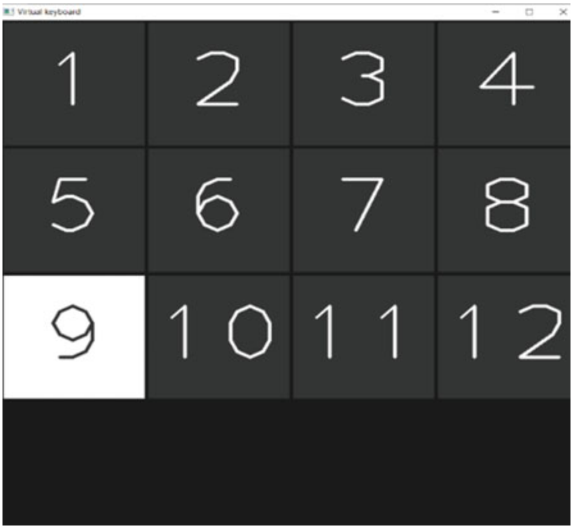


Fig. 8 Virtual console pointing at 9

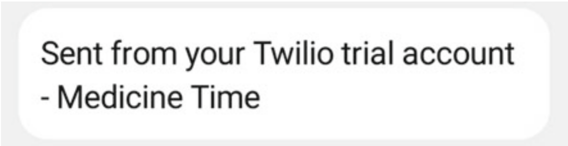


Fig. 9 Text message sent to the caretaker of the patient

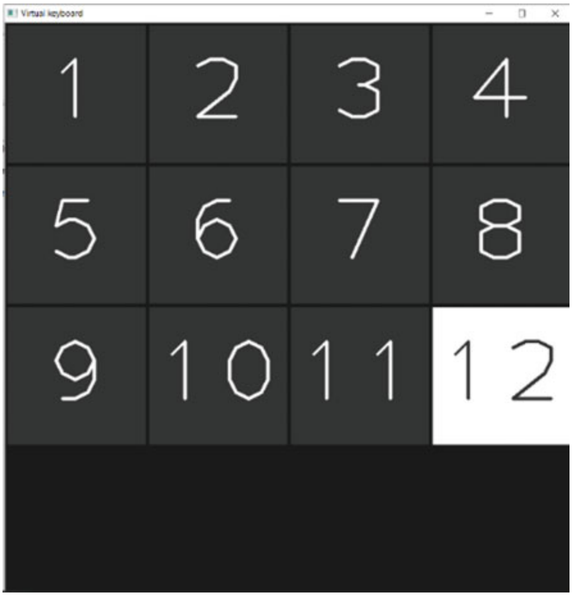


Fig. 10 Virtual console pointing at 12

virtual console should be within 1mts so that patient will not get eye strain and can use the system efficiently.

The output was tested on the human face where the webcam was focused on the eye and it was observed for eye blink or eye motion detection. A 6 s cut off was kept between cell to cell movement within the virtual console target, when the patient is visualizing the alternatives on the virtual console based on eye blink on the suitable cell, inputs of the squint and movements of eyes are sent for a calculation to process and execute the task. The system outputs the need of patients through voice message which is pre-recorded simultaneously, text messages are sent to the caretaker by using Twilio software. The project can be enhanced using Raspberry Pi and Wireless Infrared Remote Controller for multiple sensors to connect appliances like fans, light, Television which are affordable to all the common people [4].

6 Conclusion and Future Enhancement

The proposed framework empowers individuals with serious loss of motion to convey their thoughts and necessities. It additionally causes the patient to show their scholarly potential which can some of the time to pose their psychological inability analysed by the specialist. Further the project can be enhanced using Raspberry Pi and Wireless Infrared Remote Controller for multiple sensors to connect appliances like fans, light, Television which are affordable to all the common people. We can also use sensors to detect the light in the surrounding if it has low light it will enable the flashlight so that sufficient light will be available to operate.

References

1. www.cdn.shopify.com
2. Kumar A, Kaur A, Kumar M (2018) Face detection techniques: a review. Springer
3. Pandey M, Chaudhari K, Kumar R, Shinde A, Totla D, Mali ND (2018) Assistance for paralyzed patient using eye motion detection. In: Fourth international conference on computing communication control and automation
4. Nejakar SM, Pruthvi S (2013) Wireless infrared remote controller for multiple home appliances. Int J Electr Electron Res (IJEER) 2:25–35. Available at www.researchpublish.com. January–March 2014

Plant Disease Detection Using Deep Learning

K. Manjula, S. Spoorthi, R. Yashaswini, and Divyashree Sharma

Abstract Crop diseases are one of the major issues, but their identification is difficult due to the lack of required infrastructure. Plant diseases affect farmers whose livelihood depends on the crops and also it increases the vulnerability of food security at the large scale. Plant disease identification is very important because it affects the growth of the plant species. Usage of pesticides reduces the ability to fight back. The plant diseases are detected using Deep Convolutional Neural Network trained and added into the database of leaves from different plants. The dataset consists of 38 disease classes and one background class stanford open dataset [1]. The proposed model predicts from the image of a leaf if it is diseased or not and also gives the name of the disease predicted. In Deep learning, CNN is an algorithm that takes an image as an input and it assigns importance to the objects in the image and we can distinguish one image from the other. Architecture used is resnet architecture. We have many types of resnet architecture i.e. concept is same but with different number of layers, for example we have ResNet-34, ResNet-50, ResNet-18, ResNet-101 etc. Here, In this paper we are using ResNet-50 architecture, since it is a variant of Resnet model that has 48 convolution layers along with one average pool and one max pool layer. The results are deployed into the cloud (AWS) where the data can be fetched whenever required. Accuracy of the system is around 97–98%.

Keywords Deep learning · Plant disease detection · CNN

1 Introduction

Indian economy is highly dependent on agriculture therefore detection of plant diseases becomes very important. In this paper we use image classification concepts to recognize different plant diseases. Plant and also we have to halt the waste of financial and many other resources. Solution to this problem is as follows:

K. Manjula (✉) · S. Spoorthi · R. Yashaswini · D. Sharma
Department of Information Science and Engineering, Don Bosco Institute of Technology, Bangalore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_133

We have built a model that uses deep learning to detect and classify by taking an image of a leaf and classifies leaves using CNN predict the type of disease that might have affected the plant. The process proceeds as follows:

1. The leaf is detected in a given image.
2. Extracted leaf is run through a CNN classifier to identify which the leaf belongs to which plant.
3. The leaf image is then used to predict for which disease class it belongs to.

The Deep Learning (DL) approach is a subset of Machine Learning (ML), has networks capable of learning unlabeled data (Unsupervised Learning). This includes several developments like hand written text recognition, Backpropagation, chain rule etc. However, in the next phase, architectures were developed and used for many applications like healthcare sector, finance and earthquake predictions. Among those architectures, ResNet-50 architecture is used.ResNet-50 is a deep residual network and are mainly used for image classification. It uses the Skip connection from an earlier layer to a next layer to add the output. When these architectures started to evolve. The research has been done and applied to image recognition and classification Some of these architectures have been or various agricultural applications. For example classification of leaves was carried out by using CNN classifier CNN has input layer, convolution layer, pooling layer and classification layer and output layer. when compared to other algorithms diseases should be diagnosed very accurately.

2 Literature Review

1. Deep Neural Networks are used to detect and classify plant leaf diseases: In this paper, Deep learning technique is used to detect the plant diseases and classify them by examining leaf of a given plant.
2. Recognition of Plant leaf Using a Convolution Neural Network: In this paper we use YOLOv3 object detector, which is used to extract leaf from a given image and it is been analysed through ResNet-18 models.
3. Using Deep learning methods to classify Banana leaf diseases: In this paper, Deep learning techniques are used to classify the banana leaves diseases. The architecture used here in this paper is LeNet architecture that is used as a convolutional network architecture in order to classify the images.
4. Image Recognition using Deep Residual technique: In this paper, here we have a residual learning framework which is deeper than the Image Recognition using Deep Residual technique: In this paper, here we have a residual learning framework which is deeper than the framework used before. These frameworks formulate the layers again as learning residual functions with the reference to inputs of the layers, instead of learning other unexploited functions.
5. Deep Neural Networks are used for large scale plant classification: In this paper, we discuss about the possibility of deep learning techniques to apply for classification of plants and it is used for large scale biodiversity monitoring and also

the plant classification using convolutional network architectures like ResNet50 gives higher accuracy compared to other plant classification applications.

6. Detection of Plant leaf diseases using segmentation of and soft computing techniques: In this paper,an algorithm is used to segment the images i.e., segmentation technique to classify and detect the leaf diseases.
7. Leaf Recognition Algorithm to classify the Plant Using Probabilistic Neural Network: In this paper, To implement leaf recognition algorithm, we use PNN. 11 leaf characters are obtained. PNN is trained to classify 32 kinds of plants for 1800 leaves which gives accuracy greater than 90%.

3 Proposed Method

3.1 Dataset

Appropriate datasets are required for object the pre-processing required for CNN is lesser.

The process through the project PlantVillage, 26 disease(healthy) along with 54,307 images of 14 crop species are available [2]. The dataset consists of 38 disease classes and one background class stanford open dataset.

In order to differentiate between healthy leaves and disease caused ones, a new class is added which contains only leaves which are healthy and all the images of leaves are added into the dataset. An extra class which include background images was added into dataset for the classification to be more accurate. Thus, deep neural network could be trained to identify the leaves excluding the background images. Inorder to reduce the training time, the images has been resized to 256×256 (Fig. 1).

3.2 Image Preprocessing and Labelling

While collecting the images for the dataset, the images were in imprecise formats and the images were having different quality and resolution [3]. The collected images for the dataset having the recognition research, from beginning till the end of were considered as invalid to assign to the valid images for the dataset.Since Resnet has the Framework that can train DN Networks, we are able to train ultra deep neural networks, that is even if the network contains thousands of layers, it can still achieve greater efficiency (Fig. 2).

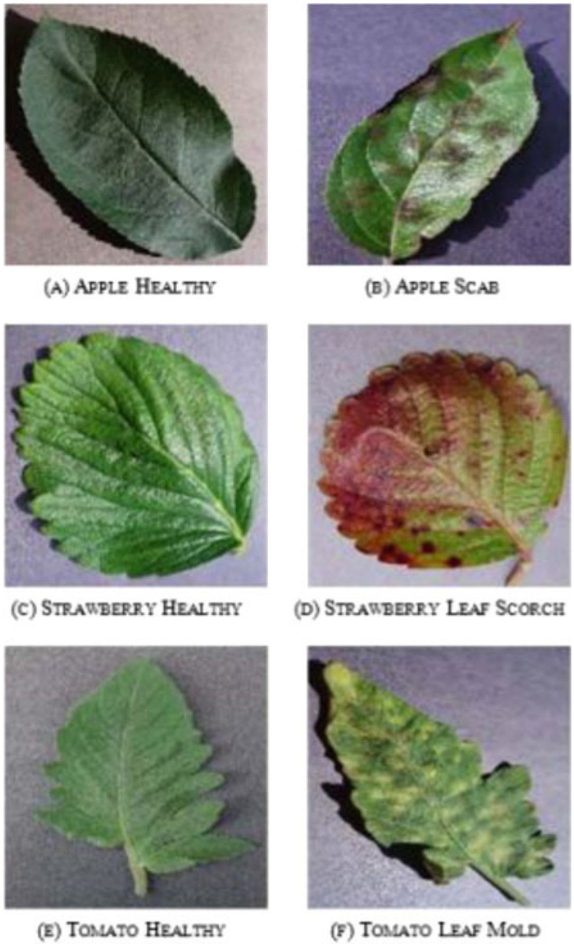


Fig. 1 Images from the dataset

3.3 Augmentation Process

The objective of using this is to introduce slight exaggeration to the images and to increase the dataset Which prevents from overfitting of the data during training stage [4–6]. In machine learning, due to random noise or statistical model describes error, overfitting appears rather than underlying relationship. And also provides the input image which are scaled by various factors. The image augmentation contains various transformation techniques which includes rotations of images, perspective transformation (Fig. 3).

Fig. 2 Example of leaf images from the plantvillage dataset

Fig. 3 Plant images

3.4 Neural Network Training

To build an image classification model, we are training the deep convolutional neural network from a dataset. Here, deep learning uses a neural network to elevate animal intelligence and helps in classification of various plant disease.

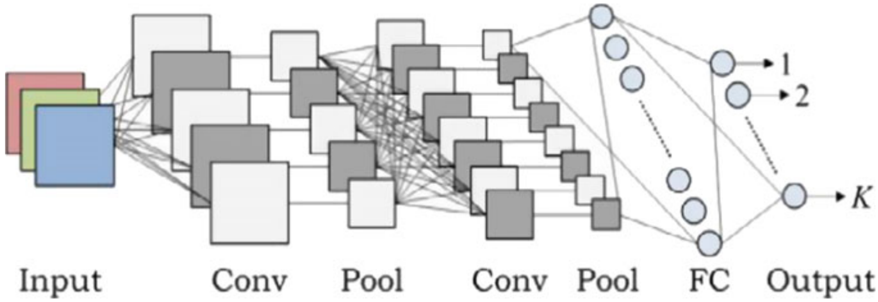


Fig. 4 Architecture of CNN

A simplified CNN architecture is made up of an input layer, multiple hidden layers and output layer, as shown in the above Fig. 4. Activation function, Pooling layer and Fully connected layers are all together form hidden layers. CNN Architecture is established when all these layers are deformed. The Architecture used in this project is Resnet50 architecture [7–9]. ResNet, used as a backbone and is a classic neural network.

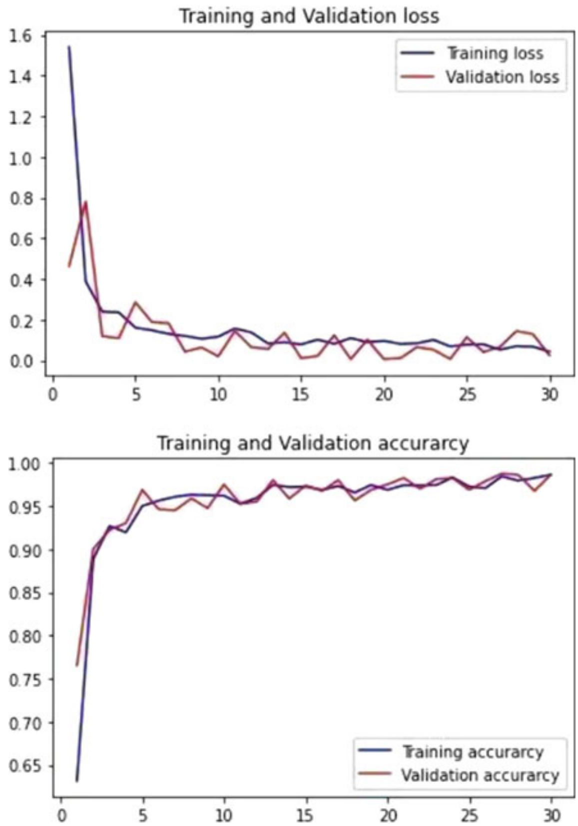
4 Results

The results are obtained by training the database which contains both augmented and original images with the complete dataset [10]. When we train the data on the large datasets the convolution neural network is able to learn new features, results will not be explored only when it is trained with original images. After the revamp of parameters of the network, overall accuracy of 97.5% is obtained (Fig. 5).

5 Conclusion

Using deep learning techniques, we executed a method to detect and classify the plant diseases from the leaf images. The obtained representational model can differentiate between healthy leaves and unhealthy leaves containing 38 different diseases, that can be visibly acknowledged. The obtained results after the experiment have achieved accuracy precision between 91 and 98%, for separate class tests. The obtained accuracy for this trained model is 97.5%.

Fig. 5 Plant images



References

1. Venkataramanan A, Kumar D, Honakeri P, Agarwal P. Plant disease detection and classification using deep neural networks. Int J Comput Sci Eng (IJCSE)
2. Jeon W-S, Rhee S-Y (2017) Using CNN to detect and classify the Plant leaf diseases. Int J Fuzzy Logic Intell Syst 17(1):26–34
3. Amara J, Bouaziz B, Algergawy A (2017) Using deep learning method to classify banana leaf diseases. BTW 2017
4. He K et al (2016) Deep residual learning approach for the recognition of images. In: Proceedings of the IEEE conference on computer vision and pattern recognition
5. Heredia I (2017) Deep neural networks are used for large scale plant classification. In: Proceedings of the computing frontiers conference. ACM
6. Riley MB, Williamson MR, Maloy O. Plant disease diagnosis. The Plant Health
7. Singh V, Misra AK (2017) Detection of plant leaf diseases using image segmentation and soft computing techniques. Inf Process Agric 4(1)
8. Wu SG, Bao FS (2007) A leaf recognition algorithm for plant classification using probabilistic neural network. IEEE international symposium on signal processing and information technology
9. Munisami T et al (2015) We use shape features and colour histogram with KNN classifiers to recognize the Plant leaf. Procedia Comput Sci 58:740–747

10. Al-Hiary H (2011) Detection and classification of plant leaf diseases will be accurate and fast. Int J Comput Appl 17(1):31–38

Personalized Emotion Recognition Utilizing Speech Signal and Linguistic Clues

A. V. Mohan Kumar, H. V. Chaitra, S. Shalini, and D. Shruthi

Abstract In this project, non-intrusive videos are taken in real-time, which detect the emotional state of an individual by analyzing the facial expression. We present the Convolution Neural Network (CNN), a machine learning algorithm used for automatic image classification. This supervised classification technique analyses and trains the classifier on the labeled images and extracts the features of the classifier. By using the learned information of the training, the newly provided image will be classified based on the features observed in the image. The prototype which we are going to develop detects whether a person is happy or sad or angry or surprise.

Keywords Deep learning · Machine learning algorithm · Sentiment · Speech recognition · Analyzing · Linguistic cues · Acoustic feature · Text analysis

1 Introduction

Human impersonation such as state of mind, emotions, and elements of emotionalism are becoming critical in strike recognition in the last few years. Affect is a neurological condition that is acquired for evaluation of responsiveness. In fact, state of mind conveys any person's immediate responses where the state of mind is more dormant towards feelings. Additionally, Sentiment is an integrated emotion that is called a collection of sentiments. Analyzing feelings implicitly to be practical for an interactivity to transpire a sufficient prompt response. Ekman and Friesen interpreted the deployed research as one or more of the six basic emotions of people with multicultural and identity, including happiness, sorrow, disgust, fear, rage and surprise. Evolving a mathematical model for the emotions deployed in human capacity is still a daunting task that needs more analysis. This is probably triggered primarily by constituents such as studying ability, environmental changes, emotional intelligence, community, etc., which are personalized in humans, releasing greater human skills From a statistical viewpoint, the associated manifestation is rendered with regard

A. V. Mohan Kumar (✉) · H. V. Chaitra · S. Shalini · D. Shruthi
Department of Information Science and Engineering, Don Bosco Institute of Technology, Bengaluru, Karnataka, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_134

to approaches to artificial intelligence and machine learning, which are prepared to impersonate the intricate neural networks into a definite field. These trigger work on gauging people's emotions and an almost critical juncture reaction like methodology of emotion recognizer. Different ways of interpreting emotions are active spaces for research in an attempt to formulate such a system, Either one or a combination of the following: (1) speech processing on the basis of a linguistic and paralinguistic model, (2) visual testing on the face, and (3) Functional Olfactory Information and Electrophysiological Signal Analysis. With the advent of HCI, Affective Computing is opening up to a wide range of fields of study and applications.

Deep learning (DL): Deep learning is a machine learning method that teaches computers to do what is common to humans: learn through analogy. It is the main technology behind driverless cars that allows them to recognize a stop sign or to distinguish between a pedestrian and a lamp post. It is a key to a range of voice control products, such as phones, laptops, televisions and hands-free speakers. There's a lot of deep learning in general Techniques such as deep neural networks, alternating neural networks, deep value networks, convolutionary neural networks, etc. This model can achieve cutting-edge precision, sometimes exceeding human performance. Models can be trained using a wide collection of labeled data and neural network architectures that include several layers.

2 Existing System

Different instruments, such as taking images, use imaging to better perceive the emotions. Here in real-time a non-intrusive image is captured which measures the emotional status of an individual by analysing facial expression. We sense an individual emotion in a picture frame and the decision is captured at stress level. We used heart beat rates to control the voltage in advance.

3 Proposed System

Using deep learning method, our project is able to detect person emotion by using their facial expression and audio. The camera module processes the image sequences captured and the pre- processing of the images for the subsequent analysis is done. Our system uses the FER dataset as a bench mark for classification of grayscale images. The facial expressions captured in the picture frame uses Haar cascade classifier to detect the emotion with higher efficiency.

4 Modules

4.1 Data Collection

We are analyzing the human voice recording dataset at zenodo.org. The same set of questions on major emotions of happiness, fear, anger, disappointment, disgust, shame, and guilt were put to students, psychologists, and non-psychologists. Currently the final data is made up of about 3000 respondents spread across 37 countries in all five continents. Each emotion is roughly 1000 responses. It is a study of culturally complex emotions. The team cites that emotions caused by the laboratory are often inefficient due to display of diffused emotional characteristics. It consists of the responses provided through questionnaires [1]. For simplicity; we consider four large emotional groups such as happiness, anger, sadness and fear.

4.2 Text Classification

A Linear Kernel machine learning algorithm with SVM consists of the text classification module [2]. The model is developed using ISEAR Emotion Database training data, and predictions are tested. Scikit-learn, a free software library with Python for machine learning, is used for implementation. The kernel parameters and the soft margin C are the parameters of the SVM that determine its effectiveness. In this paper Linear SVC was considered similar to SVC with a linear kernel but intended for large-scale learning tasks, but it is merely a practical implementation of libsvm [3].

4.3 Feature Extraction

First, the static 2-D log Mel-spectrogram is removed and then reconstructed into three log Mel-spectrogram channels (current, delta, and delta-delta). As the DCNN input, the log Mel-spectrogram derived from the utterance is divided into several overlapping data increment segments. The AlexNet DCNN model, which is pretrained on the large scale ImageNet dataset, is then used to perform fine-tuning tasks for the extraction of affective characteristics [4].

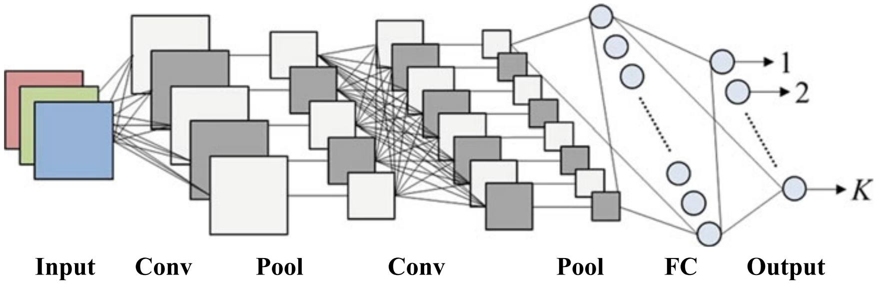


Fig. 1 Architecture of CNN

4.4 Emotion Classification

The Emotion Recogniser is developed as a combination of the Acoustic Model created in Step 1 and the text data model generated in Step 3.Problems with supervised machine learning provide a collection of historical data points that we want to use to forecast the future. Methods of neural network learning provide a robust approach to approximate real, discrete, and vector-assessed target functions. Convolutionary Neural Networks (CNN) are among the most powerful learning methods currently used for certain types of problems, such as the study of various features such as image or audio. A Convolutionary Neural Network, a subset of deep neural networks, is the most commonly used for research into visual imagery. CNNs are Multilayer Perceptron regularized models (Fig. 1).

A simplified CNN architecture is shown above. A convolutionary neural network consisting of an input and output layer, and several hidden layers. The secret layers are primarily feature of Activation, Pooling layers and completely linked layers. As these layers are stacked, the result of which was the development of CNN architecture. Multilayer perceptron usually refers to networks that are fully connected, i.e. each neuron in a single layer is linked to all neurons in the next layer. These 'full-connectivity' networks make them vulnerable to over-fitting data. There are 3 key parameters in a convolutionary neural network which need to be modified to adjust the behaviour of a convolutionary layer. These parameters are padding with philtre size, phase and zero. The size of the map generated for the output function depends on the important parameters above.

The philtre size plays a significant role in identifying the main features. Choosing an ideal philtre size is hard. It all hinges on the submission. A larger kernel size may miss the features and may skip the important details in the photos, whereas a smaller kernel may include more information leading to further uncertainty. Thus, the most appropriate size of the kernel/philtre must be calculated.

Methods such as Gaussian pyramids (set of kernels of various sizes) can be used to measure the efficiency of extraction of function and the required kernel size is calculated. Added to the filter size, it is very important to understand and decide the size of stride and the padding. Actually Stride determines the number of steps that you pass the philtre over the image data. When the stride is 1, one pixel at a time we pass the philtre. When we set the phase to 2 or 3 (uncommon), depending on the pace, we move the philtre 2 or 3 pixels at a time. The stride value also controls the size of the volume of output produced by the convolutionary layer. The bigger the phase, the smaller the volume of production. For example, if the input image is 7 server777 and stride is 1 the output volume is 5 server555 server5. On the other hand, if we raise the range to 2, the volume of output is reduced to 33.33.Stride is usually set in such a way that the volume of output is an integer and not a fraction.

The next important parameter is the padding to zero. Zero padding refers to padding the number of inputs along the boundary with zeroes. Also the zero padding helps one to monitor the output volume spatial scale. If we don't add any zero padding and we use a step of 1 the output volume spatial size will be reduced. With the first few convolutionary layers in the network however. We would like to maintain the spatial scale and ensure the volume of the input is equal to the volume of the output. This is where it's good to padd zero. If we use a stride of 1 and a zero padding of 1, then the output volume is equivalent to 7 à 77.77 in the example of the 7 quer7 input image.

The formula to be used to measure the padding value to get the spatial size of the input and output volume to be the same with stride 1 is $K - 12$ $K - 12$ where K is the filter size. Finally, the formula to calculate the output size is equal to $O = W - K + 2PS + 1O = W - K + 2PS + 1$Where O is the length of the output, W is the length of the input, K is the size of the philtre, P is the padding and S is the stride. For example, if we take $S = 1$, $P = 2$ with $W = 200$ and $K = 5$ and using 40 filters, then the output size will $200 \times 200 \times 40{,}200 \times 200 \times 40$using the above formula. On the other hand, if we use $S = 1$, $P = 1S = 1$, $P = 1$, then the output size would be$198 \times 198 \times 40$ (Fig. 2).

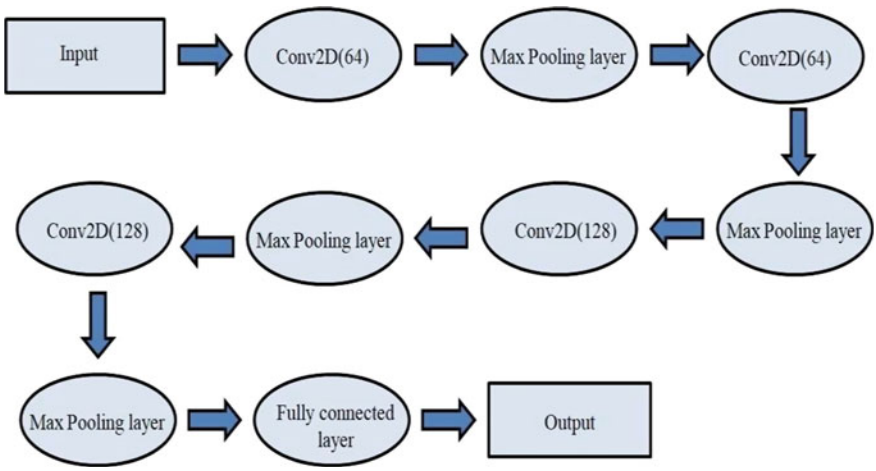


Fig. 2 Our CNN architecture, comprised of four layers

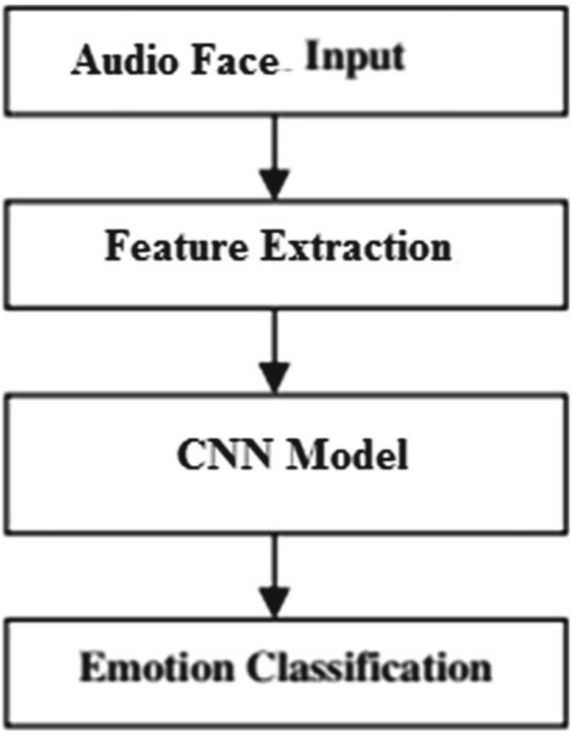
The basic functionality of the above mentioned CNN architecture can be Divide it into the following primary areas:

- Firstly, the pixel values of the image are kept in the input layer.
- The pooling layer will then simply collect samples along the spatial dimensionality of the input provided, further reducing the number of parameters in the activation.
- Completely connected layers would then incorporate the features needed to construct a model for classification purposes. It is also suggested that ReLu layer can be used to boost consistency between these layers.
- Eventually, we have an activation feature like Soft Max to arrange outputs.

5 System Architecture

The block diagram for emotion detection is shown above. The microphone reads the user voice and web camera reads face expression from that we system will predict the emotions using CNN algorithm (Fig. 3).

Fig. 3 Block diagram

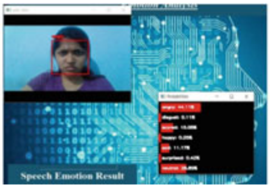


6 Results and Graphs

The main goal of this project is to recognize human emotion using their facial expression and speech, with too many of trails and error. We finally have achieved this project where it can recognize using facial expression as shown in below figure. Since, Deep learning is one of the effective ways when it comes to 2D-Mod we used CNN algorithm to detect the accuracy of the emotion a person is going through. This system is expected to be efficient and easy to handle. Convolution Neural Network Algorithm gives us the accuracy level of 76%.

Result 1: Emotion recognition for happiness, angry and Neutral

Accuracy 97.3%



Accuracy 44.11%

Accuracy 70.33%

Graph 1—Audio Accuracy

Input: X-axis contains values for the epoch, and Y-axis contains value for the loss.

Output: Graph to evaluate the loss values when the value of an epoch is given. Initially, the loss value for less epoch values is high. As the importance of an epoch increases loss decreases.

Graph 2—Emotion Accuracy

Input: X-axis contains values for epochs and Y-axis contains value for precision. Output: Chart to calculate accuracy values when the value of the epoch is given.

Input: X-axis contains values for epochs and Y-axis contains value for precision.

Output: Chart to calculate accuracy values when the value of the epoch is given.

At first, the accuracy value for less epoch values is poor. As the epoch-value increases precisions.

Graph 3—Test Results

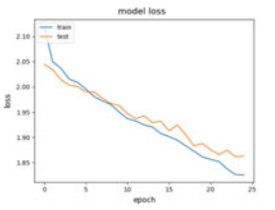
Input: X-axis contains values of emotion and Y-axis contains value of chance.

Output: Graph showing the likelihood versus the significance of the emotion.

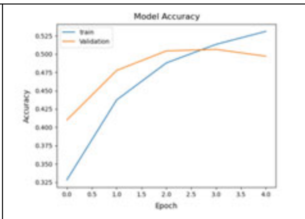
Used to represent various values of emotions at a given value of likelihood. To improve understandability different colour is used to reflect different emotions (Table 1).

Table 1 Result analysis

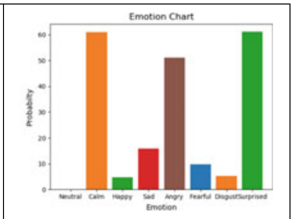
S. No.	Emotion recognition	Accuracy
1	Happiness	97.33
2	Angry	44.11
3	Neutral	70.33



GRAPH 1: Audio Accuracy



GRAPH 2: Emotion Accuracy



GRAPH 3: Test Results

7 Conclusion

Typically the technique of image processing is used by facial recognition with the assistance of webcams to detect human facial emotions, automatic detection of emotions is a key problem for ambient aided living solutions. We concentrate on evaluating research efforts through a study of text- and speech-based emotions.

Using a method of deep learning, our project is able to identify emotions of people by using their facial and audio expressions. The camera module handles the recorded image sequences, and the images are pre-processed for subsequent study.

Our framework uses the FER dataset as a reference point for grayscale image classification. The facial expressions captured in the image frame use the Haarcascade classifier to identify the emotion with higher efficiency. This project should be able to detect a person's emotion by using their facial expression and voice. By using this method, one can classify one's emotion. Our goal is to know the accuracy of one's emotion using the Convolutionary Neural Network (CNN). This system is expected to be efficient and easy to handle.

References

1. Dr. AN Nandakumar, Mohan Kumar AV. Twitter sentiment analysis using aspect-based bidirectional gated recurrent unit with self-attention mechanism. Int J Intell Eng Syst 3(5):97–110
2. Pranckeviius T, Marcinkeviius V (2017) Comparison of Naive Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for text reviews classification. Balt J Mod Comput 5(2)

3. Nandkumar AN, Mohan Kumar AV (2020) A survey on challenges and research opportunities in opinion mining. In: SN computer science. Springer, Singapore, pp 171–176
4. Nanda Kumar AN, Mohan Kumar AV. Sentiment analysis using robust hierarchical clustering algorithm for opinion mining on movie reviews-based applications. Int J Innov Technol Explor Eng (IJITEE) 8(8):452–457

A Performance Comparison of Optimization Algorithms on a Generated Dataset

Dharma Karan Reddy Gaddam, Mohd Dilshad Ansari, Sandeep Vuppala, Vinit Kumar Gunjan, and Madan Mohan Sati

Abstract Optimization Algorithms are one of the important machine learning techniques, they play a vital role in improving the performance of the model. In supervised learning, data samples are given, to train a model whose outcomes are already known. The dataset of 5000 samples with 2 features and 4 classes have been generated. We've trained the dataset with different Optimization algorithms such as Gradient descent, Mini Batch gradient descent, Momentum, NAG, RMSprop, Adagrad and Adam. To check the performance, these algorithms have been employed on the proposed model and compared the accuracy obtained by each algorithm. The presented work provides a clear understanding as well as the working of algorithms. The proposed model effectiveness can be seen by comparing different optimization algorithms. The developed model is able to obtain the maximum accuracy of 0.926 on Adam at testing phase with a learning rate of 0.5. The test result shows that Adam performs better than the other algorithms.

Keywords Machine learning · Optimization algorithm · GD (gradient descent) · MiniBatch · Momentum · NAG · Adagrad · RMSprop · Adam

1 Introduction

Deep learning plays a vital role in solving different real world problems. Neural networks and deep learning currently provide the best solutions to many problems. It makes the job easier in providing the results. Large Datasets are trained by the machine and effective results can be obtained by using certain functions such as Activation function, parameter initialization, learning rate etc. Various factors are involved when a dataset is being trained on a machine. Effective training over a

---

D. K. R. Gaddam · M. D. Ansari (✉) · S. Vuppala
CMR College of Engineering & Technology, Hyderabad, India

V. K. Gunjan
CMR Institute of Technology, Hyderabad, India

M. M. Sati
Graphic Era Hill University, Bhimtal Campus, Nainital, Uttarakhand, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_135

dataset provided better accuracy results. Moreover, many parameters are involved when an algorithm is ready to run. Dataset is one of the important elements in deep-learning, where the data is used for training purposes. Furthermore, Dataset could be in any form such as Pictures, text, audio and video etc. CNN's are used especially for the datasets comprising pictures. Some of the most popular CNN architectures used are Resnet, Googlenet and VGG etc. Additionally, a dataset can also be generated by machine. There are certain functions and packages which help in generating a dataset. Further, optimization algorithms are helpful in providing different solutions. Each algorithm has got its explicit speciality of giving out results. Training the model involves various factors such as activation function, initialization method (used for weights and bias) and optimization algorithms. Moreover, some of the activation functions used are Sigmoid, Tanh, Relu, LeakyRelu etc. Similarly, there are certain approaches to initialize the parameters i.e.—weights and bias. Additional to that, most popular initialization methods are Xavier and He. We have applied 'leaky Relu' as an activation function and 'He' as initialization method for (weights and bias) on different algorithms. A brief description about the past work, model and algorithms are provided in the next section.

2 Related Work

Numerous researches have worked on optimization algorithms of deep learning models. The term 'Deep Learning' was introduced to the machine learning community by Rina Dechter in 1986 [1]. But it was not that much popular that time because of limited availability of data and processor. Nowadays data and processor speed both are easily available. So most of the experiments dealt with variants of optimization algorithms and these algorithms played an important role in improving the deep learning model architectures. Convolutional Neural Network (CNN) has been predominant among various deep learning models. Moreover, most of the CNN models performed well with Adam Optimizer. Furthermore, to improve the performance of the model, it is trained on different datasets with diverse optimization algorithms. Additionally, learning rate is also an important hyper parameter in determining as well as improving the performance of the model. Various researchers have been trying different optimization techniques using different activation functions to improve the performance of the model. Extensive research has been carried out to understand the hyper-parameter effect on the performance of these algorithms [2–4]. So, it is quite sure that several improvised algorithms will be discovered to improve the performance of the model. Zaheer and Shaziya [5] has worked on optimizers with different datasets. They have selected some datasets namely MNIST, Cifar10, Fashion MNIST, Cifar100 etc. to check the performances. They have trained the model using different optimizers and Adam has given a better result in the testing phase. Moreover, E. M. Dogo has done the analysis of gradient descent algorithms and studied the performance of optimizers on different datasets [6–9]. A comparative effect of seven optimization algorithms on image classification datasets using simple convolution architecture was experimented. It is observed that Nadam, consisting of NAG and Adam, performed better for all the datasets used. Performance of each optimizer was displayed using plots and a deep analysis has been done on Adagrad optimizer. Lastly, an improvised Adagrad optimizer was developed to get better results, a certain improvement was observed for each optimizer, and optimizers were developed to improve the accuracy as well as the performance of the Model [10–14].

2.1 Optimization Techniques

These are one of the most important techniques in machine learning, which are useful in improving the performance of the model. Every algorithm works differently and thus it's very interesting to know the difference in the algorithms. We have used seven optimization algorithms to train the proposed model. The algorithms used for training are given below:

1. Gradient Descent
2. Mini-Batch gradient descent
3. Momentum
4. NAG
5. Adagrad
6. RMSProp
7. Adam.

Every optimization algorithm is having its own specialty and brief description for every optimizer is given as follows:

1. **Gradient Descent Optimizer**

Gradient Descent is the basic algorithm used for any model. The main functionality of this technique is to improve the parameters on effective training. Moreover, parameters play a vital role in producing efficient results. Additionally, loss should be as much as minimum to get better results. The main disadvantage of gradient descent optimizer is slow learning and used only for low-level datasets.

2. **Mini-Batch Gradient Descent Optimizer**

This algorithm is the advanced version of the gradient descent optimizer, where an input batch size is taken and effective batch size is provided to get the efficient results. It works well for low level datasets. As we increase the batch size, the stability of the model also improves, which resulted in better estimates of the gradient. Recommended batch sizes are 32, 64, 128 etc.

3. **Momentum Based Gradient Descent optimizer**

It is developed to overcome the limitation of gradient descent. Gradient Descent takes a lot of training time to navigate the regions with gentle slopes, as a result, gradient is quite small. This optimizer works confining to the past learning, which helps in running faster and provides better results compared to Gradient Descent optimizer.

4. **NAG Optimizer**

NAG is Also known as nesterov accelerated gradient descent and is used in dismissing the excessive oscillations produced by the momentum based gradient descent. Overfitting of a model may happen excessively in momentum based gradient descent, Hence the chances of graph escaping the minima valley in a graph is smaller.

5. **Adagrad Optimizer**

Adagrad optimizer works well for high level datasets. The Adagrad (Adaptive Gradient) is an algorithm which adapts the learning rate of parameters. Also, it is useful in giving out improvised results. Moreover, Decay of learning rate of parameters is proportional to their updated history. The advantage of this optimizer is that it eliminates the need to manually tune the learning rate.

6. **RMSProp Optimizer**

This Optimizer works well for high level datasets. The key advantage of RMSProp over Adagrad is that RMS provides a weighted accumulation of old and current updates of parameters leading to better results.

7. **Adam Optimizer**

Adam Optimizer is considered to be the best algorithm confining to time, updating parameters or learning rate. It is the combination of RMSProp and Adagrad, This optimizer follows the principles of both RMSProp and Adagrad Moreover, better accuracy can be achieved in less training time. The bias correction is done in a smoother manner resulting in brighter performance of the model.

3 Dataset

We have generated a dataset of 5000 samples with 4 centers and 2 features. We have trained this dataset using all the optimization algorithms and later compared their accuracies. Below is a plot of the dataset generated.

We have used four colors to represent 4 classes of the dataset, where the labels are classified as 0, 1, 2, 3.

4 Proposed Model

The proposed model is trained on a created dataset of 5000 samples comprising 2 features and 4 classes. As shown in Fig. 1, the plot consists of blobs of the data generated, where the data is classified into four labels 0, 1, 2, 3. We have considered many activation functions and initialization methods. On the basis of obtained results, we observed 'leakyRelu' along with 'He' initialization performed well compared to

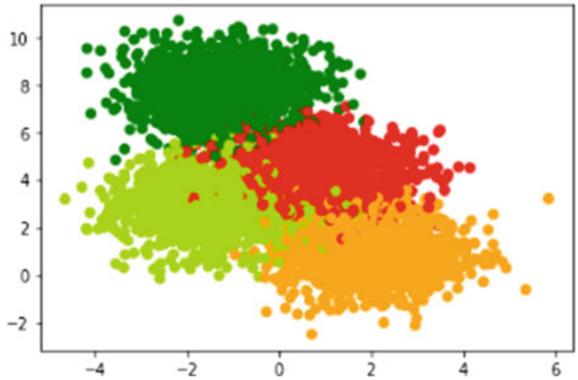


Fig. 1 Plot of generated dataset of 4 classes

the other activation functions and initialization methods. A learning rate of 0.5 was taken to train the model.

We have trained the model from 0 to 100th epochs, to see the trend of the loss of the model over every 10 epochs. Through our analysis of training the model, a major decline of graph was seen between 30th-40th epoch and loss continued to change minutely(less) from 50th epoch. Moreover, there was no bigger fluctuation in loss. We have continued to train the model using different optimization algorithms. We have created the dataset using the make_blob function. Usually this function is used in creating blogs of data, which takes labels, centres, features as arguments. We have fixed the random_state to be 0.

$$\textbf{make\_blobs}(\textbf{n\_samples} = \textbf{5000}, \textbf{centers} = \textbf{4}, \textbf{n\_features} = \textbf{2}, \textbf{random\_state} = \textbf{0})$$

The Above function is used for generating data of 5000 samples of 2 features and 4 centres.

The Model is trained using the given constraints, where 'He' and 'leakyRelu were used as initialization and activation functions. The model produced significant results with Adam optimizer. Activation function plays a vital role in producing the predicted values using the input(X).The formulae of leakyRelu is given as **:max(leaky_slope*X,X),**where X is considered as the input and leaky_slope is the value given by the user, a leaky_slope of 0.1 is taken to calculate the outcome. Training is done with different optimizers and Adam performed better in giving out a better test accuracy than others. Given below is the pseudo code of Adam optimizer, which is used in updating the parameters in a efficient manner.Beta1 and Beta2 are initialized with the values of 0.9 in the proposed model.

Pseudocode of Adam Optimizer
Require: α: Stepsize
Require: $\beta_1, \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates
Require: $f(\theta)$: Stochastic objective function with parameters θ
Require: θ_0: Initial parameter vector
$m_0 \leftarrow 0$ (Initialize 1st moment vector)
$v_0 \leftarrow 0$ (Initialize 2nd moment vector)
$t \leftarrow 0$ (Initialize timestep)
while θ_t not converged **do**
$t \leftarrow t + 1$
$g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep t)
$m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)
$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
$\widehat{m}_t \leftarrow m_t/(1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)
$\widehat{v}_t \leftarrow v_t/(1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)
$\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t/(\sqrt{\widehat{v}_t} + \epsilon)$ (Update parameters)
end while
return θ_t (Resulting parameters)

Above is the pseudo code for the Adam optimizer, where weights are updated in an efficient manner, which ultimately helps in reducing the loss and increasing the accuracy.

5 Results and Analysis

We have performed the result on the created dataset and obtained results can be seen in Tables 1 and 2. Moreover, we have trained the model using 'He' as the initialization method and 'leakyRelu' as the activation function. Training has been done with a learning rate of 0.5 for all the algorithms. We have run the model for all the algorithms and compared their accuracies for every 10th epoch. Additionally, Epoch is one such pass where the entire model gets trained and for every epoch parameters i.e. (weights and bias) are updated, which help in improving the accuracy. Table 1 show all the values, when the model is trained across all the optimization algorithms. All the algorithms gave a positive impact from the 40th epoch, and loss continued to remain same i.e.(didn't get much fluctuated). All the Bold accuracies denote the maximum accuracy produced by each algorithm. As we have trained the model from 0-100th epochs. Loss was observed to remain same from 50th epoch. Observations were made and accuracies were noted at every 10th epoch to see the trend of accuracies.

$$\textbf{Accuracy} = \textbf{Correct prediction}/\textbf{Total predictions}$$

Table 1 Obtained results by different optimizers at training phase

EPOCH	GD	MiniBatch	Momentum	NAG	Adagrad	RMS	Adam
00	0.04	0.04	0.04	0.04	0.04	0.04	0.04
10	0.56	0.915	0.350	0.651	0.411	0.778	0.410
20	0.873	0.932	0.481	0.712	0.487	0.461	0.518
30	0.881	0.931	0.661	0.824	0.569	0.516	0.694
40	0.889	0.931	0.851	0.920	0.716	0.634	0.891
50	0.894	0.932	0.810	0.926	0.832	0.905	0.919
60	0.895	**0.933**	0.912	**0.927**	0.878	0.921	0.924
70	0.898	0.933	0.917	0.926	0.891	0.921	0.925
80	0.900	0.933	0.913	0.926	0.901	0.898	0.931
90	0.890	0.933	0.921	0.926	0.905	0.666	**0.933**
100	**0.901**	0.932	**0.923**	0.926	**0.926**	**0.926**	0.932

Table 2 Obtained test accuracies by different optimizers

Optimization algorithm	Test accuracies
GD	0.893
MiniBatch	0.922
Momentum	0.907
NAG	0.922
Adagrad	0.897
RMSProp	0.657
Adam	0.926

From the above formulae, the accuracy of the model can be calculated. Evaluation of the model is done using log loss function. Loss is calculated using the log loss function, where parameters are developed or updated in such a way that, the loss gets minimized. Given is the equation representing the log loss function.

$$Log - loss = \frac{1}{n}\sum_{i=1}^{n}(\log(pi) * yi) + ((1 - yi) * \log(1 - pi))$$

n Total count
yi Actual label
Pi Predicted probabilities.

Training accuracies are given in Table 1.For Mini Batch optimization algorithm training has been done using a batch size of 64.

Figure 2 represents the loss of model on y -axis and epoch count on x-axis using a Adam Optimizer, which produced a best accuracy of 0.926 on testing phase. We can also observe that there is no great difference between the testing accuracy and

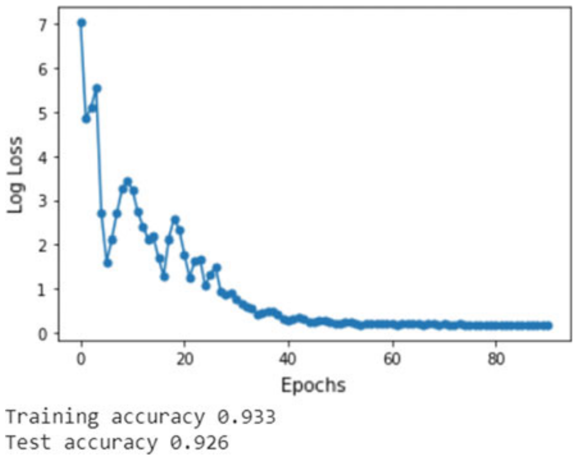


Fig. 2 Plot of loss done for Adam optimizer

training accuracy. It indicates that there is a good learning of parameters and over-fitting is prevented. Moreover Over-fitting is a major problem in any model, so it is considered as one of the disadvantages of a model, Over-fitting can be observed when there is a large difference in the training and testing accuracy. A model with low bias and variance is considered to be a good model. Hence using the right activation function, initialization method, optimization algorithm as well as learning rate, will give comparatively better results. Activation function may also depend on the dataset that the user takes. Thus, we have trained the proposed model and evaluated the accuracies on each optimizer from 0th-100th epoch. The obtained results are shown in Table 1.

We can observe from Tables 1 and 2 that each optimization algorithm provides different accuracies, which means selection of optimization algorithms to particular model and dataset is really important. It is observed that both training and testing accuracies are quite similar (i.e. very less difference). Plotting has been done for every 10 epochs, so that the trend for every optimization algorithm can be seen.

It is observed that MiniBatch and Adam optimizer produced better results in training phase, where Minibatch with batch size of 64 produced significant results in training with a best accuracy of 0.933 at 60th epoch. Similarly, Adam produced an accuracy of 0.933 at 90th epoch, whereas Adam performed better in testing phase with 0.926 accuracy. We have chosen these seven optimizers to experiment and gain insights of how each optimizer performs.

Table 2 comprises the test accuracies obtained by each optimizer in the 90th epoch. From Table 2, Adam provided the best accuracy of 0.926 in testing phase with a learning rate of 0.5

6 Conclusion and Future Scope

This paper gives us a clear understanding of how each algorithm performs when trained on a generated dataset. We have observed that the Minibatch and Adam optimizer produced better results in the training phase and achieved a maximum accuracy of 0.933 whereas Adam performed better in the testing phase with 0.926 as test accuracy. We have chosen these seven algorithms to experiment and gain insights of how each algorithm performs, when trained on a generated dataset. This topic has got a big impact in research area through successive research and practice; many optimizers can be developed in future to improve the performance of the model.

References

1. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
2. Lee JD, Simchowitz M, Jordan MI, Recht B (2016) Gradient descent converges to minimizers. In: Conference on learning theory, pp 1246–1257. PMLR
3. Mukkamala MC, Hein M (2017) Variants of RMSProp and Adagrad with logarithmic regret bounds. In: International conference on machine learning, pp 2545–2553. PMLR
4. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv:1412.6980
5. Zaheer R, Shaziya H (2019) A study of the optimization algorithms in deep learning. In 2019 Third international conference on inventive systems and control (ICISC). IEEE, pp 536–539
6. Zhang N, Lei D, Zhao JF (2018) An improved Adagrad gradient descent optimization algorithm. In 2018 Chinese automation congress (CAC). IEEE, pp 2359–2362.
7. Mercier Q, Poirion F, Désidéri JA (2018) A stochastic multiple gradient descent algorithm. Eur J Oper Res 271(3):808–817
8. Darken C, Chang J, Moody J (1992) Learning rate schedules for faster stochastic gradient search. In: Neural networks for Signal Processing II Proceedings of the 1992 IEEE Workshop, pp 1–11
9. Chollet F (2017) Deep learning with python. Simon and Schuster
10. Ketkar N (2017) Introduction to pytorch. In: Deep learning with python. Apress, Berkeley, CA, pp 195–208
11. Sethi K, Jaiswal V, Ansari MD (2020) Machine learning based support system for students to select stream (subject). Recent Adv Comput Sci Commun (Formerly: Recent Patents Comput Sci) 13(3):336–344
12. Kumar S, Ansari MD, Gunjan VK, Solanki VK (2020) On classification of BMD images using machine learning (ANN) algorithm. In: ICDSMLA 2019. Springer, Singapore, pp 1590–1599
13. Gaddam DKR, Ansari MD, Vuppala S (2020) On Sudoku problem using deep learning and image processing technique. In: ICCCE 2020. Springer, Singapore, pp 1405–1417
14. Agarwal M, Bohat VK, Ansari MD, Sinha A, Gupta SK, Garg D (2019). A convolution neural network based approach to detect the disease in Corn Crop. In: 2019 IEEE 9th international conference on advanced computing (IACC). IEEE, pp 176–181

Human Facial Emotion Detection Using Deep Learning

Dharma Karan Reddy Gaddam, Mohd Dilshad Ansari, Sandeep Vuppala, Vinit Kumar Gunjan, and Madan Mohan Sati

Abstract Human emotions are important features in human communication. It is also equally important to recognize and classify emotions using computers, which drives this research to explore in various real world domains. Deep learning techniques are rapidly growing in this era and help in tackling real world problems. In this work, deep neural networks or convolutional neural networks are used to detect human facial emotions. We have proposed a network with the help of ResNet50 to classify the human facial emotions using static images. To train the network FER2013 dataset is used and the proposed model produced significant results compared to existing models.

Keywords Deep learning · Face detection · Human emotions · VGG · ResNet50 and AlexNet

1 Introduction

Facial emotion recognition is an active research area in the fields of computer vision and artificial intelligence. Numerous researchers are working in this research area over the past few decades, and it is still challenging due to complexities in this research domain [1]. Moreover, Emotion detection and recognition play a vital role in creating a huge impact in major real world applications. Verbal and non-verbal are two major forms of expressing an emotion. In verbal form, emotions are expressed in written form whereas in non-verbal form, emotions are expressed in facial emotion gestures and vocal tone of a person. A person can exhibit many expressions, where they can be classified into 7 basic emotions and 15 compound emotions. The basic emotions are happy, sad, surprise, anger, fear, disgust and neutral. In addition, the compound

D. K. R. Gaddam · M. D. Ansari (✉) · S. Vuppala
CMR College of Engineering & Technology, Hyderabad, India

V. K. Gunjan
CMR Institute of Technology, Hyderabad, India

M. M. Sati
Graphic Era Hill University, Bhimtal Campus, Nainital, Uttarakhand, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_136

emotions are Happily surprised, Fearfully disgusted, Happily disgusted, Angrily surprised, Sadly fearful, Angrily disgusted, Sadly angry, Disgustedly surprised, Sadly surprised, Appalled, Sadly disgusted, Awed, Fearfully angry, Hatred, Fearfully surprised etc. CNN has been the most promising technique among various deep learning models. CNN is a class of deep neural networks, most commonly applied to analyze visual images and patterns from the images. Moreover, convolutional neural networks have produced noticeable results in most of the computer vision tasks and many other applications areas. CNN is able to produce higher accuracy over traditional machine learning models such as support vector machine (SVM), Neural Network etc. It can produce accurate results in different Complex conditions [1–10].

Many subsequent research works have been done in the field of computer vision. Human facial expressions are directly associated with facial muscles of humans. Features were extracted from the factors which are responsible for facial expressions and are trained on various machine learning or deep learning models to detect as well as classify human facial expressions. In this work we have proposed model using ResNet Architecture to classify the emotions, trained on the FER2013 dataset.

2 Related Work

Many approaches were developed for human facial recognition even before the outbreak of deep learning. These approaches include Local Binary Patterns (LBP), linear discriminant analysis (LDA), Principal component analysis (PCA), Locality preserving Projections (LPP) etc. Moreover, Local Binary Patterns (LBP) is used to find facial expressions automatically. This technique is employed to fix the position of eyes, eyeballs, and also focuses on face structure such as position of eyes, mouth and nearby areas on the face. After the evolution of machine learning, there is a lot of research going and different models have emerged such as support vector machines, neural networks approaches etc. [1]. Yang et al. [2] proposed Neural Network model, where vectorized facial features are used as input. The model can predict numerous emotions with the accuracy of 84.33%. Similarly, Liu [3] employed FER2013 dataset and a two and a two layer convolutional neural network model is used to identify different facial emotions. They compared their model with four different existing Methods and the proposed model achieved an accuracy of 49.8%. Additionally, Suresh et al. [4] developed a sign language recognition system which classifies six different sign languages using Deep Neural Networks. Adam and SGD optimizers are employed in two different approaches and compared them. Furthermore, the model using Adam optimizer performed good and produced significant results than SGD optimiser. Likewise, Verma and Tiwary [5] introduced an ensemble method using a support vector machine, Multilayer Perceptron, K-Nearest Neighbour and Meta-multiclass. The obtained average accuracy for SVM was 81.45%, MLP was 74.37%, KNN was 57.74% and MMC was 75.94%. Iacoviello et al. [6] presented a new hybrid classification model to detect the human emotions from EEG signals.

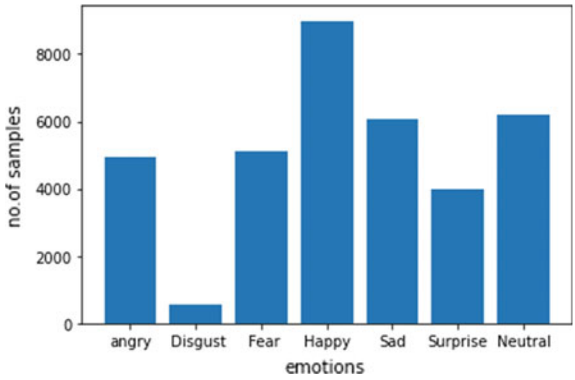


Fig. 1 No of sample images in each emotion of dataset

They combined SVM, principal component analysis (PCA) and discrete wavelet transformation [11–19].

3 Dataset

The dataset FER2013 is chosen for experimental analysis, which is an open-source dataset. It was first created for an on-going project by Pierre-Luc Carrier and Aaron Courville. Later this dataset was uploaded to kaggle for a competition. FER2013 dataset consists of 35,887 grayscale, 48 × 48 sized face images with various emotions. The emotions are labeled as (0: Angry, 1: Disgust, 2: Fear, 3: Happy, 4: Sad, 5: Surprise, 6: Neutral).

We can observe from the above Fig. 1 that there is strange variation in the number of samples of each emotion in the dataset. This variation will have an impact on the model and its efficiency. The dataset was in the form of a.csv excel sheet which comprises 48 × 48 pixel values of each image. Moreover, Images can be formed from the pixels of each image. Figure 2 is a sample of images of the FER2013 dataset, which is available at this link https://www.kaggle.com/.

In this dataset each image has less background grayscale images as color of person is not important to classify the emotion of a person and clear exposure of expression, this helps the model to extract more features with good efficiency.

4 Proposed Methodology

We proposed a deep learning model based on the convolutional neural network for classifying the human emotions based on static images. The complete network has different layers where images go through each layer of the network to classify its emotion. The layers are presented in Fig. 3.

Fig. 2 Greyscale images in fer2013

From the above diagram we can observe that input images are passed to the pre-processing stage. Moreover, Pre-processing is making data ready to pass it to the network such that the model predicts the output with high accuracy. Generally, in the Pre-processing stage noisy data is removed, uneven sizes of images are resized to a particular size, convert RGB to grayscale images.

And the next stage is Face Detection, as we are finding emotions of a human face it is required to find the face in an image. There could be multiple faces or no face in an image; therefore, it is significant to detect faces in an image. To detect faces MTCNN is used, MTCNN is an algorithm comprising 3 stages P, R, O, which detects bounding boxes of faces in an image. The bounding box is drawn based on coordinates in an image of a face. With the help of bounding boxes we can crop only faces in an image such that unnecessary parts in the image can be removed. The background is eliminated and left with only the face of an image. Now the data is good enough to pass to the model.

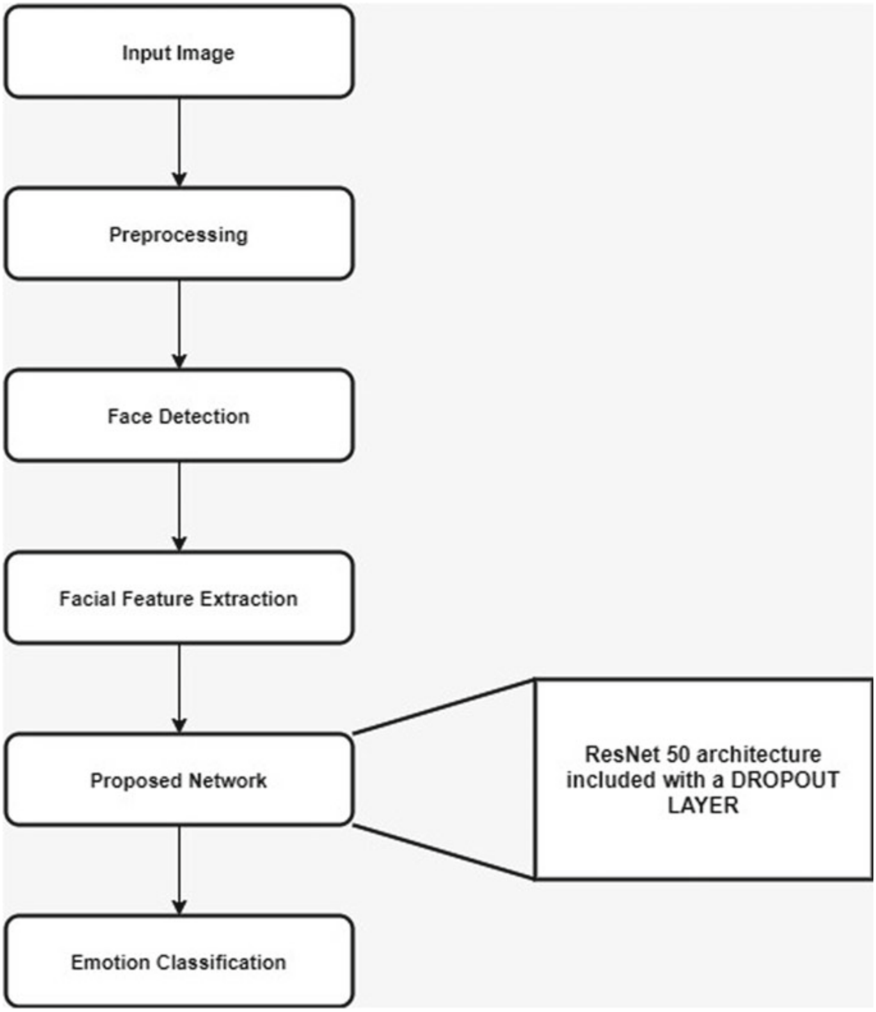


Fig. 3 Proposed model

Convolutional Neural networks are extremely good at abstract features from the face such as landmarks of face and distances between them.

4.1 Convolutional Neural Network

Convolutional Neural Network (CNN) has been pre-dominant among various deep learning models. This is a class of artificial neural networks that have produced astonishing results in the area of computer vision. A Convolutional Neural Network is a Deep learning Model used for processing, the type of data that has a grid pattern in it. These networks are built to learn the features and patterns from the images automatically and adaptively. A CNN typically comprises three major parts, they are—convolutional block, pooling layers and fully connected networks.

CNN accepts the input in the form of tensors with the shape of an image. Then the image is passed through the convolutional layer for the abstraction of features from it with a stack of multiple mathematical operations. Further, these convolutional layers convolve the inputs and move the output to the next layer. Furthermore, Pooling layers are used to minimise the dimensions of the input by aggregating the outputs of neuron clusters of a layer into a single neuron in the next layer. Additionally, the fully connected layers are used to classify the images. The neurons present in one layer are connected to each and every neuron in another layer. The hierarchy of extracted features become more complex as the layers feed their output to another layer as input. It uses back-propagation and gradient descent for the optimisation of weights and biases.

4.1.1 Convolutional Layer

Convolutional layer is responsible for feature extraction from the given image, simply known as feature extractor. The various features in the image are captured by a set of filters in a convolutional layer. Moreover, filters are simply a group of neurons which recognises the features and patterns at different places of an image. Further, feature maps are produced by a convolutional block convolving the filters across the dimensions of an image. Hyper parameters such as depth, striding and padding are used to produce feature maps from a convolutional layer.

4.1.2 Pooling Layer

Pooling layers are introduced in a convolutional network, to convolve or reduce the dimensions of each feature map produced by the convolutional layer. Moreover, pooling layer reduces the number of parameters as a result computational cost is reduced when images are too large to compute. Although, dimensions are reduced on the feature map by reducing the parameters, the important features and information is maintained by the pooling layer. Pooling can be done in three types: average pooling, max pooling, sum pooling.

4.1.3 Fully Connected Layer

The main objective of CNN is to classify the input into its class with the help of features. The previous convolutional layers along with the pooling layer help in feature extraction and reduce the dimensionality to minimize the computational cost;

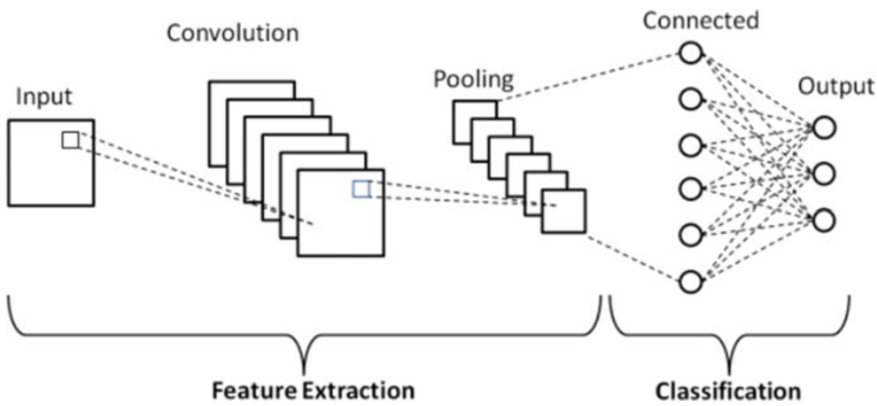


Fig. 4 Simple convolutional neural network

therefore fully connected layers are responsible for classifying the data or image into its class or category (Fig. 4).

ResNet50 architecture is employed to train the dataset ResNet, which is also known as the Residual Network. This is a convolutional neural network model which works on the concept of deep residual block. According to the universal approximation theorem, a single layer of feed forward network is used to represent any function. However, the layer might be massive and the network is prone to overfitting of the data. Therefore, our network architecture needs to go deeper. Increase in the deeper network doesn't work, resulting in a vanishing gradient problem, as the gradients are back propagated to the previous layers, rigorous multiplication operations are being conducted, gradually performance gets saturated and decreases rapidly.

A dropout is one of the regularisation techniques which prevents over fitting in the network. This dropout layer is added before the fully connected layers of ResNet50 architecture. This results in a great increase in test and train accuracy.

5 Experiments Analysis and Results

We have experimented with the proposed approach on the FER2013 Dataset. The obtained results are shown in Table 1 and Fig. 5. As the dataset is in grey-scale and size is 48×48, there is no need for pre-processing the images in the dataset. The dataset is made into batches with batch size of 64. Additionally, these batches of images were trained on the ResNet model with the help of a few hyper parameters. Cross Entropy loss function is used to calculate loss from the training result and the original dataset, loss function is given below.

Table 1 Comparison of test accuracies

Method	Test accuracies (%)
AlexNet	14.9
Vgg16	38.2
ResNet50	46.6
Proposed network	**55.6**

Highlighted value shows the higher accuracy rate

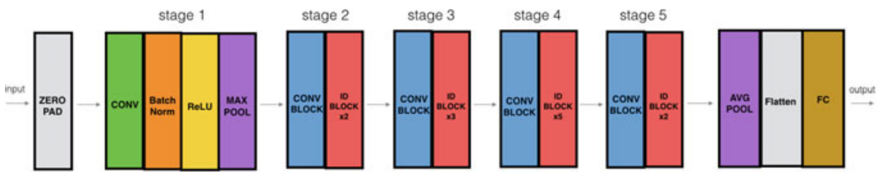


Fig. 5 ResNet50 architecture

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \log(p(y_i)) + (1 - y_i).\log(1 - p(y_i))$$

q is the input data and p is the obtained result.

The Optimization algorithm used is Adam optimization algorithm. Adam can be described as RMSprop and stochastic Gradient descent with momentum. Learning rate is scaled using squared gradients in the Adam optimiser. Moreover, the learning rate used is 0.01. Additionally, we have added a dropout layer before the fully connected layers to the ResNet50 which helped model to produce best results for that dataset. Furthermore, this dropout layer helped the problem of over fitting. These hyper-parameters are used in training the FER2013 dataset in an efficient manner. As the ResNet50 comprises trainable parameters, the network is completely trained on the FER2013 dataset. Lastly, we have obtained a training accuracy rate of 75.45% and test accuracy rate of 54.56%.

Figure 6 plotted between epochs: x-axis, loss: y-axis of training data. From the above graph we can observe that loss is minimum at 60th epoch. Though there are more epochs, the accuracies are oscillating among the same loss with less difference, which result in over fitting.

Table 1 depicted the test accuracies produced over different networks and proposed network. The selected networks are AlexNet, Vgg16 and ResNet50. The accuracy obtained after experimenting on these networks are 14.9%, 38.2% and 46.6% respectively. The test accuracies are gradually increasing on these networks and the proposed network has produced the highest accuracy is 55.6%.

Figure 7 shows the test results by passing custom images to the proposed network. The proposed network produced correct results over custom test images.

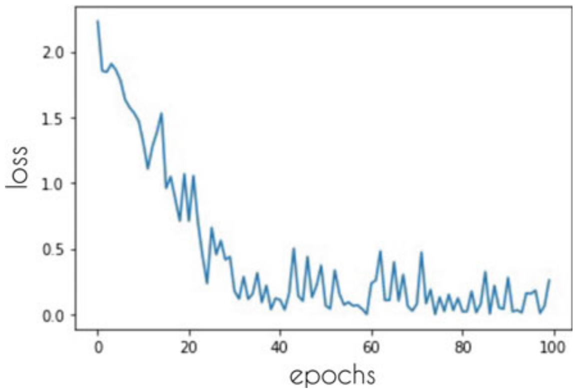


Fig. 6 Graph plot between loss and epochs

Fig. 7 Test results from proposed network

6 Conclusion

We have proposed a network by adding additional dropout layers to the actual ResNet50 architecture at required time steps. The proposed model produced better results than the other existing networks. We have achieved a training accuracy of 75% and test accuracy of 55% using the proposed network.

The future work may include collecting and adding some more images into this FER2013 dataset so that all the emotions will be balanced. It will ensure the same number of images in each emotion class with more variety of poses.

Acknowledgements I would like to thank Management of CMR College of Engineering & Technology, Hyderabad, India for providing the facility and infrastructure to carry out the research.

References

1. Ahmadvand P, Ebrahimpour R, Ahmadvand P (2016, November) How popular CNNs perform in real applications of face recognition. In 2016 24th Telecommunications Forum (TELFOR) (pp 1–4). IEEE
2. Yang G, Ortoneda JSY, Saniie J (2018, May) Emotion Recognition Using Deep Neural Network with Vectorized Facial Features. In 2018 IEEE International Conference on Electro/Information Technology (EIT) (pp 0318–0322). IEEE
3. Liu L (2019, August) Human face expression recognition based on deep learning-deep convolutional neural network. In 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA) (pp 221–224). IEEE
4. Suresh S, Mithun HTP, Supriya MH (2019) Sign language recognition system using deep neural network. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp 614–618). IEEE
5. Verma GK, Tiwary US (2014) Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals. Neuroimage 102:162–172
6. Iacoviello D, Petracca A, Spezialetti M, Placidi G (2015) A real-time classification algorithm for EEG-based BCI driven by self-induced emotions. Comput Methods Programs Biomed 122(3):293–303
7. Dachapally PR (2017) Facial emotion detection using convolutional neural networks and representational autoencoder units. arXiv preprint arXiv:1706.01509
8. Lopes AT, De Aguiar E, Oliveira-Santos T (2015 August) A facial expression recognition system using convolutional networks. In 2015 28th SIBGRAPI Conference on Graphics, Patterns and Images (pp 273–280). IEEE
9. Mehendale N (2020) Facial emotion recognition using convolutional neural networks (FERC). SN Appl Sci 2(3):1–8
10. Singh S, Nasoz F (2020, January) Facial expression recognition with convolutional neural networks. In 2020 10th Annual Computing and Communication Workshop and Conference (CCWC) (pp 0324–0328). IEEE
11. Hussain SA, Al Balushi ASA (2020) A real time face emotion classification and recognition using deep learning model. In Journal of Physics: Conference Series (vol 1432, No. 1, p 012087). IOP Publishing
12. Sagonas C, Antonakos E, Tzimiropoulos G, Zafeiriou S, Pantic M (2016) 300 faces in-the-wild challenge: database and results. Image Vis Comput 47:3–18
13. Duchenne GB (1862) Mécanisme de la physionomie humaine ou Analyse électro-physiologique de ses différents modes d'expression. P. Asselin, gendre et successeur de Labé, éditeur des Archives générales de médecine
14. Sethi K, Jaiswal V, Ansari MD (2020) Machine learning based support system for students to select stream (subject). Recent advances in computer science and communications (Formerly: Recent patents on computer science), 13(3), 336–344

15. Kumar S, Ansari MD, Gunjan VK, Solanki VK (2020) On classification of BMD images using machine learning (ANN) algorithm. In ICDSMLA 2019 (pp 1590–1599). Springer, Singapore
16. Gaddam DKR, Ansari MD, Vuppala S (2020) On sudoku problem using deep learning and image processing technique. In ICCCE 2020 (pp 1405–1417). Springer, Singapore
17. Agarwal M, Bohat VK, Ansari MD, Sinha A, Gupta SK, Garg D (2019, December) A convolution neural network based approach to detect the disease in corn crop. In 2019 IEEE 9th International Conference on Advanced Computing (IACC) (pp 176–181). IEEE
18. Rashid E, Ansari MD, Gunjan VK, Ahmed M (2020) Improvement in extended object tracking with the vision-based algorithm. In Modern approaches in machine learning and cognitive science: a walkthrough (pp 237–245). Springer, Cham
19. Vishwakarma AK, Ansari MD, Rai G (2020) SMS Spam Filtering Using Machine Learning Technique. In ICCCE 2020 (pp 689–701). Springer, Singapore

Recursive Least Squares Linear Equalizer for Spectral Efficiency Enhancement in Green Radio Communications

C. H. Nagaraju and Bharath Naga Raju

Abstract In this paper, a recursive least squares linear equalizer is used for receiving the data with less errors. In this process the data is modulated using the compressive sensing based index modulation in order to increase the spectral efficiency. This process reduces the BER and PAPR than the existing methods. This method can be mainly used in the green radio communications. This equalizes the data obtained at the receiver side. This recursive equalized based sparse index modulation can be used to reduce the noise in the receiver side. The results have been investigated using MATLAB Tool with the required number of tool boxes and blocksets.

Keywords Radio · Matlab · BER · PAPR

1 Introduction

Index modulation (IM) refers to a family of modulation techniques that rely on activation states for information integration of certain resources/building blocks. The resources/building blocks may either be physical or virtual, e.g. virtual parallel channels, signal constellation, space time matrix and antenna activation command, antenna constellations, subcontractors, time-based slots and frequency carriers. A separate characteristic of IM is the implied incorporation into the transmitted signal of part of the data.

The increase in mobile subscribers has contributed to an increase in data traffic, thus increasing the number of Base Stations (BSs), which meet customers' needs. This has led to an urgency to reduce the data traffic to improve the efficiency and accuracy of the modern-day telecommunication applications such as Video calling, internet services etc. The reference explains the rise in the number of BSs between 2007 and 2012 in development regions and forecasts that the number of BSs in this period will increase over 2 million. This survey period has made an impact on the development of infrastructures but had the limitations when it has come to the

C. H. Nagaraju (✉) · B. N. Raju
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, Andhra Pradesh, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_137

practical application and orientation due to various factors. In several past studies on this issue, both device capacity and data rates have been improved, although growing cellular energy network demand has been ignored. This the demand for energy has led to significant research into "green communications."

The OFDM is a multi-carrier digital modulation system that expands the idea of a single subcarrier by using multiple subcarrier modulation in the same single channel. OFDM uses a large number of closely spaced orthogonal subcarriers, which are distributed in parallel, rather than transmitting high-rate data streams with a single subcontractor A traditional digital modulation scheme (for example, QPSK, 16QAM, etc.) is available to each sub-contractor at a low symbol rate. However, the integration of several subcarriers makes it possible to reach data rates in comparable bandwidths close to traditional one-carrier modulations. The OFDM methods have been identified as the potential technologies in communications a way back but has been showing the fruitful results and outcomes during the recent times.

OFDM is based on the popular frequency multiplexing technique (FDM). Various information sources are mapped on different parallel frequency channels in FDM. In order to reduce interference among adjacent channels, each FDM channel is separated from the others by a frequency guard band.

In the following interrelated ways, the OFDM method varies from the conventional FDM:

1. The information stream is borne by several carriers (called sub-companies),
2. The sub-persons are mutually orthogonal and
3. To minimise the channel delay and interruption, a guard interval will be applied to each symbol.

Figure 1 shows the key concepts of an OFDM signal and the correlation between the frequency and time areas. Multiple neighbouring tones or subcontractors in the frequency domain are modulated with complex information separately. The Frequency Domain Subcarriers transform the Inverse FFT in order to generate the OFDM symbol in the time domain. Then, the ward intervals between the symbols are introduced in the time domain to avoid interference by multipath delay on the

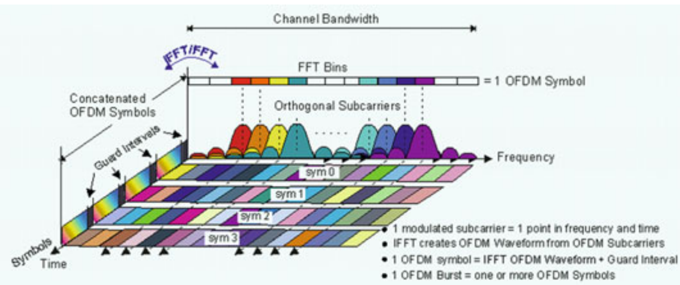


Fig. 1 OFDM channel frame

radio channel in the receiver. The final OFDM burst signal can be paired with several symbols. An OFDM symbol for retrieving the original data bits is given at the receiver.

2 Related Work

Guo et al. [1], the various methods of study have demonstrated substantial savings. In the signal transmission process, most savings are realised in some modelling conditions and assumptions. It is largely unexplored how the gains can be combined to achieve greater operational savings, especially in a practical multi-cell user setting. An integrated study of cross-layer technologies to minimise energy consumption or increase the spectral and energy quality tradeoff is used in this paper. This research is part of MVCE Green Radio's main integration process, integrating design, transmission techniques, resource management and hardware research.

Alsharif et al. [2] provides a brief overview of recent techniques for improving energy performance, such as energy-efficient power amplifier technologies, timestamping, cell switching, multi-input, multi-output (MIMO), cell zoom and relay techniques management, heterogeneous Micro-Pico-Femtocell-based Network architectures, etc. This paper further examines the advantages and drawbacks of each technique, thereby helping researchers learn more about how best to minimise energy usage in future green radio networks. In addition, it offers a concise description of each technique.

M. Wen, X, and L. Cheng. The EE of the OFDM-IM is evaluated from an information-theoretical point of view with finite alphabets input. On the basis of research, the optimal strategy is proposed to optimise the EE of OFDM-IM which involves the optimum grouping of subcontractors as well as the optimal selection of the number of active subcontractors. The analysis is validated by the Monte-Carlo simulations.

Zou et al. [5] examined the combined effects on the output of the mm-wave communication connection based on OFDM modulation on three major radio frequency impairments, namely oscillator phase noise, nonlinear power amplifier and I/Q imbalance. For each TX, RX and mm-wave communications link, general insufficiency models are first derived to explain joint effects. Based on the signal models and original air interface signals obtained from the mmMAGIC project, the channel-to-noise (CNR) and also channel fluency due to common phase error (CPE), induced by phase noise within channel coherence time, numerically assesses the impact of the RF impairments on the channel's estimate. The effect on connection performance is then assessed by detailed computer simulations in terms of the maximum sum rate. In the case of RF impairments, the simulation results show that the air interface architecture typically is robust. In the use of high-end modulation alphabets and the implementation in mm-wave communication of low power and low-cost RF transceivers, special emphasis should be put on phase noise when inter-carry interference (ICI) is the key limiting factor [3, 4].

Wetz et al. [6] presented an OFDM-based transmission system which is suitable for robust transmission in fast fading environments where it is impossible or very difficult to obtain accurate channel estimates. Our software is based on the combination of the MFSK and orthogonal frequency division multiplexing, which have not been reliably detected (OFDM). The inconsistent OFDM-MFSK detection permits the arbitrary selection of phases for all transmitter subcontractors. One alternative for this degree of freedom is to choose the subcontractor phases in a way that reduces the maximum to average power ratio (PARR). Another approach is to relay additional data using the subcarrier phases. The modulations for the subcarriers occupied by the OFDM-MFSK system will do this differentially. The robustness of the underlying OFDM-MFSK modulation is not affected by both possibilities. This is not consistent.

3 Methodology

The new concept of the sparse index modulation is clarified. The main design point resides in pursuing the upper bound of energy efficiency without losing the spectral efficiency. Based on complexity validation, sparsely indexing modulation (SIM) enables frequency indexing on the whole subcarrier domain without grouping. The block diagram of proposed method is given in Fig. 2.

To boost spectral efficiency, the combinatorial field is extended by the assumption that it is indexed in a virtual digital domain that is higher than the subsidiary domain. This is slightly assumed in the higher dimensional interactive digital domain. The CS mapping is then carried out into the lower-dimensional frequency domain. In this case, the true frequency domain is not sparsely present; in any group all subcarriers are allowed.

While the high energy efficiency is a driving force behind IM science, the SIM proposed to explore the limiting edge of the EE maximisation without the loss of the SE was added. Contrary to the IM, the proposed system of SIMs is based on the combinational mapping and unmapping of the entire number of subcomponents without grouping of analogous combinatorial states.

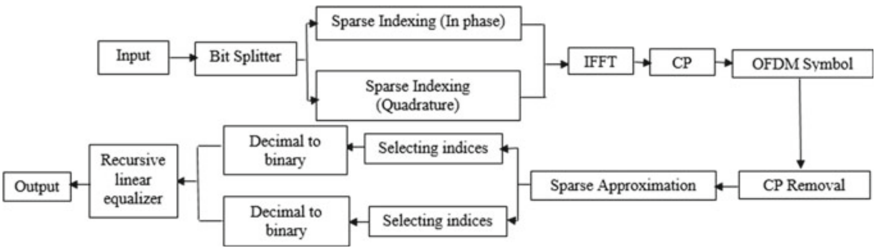


Fig. 2 Block diagram of proposed method

Therefore for increasing the spectrum efficiency and reducing the bit error rate using the recursive least squares linear equalizer.

Recursive least squares (RLS) is an adaptive filter algorithm that remedially identifies coefficients to minimise a weighted linear least square input signal cost function. This method is opposed to other algorithms, such as the LMS, aimed at minimising the mean square mistake. The input signals are regarded as deterministic for the derived RLS, while they are regarded as stochastic for the LMS and related algorithms. The RLS demonstrates extremely rapid convergence in comparison with most of its competitors.

- **Procedure**

1. The input bit stream is taken, then by using the bit splitter the random bit stream is splitted into different parts.
2. Then for this QAM modulation is applied with modulation order M = 512.
3. Then it is applied with IFFT followed by cyclic prefix.
4. Then the OFDM symbols are added to it.
5. Then this is transformed through the channel.
6. Then at the receiver side the cyclic prefix is removed.
7. Then the sparse approximation is done at the receiver side.
8. The using index modulation technique the indices are obtained.
9. Then recursive equalizer is used to obtain the data with better results.
10. Then finally the decimal data is converted into binary.
11. Finally the bit error rate and the peak average to power ratio is calculated.

4 Results and Discussion

Figure 3 is plotted between Amplitude and number of subscribers via simulation procedure. This graphical representation explains the sparsity in obtaining the sub carriers which are activated in the total number of obtained sub carriers.

The procedure explained in the above methodology is carried out and the output at the receiver is compared with the input given at the transmitter in order to calculate the peak average to power ratio. Figure 4 is obtained with the graphical representation between PAPR and Ao. This graph explains that the proposed method with recursive equalizer obtain a value of 10^0 which is low when compared with the existing method without equalizer.

The procedure explained in the above methodology is done and the output at the receiver is compared with the input given at the transmitter in order to calculate the bit error rate. Figure 5 is plotted between SNR and BER parameters and it depicts that the proposed method with recursive equalizer obtain a value less than 10^{-1} which is low when compared with the existing method without equalizer.

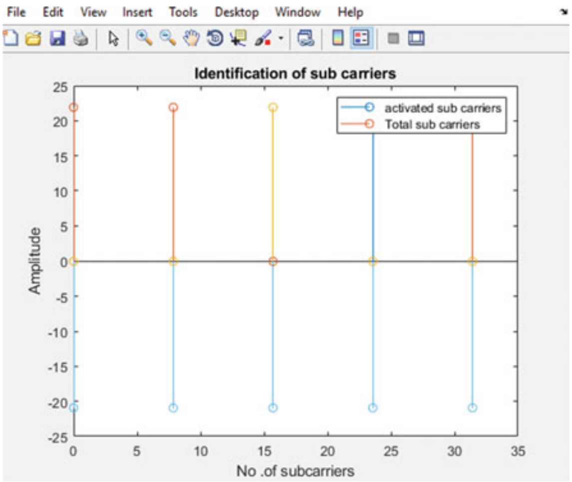


Fig. 3 Identification of subcarriers

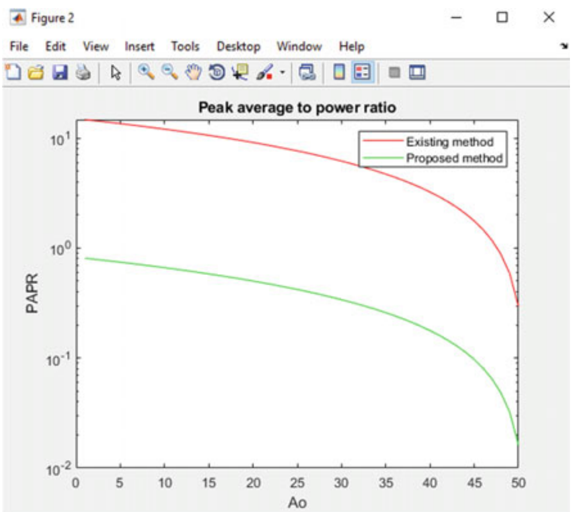
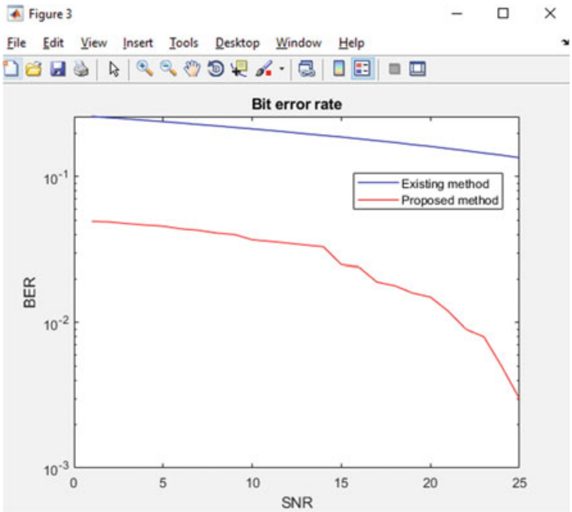


Fig. 4 Peak average to power ratio

5 Conclusion

It is concluded that the proposed algorithm reduces the bit error rate when compared to other existing methods by using the recursive least squares equalizers. This method has proven that the sparse index modulation produces better results with the recursive equaliser by depicting the various parameters by means of graphical representations. And it is also observed that the computational complexity of this equalizer is very less when implemented with the sparse index modulation.

Fig. 5 Bit error rate



References

1. Guo W, Wang S, Turyagyenda C, O'Farrell T (2012) Integrated cross layer energy savings in a smart and exible cellular network. In: Proceedings of 1st IEEE international conference on communications. China Workshops (ICCC), Aug 2012, pp 7984
2. Alsharif MH, Nordin R, Ismail M (2013) Survey of green radio communications networks: techniques and recent advances. J Comput Netw Commun 2013:113
3. Wu Y et al (2014) Green transmission technologies for balancing the energy efficiency and spectrum efficiency trade-off. In: IEEE communication magazine vol 52, no 11, pp 112120
4. Wen M, Cheng X, Yang L (2014) Optimizing the energy efficiency of OFDM with index modulation. In: Proceedings of IEEE international conference on communication system (ICCS), Nov 2014, pp 3135
5. Zou Y et al (2016) Impact of major RF impairments on mm-Wave communications using OFDM waveforms. In: Proceedings of IEEE Globecom workshops (GC Wkshps), Dec 2016, p 17
6. Wetz M, Peria I, Teich WG, Lindner J (2008) Robust transmission over fast fading channels on the basis of OFDM-MFSK. Wireless Pers Commun 47(1):113123

Intelligent Medical System for Automatic Medicine Recognition by a Novel Deep Learning Algorithm

Mannuru Srikanth Reddy and K. Naganarasaiah Goud

Abstract This project is based on intelligent medicine recognition using proposed deep learning algorithm. Proposed work is designed for assistance to chronic patients in taking all medicines on time and avoiding taking wrong medicines. This project also stores information of medication as well as chronic patient information. Depending on the patients input medicine the total application works with audio-visual alarm. We prepared MATLAB GUI (graphical user interface) which is easy interface with patients. In this we used 8 types of medicines which are recognized correctly by proposed algorithm.

Keywords Deep learning algorithm · Automatic medicine recognition · Chronicle patient · Audio-visual alarm

1 Introduction

Currently, the world's society is aging. Among the 7.5 billion people in the world, the elderly population accounts for 600 million, including 480 million people with chronic diseases. According to statistics from the World Health organization (WHO), the average elderly persons suffer from 1.4 chronic diseases, and the typical medication dosage of an elderly person is five times that of a younger person. Elderly people are also seven times more likely to take the wrong medicine because of declining physiological functions. The WHO also indicates that one-third of the world's deaths are caused not by diseases themselves but by the incorrect use of drugs, and the costs associated with such improper drug use amount to nearly 28.5 billion U.S. dollars every year.

Healthcare systems consist of patients, doctors, hospitals as well as research institutions. There are many terms which included in this field as diagnosis, treatment, monitoring and prevention with smart hospital management. The data will be plotted

M. S. Reddy (✉) · K. Naganarasaiah Goud
Department of ECE, Annamacharya Institute of Technology and Sciences (Autonomous), Rajampet, Kadapa, Andhra Pradesh, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_138

on thing Speak which is IOT for Matlab projects and which can be accessed privately or publicly.

There are many applications designed to make healthcare system even smarter than the existing healthcare systems. Many clinical researchers and scientific institutions are working on automatic decision-making systems which help doctors to avoid the laborious work and get intelligent system. The areas the researchers and scientists are working are as below,

1. Automatic Assistance for diagnosis as well as treatment
2. Management of Health using Android applications with sensors
3. Precaution for Disease prevention with its risk monitoring
4. Doctors virtual assistants to their patient
5. Making hospitals smarter
6. Assisting drug research.

Further in the topics it is mentioned about objectives of research work, literature survey, existing methods with the drawbacks, proposed work, result analysis, conclusion and references.

Objectives of the research is

- Read all existing techniques and their drawbacks for automatic medicine recognition system
- Using proposed deep learning concept design automatic medicine recognition system which helps chronicle patient to get their doses on time.
- Implement the proposed results using MATLAB software 2016a version and analyze final results using GUI (MATLAB app).

2 Literature Survey

'PillDrill-Smart medication tracking that simply works' shown at pilldrill.com [1], on the website author mentioned details of PillDrill product and how it works. There are 4 steps to get alarm for patients like plug in the hub, download the app, connect to wifi and schedule your meds. This hub mentions about which medicine is due. There are audio visual details for any due of medicines. There is need to scan pill container to track the medicines. It gives notification in real time by giving the chances of edit medication details [1].

A high-tech medication-tracking system by M. Brown shows that with the help of intelligent medication system its easy to remind the details of medicines for different types of patients. The author discussed PillDrill tool for elder person as well as small child with the help of prescription provided by the doctors. This tool used wifi model and set of some tags with the help of RFID we need to get details while taking tablets [2].

Learning objective of the author is to recognize health care functions and human activity related by ADRs (adverse Drug Reactions). Importance of the design of

ADRs and its reporting is mentioned in this application. This application also includes the details of prediction as well as prevention of the drug interactions using different methods and systems [3].

Existing Work and Drawbacks

To avoid the death because of improper drug intake there are many applications designed such as smart medicine box as PillDrill mentioned in [1, 2]. It helps patients to store, distribute as well as remind patients to take tablets on time. Most of the patients who are taking many tablets such as chronic disease patients.

As there are many tablets for some patients, some patients can't remind the tablets taken by them are correct or no as there are different variety of drugs. As it is somewhat complicated task there is much chance of misidentification of the drugs or medicine which causes to take wrong medicine for patient. Wrong medicine dose may cause complicated and serious offset. To avoid this type of problems we designed deep learning algorithm for detection of medicines intelligently by recognizing medicines correctly and reminding patient on time. From first use of medicine by the patient till completion of course, the probability of taking wrong medicine is calculated to avoid misuse of medicines.

3 Proposed Work

Scripting part uses deep learning toolboxes from MATLAB software for recognition of medicine from the 8 type of different medicines for chronicle patient. With Scanning QR code also get details of a medicine for chronicle patient. We are using Speaker for reminder for particular tablet for patient so that patient should not get confuse while selecting tablet at particular time as there are many tablets for chronicle patient. We prepared GUI (Graphical User Interface) to analyze the results in easy way or effective format (Fig. 1).

Proposed work consists of 4 main steps.

1. Selection of the proper input and training
2. Apply Deep Learning and Correct Identification of Medicines
3. Apply post processing after recognition of Medicines (Audio-Visual Alarms)
4. Display results on Thing Speak, Matlab IOT.

For this project we selected different 8 types of medicines for recognition using deep learning algorithm. We selected chronical medicines for this project as there is high dose for patients with chronical disease. As there is high dose for chronical patient there may be chances of forgetting the time to take tablet so there is need of such a system which help patients to remind the details of tablets and identification. Input we can take as an image or barcode.

For correct identification of medicines and the details provided by the doctors we used here deep learning algorithm. The input either image or barcode is processed by deep learning algorithm to identify the details of medicines. Deep learning uses CNN

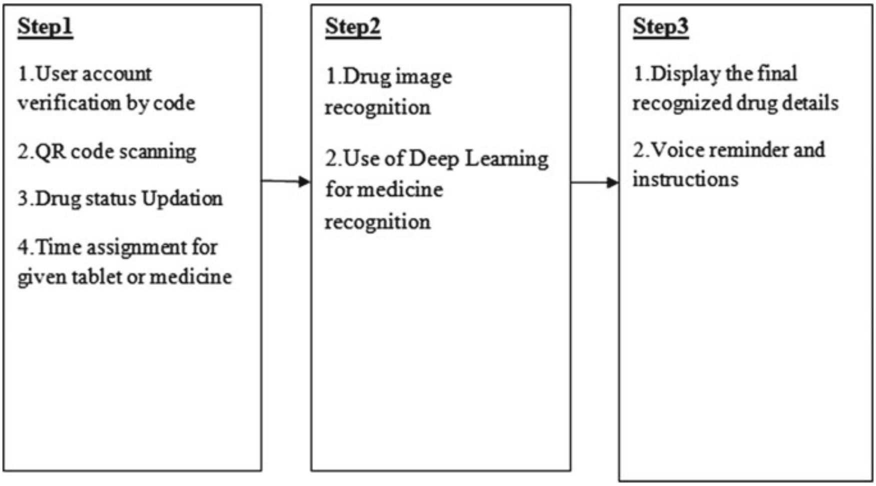


Fig. 1 Flowchart of the proposed work

(Convolutional Neural Network) for intelligent decisions. Deep learning is the type of machine learning algorithm which consist of supervised machine learning for this application. The use of CNN can be for image classification or image recognition. Also CNN wide uses in different fields because of its intelligent behavior. There are many hidden layers which helps algorithm to simplify and provide an intelligent output. Hidden layers such as convolution layer, ReLU layer, max pooling layer and Soft Max layers, etc. Convolution layer provide input image by applying filters, by padding the input. ReLU layer is used for activating the matrix. Max pooling layer is mostly used for dimension reduction. And the Soft Max layer is used for final stage of classification.

Audio-Visual alarms are used for reminding the chronic patient to take medicines on time without forgetting. Audio toolboxes are used for audio alarm and warning dialogue boxes are used for visual alarm.

Final results can be displayed on the ThingSpeak which is IOT for Matlab projects. As per the developer's information ThingSpeak is a open source IOT application for storage as well as to retrieve data from the HTTP as well as MQTT protocols over LAN and internet. We created private channel for plotting the data and retrieving information. We plotted the different charts for easy understanding and analysis by both patients and doctors.

4 Result Analysis

In the result analysis part, we designed a GUI (Graphical User Interface) for easy understanding by the patient as well as doctors (Fig. 2).

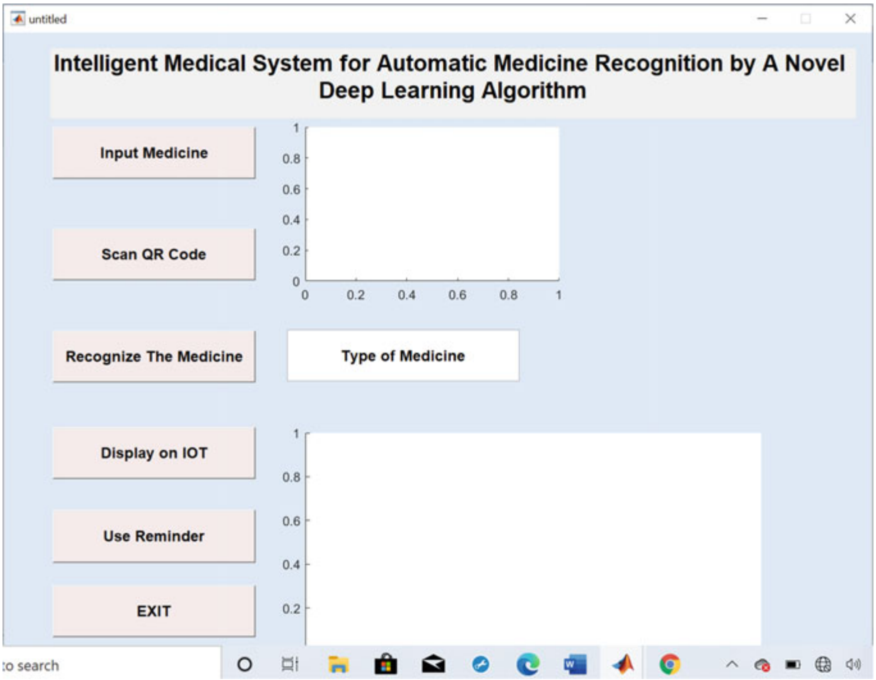


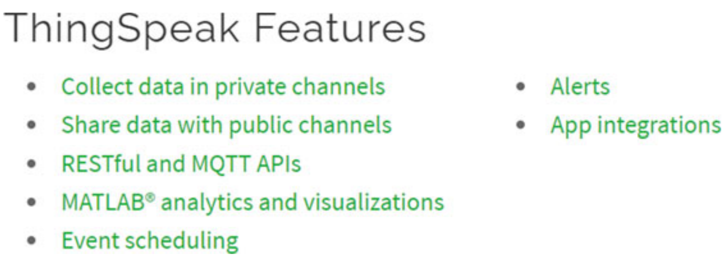
Fig. 2 GUI (Graphical User Interface) for proposed work

The GUI proposed for easy understanding of both patients and doctors. GUI consists of radio buttons, edit boxes, axes for plotting and static texts (Fig. 3).

Either we can import the input image or QR code for scanning for identification of the details of medicines. There are 8 different types of medicines we trained and can be identified using proposed algorithm (Fig. 4).

In this project we even generated QR code scanner to identify the type of medicine and warn to the patient more details (Fig. 5).

Applied deep learning algorithm for identification of medicine type and given for thingSpeak as well as for audio-visual alarms to make alert to patient for particular medicine at particular time.



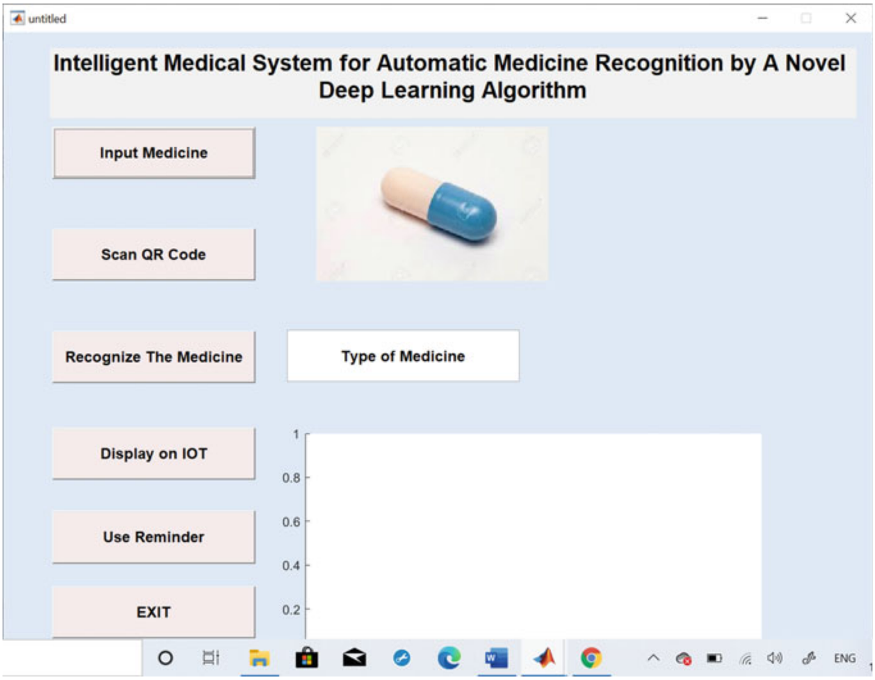


Fig. 3 Input medicine for identification

For IOT (Internet of Things) we used ThingSpeak to collect the data as well as analysis is done by Matlab (Fig. 6).

Data can be plotted on personal (private) channel like number of tables per time and some other information. Depending on it audio-visual alerts are provided (Fig. 7).

Alert box for patient which is indicating the tablet name which he or she has to take.

5 Conclusion

Chronic patients, including 480 million elderly people in the world today, suffer from a variety of diseases. In the treatment of multiple chronic diseases, many drugs are needed, and physiological functions decline. Cognitive ability is reduced, possibly causing patients to take the wrong medicine. Therefore, elderly people have become a high-risk group for adverse drug events. To solve the problem of taking the wrong medicine, in this paper, we have successfully developed an intelligent medicine recognition system. The proposed system can automatically provide notifications stating the names of drugs and indicating medication times to address the problem of lapses in human judgment.

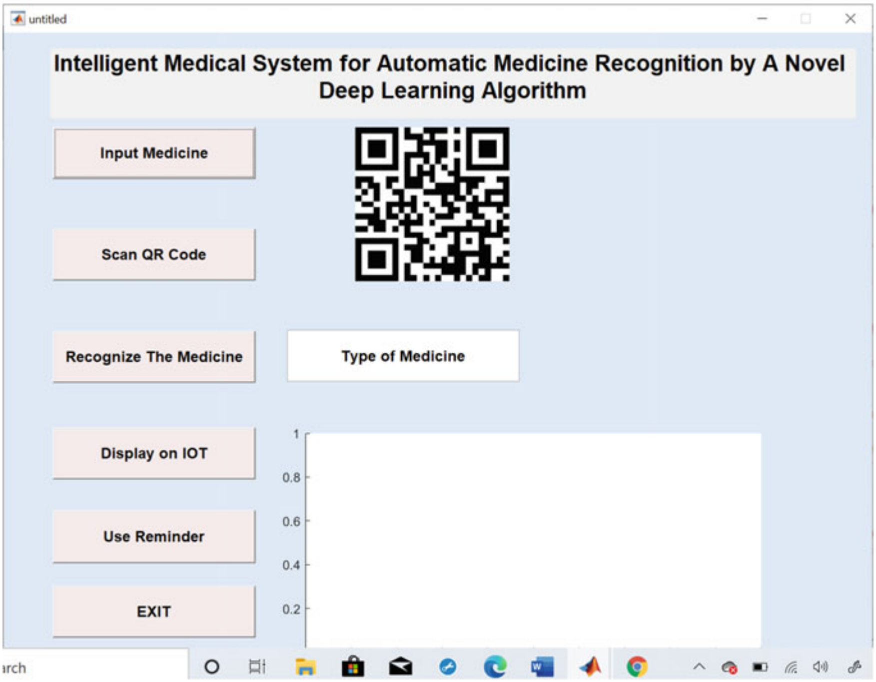


Fig. 4 Input QR code for identification

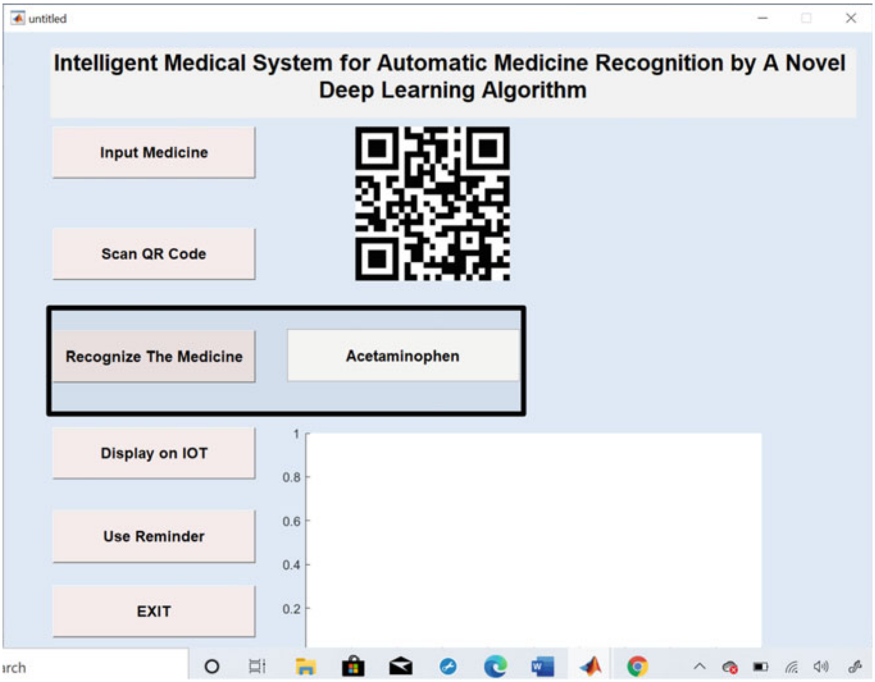


Fig. 5 Recognized medicine type by deep learning algorithm

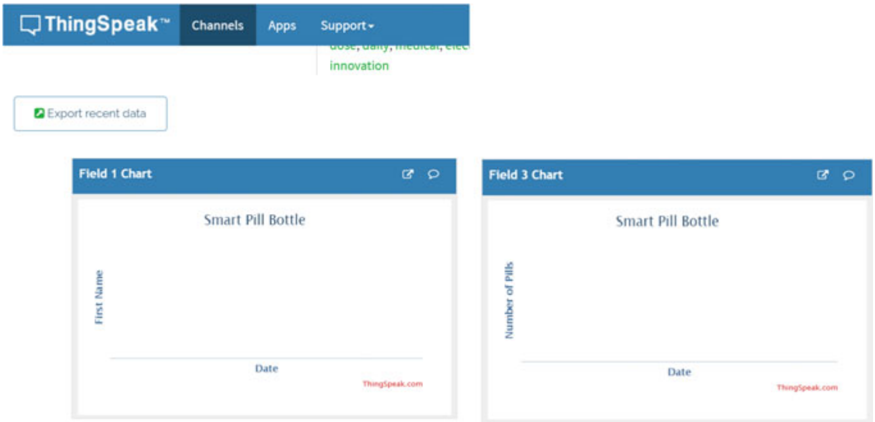
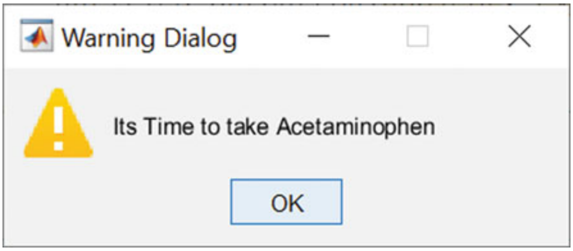


Fig. 6 Plotted figures on ThingSpeak to get details of patient name and the number of medicines per day

Fig. 7 Alert box for visual alert type



References

1. PillDrill-smart medication tracking that simply works (2018). Available https://www.pilldrill.com/
2. Brown M (2018) PillDrill review: a high-tech medication-tracking system. Available https://www.techhive.com/article/3270733/health-fitness/pilldrill-review.html
3. Dev S, Zhang S, Voyles J, Rao AS (2017) Automated classification of adverse events in pharmacovigilance. In: Proceedings of the 2017 IEEE international conference on bioinformatics and biomedicine (BIBM), pp 1562–1566

Novel ANN Based MPPT Control Strategy for Hybrid PV/Wind and Diesel Generator System

Jarugumalli Ramya Grahitha, M. Padma Lalitha, and Suresh Srinivasan

Abstract A novel single maximum power point tracking (MPPT) strategy for Hybrid PV and wind energy system (WES) to track the maximum power is proposed in this paper. The proposed MPPT technique constitutes of artificial neural network (ANN) based radial basis function network (RBFN) control strategy. The hybrid PV and WES is coupled to the grid through boost converter and voltage source inverter. Diesel generator is also inter-connected with grid to provide sufficient back-up during untoward condition. The proposed methodology uses a single MPPT control unit for both PV and WES simultaneously rather than individual MPPT techniques. The proposed MPPT strategy is validated for both standalone and grid connected system and the results are validated with single P&O method to validate its efficiency and computational speed. The proposed configuration is developed in MATLAB/Simulink platform with varying input conditions for 560 W PV system, 500 W wind system and 1 kW diesel generator system.

Keyword Hybrid system with ANN control strategy

1 Introduction

The increase in the usage of fossil fuel based generation of power has been increased tremendously in the recent year due to Industrialization and population. With the increase in the coal based power generation, the negative impact on the global environment is also increasing day by day. Thus, a reliable and effective alternative power generation techniques is highly recommended to meet the demand. Renewable energy sources (RES) based power generation technique provides a better alternative to thermal power generation techniques. RES techniques such as photo-voltaic (PV) based generation, wind energy conversion system (WECS), hydro power plant and ocean wave energy conversion techniques are most promising alternative energy [1].

J. R. Grahitha (✉) · M. Padma Lalitha · S. Srinivasan
Department of Electrical and Electronics Engineering, Annamacharya Institute of Technology and Science, Rajampet 516126, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_139

Among all, PV and WECS are widely preferable for its availability and sustainability [2].

The power generation from PV and WECS rely on the factors such as solar irradiance and wind velocity respectively. Both the parameters are highly non-linear in nature and causes uncertain voltage level in the distribution level [3]. Hence to overcome this, hybrid renewable energy system (HRES) based generation techniques are generally preferred. HRES provides better solution in terms of reliability. HRES are generally preferred for grid integration or with energy storage system because though they have reliable generation, chances of unpredictable extreme weather condition such as cyclone, heavy rain fall etc. may interrupt the power generation [4]. But energy storage system requires a high capital investment and regular maintenance. And also the energy storage system needs to be replaced once in a while which also increases the burden to the consumers. Thus, HRES are generally combined with a diesel generator (DG) when a critical load is associated with the grid. DG provides a better reliability to the system during high load demand and also during natural calamities [5].

The power generation from PV and WECS have a high fluctuation due to variable nature of irradiance and wind velocity. Also the power produced from the RES cannot be directly integrated to the grid. RES requires a power electronic converter interface and also an advanced control technique before integrating it with the grid. The best known technique to prevail over the intermittent nature of the RES and also to obtain maximum power peak (MPP) from varying condition is MPPT [6]. The MPPT strategy tracks the maximum power from available environmental parameter and delivers to the power electronic converter (PEC). They are also responsible for reducing the fluctuations of the output voltage.

There are numerous MPPT strategies present in the literature. Conventional technique, Soft-computing based technique, optimization based techniques like particle swarm optimization (PSO), flower pollination technique are few well known MPPT strategies which are present in the literature. Each technique has it own advantages and disadvantages [7]. The type of MPPT strategies are chosen based on the application and the availability of the data with the user.

P&O technique is the conventional technique with simple structure and easy implementation. P&O method are generally preferred for low rating application. It works on the principle of altering the firing pulse every cycle based on the obtained power. The P&O method observes the power and changes the firing angle value based on the previous cycle. The P&O methodology is the simplest MPPT strategy but fails to follow the rapid variations of solar irradiance and wind velocity. Thus to track the maximum power, an advanced control technique is required [8].

To overcome the above limitations an ANN based MPPT technique is implemented. The ANN technique is rapid and robust in nature. The ANN based advanced control technique is developed based biological neurons. Input layer, hidden layer and output layer are three primary layers of the ANN system. For renewable energy based application, input layer generally consists of voltage, current & parameters such as irradiance and wind speed and converters the above actual parameters to network parameters. The hidden layer consists of weight parameters where the weight elements of the network is analyzed. The output layer converts the network parameters to actual parameters such as duty cycle [9].

There exist numerous literatures about the individual MPPT technique for both PV and WECS. The individual controller in the hybrid system increases the complexity, volume and expenditure of the overall system. Hence to overcome the above problem, many authors attempted single MPPT technique for both PV and WECS [10]. Though the attempt was to enhance the controllability of HRES, it fails to track the MPP as both the sources totally differ in terms of input parameters availability. Hence to overcome this, a dedicated single MPPT technique is designed for both PV and WECS by considering the environmental parameters in this paper. The proposed MPPT strategy is based on the artificial neural network strategy and implemented for 560 and 500 W PV and WECS sources with boost converter as the PEC as shown in Fig. 1. MATLAB/Simulink platform is used to implement the proposed system and the results are validated using single P&O based MPPT technique for same hybrid system.

The organization of the paper is as follows, the modelling of the PV, WES, Diesel generator and boost converter is explained in Sect. 2. The proposed based MPPT techniques are detailed with the input parameters is explained in Sect. 3. Section 4 analyses the implementation of the topology in MATLAB/Simulink using the results obtained. Section 5 concludes the findings.

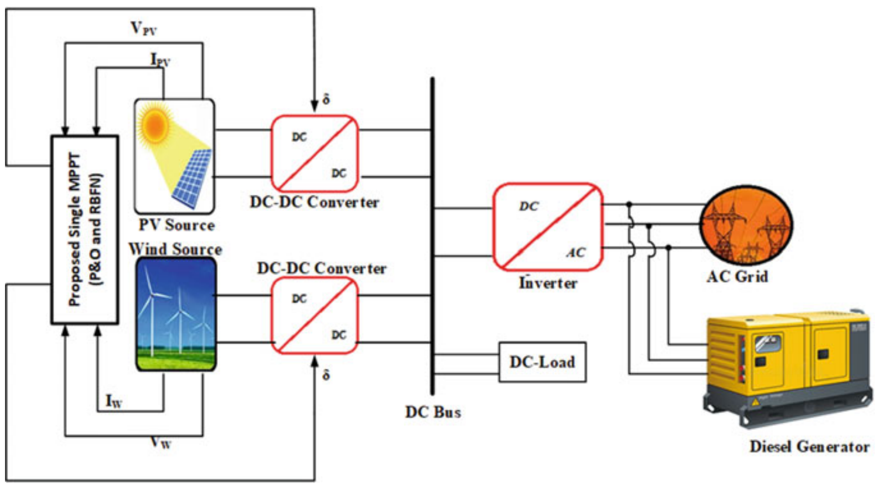


Fig. 1 Proposed hybrid PV/wind/diesel generator system with single MPPT strategy

2 Modelling of Hybrid System

2.1 PV System—Modelling

The main constraint of PV system is PV array which are arranged in various configuration to achieve the desired power. A PV system with single diode based model is illustrated in Fig. 2 [11].

The PV output relies on the specifications of the PV module. In proposed system, Kyocera KD135GX-LP PV module is considered to design a 400 W PV system. The detailed parameter of the PV system is represented in Table 1. The PV output relies on the solar irradiations and temperature. The modelling of the PV system in terms of I_{PV}–V_{PV} characteristics are as follows.

The output voltage and current of PV is given by (V_{opv}) and (I_{opv}) and it is expressed as follows,

$$V_{opv} = \frac{\eta K T}{q} \ln\left(\frac{I_{ph}}{I_{opv}} + 1\right) \tag{1}$$

$$I_{pv} = I_{ph} - I_{pvrsc}\left(e^{\frac{q(V_{opv}+I_{opv} R_{se})}{\eta K T}} - 1\right) - \frac{V_{pv} + I_{pv} R_{se}}{R_{sh}} \tag{2}$$

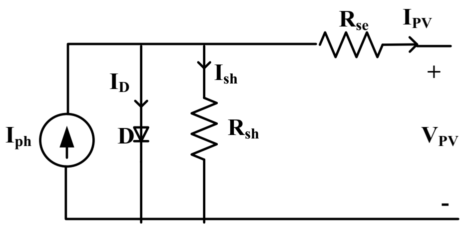


Fig. 2 Single diode PV cell—equivalent circuit

Table 1 PV System rating and specifications

PV parameters	Specifications at STC
Module type	Kyocera KD135GX-LP
Open circuit voltage, V_{oc}	22.099 V
Series string	1
Short circuit current, I_{sc}	8.369 A
Parallel strings	3
Resistance, R_{se}	0.105 Ω
Parallel resistance, R_p	142.84 Ω
Saturation current, I_{sat}	$9.845e^{-7}$ A
Phase current, I_{ph}	8.37 A

where, K is Boltzmann's constant (1.381×10^{-23} J/K), elementary charge (q) (1.603×10^{-19} C), shunt current as $I_{pvrsc} = (V_{pv} + IR_{Se})/R_{sh.}$ R_{se}: series resistance (Ω) R_{sh}: shunt resistance (Ω). I_{ph} represents Phase current (A) and T represents Temperature (°C). The modelling parameters of PV system are available in literature for different irradiance and temperature.

2.2 Wind Energy System—Modelling

Mechanical power of the WES (P_m) is expressed defined as,

$$P_m = \frac{1}{2}\pi \rho R_b^2 V_v^3 \tag{3}$$

where, ρ represents the air density, R_b represents wind blade radius, and V_v denotes wind velocity.

Power coefficient (C_p), the non-linear parameter of tip speed ratio and pitch angle β and is represented as [12],

$$C_p(\lambda, \beta) = 0.5176 \left(116\frac{1}{\lambda_i} - 0.4\beta - 5\right)^{-21\left(\frac{1}{\lambda_i}\right)} + 0.0068\lambda \tag{4}$$

Mechanical torque T_{mec} generated by the wind turbine is expressed as,

$$T_{mec} = \frac{\frac{1}{2}\pi \rho R^2 V_s^3 C_p(\beta, \lambda)}{\omega_r} \tag{5}$$

The mechanical torque (T_m) *and electrical torque* (T_e) *is represented as*,

$$T_m = \frac{P_m}{\omega_r} \tag{6}$$

$$T_e = \frac{P_e}{\omega_r n_p} * 2 \tag{7}$$

where, P_e is the electrical power. ω_r symbolize rotational speed and n_p denotes the pole pair used [12].

In order to develop a hybrid system, a 500 W wind system is designed by considering the parameters of Aeolos-H 500 W model. The detailed parameters are represented in Table 2.

Table 2 Wind system—parameters

Specification	Rating
Power (P)	500 W
Impedance (R_a)	0.699 Ω
Inductance (L_q and L_d)	6.99 mH
Magnetizing flux ($Ø_m$)	0.4023 wb
Coefficient (B)	0
Number of poles (P_P)	2
Torque per A (T/A)	1.1216 Nm/A
Cut-in wind speed (V_d)	4 m/s
Moment of inertia (J)	0.00126811 kg/m^2
Rated wind velocity (V_n)	12 m/s (26.8 mph)

2.3 Modelling of Diesel Generator

Diesel generator (DG) finds a better alternative to the battery management system in the remote locations. They provide required power during deficit from renewable sources. The DG consists of Engine—Generator set module where the fuel is converter to mechanical energy in engine and then the mechanical energy is converter to electrical energy in generator. A governor model is present to control the speed of the engine based on the fuel intake which keeps the engine speed to the desired level. The mechanical power of the diesel generator (P_{Dm}) is given as [13],

$$P_{Dm} = V_H \frac{\omega_{gm}}{\pi\ K} P_k \tag{8}$$

where, V_H is the cubic capacity of the engine,ω_{gm} represents generator rotor speed, P_k denotes effective power of the engine and K represents the characteristic parameter of the generator.

The torque of the engine (T_{De}) increases with the increase in the engine speed. After the engine speed reached to the specific point, the torque starts reducing to maintain the equilibrium. The engine torque of the DG (T_{De}) is given by,

$$T_{De} = \frac{P_{Dm\,k}}{\omega_m T_b} \tag{9}$$

The detailed specifications of diesel generator implemented is listed in Table 3. The DG is considered to be of 1 kW which can provide a complete power during deficit.

The droop control is generally used in the DG model to transfer the power based on the load and source demand. The droop control is a model of automatic voltage restorer (AVR) where the rotor angle of the generator is controlled to normalize the voltage of the generator system.

Table 3 Parameter specifications of diesel generator

Description	Rating
Power (P)	1 kVA
Number of cylinder	4
Displacement	4.87
Rated speed	1500 rpm
Governor type	Mechanical
Fuel consumption at 100% load	23.6 LPH
Rated voltage	230–380 V
Frequency	50 Hz
Alternator type	Permanent magnet synchronous
Number of poles	4

2.4 Modelling of Boost Converter

The boost converter (BC) is employed to integrate the HRES to the grid. The BC is implemented to improve the voltage to the required level. BC utilises single switch for conduction. The primary structure of the BC constitutes of an inductor, a capacitor, diode and switching device for the switch as illustrated in Fig. 3. When the switch is in ON state, the diode is reversed biased and the current starts increasing and delivers to load through inductor and switch. When the switch is in OFF state, the diode becomes forward biased and the current starts discharging from the inductor and reaches load through the capacitor. The modelling of BC are as follows [14],

$$V_o = \frac{V_i}{1 - D} \tag{10}$$

where, V_o is the output voltage and V_i represents input voltage respectively. D is the duty ratio.

The L and C parameters of the BC is analysed by,

$$L = \frac{V_{i\,\min}}{\Delta I_L * f_s} * D \tag{11}$$

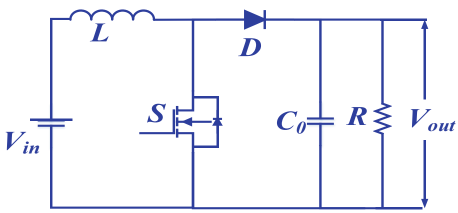


Fig. 3 Configuration of boost converter

$$C_s = \frac{I_o * D}{\Delta V_{cs} * f_s} \tag{12}$$

The boost converter is simple in designing and the implementation of the control strategy is also simple. Thus, it is widely preferred in renewable energy based application.

3 Single MPPT Technique

The obtaining of maximum power from highly non-linear renewable PV and WECS are the major challenges faced by the researchers. To overcome this many MPPT technique is proposed in the literature. Among all, P&O strategy is most traditional way to obtain the maximum power. But during rapid changes in the wind velocity it fails to follow the maximum path hence to rise above this, an advanced ANN—MPPT technique is implemented in this paper.

4 Proposed Modified Single P&O MPPT Algorithm

A single P&O MPPT control algorithm is proposed for the HRES. Generally, an individual MPPT technique is required for both PV and WES resources to obtain the peak power which in turn increases the complexity of overall configuration. Hence, to overcome the problem, two individual MPPT control techniques are fused together to form a single P&O control methodology. The structural outline of the proposed single P&O MPPT control technique is illustrated in Fig. 4.The input parameters are V_{PV}, V_W, I_{PV}, and I_W. In proposed methodology, the voltages and currents of the sources are stored in the single voltage variable (V) and current variable (I) in the form of a matrix. In this method, voltage is considered as perturbing element and the power is calculated by analyzing the value of sensed voltage (V) and current (I) value from the RES [15].

The change in voltage (ΔV) and change in power (ΔP) are calculated by evaluating it with the existing values. If ΔP is positive, then it verifies the ΔV value if it is also positive then it decreases the duty cycle (D) value with fixed step size (ΔD) and if ΔV is negative, then it increases D with ΔD. Similarly, if ΔP is negative, then it verifies the ΔV value if it is positive then it increases the D value with ΔD and if ΔV is negative, then it decreases the D with ΔD and repeats the procedure until it reaches the maximum power point (MPP) and gives the consequent duty ratios for each source.

The PV output power is expressed as,

$$P_{OPV} = V_{OPV} * I_{OPV} \tag{13}$$

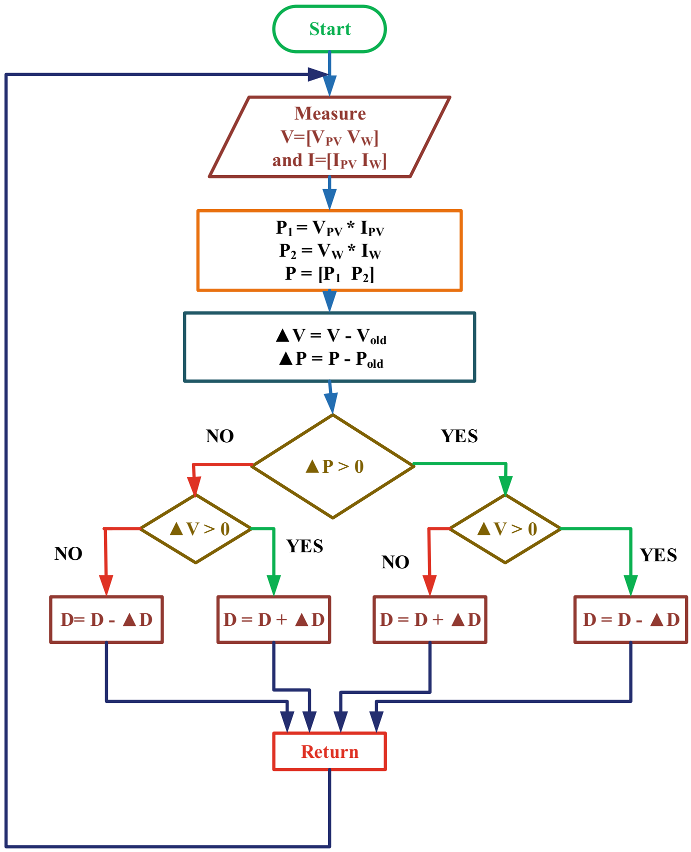


Fig. 4 Flow chart of single P&O MPPT for HRES

The relation between DC link voltage (V_{dc}) and PV output voltages is given in Eq. (14)

$$V_{PV} = \left(\frac{1}{1-D}\right)V_{dc} \tag{14}$$

Similarly, the WECS is computed from Eq. (15) as

$$P_W = V_W * I_W \tag{15}$$

The V_{dc} and wind output voltages is given in Eq. (16)

$$V_W = \left(\frac{1}{1-D}\right)V_{dc} \tag{16}$$

The step by step procedure to obtain the peak power by using the proposed single P&O MPPT control strategy from the renewable energy sources is presented.

Step-1: Start
Step-2: Sense Voltages V_{PV}, V_W and Currents I_{PV}, I_W for initial conditions
Step-3: Stores Voltages in $[V] = [V_{PV}\ V_W]$ and current in $[I] = [I_{PV}\ I_W]$
Step-4: Calculates powers $P_{PV} = V_{PV} * I_{PV}$ and $P_W = V_W * I_W$ and stores in $[P] = [P_{PV}\ P_W]$
Step-5: Calculates vary in power ΔP and vary in voltage ΔV
For $i=k$ (Maximum number of iterations)
If change in power, ΔP is positive;
If vary in voltage, ΔV is positive;
Decreases the duty ratio, $D = D - \Delta D$
Else if vary in voltage, ΔV is negative;
Increase the duty ratio, $D = D + \Delta D$
End
Else
If vary in voltage, ΔV is positive;
Increase the duty ratio, $D = D + \Delta D$
Else if vary in voltage, ΔV is negative;
Decreases the duty ratio, $D = D - \Delta D$
End
Step-6: Update $V(k-1)=V(k)$; $I(k-1)=I(k)$
Step-7: End
Step-8: Stop the process

PROPOSED RBFN BASED INTELLIGENT SINGLE MPPT ALGORITHM

The P&O technique fails to trace the intermittent nature of the sources. The P&O requires high computational time to obtain the maximum power. Thus to overcome that, a single intelligent MPPT using ANN technique is projected to track the peak power from the HRES. An RBFN based network is designed by considering four input variables to the input layer (V_{PV}, I_{PV}, V_W, and I_W) same as P&O method and 529 hidden neurons in the hidden layer and at output layer two different firing angle for two different sources exists as shown in Fig. 5.

The design of RBFN system is depicted in Fig. 6. The functioning of RBFN relies on connection pattern, weights and activation function of the network. The configuration involves an individual MPPT strategy for each source, which augments the overall complexity of system [16]. Hence, two MPPT control system are used to form RBFN based single MPPT strategy. The RBFN network is trained using gradient descent (GD) learning algorithm. The V_{PV}, V_W, $I_{PV,}$ and I_W are considered as input variable for RBFN network.

The single RBFN based MPPT controller undergoes three steps to generate duty cycle for the renewable energy sources simultaneously.

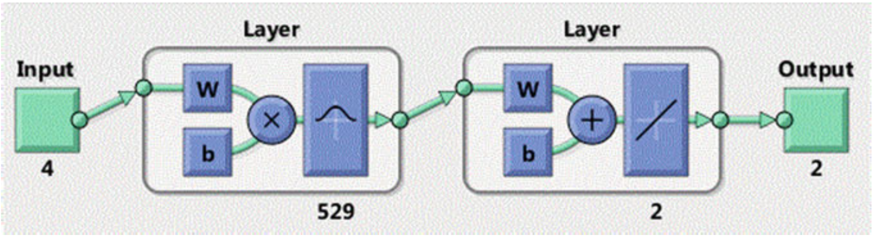


Fig. 5 RBFN based single MPPT architecture for a hybrid system

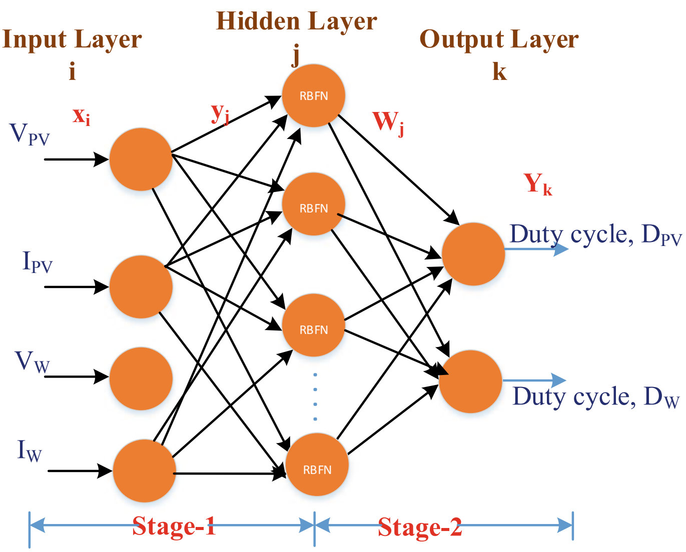


Fig. 6 Structure of three layer RBFN based single MPPT strategy

RBFN based single MPPT controller basic nodal equations.

Layer-1: Input layer

The nodes in the input layer are used to read the input data (V_{PV}, I_{PV}, V_W, and I_W) and directly transmit to the next neuron layer. The variable parameters of the first layer are represented in Eqs. (17) and (18).

$$x_i^{(I)}(n) = net_i^{(I)} \tag{17}$$

$$y_i^{(I)}(n) = f_i^{(I)}\left(net_i^{I}(n)\right) = net_i^{(I)}(n),\ i = 1, 2, 3, 4 \tag{18}$$

Layer-2: Hidden layer

Each node in the hiddenlayer performs a Gaussian membership function the I variable parameters of the second layer is represented in Eqs. 19 and 20.

$$net_j^{(H)}(n) = -\left(X - M_j\right)^T \sum_j \left(X - M_j\right) \tag{19}$$

$$\begin{aligned} y_j^{(H)}(n) &= f_j^{(H)}\left(net_j^{(H)}(n)\right) \\ &= \exp\left(net_j^{(H)}(n)\right), \ j = 1, 2, 3, \ldots\ldots \end{aligned} \tag{20}$$

Layer-3: Output layer

Two neurons are present in the third layer, which composes overall output of the controller and it generates two different duty cycles D_{pv} and D_w for PV and WECS, The input and output parameter in third layer as follows

$$net_k^{(O)}(n) = \sum_j w_j y_j^{(H)}(n) \tag{21}$$

$$y_k^{(O)}(n) = f_k^{(O)}\left(net_k^{(O)}(n)\right) = net_k^{(O)}(n) \tag{22}$$

where, $net_i^{(I)}(n)$ is the sum of the input layer, $net_i^{(H)}(n)$ is sum of the hidden layer, $net_i^{(O)}(n)$ is the sum of the output layer, W_j is the weights between hidden and output layer, M_j is the Mean and standard deviation of the output layer.

The single RBFN system constitutes four input variable such as PV voltage and current to manage PV and Wind voltage and current to manage the WECS. The corresponding duty ratio for each system is obtained based on the input PV irradiance and wind speed.

The parameter configuration chosen for the RBFN based intelligent single MPPT technique are given in Table 4.

Table 4 RBFN parameter for single MPPT controller

Parameters	Value/method
Input variable	V_{PV}, I_{PV}, V_W, I_W
Output variable	Duty cycles D_{PV} and D_W
Hidden neurons (max. limit)	2500
Training algorithm	OLS
Spread factor	0.03

5 Results and Discussion

To validate the proposed hybrid system in grid connected mode, a Simulink model of 560 W PV system, 500 W wind system and 1 kW diesel generator with BC and single MPPT strategy are implemented in MATLAB with AC grid of rating 440 V, 50 Hz as shown in Fig. 7. To validate the efficiency of the proposed controller, the system is validated in both standalone and grid connected system. In standalone system, the diesel generator is not connected to the system and the input sources PV and Wind is connected to the DC micro grid. The diesel generator is connected for grid based application purpose.

Standalone System
To validate the effectiveness of the proposed controller an different irradiance and wind velocity is considered in this study as shown in Figs. 8 and 9 respectively for PV and Wind system.

The performance of the HRES with BC along with single P&O based MPPT topology is depicted in Fig. 10. The DC link voltage, current, and power obtained using single P&O based MPPT technique employed for the hybrid system is depicted

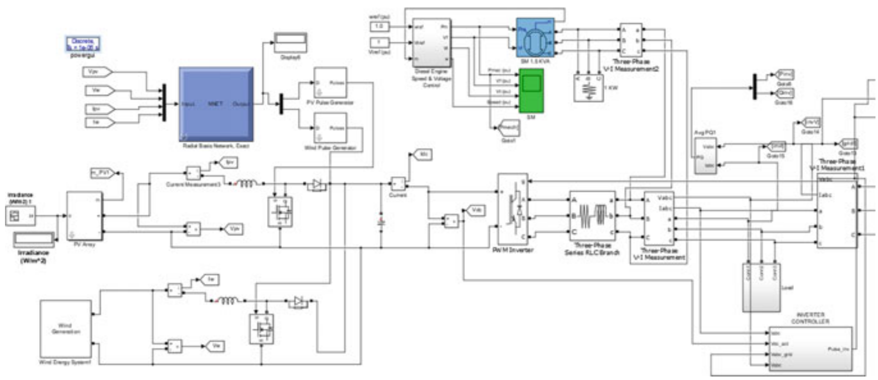
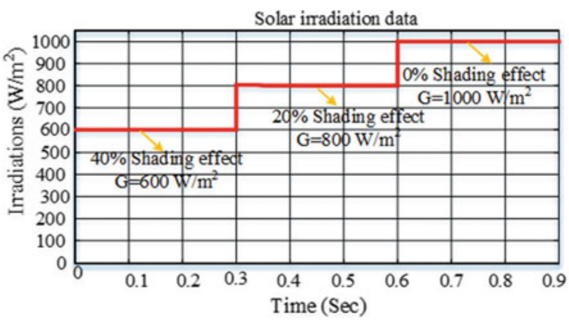


Fig. 7 Proposed hybrid system configuration in MATLAB/Simulink

Fig. 8 PV irradiance—input pattern



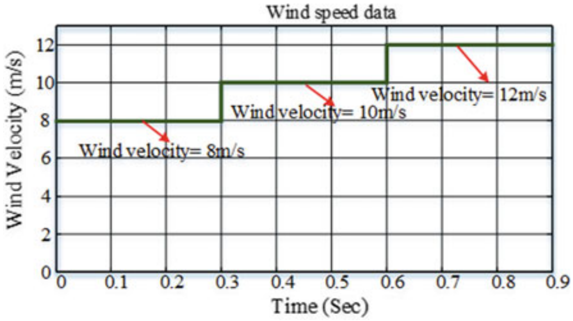


Fig. 9 Wind velocity—input pattern

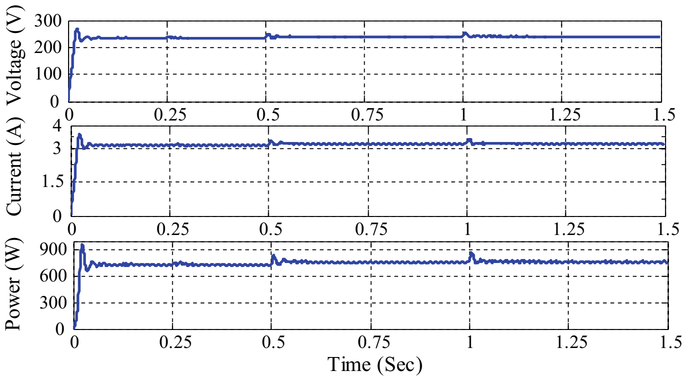


Fig. 10 Hybrid system with boost converter DC link output voltage, current and power with single P&O MPPT

in Fig. 10. The P&O technique fails to track the peak power obtained from the solar and wind due to the highly variable characteristic of PV irradiance and Wind speed.

The performance of the hybrid structure with Boost converter along with single RBFN based MPPT topology is shown in Fig. 11. The DC link voltage, current and power obtained using single RBFN based MPPT technique employed for the HRES is shown in Fig. 11. The RBFN MPPT technique not only tracks the maximum power but also provides better stability during the parameter changes of irradiance and wind speed. A constant DC link voltage is obtained because of faster convergence speed of RBFN based MPPT technique. The DC link voltage is the primary necessity of micro grid for an efficient performance.

Table 5 show the summary of the proposed HRES with Boost converter and modified single MPPT controller techniques for varying solar irradiance and wind velocity. For period 0–0.5 s, with 600 W/m^2 and 12 m/s input data, the developed power is 712.3 W in P&O MPPT method and 716.8 W in RBFN based MPPT method. Similarly, for a period of 0.5–1 s the developed power is 736 W in P&O MPPT method and 742.4 W in RBFN based MPPT method and 1 to 1.5 s, the developed power is

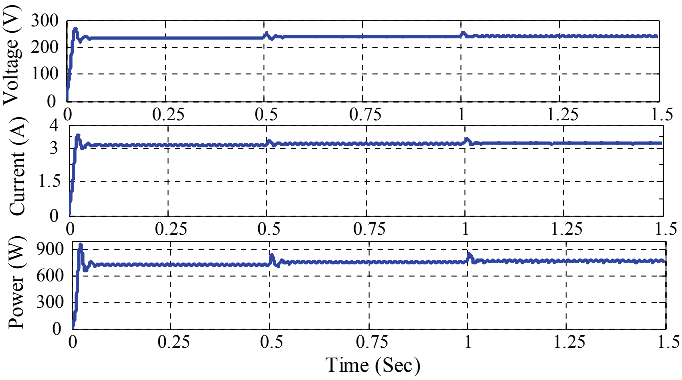


Fig. 11 Hybrid system with boost converter DC link output voltage, current and power with single intelligent RBFN MPPT

Table 5 The summary of the simulated hybrid system with boost converter output power

HRES output power interconnected with boost converter			
Period (s)	0–0.5	0.5–1	1–1.5
Solar irradiations (W/m^2)	600	800	1000
Wind velocity (m/s)	12	10	8
Single P&O MPPT controller (W)	712.3	736.0	758.8
Single RBFN based MPPT controller (W)	716.8	742.4	764.8

758.8 W in P&O MPPT method and 764.8 W in RBFN based MPPT method. From the Table 5, it is clear that the proposed single RBFN based MPPT controllers give the better results compared to the single P&O MPPT controller.

Grid Connected System

To test the performance of the proposed configuration for grid based application and also to validate its performance with diesel generator the hybrid PV/Wind/Diesel generator are connected as shown in Fig. 7. The benefit of integration of HRES with the diesel generator increases the compatibility, peak hour management, and reduces the impact of power outages.

The common DC link voltage obtained from the three different regions based on the input source data is fed to the 3-ϕ VSI and its control methodology is shown in Fig. 12 by using sinusoidal pulse width modulation technique pulses are generated to control the voltage source inverter.

Figures 13, 14, 15 and 16 represents the Inverter voltage—current, load voltage—current and grid voltage—current.

Figure 16 shows the power of the proposed hybrid system, under three different load conditions, based on the inverter output power, the load demand is met by sharing

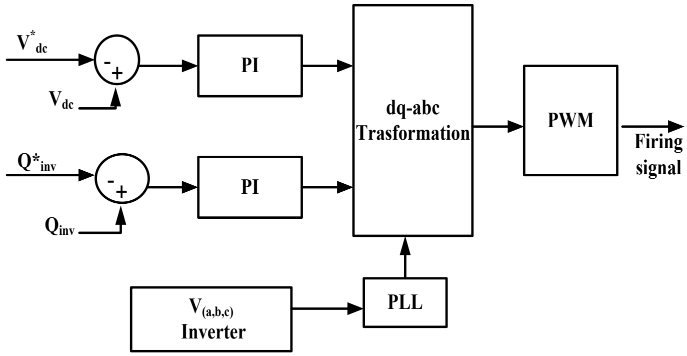


Fig. 12 Grid side current regulated controller

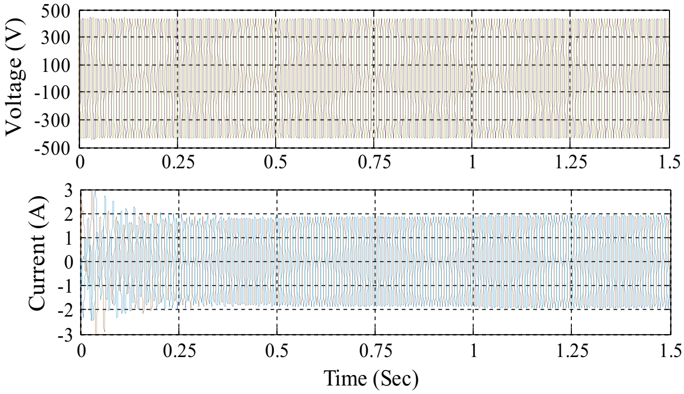


Fig. 13 Hybrid system with boost converter inverter output voltage and current

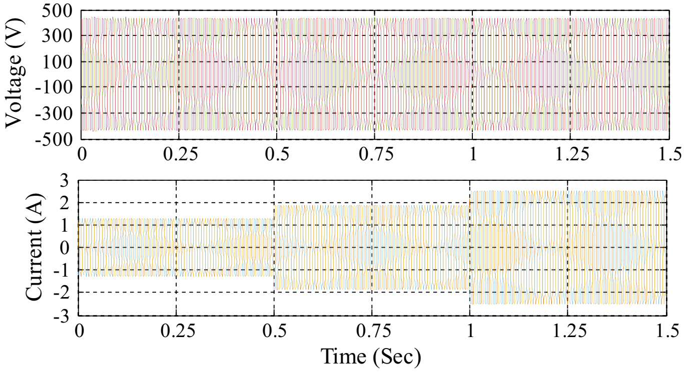


Fig. 14 Hybrid system with boost converter load voltage and current

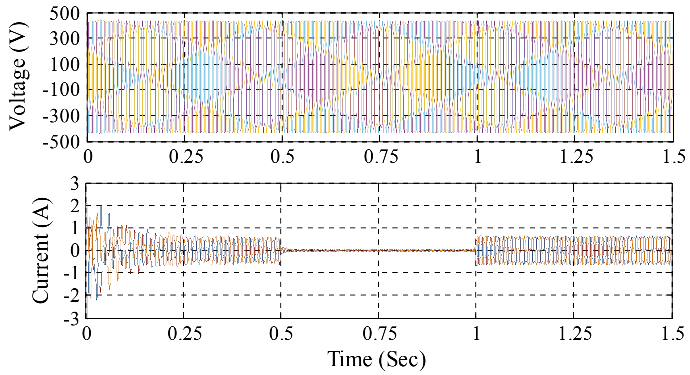


Fig. 15 Hybrid system with boost converter grid voltage and current

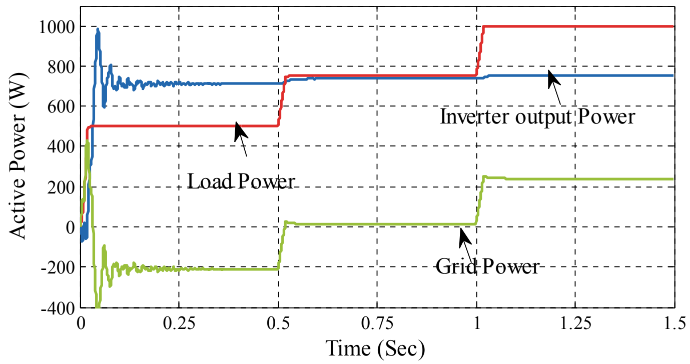


Fig. 16 Hybrid system with boost converter load, inverter, and grid active power profile

with the grid. During the period 0–0.5 s, the net inverter output power is 754.1 W, for the same period load demand is 500 W, and then the excessive 254.1 W power is fed to the grid. Similarly for period 0.5 to 1 s, the load demand is 750 W but the generated power is 780.4 W, then the excessive 30.4 W power is injected the grid and for the period 1–1.5 s, the load demand is 1000 W, for this period generated power is 802.2 W, thus to compensate the load demand the remaining 197.8 W power is

Table 6 Summary grid connected HRES with boost converter active power

Grid connected HRES with boost converter active power (W)			
Period (s)	0–0.5	0.5–1	1–1.5
Inverter output (W)	715.2	739.4	761.9
Diesel Generator (W)	0	10.6	238.1
Load (W)	500	750	1000

obtained from diesel generator. Table 6 depict the results obtained from the proposed configuration.

6 Conclusion

In this paper, a hybrid PV, wind and diesel generator system was designed along with the proposed single P&O and single RBFN based MPPT controllers. HRES is designed by for 560 W PV system and 500 W wind system. The performance of the developed hybrid system with boost converter and single MPPT is analyzed for grid connected modes under three different load conditions with three different regions by considering the different renewable energy source data with respective time. The simulation results are presented to validate the developed system output. The developed HRES with single RBFN based MPPT gives the average power 715.2 W in first region with wind 12 m/s and PV 600 W/m^2, similarly, 739.4 W in second region with wind 10 m/s and PV 800 W/m^2 and 761.9 W in third region with wind 8 m/s and PV 1000 W/m^2. The balance powers in second and third region are supported using diesel generator. The obtained results reveals that proposed RBFN MPPT strategy, gives the most excellent result. To test the grid capability of the proposed configuration, the active powers profile is analyzed in this paper under three different load conditions. The modified RBFN based single MPPT efficiently tracks the MPP from both sources simultaneously, which reduces the hybrid system implementation difficulty.

References

1. Liu X, Loh PC, Wang P et al (2013) A direct power conversion topology for grid integration of hybrid AC/DC energy resources. IEEE Trans Ind Electron 60:5696–5707
2. Cumulative development of various renewable energy system/devices in country. Retrieved from http://mnre.gov.in/mission-and-vision-2/achievements. Accessed 30 June 2017
3. Chien LJ, Chen CC, Chen JF et al (2014) Novel three-port converter with high-voltage gain. IEEE Trans Power Electron 29:4693–4703
4. Nejabatkhah F, Danyali S, Hosseini SH et al (2012) Modelling and control of a new three-input DC–DC boost converter for hybrid PV/FC/battery power system. IEEE Trans Power Electron 27:2309–2324
5. Wai RJ, Lin CY, Liaw JJ et al (2011) Newly designed ZVS multi-input converter. IEEE Trans Ind Electron 58:555–566
6. Kumar K, Babu NR, Prabhu KR (2017) Design and analysis of an integrated Cuk-SEPIC converter with MPPT for standalone wind/PV hybrid system. Int J Renew Energy Res 7:96–106
7. Fathabadi H (2016) Novel highly accurate universal maximum power point tracker for maximum power extraction from hybrid fuel cell/photovoltaic/wind power generation systems. Energy 116:402–416
8. Fathabadi H (2017) Novel fast and high accuracy maximum power point tracking method for hybrid photovoltaic/fuel cell energy conversion systems. Renew Energy 106:232–242
9. Fathabadi H (2016) Novel high-efficient unified maximum power point tracking controller for hybrid fuel cell/wind systems. Appl Energy 183:1498–1510

10. Luo FL (1999) Positive output Luo converters: voltage lift technique. IEE Proc-Electr Power Appl 146:415–432
11. Tiwari R, Babu NR (2016) Recent developments of control strategies for wind energy conversion system. Renew Sustain Energy Rev 66:268–285
12. Hong CM, Ou TC, Lu KH (2013) Development of intelligent MPPT (maximum power point tracking) control for a grid-connected hybrid power generation system. Energy 50:270–279
13. Tiwari R, Babu NR (2016) Fuzzy logic based MPPT for permanent magnet synchronous generator in wind energy conversion system. IFAC-Papers OnLine 49:462–467
14. Wu G, Xinbo R, Zhihong Y (2015) Non isolated high step-up dc–dc converters adopting switched-capacitor cell. IEEE Trans Ind Electron 62:383–393
15. Akar F (2016) A bidirectional nonisolated multi-input DC–DC converter for hybrid energy storage systems in electric vehicles. IEEE Trans Veh Technol 65:7944–7955
16. Tiwari R, Padmanaban S, Neelakandan RB (2017) Coordinated control strategies for a permanent magnet synchronous generator based wind energy conversion system. Energies 10:1493–1507

Keen Energy Efficient Street Light Controlling Framework Dependent on IoT for Smart Campus MRRITS

M. Malapati, Nagadasari Merrin Prasanna, Naveen Akula, Ravi Policherla, Santhosh Kumari Chejarla, and Dharani Lekkala

Abstract Automatic Electrification of Street Lights can eliminate different restrictions and difficulties identified with traditional (manual) lightning system. A common thing that can be observed in both urban and rural areas is the street lights that glow not only during night but also all day long. The cumulative amount of power that is wasted by all the lights due to improper switching is a lot. Also, in the present world of automation, manual switching of lights (Deng Huaqiu, Cao Yong, Xue Huaqiang, "An Intelligent Control of LED Street Lamp Based on Illumination Sensing" (Merrin Prasanna and Polaiah in Int J Control Autom 13(2):416–427, 2020 [1]), 2015 Fifth International Conference on Communication Systems and Network Technologies (CSNT)) seems to be outdated. Our idea is to build a circuit that can detect whether it is Day/Night by sensing the sunlight and turn the lights ON during day time, OFF during night time automatically without any human intervention. Besides automatic switching we have included a WiFi switch with which the lights can be controlled remotely. This device doesn't need any modifications in the existing system, it is just connected in series between the power source and the lights.

Keywords Sunlight sensing · Automatic switching of lights · Remote control

1 Introduction

This project provides solution for the downsides in the existing, i.e., improper manual switching of street lights. In most of the semi urban areas, there is no centralized control of street lights. A person is assigned to switch the lights ON during night and

M. Malapati
Mekapati Rajamohan Reddy Institute of Technology and Science, Udayagiri, A.P., India

N. Merrin Prasanna (✉)
Department of ECE, Mekapati Rajamohan Reddy Institute of Technology and Science, Udayagiri, A.P., India

N. Akula · R. Policherla · S. K. Chejarla · D. Lekkala
Department of Electronics and Communication Engineering, Mekapati Rajamohan Reddy Institute of Technology and Science, Udayagiri, A.P., India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_140

OFF during day by physically visiting each area. If that person is unable to switch the lights at required time due to any reason, a lot of electricity is being is wasted at day and there'll be no light on the streets during night. And the worst situation is that the lights of few remote areas in a town are directly connected to the supply lines and are left out to glow all day long. Due to this existing system hundreds of megawatts of electricity is wasted on a daily basis, which we should avoid right now. Few studies reveal that, around 30–40% of a city's energy consumption is going for the street lights. It can also be noted that the sun rise and sun set timings are not constant the whole year. During summer the day is long and the night is short. On the other hand, during winter the day is short and the night is long. As the system which we are proposing controls the lights [2] depending on the amount of sunlight present in the particular area, the switching will be a lot advantageous and the wastage of electricity would be very less [3]. This method of automating the street lights can be highly useful and reliable as the manual switching of street lights is proved to be inefficient. Not only reducing the wastage of power, this in turn will reduce crores of rupees of public money spent for the wasted electric power. The life of the bulbs is also being reduced due to the continuous glowing of lights. The replacement of the damaged lights with new ones costs unnecessary money for the municipalities. By installing this device, many such undesirable wastages can be completely Eradicated. Also, the components which are used in building this circuit such as Light Dependent Resistor (LDR), Transistor, Diode, etc. are capable of withstanding harsh climatic conditions [4], so there would be no additional costs for maintaining the devices often.

2 Aim and Objective

In the existing system of manual switching of street lights, it is highly evident that huge amount of electricity and public resources are being wasted. The aim of this system is to stop this wastage by bringing automation in the manual system. Prevent the restrictions and setbacks in existing system. Bring automation to the existing manual control of street lights which hasn't changed since ages. The below stated are the objectives of smart outdoor lighting system. The device would be to detect the amount of sunlight present accurately. The lights will be turned ON when there's no sunlight and turned OFF if there's enough sunlight. The lights will be controlled remotely through IOT [5] if necessary and this is optional. There should not be any human requirement in the whole process of switching unless there's a requirement to be controlled through IOT. The cost of building the device should be low so that it is cost efficient.

3 Problem Statement

The street lights are an integral part of any rural or urban area to provide visibility and safety to the people and these should not be improperly maintained. Many studies reveal that the cost of maintaining the street lights [6] is way more than what it is supposed to be. The increase in maintenance cost is only due to negligence in manual switching and this should be changed right away. Due to improper switching, the life time of the lights is reduced and they stop glowing sooner. Until the concerned authorities replace them with new lights, the streets would be dark which compromises the requirement and safety of citizens. A person needs to physically visit all the control points in a town to switch the lights ON/OFF. If due to any reason the lights not turned OFF, which happens all the time, huge amount of electric power is wasted. It is an outdated process in the modern world of automation. The lights are getting damaged and being replaced more often than they are supposed to be. A huge amount of public money is wasted if we consider the cumulative amount spent on unnecessary spending due to streetlights.

4 Proposed System

An overview of the proposed system is stated as The lights are controlled in two ways, i.e., automatically based on sunlight and remote switching through IOT. Either way, the lights cannot be turned on during day time as the light sensing circuit triggers only at night. The primary part of the system is to sense the amount of sunlight present in the surroundings. LDR is one of the best photo sensing components, Hence it is employed to detect the sunlight. The change in resistance of the LDR is fed to the TRIAC-DIAC pair which acts as the switch. The WiFi switch is employed to connect the device to the internet and enables us to control remotely.

We used Sonoff Wifi switch which has an android and IOS [1, 5] application that helps users to interface with the lights easily. This system has the potential to bring revolution in the way that the outdoor lights are controlled by the municipalities. A single device is capable of controlling tens of lights at present. In the future [7, 8] we can increase the number of lights that can be switched by balancing the incoming load from the transformers. By this we can reduce the amount spent by the municipalities on street lights management to a very low level.

5 Methodology

Methodology implemented in developing the device is the property of variation of resistance of LDR which can be used as a photo sensing element [9, 10]. When lights intensity falling on the LDR changes [11], the resistance of the LDR also changes.

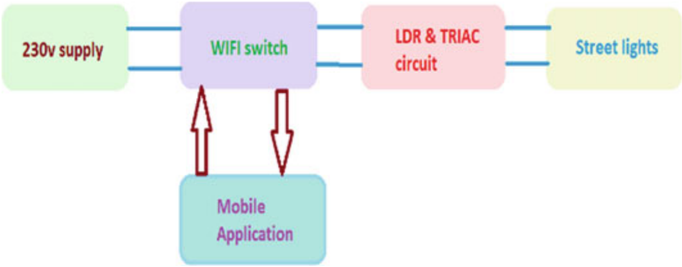


Fig. 1 Block diagram of smart outdoor lighting system

During day the value of resistance is less than a Kilo ohm, on the other hand, it is more than few Megaohms. This variation in the resistance is fed to the gate of a TRIAC-DIAC. TRIAC is a thyristor which has three terminals, it can switch up to 100A of current with less than 50 mA of Gate current. The DIAC is used to reduce the harmonics occur in the circuitry. Sonoff WIFI smart switch is connected between the power source and light sensing device. Once connected to local WIFI the lights can be controlled remotely through free android or IOS application.

6 Block Diagram

See Fig. 1.

7 List of Components

7.1 Light Dependent Resistor

A photoresistor or LDR is a resistor whose resistance decreases with increasing incident light intensity; in other words, it exhibits photo conductivity [12]. A photoresistor is made of a high resistance semiconductor. If light falling on the device is of high enough frequency, photons absorbed by the semiconductor give bound electrons enough energy to jump into the conduction band. The resulting free electron (and its hole partner) conduct electricity, thereby lowering resistance.

The spectral response of an LDR is similar to the human eye, which makes it ideal for this project (Fig. 2).

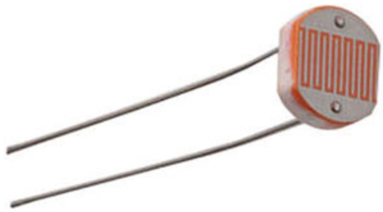


Fig. 2 Light dependent resistor (LDR)

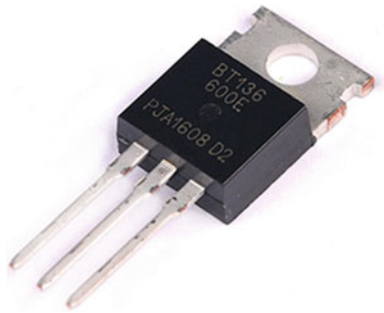


Fig. 3 TRIAC BT136

7.2 TRIAC

TRIAC is an electronic component that is widely used in AC power control applications. They are able to switch high voltages and hight levels of current, and over both parts of an AC waveform. This makes TRIAC circuits ideal for use in variety of applications where power switching is needed. One particular use of TRIAC circuits is in light dimmers for domestic lighting, and they are also used in many other power control situations including motor control and electronic switches. As a result of their performance, TRIACs tend to be used for low to medium power electronic switching applications, which meets the requirements of the proposed system (Fig. 3).

7.3 DIAC

The DIAC is a bi-directional semiconductor switch that can be switched ON in both polarities. The full form of the DIAC is diode alternating current. DIAC is connected back to back using two zener diodes and the main application of this DIAC is, it is widely used to help even activating of a TRIAC when used in AC switches [13], dimmer applications and starter circuits for florescent lamps (Fig. 4).

Fig. 4 DIAC DB3

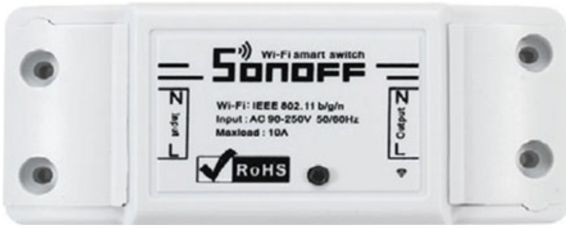


Fig. 5 Sonoff WiFi switch

7.4 *Sonoff WIFI Switch*

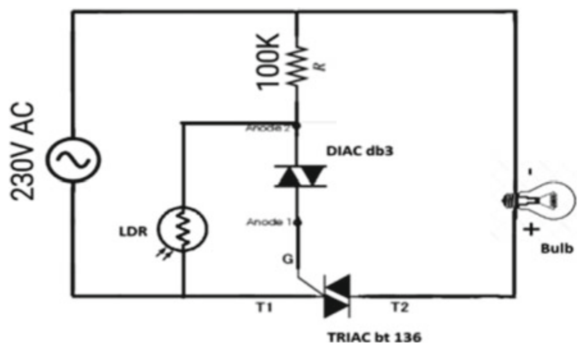
Sonoff Basic R2 is an affordable WiFi smart switch that provides users with smart home control. It is a remotely controlled power switch that can connect to a wide range of appliances. Sonoff Basic WiFi electrical switch transmits data to cloud platform through the WiFi Router, which enables users to remotely control all the connected appliances, via the mobile application 'eWeLink'. The server is Amazon AWS global server.

Sonoff WiFi controlled switch makes all home appliances smart. As long as the mobile phone has network(2G/3G/4G/WiFi), users can remotely control the appliances by turning them on off from anywhere at any time. Another feature available is to set timers for the appliances, which can include countdown/scheduled/loop timers, and can thus, help users maintain an easy life. The mobile application 'eWeLink' enables users to control the appliances easily. The iOS version of the application can be downloaded in App Store while the Android version in Google Play (Fig. 5).

8 Working Principle

The device basically does two things in order to achieve the automatic control of lights. First it detects whether it is day/night i.e., detects the amount of sunlight present in the surroundings with the help of LDR which is the best photo resistor. Then, based on the change in resistance value of LDR [14], TRIAC-DIAC pair

Fig. 6 Circuit diagram



switches the current supply to the lights. The below diagram shows the circuit of the device (Fig. 6).

As shown in the above circuit, the 230 V supply is fed to the terminal 1 of the TRIAC and LDR. DIAC is connected to the gate of TRIAC to promote even triggering and reduce harmonics.100 k ohm resistor is connected to prevent the reverse flow of current to the gate of TRIAC. During day resistance of LDR is very low (few ohms) hence, the current flows through the LDR but not through T1 of TRIAC. During night resistance of the LDR is very high (few mega ohms) hence, the current flows through T1 of TRIAC. The threshold voltage is exceeded T1 and TRIAC [15] is triggered by minimum current at the gate terminal. Now the current passes through TRIAC to the series of lights connected as the load. This process repeats continuously every day achieving the automatic control of the lights without any human interference. Sonoff WiFi switch can be optionally connected in between the power source and device. If in any case the lights are needed to the controlled remotely, we can control them with the mobile application. Even though we forget to turn the lights off in the mobile application, the device automatically turns OFF the lights when there is enough sunlight in the surroundings. By this our goal of zero electric power wastage is achieved.

9 Power Saved by the Device

Power consumed by improper manual switching:

Generally, the bulb's power rating is 9 W. Considering there are 500 bulbs in a town
Amount of power consumed by the bulbs = (9/1000) * 500 = 4.5 kW
Total number of hours operated in one month = 720 h
Amount of power consumed = 4.5 * 720 = 3240 kWh.
Total bill for 3240 units = **Rs. 15,000** (Rs. 4.5 per KWH).

Power consumed after installing the circuit:

All the 30 lights will be turned ON somewhere at 7 pm and OFF at 6 am. The total number of hours they are utilized is only 11 h a day
Amount of power consumed by the bulbs = (9/1000) * 500 = 4.5 kW
Total number of hours operated in one month = 330 h
Amount of power consumed in kilo watt hours = 4.5 * 330 = 1485 kWh
Total bill for 1485 units = **Rs. 7000**.

Percentage of power and money saved by the device:

The percentage of power saved = (3240–1485)/3240 * 100 = **54.2%**
The percentage of money saved = (15,000–7000)/15,000 * 100 = **53.2%**
It is proven that almost **Rs.8000/-** per month can be saved by installing these circuits in one town.

10 Total Public Money Saved

Amount that panchayats (considering Udayagiri as example) could save by using the device (per annum):

Salary of a person who controls the lights → ₹ 60,000 (₹ 5,000 per month)

Amount purchasing lights which gets damaged due to→ ₹ 25,000 (₹ 2000 pm for 13 lights on average)Uneven controlling
Electricity bill for one month→ ₹ 35,000 (₹ 3000 pm approx.)

Total expenditure before installing device→ ₹ 1,20,000

Amount for buying 30 circuits→ ₹ 30,000 (₹ 1,000 per circuit)

Total savings per year → ₹ 90,000

11 Implementation

In the start of year 2019 we have implemented this circuit in our campus to test the performance of the device in real time. Our campus has around 50 lights on the roads spread across an area of 100 acre. All the lights are controlled from a single point hence, we installed our device at the control point between supply and the lights. The device worked successfully as expected in switching the lights automatically [16] depending on sunlight, with zero human interference. Below is the picture of installation in our campus (Fig. 7).

We kept the device under observation for about eight months to verify lifetime and reliability. The results were very encouraging and the device had withstood all the climatic conditions. Later, as a first stage of implementing the circuit into the public street lights, we have contacted the panchayat officials of Udayagiri town and

Fig.7 Our device connected in MeRITS campus

ప్రాజెక్టు అమలుకు ఎంపీడీఓ ఆమోదం

ఉదయగిరి: స్థానిక మెరిట్స్ కళాశాలలో ఈసీఈ తృతీయ సంవత్సరం చదువుతున్న విద్యార్థులు నవీన్, సంతోష్‌కుమార్, హిమబిందు, రవి, ధరణి, విష్ణు రూపొందించిన 'ఆటోమేటిక్ ఎలక్ట్రిఫికేషన్ ఆఫ్ స్ట్రీట్ లైట్స్' ప్రాజెక్టును ఉదయగిరి మండలంలోని పలు పంచాయతీల్లో పూర్తిస్థాయిలో అమలు పరిచేందుకు మంగళవారం ఎంపీడీఓ వీరాస్వామి, ఇన్‌చార్జ్ ఈఓపీఆర్డీ అమర్‌నాథ్‌రెడ్డి ఆమోదం తెలిపారు. ఈ సందర్భంగా కళాశాలలో ప్రాజెక్టు తయారీ విధానం, పనితీరు, ఉపయోగాలను కళాశాల అధ్యాపకులు ఎస్.మెరిన్ ప్రసన్నకుమార్, ఎ.దస్తగిరయ్య పంచాయతీ కార్యదర్శులకు వివరించారు. అనంతరం కళాశాల ప్రిన్సిపల్ డాక్టర్ ఎం.మనోజ్‌కుమార్‌రెడ్డి మాట్లాడుతూ ప్రాజెక్టును పంచాయతీ స్థాయి నుంచి మండల, జిల్లాస్థాయి వరకు అభివృద్ధి చేసి విద్యుత్ వినియోగాన్ని 50 శాతం మేర తగ్గించి పంచాయతీలకు ఆర్థికంగా లాభదాయకంగా ఉండేలా తీర్చిదిద్దుతామన్నారు. కార్యక్రమంలో ప్లేస్‌మెంట్ అధికారి పాద్భాఖాన్, అన్ని విభాగాల అధ్యాపకులు పాల్గొన్నారు.

సాక్షి Wed, 04 March 2020 https://epaper.sakshi.com/c/49616004

విద్యుత్తు ఆదాకు విద్యార్థుల ప్రయోగం

ఉదయగిరి, న్యూస్‌టుడే : స్థానిక మెరిట్స్ ఇంజినీరింగ్ కళాశాల విద్యార్థులు తయారుచేసిన ఆటోమేటిక్ ఎలక్ట్రిఫికేషన్ ఆఫ్ స్ట్రీట్ లైట్స్ ప్రాజెక్టు పరికరాన్ని మండలంలోని అన్ని గ్రామ పంచాయతీల్లో ఏర్పాటు చేసేలా ఎంపీడీవో వీరాస్వామి ఆమోదం తెలిపారు. స్థానిక కళాశాలలో తృతీయ సంవత్సరం ఈసీఈ విద్యార్థులు నవీన్, సంతోష్‌కుమార్, హిమబిందు, రవి, విష్ణు, ధరణి ప్రాజెక్టులో భాగంగా తయారుచేసిన పరికరం పనితీరును ఎంపీడీవో వీరాస్వామి, పంచాయతీ కార్యదర్శులకు కళాశాల ప్రిన్సిపల్ డాక్టర్ మనోజ్‌కుమార్‌రెడ్డి, అధ్యాపకులు మెరిన్ ప్రసన్నకుమార్, దస్తగిరయ్యలు మంగళవారం వివరించారు. పరికరం ఏర్పాటులో విద్యుత్ వృథాను అరికట్టి ఖర్చులను తగ్గించుకునే వెసులుబాటు ఉంటుందన్నారు.

మాట్లాడుతున్న ప్రిన్సిపల్ మనోజ్‌కుమార్‌రెడ్డి

Fig. 8 News clipping of our project proposal to Panchayat officials in Udayagiri

briefed them regarding the benefits of installing this device and got permission to setup these devices all around Udayagiri. We haven't included the WIFI switch while installing the device on public street lights. The reasons being, there will not be WIFI available at every control point for the device to be connected to cloud. And as the public street lights [17] should always be turned ON during whole night, the lack of IOT control doesn't make any difference. Following pictures depict the installation of devices in Udayagiri town (Figs. 8 and 9).

12 Conclusion

The governments around the globe are in search of new and green energy solutions, which is fine. But, we are neglecting the unnecessary wastage of produced electricity which is not fine. Street lights around India are making up to a huge wastage of electric power. Hence, it is the need of the hour for us to check this wastage.

This proposed project is very cost efficient [18] and highly reliable in current scenario in reducing the wastage of electricity and saving [19] the public money

Fig. 9 Device installation in Udayagiri

which the municipalities spend on the maintenance of the street lights. We look forward to install this device in as many municipalities possible around India.

References

1. Merrin Prasanna N, Polaiah B (2020) Development of mathematical model based ROI for rotary kiln to measure burning zone temperature in cement industry by digital image processing. Int J Control Autom 13(2):416–427
2. Yang C-W, Nefedov E, Maksimainen M, Sierla S, Flikkema P, Kosonen I, Luttinen T (2014) Energy efficient traffic-based street lighting automation. In: IEEE conference, June 2014
3. Viraktamath SV, Attimarad GV (2011) Power saving mechanism for street lights using wireless communication, signal processing, communication, computing and networking technologies (ICSCCN). In: 2011 International conference-IEEE, pp 282–285
4. Ke Z (2015) Smart street lighting system: a new platform for smart city applications and a new frontier of cyber security. In: 4th Greater Chicago Area system research workshop (GCASR), 2015
5. Merrin Prasanna N, Polaiah B (2017) Computational intelligence survey analysis for advanced modeling and development of intelligent controller based on wireless sensor networks for industrial application. J Adv Res Dyn Control Syst 9(6):264–280
6. Zhang X, Jin J, Meng H, Wang Z (2011) A sensing optimal proposal based on intelligent street lighting system. In: Proceedings of ICCTA2011
7. Huaqiu D, Li X, Huaqiang X (2015) Wireless network node of LED lamps. In: Proceedings of EIT, Dekalb, IL, 2015, pp 207–209
8. Fujii Y, Yoshiura N, Takito A, Ohta N Smart street light system with energy saving function based on the sensor network. In: Conference on communications and information technologies (ISCIT)
9. Badgelwar SS, Pande HM (2017) Survey on energy efficient smart street light system. In: International conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC 2017)

10. Merrin Prasanna N, Bojja P (2019) Optimization of rotary kiln in cement industry using conventional control systems, ESCI, (helix–the scientific explorer), pp 4843–4849
11. Arthi J, Lydiapreethi W, Gunasundari B (2017) IOT based smart LED street lighting system. Int J Res Trends Innov (IJRTI) 2(4) ISSN: 2456-3315
12. Merrin Prasanna N, Polaiah B (2020) Design and real time implementation of Anfis controller for optimizing the burning zone temperature in rotary kiln in cement industry. The J Res Lepidoptera 51(2):480–498
13. Merrin Prasanna N, Polaiah B (2019) Optimization of rotary kiln in cement industry using conventional control systems. Helix, Web Sci 9(1):4843–4849
14. Polaiah B, Merrin Prasanna N, Pamula Raja K (2018) Development of mathematical model of rotary kiln which are used for design conventional control systems. Int J Eng Technol 7(3):1182–1186
15. Merrin Prasanna N, Bojja P (2019) Development and evaluation of mathematical model based region of interest for rotary kiln burning zone temperature in cement industry by digital image processing. Int J Innov Technol Explor Eng (IJITEE) 8(9), ISSN: 2278-3075
16. Deo S, Prakash S, Patil A (2014) Zigbee based intelligent street lighting system. In: 2014 Second international conference on devices, circuits and systems (ICDCS)
17. Huaqiu D, Yong C, Huaqiang X (2015) An intelligent control of LED street lamp based on illumination sensing. In: 2015 Fifth international conference on communication systems and network technologies (CSNT)
18. Knobloch F, Braunschweig N (2015) A trafficaware moving light system featuring optimal energy efficiency. Published in IEEE Sens J, pp 99, 1–1, February
19. Fan C-L, Guo Y, The application of a Zigbee based wireless sensor network in the LED street lamp control system. In: Proceedings IASP, Hubei, 2011, pp 122–125

Mathematical Model of Alternate Arm Converter

S. T. Rama and V. Rajini

Abstract The Modular Multilevel Converter has been one of the frontrunner in the emerging topology of VSC implemented in High Voltage Direct Current applications for a decade because of its innumerable advantages like excellent performance, reduced level of harmonics in the converter output without the necessity of filters, controllability, scalability, high power handling and high voltage capability and the foremost application in HVDC, the DC fault blocking capability and so on. A new hybrid topology of MMC called the Alternate arm converter which is designed with the combination of MMC and a two level converter possessing the ability to generate higher AC output voltage than the DC input voltage at a critical point of the converter called the sweet spot where the energy transfer between the AC and DC side of the converter is achieved is discussed in this paper. The reduction in the number of switching devices is achieved in AAC due to the alternate conduction of the director switches (two level converter) in the arms leading to a significant reduction in power loss. This paper presents a review on the operation principles and theoretical analyses of the Alternate Arm Converter (AAC). The mathematical model of the AAC has been designed and simulated with the help of Matlab for three phase Converter. The THD for the reduced model of AAC was also calculated and analyzed.

Keywords Voltage source converter (VSC) · HVDC · Modular multilevel converter · Alternate arm converter · Mathemathical model · Total harmonic distortion (THD)

S. T. Rama (✉)
Research Scholar, Department of EEE, Sathyabama Institute of Science and Technology, Jeppiaar Nagar, Chennai, India

Assistant Professor, Department of EEE, Dr.M.G.R Educational and Research Institute, Maduravoyal, Chennai, India

V. Rajini
Professor, Department of EEE, SSN College of Engineering, Kalavakkam, Tamilnadu, India
e-mail: rajiniv@ssn.edu

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_141

1 Introduction

Transmission systems have to face numerous challenges like de-carbonization, decentralization, and digitalization to the fast growing demand of power from domestic and commercial sectors [1]. Even though DC transmission system was invented first before 100 years, AC transmission was preferred globally as the voltage can be increased or decreased to any level for long distance transmission in order to avoid losses. Due to innovations in power electronic devices, researchers have again reconsidered the DC transmission for long distances where the sources are mainly renewable. To meet the new energy revolution and to build a smart grid HVDC has become the forerunner of the grid of the future. So, the present power system has an increase in Underground/sea and off shore electrical transmission to achieve stable transmission of efficient power in bulk quantities over long distances with less electrical losses. The interconnection of AC grids operated with different frequencies and controlling capability which may be the limitation of high voltage AC transmission with transmission capacity and connecting asynchronous AC grids was difficult. All these features are accomplished by the choice of HVDC systems which is environmentally superior and it is going to be the key enabler in the future grid [2].

HVDC power transmission using VSC technology has taken a significant proportion of future DC transmission systems. Different topologies of VSC has been developed and implemented for various applications in power system [3]. VSC-HVDC technology has the ability to control both active and reactive power separately which helps in downsizing equipments necessary for adjusting the voltage and reduce the voltage losses. This technology is expected to have a stable effect on the power systems [4, 5]. Additionally, an important feature of HVDC transmission using VSC is that it is capable to restore the power supply in an already existing de-energized AC network which is not possible in case of traditional LCC topology [6]. Hybrid converters was constructed by combining the advantages of classical VSC and MMC. This topology was designed to achieve a better output voltage than with classical VSC combined with using fewer semiconductor devices than with MMCs alone. MMCs are used in almost all the HVDC projects [3]. The first MMC was commissioned in Trans Bay Cable Project in USA by Siemens [7]. So, this paper discusses about the working operation of the latest VSC topology, the Alternate Arm Converter that is to be incorporated in the future power grids. The mathematical model of the AAC is also designed and analyzed using Matlab/Simulink.

2 AAC Topology

2.1 Description and Operation of AAC

A new modified topology of the Modular Multilevel Converter (MMC) is designed with the combination feature of MMC and the two level VSC in order to enhance various characteristics like reduction in overall cells and losses in the converter station, improve power efficiency, the temperature distribution between the IGBT modules in Alternate Arm Converter (AAC) is shown better than MMC and the most important feature of AAC is that it has the ability to block DC faults which means it has the capability to control the phase currents during the loss of the dc-bus voltage [8]. The hybrid MMC comprises of the upper and the lower arm in each phase of the converter. Each arm consists of a Stack of full H Bridge cells, a small inductor and a series of IGBTs as director switches to achieve a modified arm current waveforms called the Alternate arm converter (AAC) is shown in Fig. 1 [9]. This topology has been named because of its alternate use of the upper and lower arm of each phase to achieve the AC output which is discussed in [8–12]. The topology of the AAC has almost the same topology as MMC except that AAC has an additional design feature

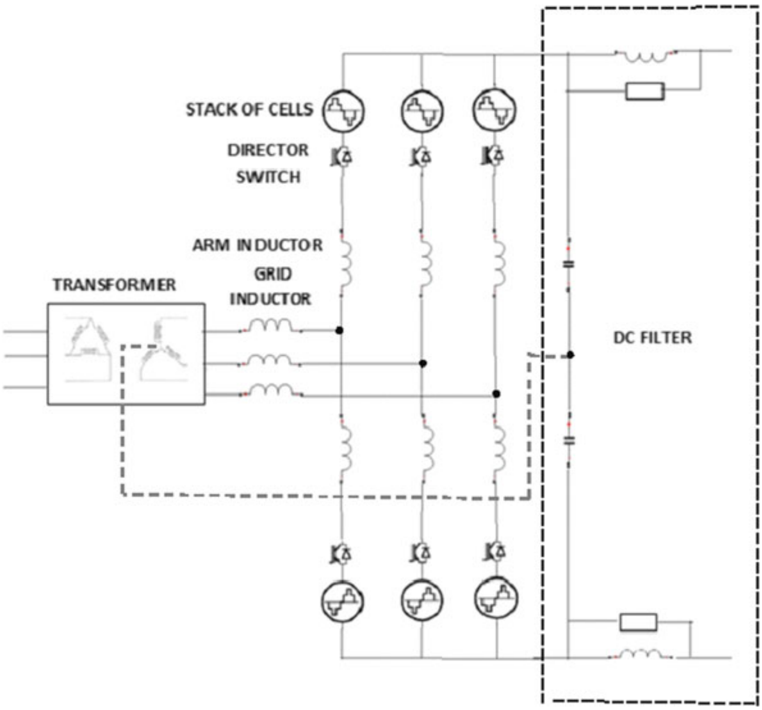


Fig. 1 Topology of three phase alternate arm converter

called the director switches which is a series of IGBTs connected along with the stack of H bridge cells which determines the switching state of the arms that causes the converter to generate the less harmonic AC voltage at the output and also carrying the AC current to the respective DC terminal. The director switches are placed on the upper and lower arm of the AAC to generate positive and negative voltage as shown in Fig. 2 [9, 12, 13].

Due to the presence of stack of H Bridge cells, the converter is able to produce higher AC voltage than the DC terminal voltage. The selection of the number and rating of IGBTs connected in series to form the director switch is based on the maximum voltage applied across the director switch in open state in order to avoid in damaging of the IGBTs. Each stack of cells will produce a maximum voltage that is equal to half of the input DC voltage, that is approximately half the rating of the arm of the MMC [9]. The idealized voltage and current waveforms for the upper and

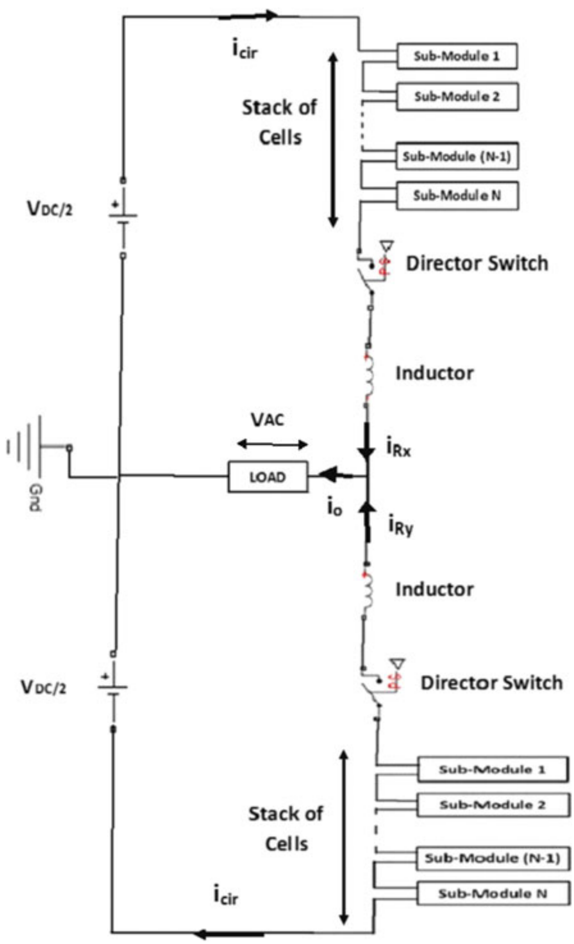


Fig. 2 Operation of single phase AAC

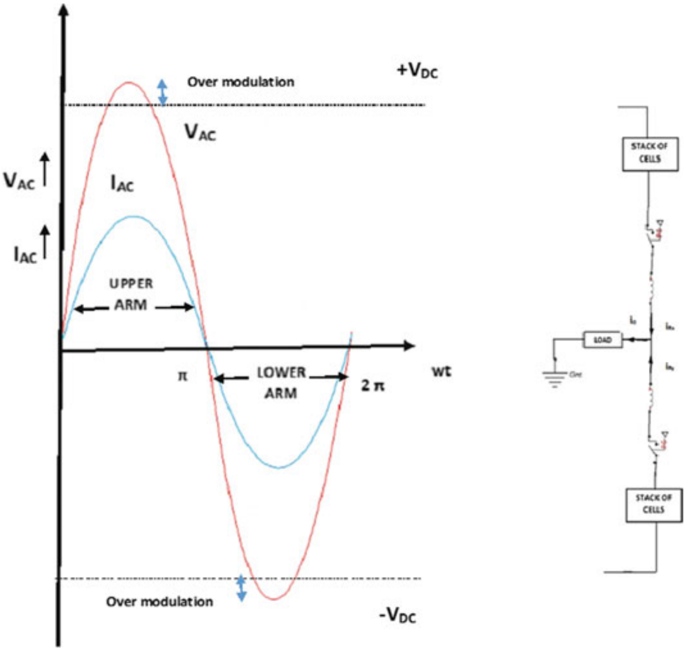


Fig. 3 Idealized voltage and current waveforms in a single phase converter of the AAC, showing the working period of each arm

lower arm in a single phase converter of the AAC is shown in Fig. 3 [9]. The reduction in the number of cells, which decrease the volume and losses of the converter station is the main achievement of the AAC [9, 14]. The presence of Stack of cells in the upper and lower arm produces a maximum AC voltage of 4/π times greater than half the DC bus voltage or half the DC terminal voltage which is represented in Eq. 1 [9, 13].

$$V_{AC} = \frac{4}{\pi}\frac{\mathrm{V_{DC}}}{2} \tag{1}$$

At this operating point called the sweet spot of the AAC, the net energy exchange between the arms is equal to zero. If non sweet spot operation occurs in the converter then the net energy exchange between the arms will lead to deviations in the capacitor voltages connected across the H bridge in steady state condition [13]. Different methods are adopted for balancing the capacitor voltage like the 'sort and select' algorithm, individual capacitor control schemes and energy balancing control schemes using increase in overlap period or current injection methods within the arms in order to achieve proper operation of the AAC [13, 15–18].

Furthermore, the significant advantage of the AAC is that it performs a STATCOM operation during DC faults in order to support the AC grid during outage especially in multi-terminal HVDC applications [9]. Alternate arm Converter is best suitable for HVDC transmission and industrial applications because of its effectiveness and reliable operation, so it the emerging choice for the future DC grids [19].

The next section deals with the average dynamic model of the Alternate Arm Converter, its methodology and analysis.

3 Reduced Model of AAC Topology

The AAC configuration consists of 3 legs representing the 3 phases comprising of 2 arms in each leg. Each arm consists a stack of cells (wave shaping cells or sub modules) which has the ability to generate the positive and negative voltage. To achieve the output alternatively across each arm of a phase, the main component called the director switches which is nothing but a two level VSC present in AAC. This design only differentiates the AAC topology from the Modular Multilevel Converter [20].

As the converter topology design comprises of semiconductor devices like thyristors, IGBTs, MOSFETs where the main controlling factor is the triggering or firing angle which makes the modeling of the VSC or LCC more easy in representation theoretically and system study which is discussed in [21]. As in case of VSC topology is concerned, the design possesses series connected IGBTs depending upon the rating of the project. These IGBTs are either conducting or blocking depending upon the firing angle. But the concept that is incorporated in modeling the AAC is different from VSC as this topology contains sub-modules that has stack of H bridge cells which may be inserted or by passed depending upon the capacitor voltages. The selection of the sub-modules is based on various selection algorithms which are mainly based on the direction of the current either flowing inside or outside the converter which causes the capacitor to either charge or discharge [22]. The methodology and equations to develop the Reduced Dynamic Model (RDM) are described in this section.

3.1 Mathematical Model

Research is being carried out to explore the different mathematical modeling approaches of the emerging multilevel converters and it was inferred that the computational speed is faster with acceptable accuracy and other aspects like behavior on fault conditions in Reduced Dynamic model was almost close to the full scale model (original topology) of AAC which was proven from the simulation results which was discussed in [20, 23]. The switching effects of the IGBTs in the stack cells and the director switches which causes high frequency effects and harmonics are not taken into consideration in this model. In order to obtain the energy balance between the cells within the arms various algorithms are incorporated [20]. The Reduced Dynamic Model of AAC is designed by considering the N sub-modules present in each arm of a phase by a controllable voltage source due to the presence of capacitors in the sub-modules with a maximum voltage equal to the product of the cell capacitor voltage and N number of sub-modules in each arm depending upon the rating of the project is represented in Eq. 2, equivalent or on resistance which is associated with the IGBT switches in the stack, director switch is represented by an ideal switch and finally inductor for each arm are all connected in series along with the voltage source which is shown in Fig. 4 [20, 21, 24]. The purpose of the inductor in each arm is to reduce or lessen the voltage difference caused due to the insertion or bypassing of the module [22].

$$Varm(max) = N * Vc \tag{2}$$

where Varm(max) is the maximum voltage appearing across the arm when all the capacitors are inserted, N is the number of sub-modules in each arm of a phase, Vc is the cell capacitor voltage.

In the Converter, each arm of a phase/leg operation is decided by a controlling number called the insertion index “n”. If $n_j = 0$ is assumed, then the sub-module in the respective arm is bypassed while if $n_j = 1$, then the respective sub-module is inserted or activated to generate the output voltage and j = 0, 1, 2, …, N represents

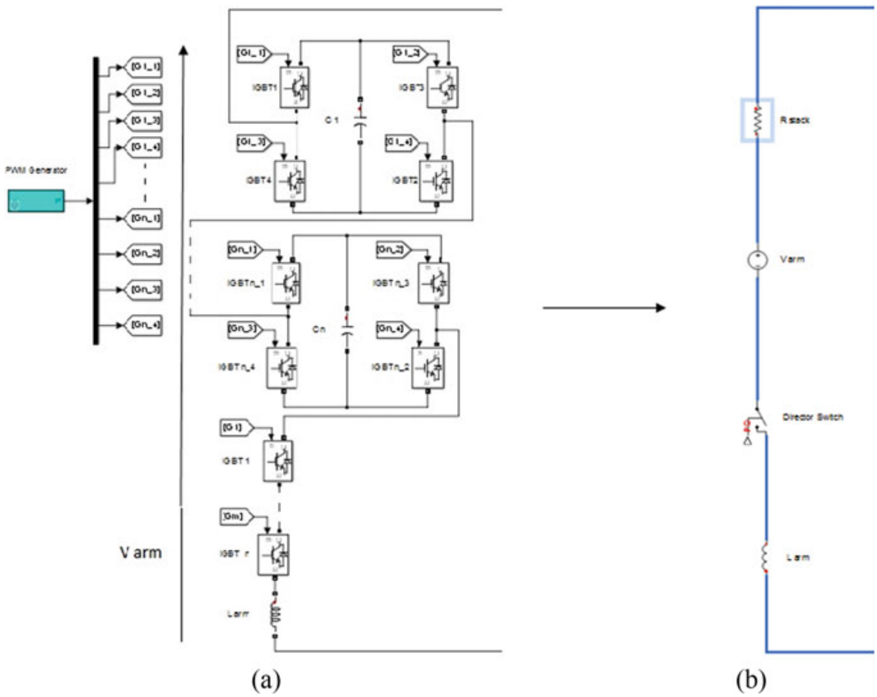


Fig. 4 Topology of AAC **(a)** Full model of each arm representation in a phase **(b)** Reduced Dynamic Model of each arm representation

the number of sub-modules in an arm. The voltage across each sub-module Vsmj is determined by Eq. 3. Upper arm voltage and Lower arm voltage are represented as Vax and Vay and is given by Eqs. 4 and 5, where a Є {R, Y, B} [24]. Parameters of upper and lower arm are represented by subscript 'x' and 'y'.

$$\mathrm{Vsmj} = \mathrm{nj} * Vcj \tag{3}$$

$$Vax = \sum_{j=1}^{N} \mathrm{Vsmj} \tag{4}$$

$$Vay = \sum_{j=1}^{N} \mathrm{Vsmj} \tag{5}$$

When the two insertion indices are added, its additive should be equal to 1. Because if one of the arms in a phase are inserted then the other will be bypassed, so the expression will be given as in Eq. 5 [21].

$$nx + ny = 1 \tag{6}$$

The behavior or characteristics of the capacitor can be found using Eq. 7. Different algorithms are implemented for equal voltage sharing across capacitors in each arm for insertion and bypassing the modules [22, 24].

$$C\frac{dVcj}{dt} = nj * im \tag{7}$$

where C is capacitance and i_m is the current flowing in the sub-module. "m" can be represented for current in upper or lower arm.

The mathematical model for a single phase is derived and analyzed for the upper and lower arm of R phase. The model can be derived for other 2 phases by 120° phase shift from each other. As the capacitors are connected in series for upper and lower arms in the AAC topology, the capacitor voltages can be represented by Eqs. 8 and 9 for R phase [22, 24, 25].

$$C\frac{d\sum_{j=1}^{N} Vcxj}{dt} = \sum_{j=1}^{N} nxj * iRx \tag{8}$$

$$C\frac{d\sum_{j=1}^{N} Vcyj}{dt} = \sum_{j=1}^{N} nyj * iRy \tag{9}$$

The total voltage across the sub-modules for the upper and lower arm are represented by Eqs. 10 and 11.

$$\sum_{j=1}^{N} Vcxj = Vcx \sum \tag{10}$$

$$\sum_{j=1}^{N} Vcyj = Vcy \sum \tag{11}$$

The nx and ny represents the insertion indices for the upper and lower arm of R phase which is given by Eqs. 12 and 13 [22, 24, 25].

$$nx = \frac{1}{N} \sum_{j=1}^{N} nxj \tag{12}$$

$$ny = \frac{1}{N} \sum_{j=1}^{N} nyj \tag{13}$$

The total voltage dynamics across the capacitors can be obtained from Eqs. 10–13 by substituting in 8 and 9 [22, 24, 25].

The currents in the upper and lower arm, circulating currents of the Reduced Dynamic model of AAC can be found by applying Kirchoff's current law which is shown in Fig. 5 [21, 22]. The Reduced dynamic model is derived by representing the

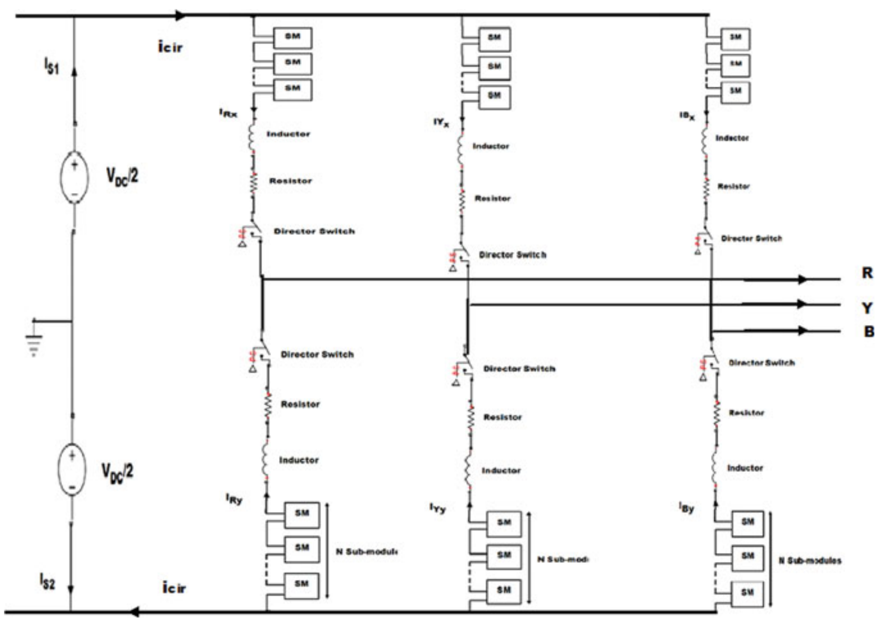


Fig. 5 Topology of three phase AAC with various current representation in each phase

N sub-modules in each arm by a controllable voltage source. The analysis is done for R phase which is the same for other phases by a phase shift of 120° from each other.

$$iRx + iRy = io \tag{14}$$

$$iRx = icir + Is1 \tag{15}$$

$$iRy = Is2 - icir \tag{16}$$

So, i_0 can be determined by inserting Eqs. 15 and 16 in Eq. 14 which is given by Eq. 17.

$$io = Is1 + Is2 \tag{17}$$

where $icir = \frac{iRx - iRy}{2}$ is the circulating current, Is1, Is2 are the source currents and i_0 is the load current. The difference between the arm currents of R phase is given by Eq. 18 [21, 22, 24, 25].

$$iRx - iRy = Is1 - Is2 + 2icir \tag{18}$$

The voltage dynamics or the characteristics of the R phase of AAC can be determined by applying Kirchoff's voltage law. Voltage of the upper and lower arm stacks of R phase are given by V_{Rx} and V_{Ry} and is represented in Eqs. 19 and 20.

$$Vcx * N * nx = Vcx \sum nx = \mathrm{V_{Rx}} \tag{19}$$

$$Vcy * N * ny = Vcy \sum ny = \mathrm{V_{Ry}} \tag{20}$$

The voltage drops across each component in R phase is represented in Eqs. 21 and 22 [21, 22, 24, 25].

$$\frac{VDC}{2} - VAC - Lx * \frac{dix}{dt} - Rx * ix - Vx = 0 \tag{21}$$

$$\frac{VDC}{2} + VAC - Ly * \frac{diy}{dt} - Ry * iy - Vy = 0 \tag{22}$$

where V_{DC} is the DC voltage, L_x, L_y are the inductor values across the upper and lower arm. R_x and R_y are the equivalent resistances of the stack. The V_{AC} output voltage which is AC is found across the load connected to the 3 phases of the AAC. The analysis of the RDM of AAC is not modeled for high frequency signals. The mathematical model of the AAC was framed and analyzed using Matlab Simulink [20, 21, 24].

4 Simulation Results

The computational time of the solid state semiconductor devices used in the complex design of the converters is high. So, Reduced Dynamic model of the semiconductor devices are implemented in the design of various converters topologies at the system level analysis of switching studies [26]. In this paper, RDM of the AAC topology was designed under normal conditions by representing the stack of H bridge cells by a controllable voltage depending upon the rating of the converter, an equivalent or on resistance of the IGBT switches in the stack represented by R_{stack}, Director switch represented by an ideal semiconductor switch which alternates the output between the upper and lower arm to achieve positive and negative cycle at the AC output voltage across the load and an inductor are connected in series. The combination of the components form the dynamic model of each arm of a phase which is represented in Fig. 6 [20, 24]. The RDM of AAC topology has been designed using Matlab/Simulink under normal conditions. The characteristics/parameters used in the design of the AAC model using Matlab/Simulink is given in Table 1 [20, 24].

The number of sub-modules in each arm is represented by N and the number of IGBTs to be connected in series as Director switch is given by N_{DS} [9, 20, 24]. The number of sub-modules per stack and the number of IGBTs to form a Director switch is given by Eqs. 23 and 24.

$$N = \frac{Vx}{\mathrm{Vsm}} \tag{23}$$

$$\mathrm{NDS} = \frac{Vdir}{\mathrm{Vsm}} \tag{24}$$

where $Vdir$ is the voltage peak across the director switch. The Average model of the AAC has been simulated using Matlab software. The conducting of either upper or lower arm in each phasehas been determined or decided by n_x and n_y which is represented in Eqs. 25 and 26.

$$nx = \frac{-1}{2}\frac{VAC}{VDC} + \frac{1}{2} \tag{25}$$

$$ny = \frac{1}{2}\frac{VAC}{VDC} + \frac{1}{2} \tag{26}$$

When nx = 1, the upper arm conducts otherwise if ny = 1, the lower arm conducts alternatively and so the load current flows alternatively in the upper and lower arm

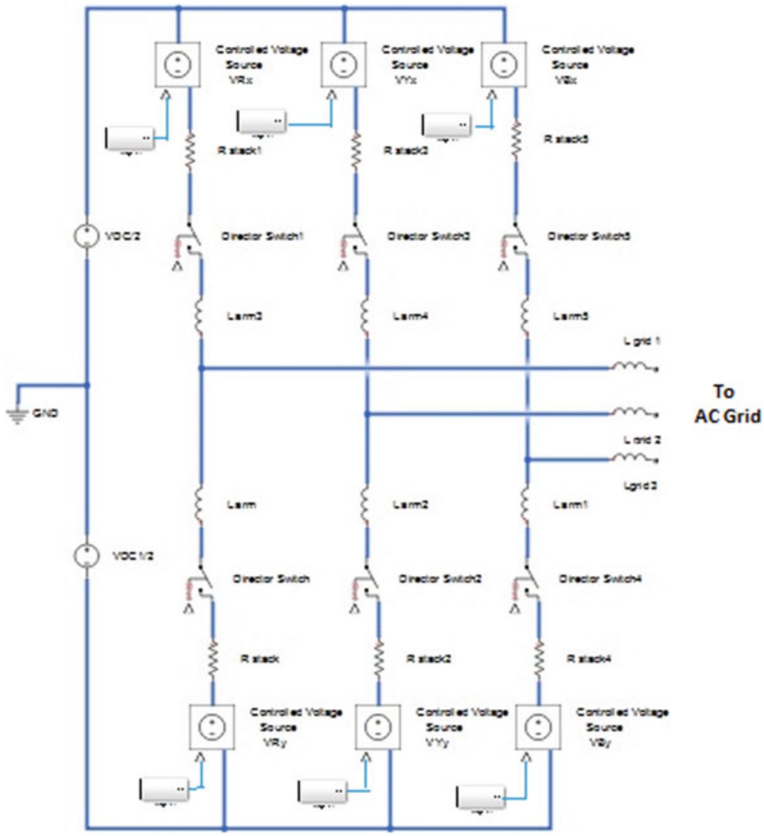


Fig. 6 RDM of AAC topology

Table 1 Model parameters of AAC topology

Parameter	Symbol	Value
Rated power	V_{GRID}	20 MW
DC voltage	V_{DC}	±10 kV
Converter AC voltage	V_{AC}	10 kV
Operating frequency	f	50 Hz
No. of cells per arm	N	9
No. of series IGBTs in director switch	N_{DS}	5
Sub-module voltage	V_{sm}	1.5 kV
Sub-module capacitor	C	4 mF
Arm inductor	L_{arm}	100 μH
Load resistance	R_{load}	2 Ω

for a phase/leg. The simulated result of the insertion indices, nx and ny is shown in the Fig. 7.

The RDM of AAC topology was designed and simulated using Matlab/Simulink software using the information from Table 1. The three phase AC output voltage and AC output current for the AAC is illustrated in Figs. 8 and 9.

The AC output current waveform obtained from the simulation infers that there is an unbalance at the initial computation time period and then settles down to obtain the steady state rated AC output voltage in the load. The imbalance in the output current and voltage is due to voltage fluctuations in the sub-module capacitors of charging and discharging at different times. In order to maintain the voltage balance, various

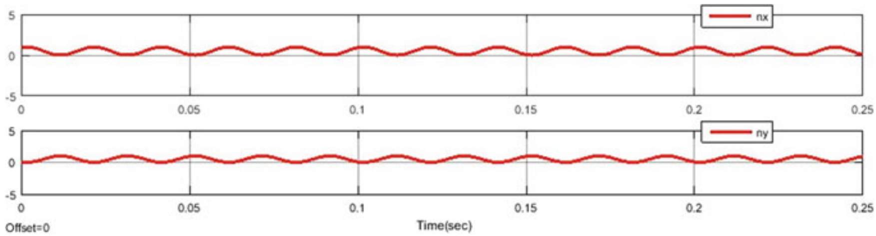


Fig. 7 Waveform of indices—nx and ny

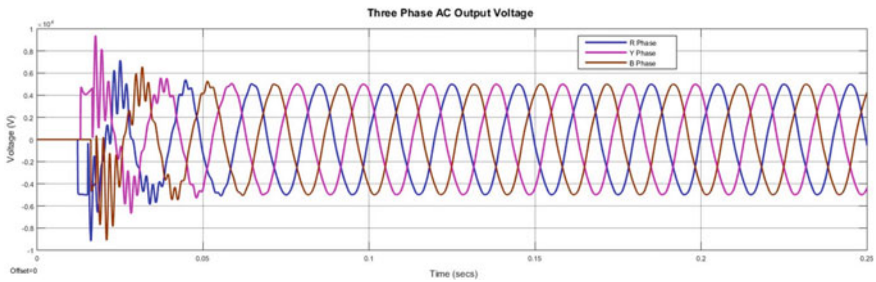


Fig. 8 Waveform of AC output voltage

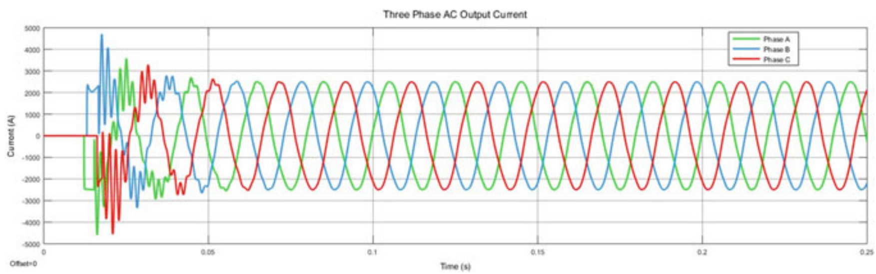


Fig. 9 Waveform of AC output current

balancing techniques like the sweet spot derivation implementation, extension of overlap period, circulating current controlling strategies, capacitor resizing, etc. that has been incorporated in the AAC topology was discussed in [6, 8, 9, 13–35]. The simplest technique, the sweet spot operating point where the energy is balanced on both the AC and DC side of the converter which has been implemented by the excess IGBT's present in the half bridge or full bridge configuration of the AAC due to which DC fault blocking capability has been achieved which is the main advantage of AAC. The important requirement to obtain the sweet spot is that, the AC voltage should be 1.27 times higher than DC voltage.

The THD analysis was performed for the generated AC output current using the simulink FFT analysis tool in the powergui and is found to be 4.26% [36] which is shown in Figs. 10 and 11.

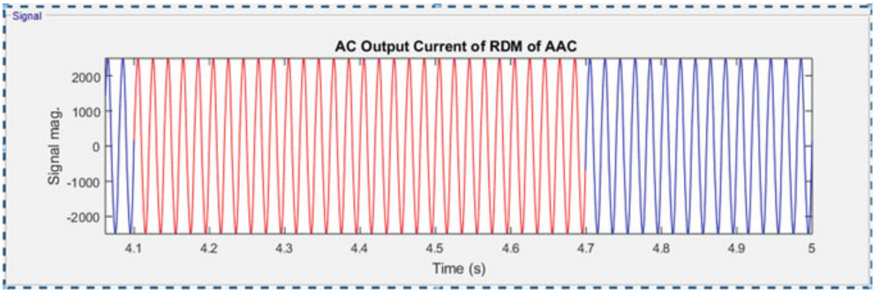


Fig. 10 Selected window of the generated AC output current in the FFT analysis

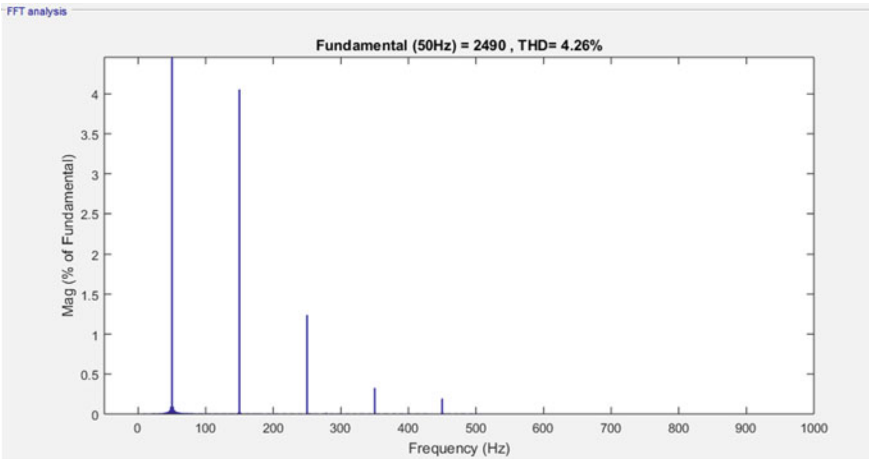


Fig. 11 FFT analysis of the generated AC output current

5 Conclusion

The Alternate arm converter configuration is composed of various characteristics where the topology design possess reduction in IGBT switches which resulted in less loss at the converter station thereby improving the efficiency of the power and the notable feature is the ability to block DC faults. The RDM of AAC was designed and simulated using Matlab/Simulink. The presence of voltage imbalance across the sub-modules due to uneven charging and discharging of the capacitors resulted in voltage fluctuations, which caused distortion at the initial time and then settled down at the upper limit. The FFT analysis was performed on the generated AC output current and it was found that the THD obtained was within the IEEE STD 519.

References

1. https://new.siemens.com/global/en/products/energy/high-voltage/highvoltage-direct-current-transmission-solutions/hvdc-plus.html
2. https://new.abb.com/systems/hvdc/why-hvdc
3. Davidson C, Trainer D (2010) Innovative concepts for hybrid multi-level converters for HVDC power transmission. In: 9th IET international conference on AC and DC power transmission (ACDC 2010)
4. Lesnicar A, Marquardt R (2003) An innovative modular multilevel converter topology suitable for a wide power range. In: Power tech conference proceedings, 2003 IEEE Bologna, vol 3, pp 1–6, June 2003.
5. https://www.hitachi.com/New/cnews/month/2019/04/190423c.html
6. Trainer DR, Davidson CC, Oates CDM, MacLeod NM, Critchley DR, Crookes RW A new hybrid voltage-sourced converter for HVDC power transmission. In: CIGRE Paris Session, 2010
7. Davies M, Dommaschk M, Dorn J, Lang J (2008) HVDC PLUS, basics and principle of operation. http://www.energy.siemens.com/nl/pool/hq/powertransmission/HVDC/HVDC_Plus_Basics_and_Principle.pdf
8. Judge PD, Merlin MMC, Mitcheson PD, Green TC (2013) Power loss and thermal characterization of IGBT modules in the alternate arm converter. In: Energy conversion congress and exposition (ECCE), 2013 IEEE, pp 1725–1731, 15–19 Sept 2013
9. Merlin MMC, Green TC, Mitcheson P, Trainer DR, Critchley R, Crookes W, Hassan F (2014) The alternate arm converter: a new hybrid multi-level converter with DC-fault blocking capability. IEEE 29(1):310–317
10. Davidson CC, Lancaster AC, Totterdell AJ, Oates CDM A 24 MW level voltage source converter demonstrator to evaluate different converter topologies. CIGRÉ, 2012
11. Merlin M, Green T, Mitcheson P, Trainer D, Critchley D, Crookes R A new hybrid multi-level Voltage-Source Converter with DC fault blocking capability. In: 9th IET international conference on AC and DC power transmission, ACDC, pp 1–5, Oct 2010
12. Merlin MMC, Judge PD, Green TC, Mitcheson PD, Moreno F, Dyke K (2014) Alternate arm converter operation of the modular multilevel converter. IEEE energy conversion congress and exposition (ECCE), 2014
13. Wickramasinghe HR, Konstantinou G, Pou J, Picas R, Ceballos S, Agelidis VG (2017) Comparison of bipolar sub-modules for the alternate arm converter. Elsevier 2017
14. Abdel-Moamen MA, Shaaban SA, Jurado F (2017) France-Spain HVDC transmission system with hybrid modular multilevel converter and alternate-arm converter. In: International conference on innovations in power and advanced computing technologies [i-PACT2017]

15. Kharade JM, Thorat AR (2015) Simulation of an alternate arm modular multilevel converter with overlap angle control for capacitor voltage balancing. In: Proceedings international conference on industrial instrumentation and control (ICIC), pp 502–506, 2015
16. Farr E, Feldman R, Watson A, Clare J, Wheeler P (2013) A sub-module capacitor voltage balancing scheme for the alternate arm converter (AAC). In: Proceedings European conference on power electronics and applications (EPE), 2013, pp 1–10
17. Moreno FJ, Merlin MMC, Trainer DR, Green TC, Dyke KJ (2015) Zero phase sequence voltage injection for the alternate arm converter. In: 11th IET International conference on AC and DC power transmission, pp 1–6
18. Moreno FJ, Merlin MMC, Trainer DR, Dyke KJ, Green TC (2014) Control of an alternate arm converter connected to a star transformer. In: Proceedings European conference on power electronics and applications (EPE), pp 1–10
19. Kharade JM (2015) Simulation of an alternate arm modular multilevel converter with overlap angle control for capacitor voltage balancing. In: International conference on industrial instrumentation and control, pp 502–506
20. Sheridan CE, Merlin MMC, Green TC (2014)Reduced dynamic model of the alternate arm converter. In: IEEE 15th workshop on control and modeling for power electronics (COMPEL), 2014
21. Abildgaard EN, Molinas M Modelling and control of the modular multilevel converter (MMC). Energy Procedia, pp 227–236. https://doi.org/10.1016/j.egypro.2012.03.023,2012
22. Antonopoulos A, Angquist L, Nee H-P (2009) On dynamics and voltage control of the modular multilevel converter. In: IEEE, 13th European conference on power electronics and applications, 2009
23. Perkar YR, Ramamohan KV (2016) Alternate arm converter with MMC. Int J Innov Res Sci, Eng Technol, pp 16749–16758. https://doi.org/10.15680/IJIRSET.2016.050918
24. Dembinskas D (2017) Modeling, simulation and control of the alternate arm converter. Thesis report
25. Ishfaq M, Uddin W, Zeb K, Khan I, Islam SU, Khan MA, Kim HJ (2019) A new adaptive approach to control circulating and output current of modular multilevel converter. Energies 2019. https://doi.org/10.3390/en12214084. https://doi.org/10.3390/en12061118,2019
26. Spallarossa C, Merlin MMC, Pipelzadeh Y, Green TC (2015) Reduced dynamic model of a modular multilevel converter in powerfactory. https://doi.org/10.1109/COMPEL.2015.7236482
27. Najmi V, Burgos R, Boroyevich D (2015) Design and control of modular multilevel alternate arm converter (AAC) with zero current switching of director switches. In: IEEE energy conversion congress and exposition, ECCE, pp 6790–6797. https://doi.org/10.1109/ECCE.2015.7310610,2015.
28. Mathew EC, Shukla A (2014) Modulation, control and capacitor voltage balancing of alternate arm modular multilevel converter with dc fault blocking capability. In: Proceedings IEEE applied power electronics conference and exposition (APEC) pp 3329–3336. https://doi.org/10.1109/APEC.2014.6803784
29. Vozikisa D, Adama GP, Raultb P, Tzelepisa D, Hollidaya D, Finneyc S (2018) Steady-state performance of state-of-the-art modular multilevel and alternate arm converters with DC fault-blocking capability. Electr Power Energy Syst 99:618–629
30. Vozikis D, Adam G, Holliday D, Finney S (2018) An improved alternate arm converter for HVDC applications. IECON 2018—44th annual conference of the IEEE industrial electronics society. IEEE, Piscataway, NJ. ISBN 9781509066858
31. Jasim OF, Moreno FJ, Trainer DR, Feldman R, Farr EM, Clare JC (2017) Hybrid experimental setup for alternate arm converter and modular multilevel converter. In: 13th IET international conference on AC and DC power transmission (ACDC 2017), pp 1–6
32. Judge PD, Chaffey G, Clemow P, Merlin MMC, Green TC (2015) Hardware testing of the alternate arm converter operating in its extended overlap mode. In: International conference on high voltage direct current 2015 conference (HVDC 2015), Oct 2015

33. Farr EM, Feldman R, Clare JC, Watson AJ, Wheeler PW Alternate arm converter (AAC)—"short-overlap" mode operation—analysis and design parameter selection
34. Moreno FJ, Merlin M, Trainer D, Green T, Dyke K (2015) Zero phase sequence voltage injection for the alternate arm converter. In: 11th IET international conference on AC and DC transmission, pp 1–6, Feb 2015
35. Farr E, Feldman R, Watson A, Burgos R, Clare J, Wheeler P, Boroyevich D (2014) Alternate arm converter (AAC) operation under faulted AC-grid conditions. In: 7th IET international conference on power electronics, machines and drives (PEMD 2014), pp 1–6, April 2014
36. http://www.egr.unlv.edu/~eebag/IEEE_STD_519_1992vs2014.pdf

A Novel Concept of Roof Top Tip Mass in Cantilever Based Energy Harvester for Wireless Sensor Node

Vicky Butram, Abhishek Ray, Alok Naugarhiya, and Guru Prasad Subas Chandra Mishra

Abstract This paper provides an efficient design of piezoelectric energy harvester with rectangular array and the roof-top arrangement of Tip mas. This energy harvester used thin film PZT-5H on Aluminium base and proof mass of tungsten. In this design, we have optimized the area of the energy harvester with non-traditional arrangement of tip mass to achieve lesser area. The maximum output power of 230 nW is generated by the harvester at 1.0 g acceleration. Moreover, the output voltage of energy harvester is 53 mV at the resonant frequency of 611.11 Hz. The performance is measured on the basis of output power, resonant frequency, output voltage, and energy achieved. With respect to the variable load resistance from 1 k to 1 M output power is calculated and maximum power of 300 nW at 55 K. Further, due to its smaller size and efficient performance it is compatible for sensors used in Internet of Things (IoT) applications.

Keywords MEMS · Piezoelectric · WSN · IoTs · FEM

1 Introduction

With the advancement of VLSI technology and vast reduction of size and power consumption of the sensor. The sensor technology demands on-chip power supply which can recharge batteries or replace. Due to the global energy crisis and the harshness of environment researcher are more attracts towards the self-powered power supply [1]. The ceaseless operation of self powered system without any disturbances is only possible by adopting the energy harvesting concept. There are various ambient

V. Butram (✉) · A. Ray · A. Naugarhiya · G. P. S. C. Mishra
Department of ECE, NIT Raipur, Raipur, India
e-mail: vbutram.phd2017.etc@nitrr.ac.in

A. Ray
e-mail: aray.phd2019.etc@nitrr.ac.in

A. Naugarhiya
e-mail: anaugarhiya.etc@nitrr.ac.in

G. P. S. C. Mishra
e-mail: gpscmishra.etc@nitrr.ac.in

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_142

sources are present in our surrounding like sunlight, fluid flow, mechanical vibration, and human activities used to obtain energy and replace conventional batteries [2]. Energy harvesting is the enabling advancement for different remote devices, e.g., wireless sensor nodes, having application in various domains such threat detection, vehicle health monitoring, smart cities/infrastructure, and condition based monitoring of manufacturing equipment. Among the various, mechanical excitations are the endemic form of ambient energy. On the basis of conversion mechanism, techniques which used for the generating output electrical from ambient vibration are being capable to harvest energy in the range of microwatts to mill watts [3]. The vibration-based energy harvester is being capable of driving low-power electronic devices without any interruption. Vibration presents in common household environments are classified as low-level vibration and vibration found in the large industrial equipment areas classified as high-level vibration. Low-level vibrations are of main interest and targeted in device design because of a broad range of potential applications that can be powered by harvesting vibrations at this level [4].

In the environment a kind of ubiquitous energy exists known as vibration energy, which can be directly transformed into electrical energy. Piezoelectric, electromagnetic and electrostatic techniques are the conversion mechanism [5]. Each approach has its own advantages and disadvantages. Electromagnetic energy scavenger mainly consists of a resonating cantilever, permanent magnets, and a coil. Essential burdens of electromagnetic energy harvesting are their low yield voltages, normally well beneath 1 V in size, and the troubles in lessening down to the micron-scale. The idea behind the electrostatic energy harvester is a variable capacitor that comprises of a charged capacitor with moving plates. As the initially charged capacitor plates found the relative movement between the charge plates it harvests energy [6]. Initial voltage requirement, high output impedance makes the electrostatic scavenger less appropriate [7]. Piezoelectric energy harvester received more attention because of piezoelectric property, MEMS compatibility, and simple structure. Due to its simpler structure fabrication process of this harvester is same as CMOS Fabrication [8]. Most precious advantage of these harvesters is well suited for the macro and micro-scale application. Piezoelectric based harvesters are additionally fit for conveying generally high output voltage (low output current), giving the certain voltage level (0.3–4 V) to charge an auxiliary battery or run a sensor legitimately.

In the journey vibration based piezoelectric energy harvester, Parshant et al. [9] have demostrated the spiral shape cantilever structure, harvested power of 23.3 nW which less than the proposed structure. Saxena et al. [10] illustrated a guided beam shape cantilever, which have complex geometry compared to the proposed work. In 2018, Tian et al. [11] demonstrated a rectangular beam with hole over it. Moreover, this structure generates a good amount of power but dimension of the structure is larger. Bhatia et al. [12] have discussed about Novel piezoelectric zinc oxide material fabrication technique. Authors stated the dynamic characterization of the material has a linear response but the piezoelectricity of ZnO is lesser than the PZT. Alperen Toprak, and Onur Tigli [13] reported a rectangular beam with the harvested power of 35.1 pW for 500 μm displacement. The Introduction of the MEMS based piezoelectric energy harvesters have been discussed in Sects. 1 and 2 contains description of the proposed work. This work consists of nonconventional cantilever structure, which is used in the proposed design of rectangular array cantilever beam. Simulation results are discussed in Sect. 3. Finally, this paper concludes in Sect. 4.

2 Design of Rectangular Array Cantilever with Roof Top Tip Mass

In this work, proposed a superior design of cantilever to provide enhanced energy harvester. The proposed design utilizes a unimorph type cantilever with the array of the rectangular beam. The structure has a non-traditional arrangement of tip mass like a roof-top arrangement as shown in Fig. 1. The structure comprises of the piezoelectric material of Lead zirconate titanate (PZT-5H). The base and tip mass of the proposed cantilever is made of Aluminium (Al) and Tungsten (W) respectively. Properties chosen for the designing of the structure is presented in Table 1. The dimension of the proposed beam is presented in Table 2. The beam vibration at resonant frequency gives maximum output power. Piezoelectric material has an ability to

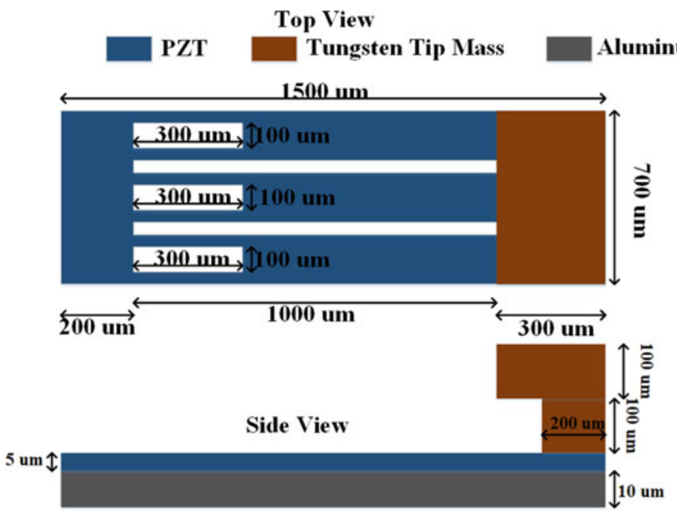


Fig.1 Schematic view of proposed cantilever

Table 1 Material properties proposed design

Parameters	PZT	Al	Tungsten
Young's modulus (GPa)	6.6	70	411
Poisson ratio	0.31	0.33	0.28
Density (kg/m^3)	7500	2700	19,350

Table 2 Dimension of proposed cantilever

Material	Height (μm)	Width (μm)	Thickness (μm)
PZT film	1500	700	5
Al substrate	1500	700	10
Hole over the beam	300	100	15
Tip mass	200	700	100
Roof top tip mass	300	700	100

convert stress or strain over it into charge and voltage or vice versa. The two constitutive Eqs. (1) and (2) explain mechanical and electrical properties of piezoelectric material [14].

$$S = s^E T + d^t E \tag{1}$$

$$D_i = dT + \varepsilon^T E \tag{2}$$

where,

S = mechanical strain;
T = mechanical stress;
D = electrical displacement;
E = electrical field;
sE = compliance under constant electric field;
εT = dielectric permittivity;
dt = reverse piezoelectric effect;
d = direct piezoelectric effect.

3 Simulation Results

In COMSOL Multiphysics 5.4a Finite Element Model (FEM) simulation is done. In the MEMS module of the simulator, solid mechanics and electrostatic are the two physics added to achieve eigen frequency and displacement of free end.

Figure 2 shows the displacement of beam at resonant frequency and Fig. 3 shows the stress developed over the beam at 1.0 g acceleration. Power of the designed structure is calculated on various aspects: with respect variable load and frequency. This aspect signifies the structure gives the maximum power of 230 nW at a frequency of 611.61 Hz and with variable load resistance gives 300 nW at a load of 40 KΩ as shown in Figs. 4 and 5 respectively. Comparison of proposed device with the existing work presented in Table 3. It is observed from the Table that the proposed device

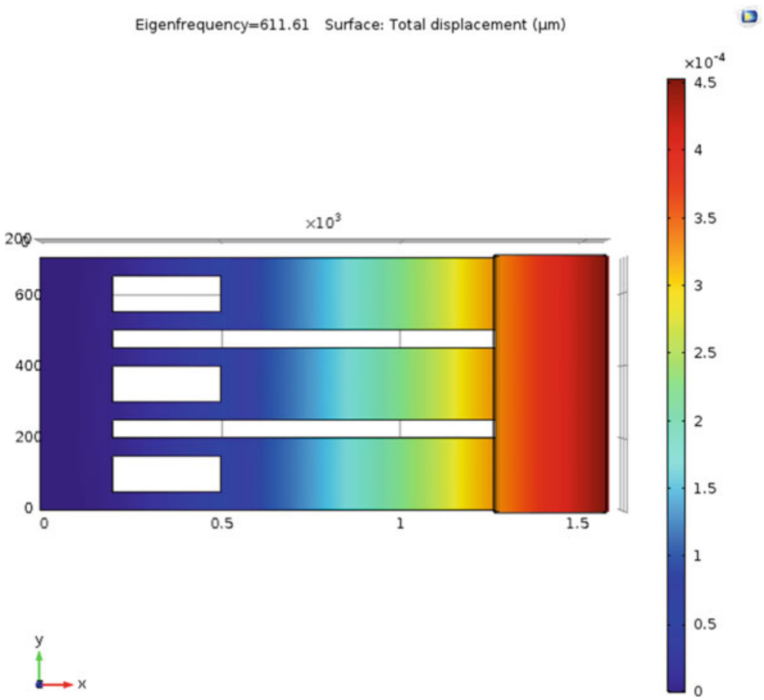


Fig. 2 Displacement at resonant frequency of the proposed structure

exhibits improved performance parameters (dimension, acceleration, frequency and power) in comparison to the existing work.

4 Summary/Conclusion

An array of the rectangular cantilever with the roof top tip mass is designed. Dimension of the beam is 1500 µm × 700 µm × 5 µm is simulated in COMSOL Multiphysics. The proposed design vibrates at resonant frequency of 611.61 Hz. At the resonant frequency maximum power of 230 nW and RMS voltage of 53 mV is harvested. Due to its simple structure fabrication complexities is minimize. In light of a few written works, it is discovered that the acceleration, excitation frequency for the harvester ought to be known before designing the structures.

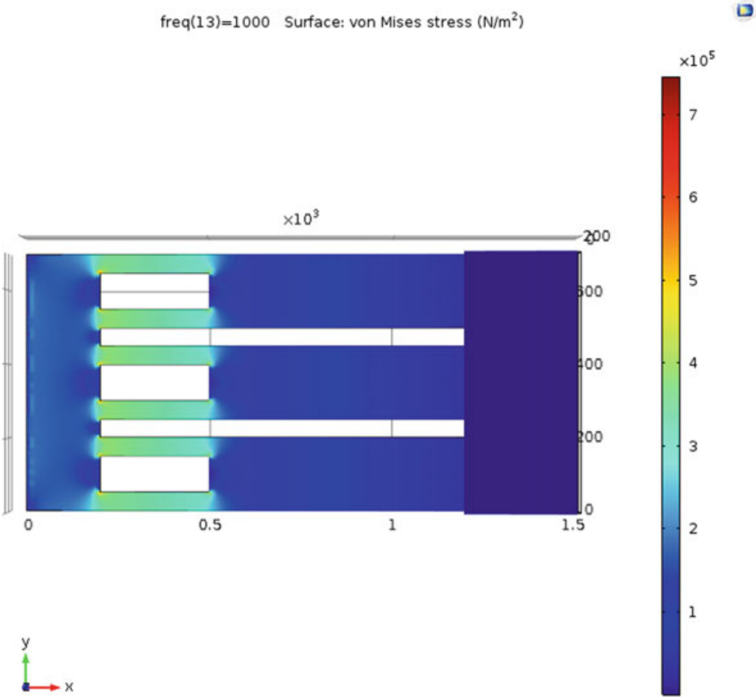


Fig. 3 Stress developed over the beam at 1.0 g input acceleration

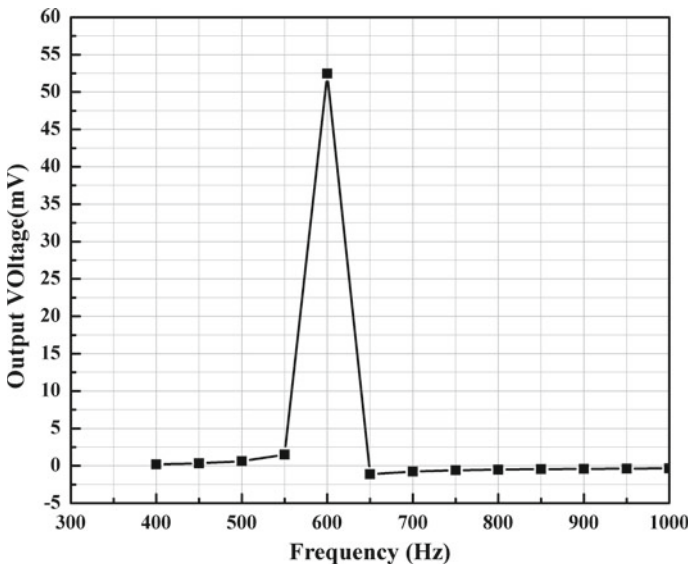


Fig.4 Output RMS voltage generated by harvester

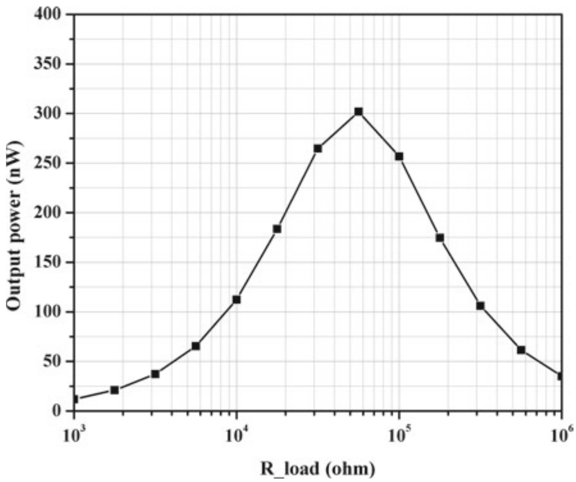


Fig. 5 Output power with respect to varying load resistance

Table 3 Comparison of result with existing work

Author (s)	Dimension (mm^2)	Resonant frequency (Hz)	Acceleration (g)	Harvested power
Hajati et al. [15]	120	1300	4.0	22 μW
Morimoto et al. [16]	76.6	126	0.5	5.3 μW
Fang et al. [17]	2.65	608	1.0	2.16 μW
Saxena et al. [10]	5	466	5.0	–
Proposed work	1.05	611.61	1.0	0.230 μW

References

1. Priya S, Song H-C, Zhou Y, Varghese R, Chopra A, Kim S-G, Kanno I, Wu L, Ha DS, Ryu J et al (2017) Areview on piezoelectric energy harvesting: materials, methods, and circuits. Energy Harvest Syst 4(1):3–39
2. Guo H, Wen Z, Zi Y, Yeh M-H, Wang J, Zhu L, Hu C, Wang ZL (2016) A water-proof triboelectricelectromagnetic hybrid generator for energy harvesting in harsh environments. Adv Energy Mater 6(6):1501593
3. Cook-Chennault KA, Thambi N, Sastry AM (2008) Powering mems portable devicesa review of non-regenerative and regenerative power supply systems with special emphasis on piezoelectric energy harvesting systems. Smart Mater Struct 17(4):043001
4. Kim M (2012) Materials and device design for mems piezoelectric mechanical vibration energy harvesters. Ph.D. dissertation, Massachusetts Institute of Technology
5. Todaro MT, Guido F, Mastronardi V, Desmaele D, Epifani G, Algieri L, De Vittorio M (2017) Piezoelectric mems vibrational energy harvesters: advances and outlook. Microelectron Eng 183:23–36
6. Beeby SP, Tudor MJ, White N (2006) Energy harvesting vibration sources for microsystems applications. Meas Sci Technol 17(12):R175

7. Roundy S, Leland ES, Baker J, Carleton E, Reilly E, Lai E, Otis B, Rabaey JM, Wright PK, Sundararajan V (2005) Improving power output for vibration-based energy scavengers. IEEE Pervasive Comput 4(1):28–36
8. Butram V, Naugarhiya A, Non-traditional proof mass arrangement in cantilever based piezoelectric energy harvester. In: 2018 IEEE 3rd international conference on computing, communication and security (ICCCS). IEEE, 2018, pp 100–103
9. Song H-C, Kumar P, Maurya D, Kang M-G, Reynolds WT,. Jeong D-Y, Kang C-Y, Priya S (2017) Ultra-low resonant piezoelectric mems energy harvester with high power density. J Microelectromech Syst 26(6):1226–1234
10. Saxena S, Sharma R, Pant B (2018) Fabrication and sensitivity analysis of guided beam piezoelectric energy harvester. IEEE Trans Electron Devices 65(11):5123–5129
11. Tian Y, Li G, Yi Z, Liu J, Yang B (2018) A low-frequency mems piezoelectric energy harvester with a rectangular hole based on bulk pzt film. J Phys Chem Solids 117:21–27
12. Bhatia D, Sharma H, Meena R, Palkar V (2016) A novel zno piezoelectric micro cantilever energy scavenger: fabrication and characterization. Sens Bio-sens Res 9:45–52
13. Toprak A, Tigli O (2015) Mems scale pvdf-trfe-based piezoelectric energy harvesters. J Microelectromech Syst 24(6):1989–1997
14. Hehn T, Manoli Y (2015) Cmos circuits for piezoelectric energy harvesters. Springer Ser Adv Microelectron 38:21–40
15. Hajati A, Kim S-G (2011) Ultra-wide bandwidth piezoelectric energy harvesting. Appl Phys Lett 99(8):083105
16. Morimoto K, Kanno I, Wasa K, Kotera H (2010) High-efficiency piezoelectric energy harvesters of c-axis-oriented epitaxial pzt films transferred onto stainless steel cantilevers. Sens Actuators, A 163(1):428–432
17. Fang H-B, Liu J-Q, Xu Z-Y, Dong L, Wang L, Chen D, Cai B-C, Liu Y (2006) Fabrication and performance of mems-based piezoelectric power generator for vibration energy harvesting. Microelectron J 37(11):1280–1284

Cost Optimization for Dynamic Data Migration and Re-placement with Load-Balancing in Geo-Distributed Systems

T. V. Rohini and M. V. Ramakrishna

Abstract With wide usage of smart digital devices and social media, enormous data is being continuously produced. Big data techniques are evolving to fulfill the ever-growing data processing demands, for which heterogeneous techniques are being used. Heterogeneous data processing consists of a collection of powerful, scalable and distributed nodes with different storage and processing capacity. Data replication is a strategy used to improve such system's availability and reliability. The data migration is a relocation scheme, in which data is moved from one node to another, either in the same or different data centers. To minimize the cost of access latency, migration cost and maximize the availability of the data, load balancing among the data nodes is done. Towards this goal, in this paper we formulate an optimization problem for data replica placement and migration. We have proposed a dynamic data migration scheme which moves the data nearer to the users accessing the data. If a node is overloaded or removed, we relocate the data partitions to the lightly loaded or new node. Our simulation results show that, the proposed dynamic data migration strategy balances the load across the nodes and improves the system performance as compared to the existing methods.

Keywords Data migration · Load-balancing · Optimization · Replication

1 Introduction

Social media has become a part of our daily life. With the increasing use of smart digital devices and social media applications such as Facebook, Whatsapp and Twitter, large volumes of data is being produced.

To store and manage these large volumes of data big data technology has emerged. The big data analytics is being used in many fields, such as sales, inventory, online-shopping and decision making. All these applications require access to large distributed data. The access latency of data for these applications has to be minimized

T. V. Rohini (✉) · M. V. Ramakrishna
Department of ISE, SJB Institute of Technology Affiliated to Visvesaraya Technological University, Belagavi/Bengaluru, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_143

as bandwidth cost reduced [1]. This is achieved by data replication (placing the same data in more than one location). We have previously proposed adaptive dynamic replication and data placement to maximize the availability and to minimize access delay [2].

Data distribution has been dealt in [3–10]. For such applications, the data need be stored near the user application. This is known as data locality. If the access patterns changes we need to relocate the data. In this paper, we proposed a dynamic data migration scheme to move the data between or within the geo-distributed system to balance the load across nodes. The goal is to minimize the access delay and communication cost and network bottleneck. This is an extension of our work reported in [2]. The main contributions of this paper are as follows:

- We have proposed a dynamic data migration strategy to relocate the data in Geo-distributed storage systems. The proposed data migration strategy balances the load across the distributed nodes by migrating the data partitions near to the user.
- We formulate the optimization problem for the dynamic data replica placement and migration.
- We have studied the proposed data migration scheme by simulation. The simulation result shows that, the proposed scheme minimizes the access latency and communication cost as compared to existing data migration schemes and increases the availability of the data.

The remainder of the paper is presented as follows. Section 2 describes the related work; Sect. 3 elaborates problem statement, optimization problem and a proposed dynamic data migration strategy. Section 4 presents the simulation experiment results. The last section presents conclusions and directions for further work.

2 Related Work

In this section we reviewed the state-of-the-art solutions given by various researchers [11–15]. Li et al. proposed a optimal based on data migration scheme based on hotspot prediction [9]. They proposed a Data Location Optimization method based on data locality DLO-data migrate. The data locality optimization algorithm uses hotspots prediction (DLO-Predict) to reduce user data access latency in cloud computing systems. They try to reduce the waiting delay moving the tasks and predicting hotspot files. The execution time of non-node-locality task is reduced by fetching data in advance and made available in the target machine. In this method the hot files are migrated periodically to multiple machines of the cloud data centers. In the DLO-Data migration algorithm, the nodes with idle bandwidth are selected and non-node-locality tasks are identified. Data of non-node-locality tasks are migrated.

Mansouri et. al. proposed a replication and migration scheme [16]. They formulate optimization problem considering two different types of costs: residential cost and migration cost. Residential cost means read and write cost and second, the migration cost is network cost. They have proposed two algorithms: offline optimal algorithm and online algorithm. The offline algorithm gives the knowledge about available workload on objects. The two online algorithm that consider the trade-off between residential and migration cost. The data migration has been classified into two types: live and non-live. In case of live data migration, the data is available to users while data migration is under progress. In the second case, write operation is stopped when data migration is under progress but available for read.

Mohamed Faten et al. proposed a data migration technique called CRANE, to reduce the cost of creating new replicas [17]. This approach reduces the time required to copy the replica to the new location and network bottleneck. This scheme complements replica creation and replica storage management is performed effectively. The main contribution of this work is to minimize the time needed for replica placement. Replica placement involves creation or migration of data replica between or within data centers. This process incurs traffic between those nodes which involves in data exchange. This scheme performs replication and creates replicas in data centers with the reduction in time required to copy the data to the new location. This eliminates the traffic in network and ensures availability of the data replicas.

Denis et al. proposed a scheme for replica placement using distributed key value store [18]. They describe a multi-objective optimization and popRing approach based on genetic algorithm for replica placement. A distributed replica placement on KVS system uses consistent hashing with virtual nodes. They compared popRing with OpenStack-Swift replica placement and found that, KVS approach balances the load.

Jingya et al. proposed a lightweight replica placement algorithm [19]. They formulated optimization and which is adapted dynamically. They evaluated the LRP strategy with two dataset from Facebook and Twitter.

Kitae Choi et al. [20] proposed a load balancing technique in distributed systems that handles efficiently data replication and migration. A big data technology has come into view for the processing and storing of large volume of data. In this work, authors have considered performance of the storage nodes and check the load cost of the nodes and perform the allocation of initial data. The strategy selects the heavy loaded nodes and hot data based on the load cost.

Edwin et al. proposed a dynamic data distribution and migration algorithm based on user request patterns [21]. In this work, data migration strategy consists of three steps. Initialization, profiling and migration. In this work, authors analyzed the get/put operation of several enterprises and try to reduce write operation on solid disks. They investigated several objects are highly involved in read and write operations and migrated write-intensive objects from solid state disks to Hard disk drive. Read intensive objects from hard disk drive to solid disk drive. The objects which are frequently accessed are migrated out of archive Hard disk drives. A very small number of data objects are migrated and improved the performance.

3 Problem Statement

A geo-distributed system consists of set of data centers. $DC1$, $DC2$, …, DCn. Each data center consists of commodity nodes with different storage and processing capacities. The data nodes are connected with communication links. These links are characterized by a bandwidth capacity BWc. Let $D1$, $D2$, $D3$, …, Dn denotes the data objects stored across the data centers. When a user query arrived at the node, the query processing takes place. If the associated data is present in the remote node, then the multiple associated data from various other nodes are accessed and given to the user. In this scenario, cost of accessing the data is given by, cost of accessing the host node, where the user request has arrived and cost of accessing the remote node. Hence to minimize this cost, we need to move the data objects near to the user location. Also, as the load on some node increases, then data must be migrate to the less overloaded neighbor node. When data is migrating between the nodes, bandwidth is the most important cost factor to be considered. As the new nodes are added we need to migrate the data to the newly added to balance the load. The main objective of this paper is to optimize the cost of data migration and locating the data partitions at the appropriate storage node to minimize the migration cost and access latency. We model the cost optimization problem for data migration and replica placement Table 1. Summarizes the notations used in the paper.

Table 1 The notations used in this paper

S. No.	Symbol	Meaning
1	$DC1$, $DC2$, …, DCn	Set of data centers
2	$D1$, $D2$, $D3$, …, Dn	Data partitions
3	K	Total number of regions k1, $k2$, $k3$, …, $kn \in K$
4	$R(Di)$	Read request of the data object Di
5	$W(Di)$	Write request of the data object Di
6	$Size(Di)$	Storage size of the data object Di
7	$B(D)$	Binary variable. $B(D) \in (0, 1)$
8	DCS	Source data center
9	DCT	Target data center
10	DCN	New data center
11	BWc	Bandwidth cost
12	$Hostnode$	The node where user request arrived
13	$Remotenode$	The node that receives request from $Hostnode$
14	$ProcCost(Di)$	Processing cost of data Di is read/write cost
15	$AvailStorage(DCT)$	Available storage space at Target node

3.1 System Model and Cost Optimization Problem

In this section we briefly explain few important definitions: These definitions give the clarity of the system model.

Definition 1: Data Node and Data object: The Geo-distributed system shown in the figure consists of six data centers with different processing and storage capacities. $DC1$, $DC2$, $DC3$, …, DCn are data centers in which data objects $D1$, $D2$, $D3$, …, Dn are stored. For each data item, the cost of processing read is $R(Di)$, and write is $W(Di)$. The storage cost is dependent on its size $Size(Di)$. We use binary variable $B(D)$ to indicate if the data partition is present in the data center ($B(D) = 1$ if the data partition is present in the data center CDi, otherwise 0) at time slot T.

Definition 2: Access Cost of Host Node and Remote Node: In geo-distributed system, the node where the request arrives is called host node and the node that receives the request from other nodes is called remote node. If all the requested data partitions are present in the host node, then the cost of accessing the data item Di is given by storage cost and processing cost at time T. If the requested data partition is present in the remote node then cost of accessing data partition includes the bandwidth cost *BWc and* accessing the remote node.

3.2 Migration Cost

Data migration in Geo-distributed system is the moving data objects from source data center DC_S to target data center DC_T. The cost of moving data object from source data center to target data center is given by bandwidth cost BWc multiples with size of the data object and processing cost. Here the processing cost is the read request or write request.

$$C_{Migration\_cost} = BWc * Size(Di) * ProcCost(Di) \tag{1}$$

In Eq. (1), BWc is bandwidth cost, Size(Di) is storage cost of data partition Di, and ProcCost(Di) is read/write cost.

3.3 System Architecture

Figure 1 show overall architecture of distributed system and it consists of six data centers. The system contains client, scheduler, and Load balancer: which balances the load among nodes, replica manager is present in all the data centers and it is responsible for creating, deleting and maintaining replicas and replica catalog contains list of

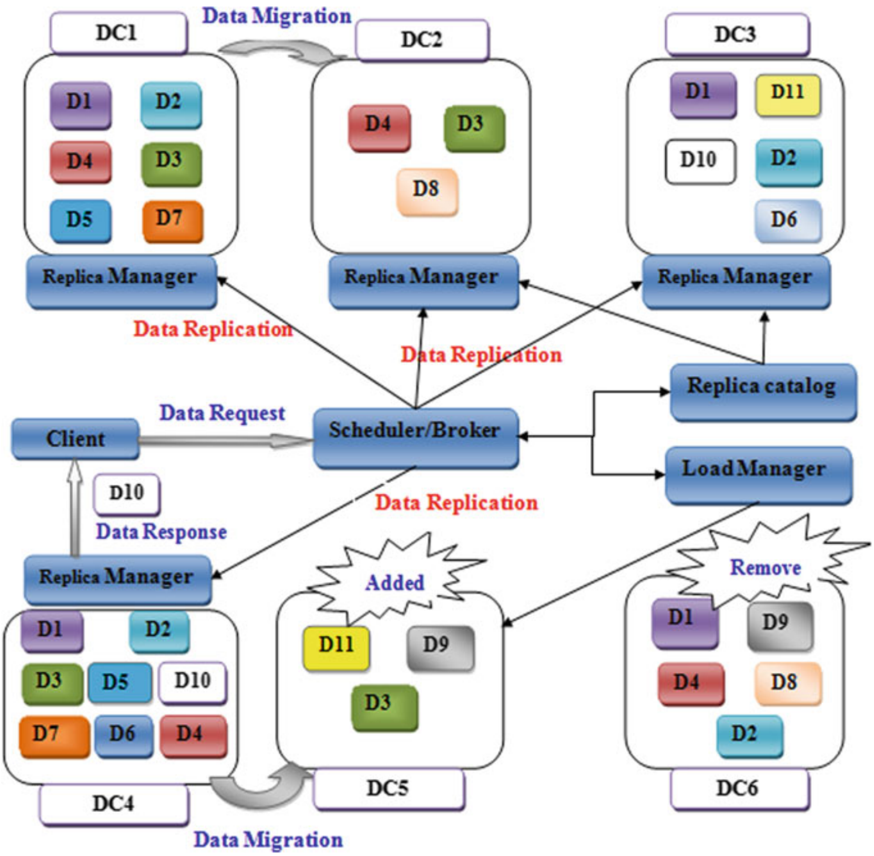


Fig. 1 Overall architecture [22]

files and location of the data replicas. The broker sends request to replica catalog and get the locations of the requested data. Replica catalog returns the reply containing the list of replicas and corresponding locations. Broker or scheduler will forward the user request to the nearest data center. In this paper, we proposed a dynamic data migration strategy. The proposed strategy re-distributes the data items based on the storage and processing capacities and access patterns of the client. As the load on some machine increases, the strategy migrates the data to the lightly loaded node. As the user request patterns are changed, the data items are migrated near to the user. If the cost of accessing the remote node is increases, then data is migrated to the host node. When new nodes are added to the system, data is migrated to the newly added node and balances the load among all the data nodes in the system.

3.4 Initial Data Placement and Data Migration

In data center storage system, the data objects are stored near to the user based on the user access patterns and node's storage and processing capacity. The data objects are replicated and placed in more than one location and this increases the data availability. The data migration in distributed systems is moving the data items from source data center to destination data center. During the process of data migration, until data object relocated at destination location, the data objects will not be available to the user. In this paper, we proposed a dynamic data migration scheme. The proposed strategy migrate the data objects from source location to destination location in the following cases:

- Whenever the user access patterns are changed, migrates the data partitions to minimize the access latency.
- When a data request arrives at the host node and associated data is present in the remote node. In this case, if the size of the data partition accessed in remote node is greater than the size of the data accessed at host node then migrate the data partition from remote node to host node. This reduces the bandwidth cost because each time when user requests the data, it avoids accessing of remote node.
- When nodes are included or evicted and when the storage node is overloaded, migrate data to the lightly-loaded node.

3.5 Optimization Problem

The cost of accessing the data partitions and the migration cost are given above. We define the overall cost optimization problem to minimize the objective function for all the data partitions at time T.

$$Minimise \sum_{k=1}^{K} \{C_{Access\_Delay} + C_{Migration\_cost} + C_{Load\_imbalance}\} \tag{2}$$

Equation (2) is subjected to following constraints:

- If a node is overloaded or if a node is removed due to some malfunction, relocate the data partitions in the other new node, where the same (data replica) copy of the data partition is not present. The replicas are placed geographically at different location to achieve high availability.
- If the data object Di is present in the source node at time T_1, then migrate the data object to the destination node, only when the overall accessing cost in new data center is less than or equal to the cost of accessing the data object if it stays in old data center at T_1 incurred to process the data. The binary variable $B(D) =$ 1 if the data partition is present in the data center otherwise 0.

$$C_{Access\_Delay(DCs)} \geq C_{Access\_Delay(DC_T)} \tag{3}$$

- When a data partition is migrating from source node *DCs* to target node DC_T, the storage free space available at destination node is greater than or equal the storage cost needed to relocate the data partition. i.e. The available storage space at DC_T is greater than or equal to storage space needed to locate the data object *Di*.

$$Size(D_i) \leq AvailStorage(DC_T) \tag{4}$$

3.6 Proposed Algorithm

In this paper, we proposed a dynamic data migration algorithm based on load and user access pattern. In Fig. 2 shows that data migration based on the load of the data center. All the nodes are heterogeneous in nature. In Fig. 3 *DCs* and DC_T represents the source data center and the target data center respectively. *Di* represents the data object. If the source node is overloaded then migrate the data to the destination node. As the user access patterns are changed then migrate the data near to the user. Let *Di* and *Dj* are the two associated data partitions located at *Remotenode and Hostnode* respectively in different data centers. When a user request arrives at *Hostnode*. If the size the data partition *Di* at *Remotenode* is greater than the size of the data partition *Dj* at *Hostnode*, then migrate the data partition *Di* from *Remotenode* to *Hostnode*. If the new data center DC_N is added, then migrate the data partition from the overloaded node to the newly added node. If some node is removed, then adjust the replica factor and relocate the data and update the metadata.

4 Experimental Results

We have simulated the proposed dynamic data migration technique using NS2 simulator and the version of NS2 used is ns 2.35. Hadoop is the tool which provides the software framework for processing large volume of data. We have implemented the proposed dynamic data migration using and Hadoop tool. Hadoop provides a software

Algorithm: 1 Load Balancing and Data Migration

Load = Storage Capacity + Processing Capacity + Network Bandwidth

1. if (Load (DC_S > DC_T) && ! overloaded (DC_T)
2. Data Migration (DC_S (D_i) à DC_T(D_i))

Fig. 2 Algorithm for data migration with load balancing

Algorithm: 2 Dynamic Data Migration (DDM) based on user access pattern and load

1. Data Migration (D_i , DC_S , DC_T)
2. if (node DC_S is overloaded)
3. Data Migration (DC_S (D_i) à DC_T(D_i))
4. if (user access pattern changed)
5. Data Migration (DC_S (D_i) à DC_T(D_i))
6. if (Remotenode (Size(D_i) > Hostnode (Size(D_j))
7. Data Migration (DC_S (D_i) à DC_T(D_i))
8. if (new node DC_N is added)
9. Data Migration (DC_S (D_i) à DC_T(D_i))
10. Update the Metadata
11. if (node is removed)
12. Adjust the Replica Factor and Data replication // Adaptive data replication [22]
13. Update the Metadata

Fig. 3 Algorithm for dynamic data migration

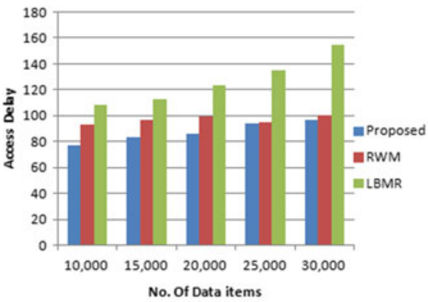
framework for storing and processing big data and distributed storage. The experimental results of proposed data migration strategy are compared with the existing data migration strategies such as Load-balanced data migration and replication LBRM [20] and RWM [23]. We conducted extensive set of experiments with various data sets to evaluate the proposed method. The experiment results are compared with the existing dynamic data migration strategies. The data sets are tabulated in Tables 2 and 3.

Table 2 Simulation results of number of nodes versus delay, throughput, communication cost

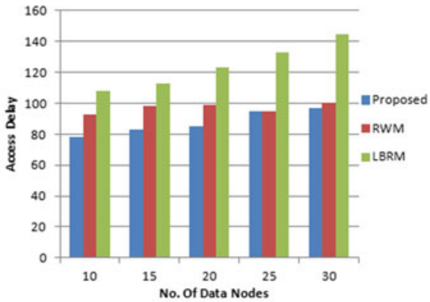
No. of nodes	Delay			Throughput			Communication cost		
	Proposed	RWM	LBRM	Proposed	RWM	LBRM	Proposed	RWM	LBRM
10	78	93	108	108	103	104	88	88	110
15	83	98	113	138	120	136	93	95	102
20	85	99	123	290	156	181	103	103	150
25	95	92	133	300	280	280	106	113	153
30	97	99	145	350	302	310	110	134	183

Table 3 Simulation results of number of data items versus delay, throughput, and communication cost

No. of data items	Delay			Throughput			Communication Cost		
	Proposed	RWM	LBRM	Proposed	RWM	LBRM	Proposed	RWM	LBRM
10,000	77	93	108	110	106	100	87	87	108
15,000	83	97	113	136	120	103	92	93	112
20,000	86	99	123	181	132	156	103	103	135
25,000	94	95	135	300	157	300	123	112	154
30,000	97	100	155	340	260	302	128	135	183



(a)



(b)

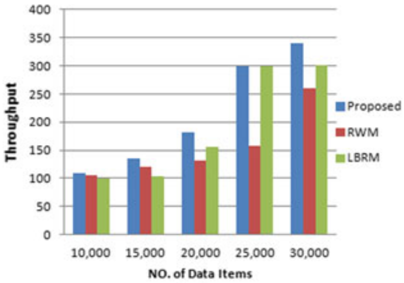
Fig. 4 Comparison graph of proposed method versus existing method by varying number of data nodes and data items

4.1 Access Latency

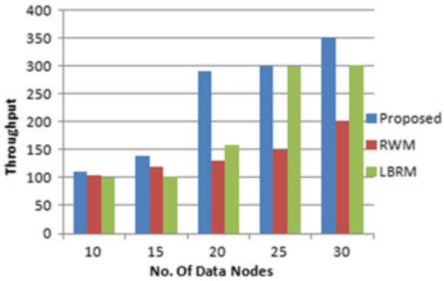
We compared the access latency of proposed data migration strategy with existing data migration strategy. The access delay of the proposed DDM method is minimized as compared to the existing methods. Because, as the user access patterns are changed the data partitions are migrated near to the user location. If the multiple associated data partitions are present in the remote node and if the cost of accessing remote node is more than data object is migrated to the host node. In Fig. 4a, we varied the number data objects. The graph shows that access delay is minimized with proposed method. In Fig. 4b, we varied the number of data nodes. The graph shows that the access delay is minimized with the proposed dynamic data migration scheme.

4.2 System Throughput

In Fig. 5a, the graph shows that system throughput is increased using proposed data migration strategy as compared to the existing data migration strategies. Because as



(a)



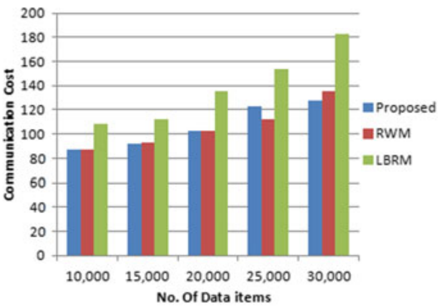
(b)

Fig. 5 Comparison graph of proposed method versus existing method by varying number of data nodes and data items

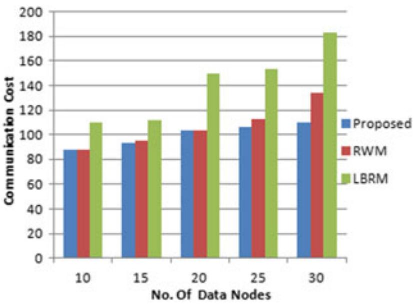
the user access patterns are changed, the data is migrated near to the user and as the nodes are overloaded the data is migrated to the lightly loaded nodes. Hence with the proposed data migration strategy we can achieve load balancing, the overall system performance is increased. The availability of the data to the user is increased. The proposed method will migrate data replicas geographically at different locations. In Fig. 5b, as we increase the data nodes the system throughput is increased as compared to the existing techniques.

4.3 *Migration Cost*

We compared the migration cost of the proposed method with existing strategies. The overall migration cost of the proposed method is minimized. In Fig. 6a, as we increase the number data items or queries, the graph shows that migration cost is minimized as compared to the existing strategies. Similarly, Fig. 6b, shows that as



(a)



(b)

Fig. 6 Comparison graph of proposed method versus existing method by varying number of data nodes and data items we vary number of data nodes the proposed method perform better compare to the existing methods.

5 Conclusion

Data migration in distributed system relocates the data objects dynamically to improve the performance of the system. Data replication strategy improves the availability of the data to user by placing replica in the various data centers. We formulated optimization problem for data replica migration and placement. The objective function of the optimization problem is to minimize the migration cost and access latency and maximize the availability. In this paper, we proposed dynamic data migration strategy. This proposed strategy moves large data partitions between the storage nodes in an efficient manner. The proposed method optimizes the migration cost, access latency and balances the load among the storage nodes. The proposed method improves the reliability and availability by not placing same (replica) data partition in the same node. The simulation result shows that, the proposed method reduces the access delay and overall communication cost and increases the system throughput. In future we plan to extend our work to maintain the consistency of replicas stored in various location of the geo-distributed system your final files.

Acknowledgements This research was supported by Visvesaraya Technological University, Jnana Sangama, Belagavi-590018. We gratefully thank for extended research work support.

References

1. Huang P-C, Chang Y-H, Chen T-Y, Tu C-H, Chen C, Wei H-W, Shih W-K (2018) Performance optimization of heterogeneous cloud storage with bandwidth & capacity considerations. In: Proceedings of the 33rd annual ACM symposium on applied computing. ACM, 2018, pp 210–213
2. Rohini T, Ramakrishna M (2019) Spectral clustering and bounded-load consistent hashing for data placement in heterogeneous geo-distributed systems. J Adv Res Dyn Control Syst 11(06), 306–315
3. Ishikawa, K (2013) Asura: scalable and uniform data distribution algorithm for storage clusters. arXiv preprint arXiv:1309.7720
4. Jiao L, Li J, Xu T, Du W, Fu X (2014) Optimizing cost for online social networks on geo-distributed clouds. IEEE/ACM Trans Netw 24(1):99–112
5. Qiu X, Li H, Wu C, Li Z, Lau FC (2014) Cost-minimizing dynamic migration of content distribution services into hybrid clouds. IEEE Trans Parallel Distrib Syst 26(12):3330–3345
6. Mseddi A, Salahuddin MA, Zhani MF, Elbiaze H, Glitho RH (2015) On optimizing replica migration in distributed cloud storage systems. In: 2015 IEEE 4th international conference on cloud networking (CloudNet). IEEE, 2015, pp 191–197
7. Mansouri Y, Buyya R (2016) To move or not to move: cost optimization in a dual cloud-based storage architecture. J Netw Comput Appl 75:223–235

8. Xu X, Zhao X, Ruan F, Zhang J, Tian W, Dou W, Liu AX (2017) Data placement for privacy-aware applications over big data in hybrid clouds. Secur Commun Netw 2017
9. Li C, Zhang J, Ma T, Tang H, Zhang L, Luo Y (2019) Data locality optimization based on data migration and hotspots prediction in geo-distributed cloud environment. Knowl-Based Syst 165:321–334
10. Erradi A, Mansouri Y (2020) Online cost optimization algorithms for tiered cloud storage services. J Syst Softw 160:110457
11. Fu X, Li J, Liu W, Deng S, Wang J (2019) Data replica placement policy based on load balance in cloud storage system. In: 2019 IEEE 3rd information technology, networking, electronic and automation control conference (ITNEC). IEEE, 2019, pp 682–685
12. Canali C, Chiaraviglio L, Lancellotti R, Shojafar M (2018) Joint minimization of the energy costs from computing, data transmission, and migrations in cloud data centers. IEEE Trans Green Commun Netw 2(2):580–595
13. Mansouri Y, Erradi A (2018) Cost optimization algorithms for hot and cool tiers cloud storage services. In: 2018 IEEE 11th international conference on cloud computing (CLOUD). IEEE, 2018, pp 622–629
14. Mansouri Y, Toosi AN, Buyya R (2017) Data storage management in cloud environments: taxonomy, survey, and future directions. ACM Comput Surv (CSUR) 50(6):1–51
15. Mansouri N, Javidi M (2018) A hybrid data replication strategy with fuzzy-based deletion for heterogeneous cloud data centers. J Supercomput 74(10):5349–5372
16. Mansouri Y, Toosi AN, Buyya R (2017) Cost optimization for dynamic replication and migration of data in cloud data centers. IEEE Trans Cloud Comput 2017
17. Mseddi A, Salahuddin MA, Zhani MF, Elbiaze H, Glitho RH (2018) Efficient replica migration scheme for distributed cloud storage systems. IEE Trans Cloud Comput 2018
18. Cavalcante DM, de Farias VA, Sousa FR, Paula MRP, Machado JC, de Souza JN (2018) Popring: a popularity-aware replica placement for distributed key-value store. In: CLOSER, 2018, pp 440–447
19. Zhou J, Fan J, Jia J, Cheng B, Liu Z (2018) Optimizing cost for geo-distributed storage systems in online social networks. J Comput Sci 26:363–374
20. Bok K, Choi K, Choi D, Lim J, Yoo J (2019) Load balancing scheme for effectively supporting distributed in-memory based computing. Electronics 8(5):546
21. Wu L, Zhuge Q, Sha EH-M, Chen X, Cheng L (2017) Boss: an efficient data distribution strategy for object storage systems with hybrid devices. IEEE Access 5:23979–23993
22. Rohini T, Ramakrishna M Adaptive dynamic data replication in distributed systems. In: International conference on chip, circuitry, current, coding, combustion & composites Bengaluru, Karnataka, 28–29 November 2019, i7C—19. Institute for Engineering Research and Publication (IFERP)
23. Mansouri Y, Buyya R (2019) Dynamic replication and migration of data objects with hot-spot and cold-spot statuses across storage data centers. J Parall Distrib Comput 126:121–133

WBAN Technology: Challenges and Security Attacks

Ananya Nandikanti and Kedar Nath Sahu

Abstract Wireless Body Area Network (WBAN) is a promising technology that is developed due to the advancement in wireless communication technologies, physiological sensors and low power Integrated Circuits (ICs). It is used as a part of e-healthcare systems for continuous monitoring of patients. Each of the WBAN system introduced so far, for a unique application, suffers from some operational challenges and security attacks. However, increase in the complexity of any WBAN application demands careful attention to multiple risks of attacks and challenges. This calls for a need of accumulated data on challenges and attacks pertinent to different applications. This paper presents a review on the challenges and probable security attacks of WBAN technology carried out based on the data reported over past eighteen years by noted databases. The review shall be useful to practical engineers to take precautionary measures in the design of complex WBAN systems.

Keywords WSN · WBAN · Challenges · Security attack

1 Introduction

Wireless sensor networks (WSNs) have less infrastructure and can function without direct human intervention. They can (i) detect and record the variation in vital signs or external conditions in particular pressure, vibration, sound, temperature, etc. and (ii) transmit this data over a sink where it is gathered and analyzed for further process. These networks contain communicating sensor nodes and the radio signals establish node level communication [1]. Various Wireless Sensor Networks include Wide Area Network [WAN], Wireless Metropolitan Access Network [WMAN], Wireless Local Area Network [WLAN], and Wireless Personal Area Network [WPAN]. Figure 1 represents the schematic view of various WSN and their classifications as reported in [2, 3].

A. Nandikanti (✉) · K. N. Sahu
Department of ECE, Stanley College of Engineering and Technology for Women, Hyderabad, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_144

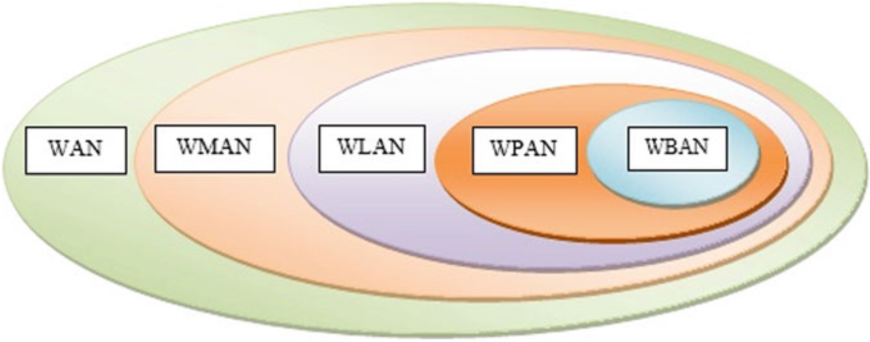


Fig. 1 WSN classification [2, 3]. *Source* Proc. Int. Conf. Complex Medical Eng. (ICME) 2012, Kobe, Japan

In recent times, yet another wireless network technology that finds growing interest among the researchers is the Wireless Body Area Network [WBAN], introduced by Van Dam et al. As reported by Balouchestani et al. [3], WBAN is a subset of WSN. This comprises of (i) wearable and (ii) implantable computing devices. Wearable devices can be mounted on the surface of the body and the implantable ones can be embedded inside the body. The network uses multiple communicating sensor nodes which communicate among themselves and transmit the data to a sink or database. This is used to serve as a part of e-healthcare administrations by using which the health of a patient can be continuously monitored. A few important definitions of WBAN coined variously in the literature are reported below. WBAN is defined as a wireless networking technology which uses radio signals to interconnect tiny nodes with sensor or actuator that can be deployed in, on, or around a human body [3]. As reported by Ramli et al. [4] WBAN detects and records the vital signs by establishing wireless link among sensor nodes operating on, in or around the human body.A personal device gathers the signals of these vital signs which acts as a sink for data of the sensor nodes and transmits them to the physician for supervision. WBAN is also defined as a collection of wireless sensor nodes that are located either inside or outside the human body to monitor the surrounding environment and functions of the body [6]. Zou et al. defined WBAN as an ultra-short range wireless networking technology [7]. It is a personal network associated with the human body with each of its parts assigned with an IP address. IEEE 802.15 Task Group 6 standard is assigned to produce energy-efficient WBAN systems with ultra-wideband (UWB).

WBAN finds extensive applications in many areas including sports, defence, consumer electronics, medicine, entertainment and so on [8–10]. The first and foremost application of WBAN is the field of medicine that provides health monitoring in sports. It is also used to record blood pressure, heartbeat rate, weight scale, motion etc. In defense, WBAN provides intercommunication between soldiers, blast dosimeter and soldier health monitor. Consumer electronics such as mobile phones, music player, head phones and hearing aids use WBAN. Emotion detection, ambient intelligence and detection of postures are few prominent lifestyle applications. In entertainment, they are used in the design and implementation of electronic games and gesture detection.

In summary, the function of WBAN involves (i) accumulation of essential health parameters of the patient using sensors and (ii) transfer of the processed data through a personal device via internet or server to a database that can be accessed by a physician. The doctor then sends necessary suggestions based on the data collected. The collection of data through sensors and transferring it over the communication channel, involves a wide variety of security and privacy issues. The objective of this paper is to review the research that exists emphatically on issues such as (a) challenges and (b) security attacks of WBAN technology. The aim of this attempt to bring together what is reported on these two issues is to provide a solid theoretical perspective on the topic.

The review is carried out by referring about fifty articles reported on WBANs in various databases namely, Elsevier, IEEE Xplore, Google Scholar, Research gate, and so on. However, twenty eight selected papers were relevant to the main objective of this work.

2 WBAN Challenges

Although WBAN technology has witnessed a tremendous development in recent times, there are many challenges, related to its real-time applications, experienced which are investigated and reported by several researchers [5, 6, 11–22]. The prominent challenges in WBAN are presented below.

Security and Privacy: As these systems deal with the exchange of patient's data, they require high level security and privacy. The design constraints of the network must be choosen in such a way that it does not combine the parameters of one patient with another by providing limited access to the data. The data transfer methods prescribed for WSN are not suitable for WBAN. Confidentiality, authentication, integrity, freshness of data along with availability and secure management contribute to the security requirements of WBAN and pose as one of the greatest challenges in providing secure WBAN system [5].

Data quality: The data accumulated by WBAN should be of high quality, as it plays a vital role in analysing the condition of the patient to provide suitable treatment [11–18].

System device: The sensors used in WBAN systems have to be power efficient and reconfigurable with low complexity, less form factor, light weight, easy usage. The design of the storage device should provide flexible access to the medical history of the patient at any instance for external processing and analysis using online tools [11, 19].

Sensor validation: At times, the pervasive sensing devices may be subjected to inherent communication and hardware constraints. This infers possible chances of manipulation of data leading to erroneous data transfer to the end user. Protection against the illegal use of data and identification of the weaknesses present in software and hardware design needs validation of the sensors used [12, 19].

Social acceptance: Use of WBAN for purposes beyond medical applications is a challenge as they involve likely apprehension owing to loss of privacy of their personal data [16].

Data consistency: As WBAN is a wireless network with more number of nodes, the data from multiple devices as well as at the patient nodes should be accumulated and analysed in a systematic manner else it might lead to degradation of treatment [14, 15].

Interference: The wireless link established between the sensors use radio signals which might interfere with other network devices during the large scale implementation of WBAN [16].

Data management: Database requirement is tremendously large to store the data collected continuously for monitoring of patient's health for easy access and retrieval. Necessary measures also must be taken to avoid collision of data [13, 14].

Cost: As consumers prefer high functionality at low cost, it remains an all-time challenge to develop cost efficient WBAN systems which usually have inherent design complexity [18].

Constrained deployment: The design of WBAN must ensure transparency to the user without being detrimental to the routine activities of the user. It must be wearable, non-intrusive and of light weight [19].

Interoperability: The WBAN systems must ensure transfer of data across standards such as Bluetooth, and Zigbee, for the exchange of information, plug and play device interaction. To achieve stable connection and efficient migration across the network the systems have to be scalable [20].

Constant monitoring: The need for continuous monitoring apart from the monitoring on the occurrence of an exclusive event demands a large need of energy for functioning of WBAN. This badly challenges in the design of energy efficient WBAN systems [20].

Consistent performance: The WBAN must ensure data consistency even when it is restarted after its switched-off state. Consistent handling of data at any circumstance throws a challenge for its assured performance [19, 21].

Scalability: The system has to expand its design to perform any number of functionalities whenever the situation demands, by adding external resources [7, 22].

3 Security Attacks

The working performance and capacity of a WBAN are greatly affected by the likely attacks from security viewpoint. The threats and attacks to various layers of WBAN as mentioned in [23–28] are briefly explained in this section.

Denial of Service Attack (DoS): The adversary may electronically interface or damage the WBAN causing transfer of unnecessary data packets that might not be handled by the WBAN system. Consequently, the authorised users too might be restricted from using the required data. This causes a great extent of disruption of WBAN system [24, 26].

Physical Layer Attacks: The common attacks on the physical layer are Jamming and Tampering as explained below [23, 25].

a. *Jamming:* In this attack, the corresponding adversary blocks the entire network with the help of a few nodes which interfere with the radio signals used by WBAN. However, large communication networks cannot be blocked by this attack. As WBANs are small networks there is high probability of network blockage.
b. *Tampering:* In this attack, the confidential data of the patient can be extracted as the intruder gets physical access to the node leading to deterioration of the network.

Data Link Layer Attacks: These attacks are categorised as collision, unfairness and resource exhaustion [25, 28] as described below.

a. *Collision*: This is caused due to the transmission of multiple nodes at same time and frequency. To prevent this error, correcting codes are used.
b. *Unfairness*: The interruption of Medium Access Control (MAC) priority schemes, deteriorates the performance of the network.
c. *Resource Exhaustion*: The transfer of extra data packets over the channel by an adversary leads to the exhaustion of the reserved energy for retransmitting nodes.

Transport Layer Attacks: The primary attacks in the transport layers include flooding and de- synchronization [25].

a. *Flooding*: When a new connection is made by the adversary, it performs repeated requests while transmitting the control information, thereby leading to exhaustion in memory resources.
b. *De-synchronization*: The adversary forges the data between the body networks, and this leads WBAN to continue as an infinite cycle.

Network Layer Attacks: Routing is difficult in WBAN. As the body network co-ordinator (BNC) is connected in star topology, it is impossible to transmit the routing packets to another WBAN. However, routing is possible when multiple WBANs communicate with each other through BNCs [27, 28]. The prescribed routing attacks are as given below.

a. *Spoofing*: In this attack the intruder deliberately complicates the network by creating routing loops.
b. *Selective Forwarding*: A node is introduced by the intruder in the path of data flow which prevents the transmission of the packets.
c. *Hello Flood Attacks*: In this, the body networks are convinced by a node, with a highly powered antenna, that, it is their neighbour. Cryptography algorithm is implemented to prevent these attacks. It also includes other attacks namely Sink Hole, Sybil and Worm Holes aimed at obtaining illegal access to the network.

4 Conclusion

In this paper, we presented various challenges and probable security attacks commonly experienced in a WBAN. The review has brought together all possible challenges and security attacks, awareness of which shall be useful to incorporate necessary preventive measures in the design of efficient WBAN systems.

The design of WBAN system involves many challenges namely, management of cost and data, interference of the wireless link, consistent performance, sensor validation, consumption of power and so on. The maintenance of privacy and security of this system has been identified as one of the most challenging tasks. There is an urge to develop a trusted secure management system, in order to overcome the threats and attacks which invade the privacy of the convalescent and cause a great damage to the system.

The implementation of advanced cryptographic algorithms, authentication through iris recognition, finger print, face recognition, physiological signals like Electro-Cardiogram (ECG) and Photoplethysmogram (PPG) and, the appropriate use of the security paradigms prescribed by IEEE 802.15.6 protocols for the development of a reliable management system can lessen the security issues in WBAN to a great extent.

References

1. Matin MA, Islam MM (2012) Overview of wireless sensor networks. In: Intec open access. https://doi.org/10.5772/49376
2. Haupt J, Bajwa WU, Rabbat M, Nowak R (2008) Compressed sensing for networked data. Signal Process Mag 25:92–101
3. Balouchestani M, Raahemifar K, Krishnan S (2012) Wireless body area networks with compressed sensing theory. In: International conference on complex medical engineering
4. Zhen B, Bang Li H, Kohno R (2009) Networking issues in medical implant communications. Int J Multim Ubiquitous Eng 4(1)
5. Ramli SN, Ahmad R (2011) Surveying the wireless body area network in the realm of wireless communication. In: 7th International conference on information assurance and security (IAS), IEEE (2011)
6. Ha I (2015) Technologies and research trends in wireless body area networks for healthcare: a systematic literature review. Int J Distrib Sens Netw 2015. Article ID 573538
7. Zou S, Xu Y, Wang H, Li Z, Chen S, Hu B (2014) A survey on secure wireless body area networks. Hindawi Secur Commun Netw (2017). Article ID 3721234
8. Evangelin E, Sam D Wireless body area networks and its emerging technologies in real time applications. Int J Eng Sci Res Technol (2014). ISSN: 2277-9655
9. Movassaghi S, Abolhasan M, Lipman J, Smith D, Jamalipour A (2013) Wireless body area networks: a survey. IEEE Commun Surv Tuts 16(3):1658–1686
10. Lewis D(2008) 802.15.6 call for applications in body area networks response summary. 15-08-0407-05-0006
11. Yuce MR (2010) Implementation of wireless body area networks for healthcare systems. Sens Actuators, pp 116–129. https://doi.org/10.1016/j.sna.2010.06.004
12. Lai D, Begg RK, Palaniswami M (eds) (2011) Healthcare sensor networks: challenges towards practical implementation. In: ISBN 978-1-4398-2181-7

13. Toorani M (2015) On vulnerabilities of the security association in the IEEE 802.15.6 Standard. In: Financial cryptography and data security, 8976 pp 245260. https://doi.org/10.1007/978-3-662-48051-9_18. ISBN 978-3-662-48050-2
14. Garcia P (2011) A methodology for the deployment of sensor networks. IEEE Trans Knowl Data Eng, 11(4), December. https://www.ucc.ie/en/media/research/misl/2009publications/pervasive09.pdf
15. Patel M, Wang J (2010) Applications, challenges, and prospective in emerging body area networking technologies. IEEE Wirel Commun
16. Mcdonald N, Atkinson D, Khmelevsky Y, McMillan S (2016) Sport wearable biometric data encrypted emulation and storage in cloud. In: Proceedings IEEE Canadian conference of electrical and computer engineering (CCECE), May (2016), pp 1–4
17. Li M, Lou W, Ren K (2010) Data security and privacy in wireless body area networks. IEEE Wirel Commun 17(1):51–58
18. Chen Z, Hu H, Yu J (2015) Privacy-preserving large-scale location monitoring using Bluetooth low energy. In: Proceedings of international conference on mobile ad-hoc and sensor networks (MSN), pp 69–78
19. Chou C-Y, Chang E-J, Li H-T, Wu A-Y (2018) Low-complexity privacy-preserving compressive analysis using subspace-based dictionary for ECG telemonitoring system. IEEE Trans Biomed Circ Syst 12(4):801–811
20. Tarin C, Traver L, Cardona N (2009) Wireless body area networks for telemedical applications. ISSN 1889-8297/Waves (2009)
21. Wang G, Lu R, Guan YL (2019) Achieve privacy-preserving priority classification on patient health data in remote E-healthcare system. In: Special section on healthcare information technology for the extreme and remote environments, IEEE Access. https://doi.org/10.1109/ACCESS.2019.2891775
22. Sun R, Shi Z, Lu R, Qiao J, Shen X (2012) A lightweight key management scheme for 60 GHz WPAN. In: Proceedings of the international conference on wireless communications and signal processing (WCSP'12), pp 1–6 (2012)
23. Han ND, Han L, Tuan DM, In HP, Jo M (2014) A scheme for data confidentiality in cloud-assisted wireless body area networks. Inf Sci 284:157–166
24. Sen J (2012) Security in wireless sensor networks. In: Wireless sensor networks: current status and future trends
25. Venkatraman K, Vijay Daniel J, Murugaboopathi G (2013) Various attacks in wireless sensor network: survey. Int J Soft Comput Eng 3(1)
26. Wood AD, Stankovic JA (2002) Denial of service in sensor networks. IEEE Comput 35(10):54–62
27. Kang DO, Lee HJ, Ko EJ, Kang K, Lee J (2006) A wearable context aware system for ubiquitous healthcare. In: 28th IEEE EMBS annual international conference, New York, USA (2006), pp 5192–5195
28. Barakah DM, Ammad-uddin M (2012) A survey of challenges and applications of wireless body area network (WBAN) and role of a virtual doctor server in existing architecture. In: IEEE third international conference on intelligent systems modelling and simulation. https://doi.org/10.1109/ISMS.2012.108

Farm Management and Resource Optimization Using IoT

Pattlola Srinivas, M. Swami Das, and Y. L. Malathi Latha

Abstract The Internet of things is smartly connected to devices for processing of information. The agriculture immerses growth in India is 58% of people's primary source. World health organization target to provide quality food without shortage. The proposed IoT based smart farming system helps the farmer to optimize profit, sustainability, protection of land resources. The methodology users to use sensor-controlled IoT applications with Raspberry Pi and analyze data. We conducted case studies for soil moisture and temperature sensors. The experiments conducted using IoT—farm management and results will recommend that data soil in irrigation and efficient utilization of resources.

Keywords IoT · Smart farming · Quality of services · Resource optimization

1 Introduction

The Internet of things is smart connected devices to communicate and process the information. It uses physical smart devices, networks, sensors, objects containing data recording, storing and processing information to users [1]. Agriculture there will be immerse growth in the field of agriculture globally of 8.6 billion by 2020 [2]. In India, 58% of the population agriculture is the primary source of living. In India agriculture has huge growth and contribution to the world. India exports the grocery & food market in the world's sixth-largest in retail and contributing 70% of sales and in fruit production that is India's second-largest country in the world [3]. The present model of agriculture not suitable for sustainability due to future demands. World Health Organization (WHO) target to provide quality food without any shortage in

P. Srinivas (✉)
Department of CSE, Malla Reddy Engineering College (A), Hyderabad, Telangana, India

M. Swami Das
Department of CSE, CVR College of Engineering, Hyderabad, Telangana, India

Y. L. Malathi Latha
Department of Information Technology, Mahatma Gandhi Institute of Technology (MGIT), Gandipet, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_145

the future. Information technology helps the farmer in agriculture to optimum profit, sustainability and protection of land resources. It is clear that the food production and agriculture to use of IoT technology with AI and machine learning methods will give revolution to farmers. This requires to enhance the quality, optimum utilization of resources, like water and natural resources that optimization of crop yield by minimizing environmental side effects. IoT based Smart agriculture gives that the optimization crops by suggestions recommendations in crop based on data moisture level, soil quality and weather conditions.

Today the agriculture process is data-centric, precise and smart. The use of IoT based technology in agriculture will improve the resources based on demand and supply. The world population in the future will increase the use of food and crops. For example, food crops in the production of fuel, bio-energy and other industrial applications and security of food. The earth is sustainable and the use of agriculture is based on climate conditions like temperature, soil quality, topography and water. Quality nutrition crop site analysis of optimal production in specific crops done with the use of IoT technology. Land and crop conditions are detected and monitored with the use of the crops and detection harvesting. The types of equipment used tractors, robotic weeders, drones and satellites with sensors to collect the data to predict the key role of applications in agriculture [4]. Smart agriculture, the architecture is equipped to monitor the crops in various stages, based on the type of crop, sensors, vehicles, robotics that used to collect the information, storage, process and analyze the information [5]. The crops are not growing due to a lack of optimal climate and crops lost every year. The IoT based smart farming huge sensor-controlled real-time analysis and crop will predict based on past and present data. In Sect. 2 describes literature survey, Sect. 3 proposed model, Sect. 4 Results and Discussions, and Sect. 5 Conclusion and future scope.

2 Literature Survey

2.1 Related Work

Agriculture in India changes rapidly the use of IoT to solve real-time problems in farming. Smart farming gives away understanding and helps to increase production and profit in Indian agriculture. World second, products like wheat, onion, etc. Ram Kumar et al. [6] proposed smart agriculture cultivation using drones. Technological improvements in technology to use IoT smart farming. Sensors are used for rainfall, soil nutrition, tracking the product and Crop monitoring. The traditional system and manual methods, vegetables, fruited in the market affected by pesticides, Shafi et al. [7] proposed food production and monitoring with the use of IoT-Arduino will improve the health crop agriculture. Muangprathub et al. [8] proposed smart agriculture with the use of hardware, web applications and mobile apps which will monitor the crops. Koksai et al. [9] proposed a smart farming quality choice in functionality design as per requirements. He also conducted case studies in turkey, Konya for smart wheat production.

Nayyar et al. [10] proposed an Agriculture stick that assists farmers in live data such as soil moisture and temperature and used agriculture stick which is attached to the Arduino technology. Prakash et al. [11] proposed crop recommendation (SmartFarmNet) using sensors to provide performance analysis. Bhatnagar et al. [12] proposed a model soil health monitoring based on soil and temperature pH values using IoT based smartphone. Raut et al. [13] proposed a soil monitoring with the use of IoT and automation of irrigation with the supply of major nutrition. Chakrabarty et al. [14] proposed an IoT agriculture model using IoT that optimizes the resources and minimizes the loss to achieve sustainability. Pathak et al [15] proposed a cuckoo search method that is used to allocate water in farming conditions using IoT.

2.2 Problem Definition

IoT based crop monitoring system goal is to help to understand the process, learn, analyze data using soil moisture sensors and temperature sensors. Optimize the resources and preventive measures to be taken to improve the crops.

3 Proposed Model

3.1 Components—Functions

Radio Frequency Identification (RFID) Tag implant on the crop which read, store and send via the communication network to the application layer.

Sensors: IoT sensors networks used in Smart Farming growing the plants. The crop environment is most important, sensor devices support applications, accuracy monitoring parameters like temperature, light and water. IoT based products for smart irrigation, smart farming use of Wireless sensors, which play the most significant role to collect crop data based on conditions and requirements and send to the system for analysis. For example, the acoustic sensor used in soil cultivation, the optical sensor is used in light presence.

The ultrasonic sensor is used in spray coverage of water, optoelectronic sensors are used in unwanted plants, electrochemical sensors are used to identify electric response, mechanical sensors are used to test soil level, water level sensor, temperature sensor and remote sensing are used to locate geographical information for analyzing and managing the spatial agriculture information soil, water levels, humidity and other measurements. Management of Aid Pest (MAP), IoT sensors in real-time when crop grows to find the crops across the large area the devices which

Table 1 Smart agriculture use of sensors, services and applications

Sensors	Temp sensor	Humidity sensor	Leaf sensor	Fruit size sensor
Services	Soil preparation	Irrigation, fertilization	Pesticides	Yield condition
Applications	Water and nutrition monitoring	Machinery	Environment	crop conditions
Sensors	Temp sensor	Humidity sensor	Leaf sensor	Fruit size sensor

use the high-resolution of cameras that will regular data collection updates, apply pest management methods according to the best outcomes.

Water usage—Control: The water is the most important resource that helps the crops to increase soil health. Using IoT sensors to reduce the water level, farmers take action according to the sensors. Sprinklers are used for the usage of water according to the instructions by IoT based systems that will help to increase profit in cultivation. Based on the demand of the crops, using IoT based sensors to harvest and get the producer into the market and increase the productivity.

Animal Monitoring: IoT technology used for sheep, with the use of sensors and captures the data and growth of animal production. Climate—weather Prediction: according to the notification by the Government of India guidelines/reports in climate rain, winter and summer seasons. To check the humidity, temperature and other information during the growth of crops with the use of IoT based controlled systems. Growth in agriculture and helping farmers to predict and test and use of smart farming that will increase productivity. Table 1 describes smart agriculture, sensors service and applications.

IoT interconnected objects to collect data, store and exchange data using sensors and protection of collected data. Smart farming, with activities planting, applying fertilizers, monitoring machines with GPS connectivity to identify the activities. Climatic-weather and Robotics, Drones are used to exact land size, soil, sprinkling and harvesting robots, sprinkling robots, soil sensors, drones, food tools, bug detectors, food safety tools are used in smart farming. Drip, solar pumps, soil, crop monitoring, GPS and IoT devices are also used in smart farming. The input sensors are connected with network and GPS, this receives the data and sends it to the service and application layer. The service layer consists of design, implementation, APIs and specific support services. The application layer contains processed information with the received data inputs and outputs in smart farming.

3.2 Proposed Architecture

The Proposed Architecture Fig. 1 has user, sensors, IoT app and analysis and prediction system.

User is the farmers use of IoT based Smart Agriculture System. The users are connected to the IoT Management system connected to sensors with a raspberry pi

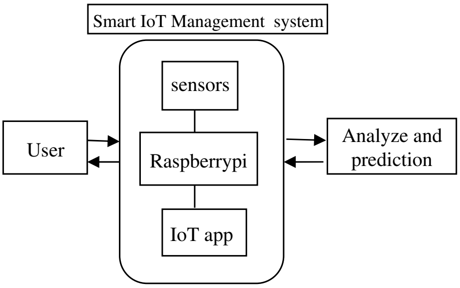


Fig. 1 Smart farming IoT-architecture

and IoT app. IoT devices connected to Wireless sensor systems and require integration for the heterogeneous system which has sensing, acting, communication and processing. IoT with machine learning applications with automating process and procedures are used for predicting the health of the crop. To predict the conditions of crop optimum type of soil, environmental conditions, the sensors and control of various activities, by reducing pesticides on crops to improve quality, prediction of soil moisture and data collected from Web Server nodes tested with data and analyze the results.

Hardware components: Raspberry Pi Board, sensors like RFID are used for identifying and tracking objects a small chip enabling communication. Bluetooth: It uses short-range devices with wireless radio connectivity, Wi-Fi provides to access wireless network access at a short distance with the use of the router. The sensors to read the humidity, temperature and soil moisture [16]. The complete system is an electronic circuit with software embedded in things and objects. The output of the sensor interfaced with Raspberry-pi and used analog to digital converter. Sensors are used to measure air quality, monitor sunlight plants. The analog to digital converter to read by raspberry pi with the implementation in Arduino. The software python is installed on the server and machine learning method to analyze the data. For example, we have used the temperature sensor LM 393. The proposed algorithm.1 describes to read the sensor data, predict the crop growth and analyze the data.

Crops: Indian crops like wheat, cotton, paddy, sugar cane, groundnut, millets, pulses. According to sensor data based on soil, weather results and predictions conditions for improvement in the crop. The data set with sensors are used with Raspberry and IoT app is connected to Smart agriculture system that is used to analyze and predictions using machine learning methods.

Algorithm:1 Read sensor data: Analyze and improve smart forming.

```
Input:  sensor data, temperature, soil data, humidity and
        water
Output: Predict the growth of the crop
  Step1. Read the data from sensors
  Step2. Data raspberry-pi to IoT app
  Step3. Analyze the data using machine learning methods
  Step4. Compute the results and appropriate actions
         according to the analysis results.
  Step5. Prediction and control signal to the farmer to
         improve smart farming.
  Step 6.End
```

The optimal solution for soil moisture content with the use of the devices Raspberry Pi set of resources with storage, ensure on/off of appliances and the model is significant to use of resources and crop status viewed with the highest priority. The application constant functionality, switches and devices are operated from home automatically information to the farmer and the appropriate decision can be taken based on results.

3.3 Smart Farming Methodology

Data aggregation, Design Analytics, Security and Maintainability are needed for the smart IoT agriculture system. Need for Transforming the traditional farming into IoT that use of interconnected objects are used to collect data, store and exchange data using sensors and protection of collected data. Smart farming with activities planting, applying fertilizers, monitoring through the machine with GPS to identify the set of activities like climatic and weather conditions.

Farm management using IoT is the best solution for sustainability agriculture in Indian Food Production. Challenges for farmers are precision farming, increasingly the productivity, preventive measures to soil degradation reduction of chemical usages in crops, Use of water resources (drip), farm quality, quantity and cost profit. Smart agriculture agribusiness predicts the climate, crops specific conditions and use of Smart agriculture contains three things hardware, mobile application, web application are shown in Fig. 2. The IoT applications daily using various domains

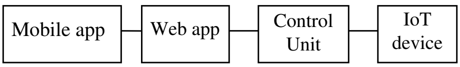


Fig. 2 Smart farming IoT-mobile app

like the hospital, and traveling. The optimization algorithms use WSN, harvest time, temperature, humidity and soil moisture data.

Agriculture data analysis analyzes, optimizes and predicts the results. In algorithm 2. describes the water leveling pseudo code.

Algorithm: 2: Water level – sensor use a Web app.

```
Input : X1, X2,.., Xm
Output: Watering on/off - manual selection or automatic
           selection
  Begin
  Step1. Read the input parameters, humidity, temperature
         and soil moisture
  Step2. Use of Web application function, crop1, crop2,
         crop3.
  Step3. Open App and select the crop, data analysis
         with a control box.
  Step4. Analyze the data and send results to the user
  Step5. The user selects the option to control the crop
         and switch on /off according to the instructions
End
```

3.4 Proposed Architecture Case Study—Example

The smartphone has an application that is used to control the water on/off with Automatic control or manual control. A web app is used to read the timely data from IoT sensors and each crop appropriate instructions. This uses Wi-Fi or GPS connectivity to manage the crop at various locations. The control unit is used to control the IoT data from sensors like waterproof hold sensors. The proposed system will notify the monitoring of water and notification. Soil Moisture program code is implemented in Raspberry Pi described in Program code1. and program code2. describes the temperature sensor data.

Program code1: Smart Farming- IoT sensor data for soil moisture.

```
#include <SoftwareSerial.h>
#include <LiquidCrystal.h>
LiquidCrystal lcd(13,12,11,10,9,8); //
rs,en,d4,d5,d6,d7
int soil = A0;
int temp =0 ; // sensor value
void setup() {
  // put your setup code here, to run once:
lcd.begin(16,2) ;
Serial.begin(9600) ;
pinMode(soil,INPUT);
lcd.print("soil moisture");
delay(1000);
}
void loop() {
  // put your main code here, to run repeatedly:
  temp= analogRead(A0);
  lcd.print("water level detection");
  lcd.setCursor(0,12);
  lcd.print("Soid =  ") ;
  delay(1000);
  lcd.clear();
  lcd.println(temp);
  Serial.print("Water level is ");;
  Serial.print(temp);
  lcd.clear();
}
Output
```

Program code2: Smart Farming- IoT sensor data for temperature // TEMPERATURE PROGRAM

```
#include <SoftwareSerial.h>
#include <LiquidCrystal.h>
LiquidCrystal lcd(13,12,11,10,9,8); //
rs,en,d4,d5,d6,d7
int temperature = A3;
int temp =0 ; // sensor value
void setup() {  // put your setup code here, to run
once:
lcd.begin(16,2) ;
Serial.begin(9600) ;
pinMode(temperature,INPUT);
lcd.print("temperature monitoring");
delay(1000);
}
void loop() {
  // put your main code here, to run repeatedly:
  temp= analogRead(A3);
  temp= temp/2.4;
   lcd.print("Room temperature");
  lcd.setCursor(0,1);
  lcd.print(" temperature is = ") ;
  delay(1000);
  lcd.clear();
  lcd.println(temp);
  Serial.print("temperature level is ");;
  Serial.print(temp);
  lcd.clear();
  if(temp > 34)
  {
     Serial.print(" Temperature is HIGH ");
    }
```

```
        if(temp < 29)
        {
          Serial.print(" Temperature is below room
    temperature ");

        }
       if (temp <34 && temp > 30)
        {
          Serial.print(" Temperature is MEDIUM ");
        }
    }
    Output:
```

4 Results and Discussions

The experimental results are conducted according to design Fig. 1. Smart farming IoT-architecture with Raspberry Pi and connected to soil moisture sensor as described in program code.1 and Fig. 3 are experimental results. In Program code2: Smart Farming-IoT sensor data for temperature and results are described in Fig. 4 and the Table 2 is the case study of input, process and output scenario.

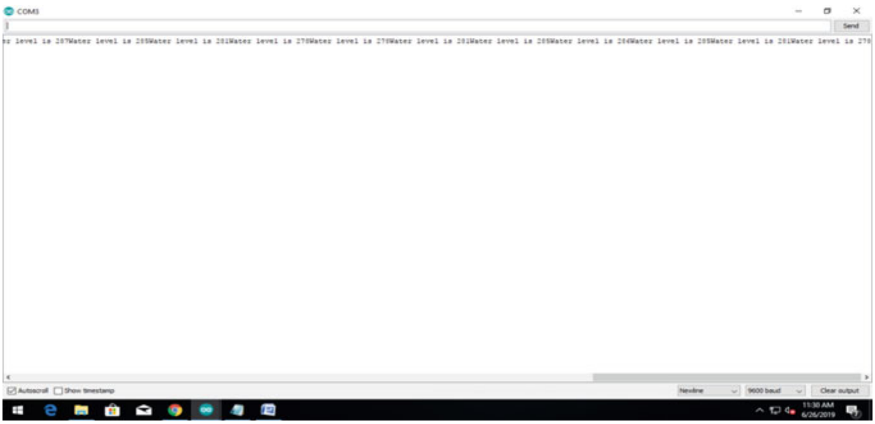


Fig. 3 The results of soil moisture and water level

Fig. 4 The results of the temperature sensor

Table 2 Inputs, processes and outputs in smart farming

Input	Process	Outputs
Crop data	Land preparation for cultivation	Quality of product crop
Data environment	Sewing with machines	Quality garment production
Data soil testing	Irrigation	Control the crops
Forecasting weather	Fertilization	Economy
Use of best practices	Without use pesticides	Future predictions of the quality of crops
Govt. policies	Guidelines to be followed to farmers	Future requirements applications

5 Conclusion and Future Scope

The advancements in technology, the use of IoT devices in every sector is easy and to improve the liveness of people. In India, most people had chosen the occupation of farming. Loss of income due to manual labour and no supervision to the farmers is provided at the right time. The proposed smart farming model will improve the crops according to data. Soil testing gives nutrients that help the growth of the crop. Moisture, temperature, water plays the most significant role in the growth of crops. The proposed system will be the best solution for smart farming using IoT and the efficient and optimize utilization of resources. In future, we can extend for smart farming-IoT to apply climatic changes, pollution and vertical farming with the use of robotic tractors.

References

1. Sucheta PSV, Vijaya Lakshmi D, Swami Das M (2018) Building an IoT web services for intelligent systems. Int J Pure Appl Math 120(6):3101–3117
2. https://www.softwebsolutions.com/resources/iot-solution-for-agriculture-industry.html
3. https://www.ibef.org/industry/agriculture-india.aspx
4. Ayaz M, Ammad-Uddin M, Sharif Z, Mansour A, Aggoune E-HM (2019) Internet-of-Things (IoT)-based smart agriculture: toward making the fields talk. Special section on new technologies for smart farming 4.0: research challenges and opportunities. IEEE Access, 7:129551–129583
5. Varghese R, Sharma S (2018) Affordable smart farming using IoT and machine learning an AI-powered cost-effective solution to improve traditional farming. In: Proceedings of the second international conference on intelligent computing and control systems (ICICCS 2018) IEEE, pp 645–650
6. Ram Kumar RP, Sanjeeva P, Vijay Kumar B (2018) Transforming the traditional farming into smart farming using drones. In: Proceedings of the second international conference on computational intelligence and informatics, advances in intelligent systems and computing, 712, Springer, pp 589–597
7. Shafi U, Mumataz R, Hassan SA, Zaidi SAR, Akhtar A, Malik MM (2020) Crop health monitoring using IoT enabled Precision agriculture. IoT Architectures, models and platforms for smart city applications. IGI Global, pp 134–152
8. Muangprathub J, Boonnam N, Kajornkasirat S, Lekbangpong N, Wanichsombat A, Nillaor P (2019) IoT and agriculture data analysis for the smart farm. Elsevier, Computers and Electronics in Agriculture 156, pp 467–474 (2019)
9. Koksal O, Tekinerdogan B (2019) Architecture design approach for IoT-based farm management information systems. Precis Agric 20:926–958
10. Nayyar A, Puri V (2019) Smart farming: IoT based smart sensors agriculture stick for live temperature and moisture monitoring using Arduino, cloud computing & solar technology. tailor and Francis, pp 1–10
11. Prakash P, Ali J, Dimitrios Y, Morshed GA, Zaslavsky A (2016) Internet of Things platform for smart farming: experiences and lessons learnt. Sensor, MDPI 16(11):1–17
12. Bhatnagar V, Chandra R (2020) IoT-based soil health monitoring and recommendation system. In: Pattnaik P, Kumar R, Pal S (eds) Internet of Things and analytics for agriculture, volume 2. Studies in Big Data, vol 67. Springer, Singapore, pp 1–21
13. Raut R, Varma H, Mulla C, Pawar VR (2018) Soil monitoring, fertigation, and irrigation system using IoT for agricultural application. In: Hu YC, Tiwari S, Mishra K, Trivedi M (eds) Intelligent communication and computational technologies. Lecture Notes in Networks and Systems, vol 19. Springer, Singapore, pp 67–73
14. Chakrabarty A, Mudang T (2020) Smart and sustainable agriculture through IoT interventions: improvisation, innovation, and implementation—an exploratory study. In: Pattnaik P, Kumar R, Pal S, Panda S (eds) IoT and analytics for agriculture. Studies in Big Data, vol 63. Springer, Singapore, pp 229–240
15. Pathak A, Uddin MA, Abedin MJ, Andersson K, Mustafa R, Hossa MS (2019) IoT based smart system to support agricultural parameters: case study A. Elsevier, Procedia Computer Science, pp 648–653
16. Rehman A, Mehmood K, Baksh A (2013) Communication technology that suits IoT—a critical review. WSN4DC 2013, Springer, pp 14–25

A Model of Women Security Using IoT

M. Swami Das, A. Govardhan, and D. Vijaya Lakshmi

Abstract India is the second-largest populated country in the world, in which probably 40–48% of people are women. Everyday women and children walk life struggling to save and protect their lives due to violence, harassment, and crimes. It is necessary to provide women security. Technological advancements in Internet of Things brings Women security app in the smart phone will automatically intimate to police when womens are in an emergency. The proposed model women security using IoT has consisted of the user, mobile app with sensors, communication and message reporting system will be the best solution. This provides women security using sensor data, analysis and message reporting and notification to police in critical conditions, alert message and save the lives of women. The proposed model is implemented and applicable for women security integrated to live monitoring system. In future the IoT based applications will be used other domains like healthcare, military, surveillance and pilotless planes applications.

Keywords IoT · Women security · Smart phone

1 Introduction

The India population in 2011 is 121 crores, in male population is 62.3 crores (51.5%) and female population is 48.6 crores (48.42%) [1]. According to Statistics time, the Indian population is 1.37 billion in 2019. India is the second most populated country in the world [2]. It is good to know that the womens are working from CAB drivers to CEOS of top MNC companies. Everyday womens all walk in their lifes struggling to save and protect themselves, due to violence, harassment against womens and

M. Swami Das (✉)
Department of Computer Science and Engineering, CVR College of Engineering , Hyderabad, Telangana, India

A. Govardhan
JNTUH, Hyderabad, Telangana, India

D. Vijaya Lakshmi
MGIT, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_146

childrens. Women safety is required in public Transportation. By considering the issues of women teasing, in critical situations, womens and girls. In order to provide and improve women safety intelligent based smart solutions are needed [3]. The government providing the right to the people, right as a human being and employment opportunities, Choice of freedom, job, security, protection and safety. IoT plays an important role in daily applications like home automation, industry, medical, etc. Applying IoT based technology to provide security to women's and children's. Intimate the location when the womens are at risk through the web-based application. Women's and children safety is the most challenging task due to the crimes. When Women are at risk, the intelligent decision-making system (or security app) in the smart phone will automatically intimate to find the friend, chosen people and police alert in emergency situations. Protector and police Internet of things is information is received and send store and process and analyze the information. The paper is organized Sect. 2, a literature survey, Sect. 3 contains the proposed model, Sect. 4 discussions and Sect. 5 conclusion and future scope.

2 Literature Survey

2.1 Related Work

Protection is provided to females using Arduino with IoT based system that pings location tracks the movements and sends the location to authorities. By maintaining safety activities when womens feels in danger [4]. According to Helen et al. [5] Women are facing harassments, teasing in the society, apps and devices like smart watch can detect the heartbeat of concerned chosen women and automatically call to police/registered mobile from their contacts. The safety smart phone, apps VithU app will alert list of contacts with the message, track location and send information to chosen people. The disadvantage of this is click button twice consecutively. According to Seelam et al. [6] Women safety-security in critical conditions, they proposed a smart security app and women security app using Arduino with sensors. The sensors track and detect the movements, locations by analysis and suggest necessary steps. According to Singh et al. [7] studied security for women on campus (University Educational institution) and provide security to students with the help RFID, vehicle number, camera and message alert systems.

According to Sogi et al. [8] proposed a women safety using Raspberry Pi and IoT. The security using wearable devices consists of sensors that used to track the location which will be sent to the communicated device using GPS connected in a smart phone. According to Monisha et al. [9] used ARM controller and Android application with smart phones using Bluetooth connectivity that sends every two minutes task the live status using a hidden camera that detects the movements in pictures. According to Harikiran et al. [10] proposed IoT App a solution for women safety and security includes the human behavior and reactions and the signal sends to GPS with an emergency message to police and public, then victim persons will be caught by the police and save a life. According to Nandini Priyanka et al. [11] surveillance, safety among children, girls and women's, where smart use of IoT application alert and send information to police in emergency and risk situations. Edna et al. [12] studied security attacks and IoT based applications that provide women security. Pampapathi et al. [13] proposed a smart security application for women and safety. Human behavior reactions of different situations, GPS, messages report the location to the nearest police station. According to Sehgal et al. [14], Human security is most important. Framework with the Internet of Things with equipment devices that is physical devices like smart watch, ear rings, cloth, smart phone and etc. which tracks the details and process the data. According to Sethi et al. [15] proposed Women safety using Arduino microcontroller, event face, in any danger situations of women and ping the location to authorizer using GPS user location.

2.2 Problem Definition

Design and develop the model, IoT based Women security, which helps the women are a safeguard in critical and emergency situations and report to the required persons like police, parents and people.

3 Proposed Model

Internet of Things is a communication of devices using sensors. IoT as a global network with configuration based on communication protocols used to analyze security devices that having storage, computing ability, and power supply. This provides features such as IoT threats, features that the impact security and privacy and IoT features to analyze the data respond accordingly. The proposed Women based IoT security model has components, functions, architecture, methodology and case study application.

3.1 Components—Functions

The framework IoT based women security consists of basic elements, Sensors with data, global network with configuring based on communication protocols, physical virtual identities, Communication, software-algorithms and goal-oriented data analysis. Hardware smart sensors like power and energy, battery, security, privacy, networks a combination of multi-key with the use of micro-operating system, Women security app with functional watch using Bluetooth and sensor. The GPS and Internet facility of the mobile phone with the use of sensor a GSM/GPRS, Motion sensors are used in IoT Women security-applications.

3.2 Women Security—Architecture

The proposed model is shown in Fig. 1 which has users holding Mobiles with a Women security app that has a motion sensor, microcontroller and analysis and this is a connected communication system that uses GSM/GPRS and reports to the required people in the emergency message like audio, SMS and location.

Users: The women are using smart phones equipped with the Women security app which provides the functions in an emergency, like sending a message to police and location.

Mobile App: The smart phone consists of a Women security app that is always running and reads the sensor data like Women's critical situations Heartbeat of a person high sensor motion to detection using Smart band integrated with Smartphone. Women in an emergency the smart phone safety—security track the history, walking and physical harassment.

Sensors: Smart Sensor which finds the heartbeat and tracking and monitors the system and based on active threshold value to send data to system for analysis. Smartphone consisting of Bluetooth device and sensors identify the behavior of persons and intelligent predicting and reporting system in emergency situations.

Communication protocols: The communication as an interface to device and phone. Continuously to monitor, pre-installed phone. The device is programmed with the hand from connected to GPS, user location to recognize the user movements and it sends the message to the family, message to police and Information to the public.

Message and Reporting: Provide security for women using Human safety. Embedded Electronic Device that Helps to find the victim. Security, Alarm and send the SMS to police. The implementation methodology is described in Sect. 3.3 and algorithm 1.

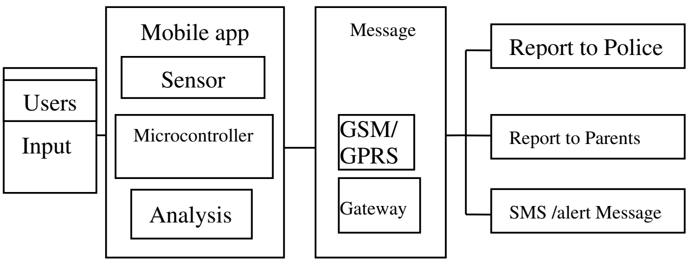


Fig. 1 A model of women security using IoT

3.3 Methodology—Pseudo Code

Algorithm:1: IoT based Women security

```
Input:
Read the parameters X1, X2,…, Xn( Sensor data)

Output: Emergency status of the condition, and send
        information based on the decision. 1. Police
         and2. Parents
Begin

Step1:  Read the input parameters, temperature, heart
        beat, voice, and location by surrounding
        people.

Step2: IoT app is embedded with rules -in IoT app
       Security decision-based on complexity and
       Urgency according to the real-time parameters.
       Based on rules and input data, decision can be
       Safe or unsafe

Step3: Based on the analysis report, send information
       and appropriate service to police and parents

Step4. The reporting system has to be saved the data
       and send it to the police

Step5. If safe track the details to parents.
       Emergency situations provide appropriate help to
       women based on data analysis
End
```

The summary of the women web security model and its functions are described in Table 1.

3.4 Women Security—Case Study Example

The smart phone is equipped with Voice recognition sensors, and the behavior of a person's sensors, and Women security app. The Women security app continuously tracks the sensor data and evaluates the analysis risk or safe using the proposed model and algorithm.1. Find the behavior of sensor and voice data assumed that find risk and emergency during analysis and then, the Women security app function automatically reports a voice call and sends live location and alert appropriate actions to police. The victim persons are tracked and caught by police and save the life of women based on the report.

Table 1 IoT based women security process

Steps	Elements	Methods applied	Insight
1. Collect sensor data	Heartbeat sensor, temperature sensor, motion sensor. garment, embedded with the sensor, body environment sensor, wrist device behavior of persons and continuously observing and monitoring	Sensors with pre-installed control unit helps to collect and information used in feature selection	To collect the data required for analysis
2. Mobile app	Women security app, easy to use, an app that integrated with the smartphone and creating and developing an app which ensures women safety in emergency	Objectives single click a girl in trouble by pressing a button or function automatically take action based on sensor data analysis in emergency	After analyzing sensor data this app sends messages to police people nearby request immediate action. Public attention. Alarm send message to Police station
3. Communication system	GPRS/GSM sim memory	Get the data from Sensor periodically track the location, when a sensor is activated according to the data analysis in Mobile app send information in an emergency to police	Send the information regular interval of time, according to the analysis of data in the Mobile app when emergency use communication protocols to police, parents and SMS
4. Reporting system	Actions by IoT women security system when an emergency	Inform to the nearby police station in critical, situation send a message, alarm, immediate notification to public alert or contact persons chosen mobile numbers	Human interactions precautions, countermeasures and trigger alarm and safety major implementation in women security track and save a life

4 Discussions

The IoT based women control security is used for detection and provides security, root cause clearly is to identify women in critical and dangerous situations and the system will give interface application with functional features, women safety media broadcast information, and preventive measures, precautions, suggestions based on input data sensors, Mobile apps, communication system and reports to police and ensure women safety in emergency. It is continually observing and monitoring persons based on behavior sensors, motion sensors, analyze and identify information. This Women security app sends message to police station and parents.

5 Conclusion and Future Scope

The Women Security app using IoT, its evaluation and model use of sensors, network communication, and processing. The framework collects the data using sensors and analyse the data and sends the results based on confirmation to appropriate services for women security. In an emergency when smart phone-Women security is a model that has been implemented using IoT. In urban and rural areas, women and children use this application. It is easy to carry and is integrated with functional location tracking, notification, image capture, location share location to police when an emergency. IoT based applications also widely used in cyber attacks, military, industry, agriculture, healthcare, twitter and pilotless airplanes.

References

1. https://en.wikipedia.org/wiki/Demographics_of_India
2. http://statisticstimes.com/demographics/population-of-india.php
3. https://timesofindia.indiatimes.com/blogs/the-rock-bottom/women-safety-in-india/
4. METIS (2020) Mobile and wireless communications enablers for the twenty-twenty information society, Online at https://www.metis2020.com
5. Helen A, Fathima M, Rijwana, Kalaiselvi (2017) A smart watch for women security based on IoT concept 'watch me'. IEEE, pp 190–194
6. Seelam K, Prasanti K A novel approach to provide protection for women by using smart security device. In: 2nd international conference on inventive systems and control (ICISC), Coimbatore, pp 351–357 (2018)
7. Singh V, Kharat V (2017) A proposed system for security in campuses using IoT platform: a case study of a women's university. ICCTCEEC-2017), IEEE, pp 305–310
8. Sogi NR, Chatterjee P, Nethra U, Suma V (2018) SMARISA: a Raspberry Pi based smart ring for women safety using IoT. In: Proceedings of the international conference on inventive research in computing applications (ICIRCA 2018), IEEE Xplore, pp 451–454
9. Monisha DG, Monisha M, Pavithra G, Subhashini R (2016) Women safety device and application-FEMME. Indian J Sci Technol 9(10):1–6
10. Harikiran GC, Menasinkai K, Menasinkai K (2016) Smart security solution for women based on Internet of Things (IoT). In: International conference on electrical, electronics, and optimization techniques (ICEEOT), IEEE, pp 3551–3554 (2016)
11. Nandini Priyanka M, Murugan S, Srinivas KNH, Sarveswararao TDS, Kusuma Kumari E (2019) Smart IOT device for child safety and tracking. Int J Innov Technol Explor Eng (IJITEE) 8(8):1791–1795. ISSN: 2278-3075
12. Conway E (2019) IoT: is it a digital highway to security attacks. Springer, Women Securing the Future with TIPPSS for IoT, Women Engineering Sciences, pp 1–11
13. Pampapathi BM, Singh K, Madhavi V, Yallaraddi MB, Desai M (2018) Smart band for women safety using Internet of Things (IoT), IJARCCE 7(3):120–123

14. Sehgal VK, Patrick A, Soni A, Rajput L (2015) Smart human security framework using the Internet of Things, cloud and fog computing. In: R. Buyya and S.M. Thampi (eds) Intelligent distributed computing, advances in intelligent systems and computing, 321, pp 251–263
15. Sethi P, Juneja L, Gupta P, Pandey KK (2017) Safe sole distress alarm system for female security using IoT. In: Somani AK et al (eds) Proceedings of first international conference on smart system, innovations and computing, smart innovation, systems and technologies 79, Springer, pp 863–874 (2017)

Analysis of Soft Computing Techniques to Forecast Mechanical Attributes of an Alloy

N. Sandhya, M. Rajasekar, and V. Sowmya

Abstract Soft computing uses data science techniques that include machine learning algorithms, which being the cutting-edge technologies in the current market aided people from different sectors to select materials or evolve the extant materials for products with a purpose of prolongedly modelling improvements. The insights of the mechanical attributes of an alloy plays a vital role in any engineering application. The mechanical attributes if predicted properly, will guide through the decision making on use of the alloys for the product, based on the suitability parameters. Stainless steel being versatile by nature and environment friendly is being widely used in industries. Its applications include architecture, building, construction, chemical, processing and oil & gas industries etc. This research targets at foreseeing the mechanical properties of an alloy named steel, by executing different data science algorithms and also this study can be further enhances to analyse the efficiency of the applied algorithms. A practical and feasible tool is developed having a Graphical User Interface integrated with the implemented algorithms that are compared to check for decent prediction accuracy. AICTE grant was sanctioned under RPS scheme ("Development of a tool to predict mechanical properties of steel by exploring data science techniques" project grant sanctioned by All India Council for Technical Education under Research Promotion Scheme, 1) and supported the research work that is being presented here. The work here targets at a user-friendly GUI based model for predicting tensile strength, yield point by specifying input considerations of stainless steel like carbon content, sectional size, temperature and manufacturing process using the data science algorithms.

Keywords Analytical study · Soft computing techniques · Graphical user interface model · Manufacturing process · Mechanical properties of steel

N. Sandhya (✉) · M. Rajasekar · V. Sowmya
Department of CSE, VNRVJIET, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_147

1 Introduction

The work here is useful in material science engineering to understand the mechanical attributes of an alloy depending on its chemical constituents and processing parameters. The researchers, students, manufacturers, are depending on the results obtained in the traditional laboratory experiments. To select a suitable material, they need to obtain some desirable properties for the material. For that they need to customize the chemical composition, processing parameters till the desirable properties are achieved [2]. This is an iterative process which demands massive expenditure and time. So the research here aims to lessen the time, cost involved in the process and automate the process of finding the mechanical properties of steel with data science algorithms through one click.

Steelmaking is done from the metallic ore pig iron holding 0.1–1.5% of carbon in it. It is manufactured by burning out carbon and some alloying elements. These alloying elements are different based on the type of steel. To obtain the desired quality of mechanical properties the carbon content is adjusted. Steel is manufactured in different processes like bessemer process, open hearth process, electric arc process, L.D process. The two types of open hearth processes are: acid open hearth and basic open hearth process [3]. When the shallow hearth in the open-hearth process is aligned with silica then the process is called the acid open hearth process. The process is said to be basic open hearth when the hearth is aligned with lime and magnesia. In our research three manufacturing processes namely acid, basic open hearth process and basic electric arc process were considered. A small steel piece with desired chemical composition is manufactured by any of the manufacturing processes with heat treatment at a certain temperature. This test piece has some sectional size i.e. diameter of the manufactured test piece. Keeping all the above context in mind, the models in this research are trained considering the features such as the carbon content in the steel, heat treatment temperature, manufacturing process, and the sectional size of the steel piece manufactured.

A methodology which foresees the mechanical properties of hydrogen charged aluminum alloys using artificial neural network in view of composition and processing parameters was proposed by Thankachan et al. [4]. Weng et al. [5] executed Single Index model to predict tensile strength, yield strength. Somkuwar et al. [6] implemented RBF and back propagation techniques for envisaging hardness of low carbon steel. Xu et al. [7] used convolutional neural network models for predicting mechanical attributes of hot rolled steel. Fragassa et al. [8] has implemented random forest, neural network, knn to predict ultimate tensile stress and strain, young's modulus, yield strength of cast alloys. Guo et al. [9] has proposed an approach where based on the alloy content tensile, yield strength and elongation of steel are predicted. Santos et al. [10] has proposed a method where the ultimate tensile strength of materials in foundry production is predicted using neural network and bayesian approaches. Shirish Deo et al. [11] has implemented a clustering regression model to predict the concrete's compressive strength.

2 Classification of Steels

Steel is an important metal that is used extensively, and manufactured in bulk considering the price too. The study here investigates the steel relating to British Standards (En 21 to En 39) to research into the mechanical properties prediction using techniques from Data Science. Iron and steel are widely used in many applications like road construction, construction of buildings, infrastructure, cars, refrigerators and other appliances. A metal that is adaptable to tune its composition and structure so as to mould the properties to satisfy the requirements is steel. Steel is a mixture of iron, carbon made from iron ore that is mined from the ground and smelt in blast furnace. The carbon content in the steel ranges from 0.05 to 2.1%. Different types of steel are available [12].

Tensile test is performed by placing a sample test specimen of material between two fixtures called grips of UTM machine. This test specimen has some specific sectional size which is its diameter. Now load is applied on the material that is gripped at one end while the other end is fixed. Gradually the load is increased and whenever the material starts stretching that point is called as yield point. Tensile strength is the point where the material can withstand maximum load before it breaks. Limitations of this approach are that we can't perform tests at different temperatures for one sample. It requires massive investment of time and expenditure. Still there will be no proper idea in material selection based on the mechanical properties. Moreover, in this tensile testing conducted by UTM processing parameters are to be adjusted and retested again until the user acquires desired mechanical properties of steel.

3 Related Work

The kind of steel selection for any kind of application is based on the mechanical properties held by it. The significant properties for any steel to be considered are tensile strength and yield point. The traditional method of conducting tensile tests in UTM were used to find the above two values of steel.

Universal Testing Machine (UTM) is the most commonly used traditional approach to find out tensile strength, yield point and strength of materials. Tensile test is performed in a universal testing machine by placing the steel test piece to a clamp and pull it apart until it breaks. In this process of pulling the steel test piece strength and stiffness of steel test piece can be determined. This tensile test helps in finding out the tensile strength and yield point of specific steel test pieces using UTM i.e. gradually the load is increased on the steel test piece the point where the material starts stretching is called yield point of the respective material and the point where the material can withstand maximum load before it breaks is called tensile strength of the material. But this traditional process is a time taking procedure which requires manpower and money. Moreover, in this tensile testing conducted by UTM processing parameters are to be adjusted and retested again until the user acquires

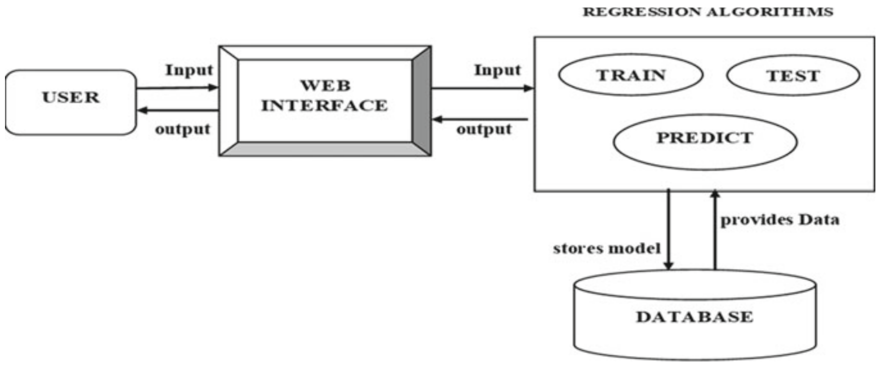


Fig. 1 System architecture of steel property predictor

desired properties of steel. The target of this work leads to conserving time, money, manpower and acquire the desired mechanical properties using the proposed tool which is integrated with data science algorithms.

4 Proposed Method

The proposed solution automates the traditional approach of finding mechanical properties of steel by incorporating various algorithms from data science. To implement the proposed methodology initially a relative study of various data science techniques experimented in the area of material science is done. Later various algorithms like random forest, neural network, support vector machine, decision tree, boosting are implemented to predict mechanical properties of British standard stainless steels with respect to its carbon content, sectional size, heat treatment temperature and manufacturing process. Later a user-friendly interface incorporated with efficient and accurate data science algorithms is developed to predict tensile strength and yield point. The features considered as the inputs are carbon content, sectional size, temperature in heat treatment, manufacturing processes Acid Open Hearth (AOH), Basic Open Hearth (BOH), Basic Electric Arc (BEA). The resultant predicted properties are tensile strength and yield point. Figure 1 describes the system architecture of the stated method.

5 Data Science Algorithms

Models can be built to envisage the steel mechanical properties considering carbon content, sectional size, temperature and manufacture process using various predictive methods. Random forest, Support Vector Machine (SVM), Artificial Neural Network

(ANN), Decision Tree, Ensemble Method (boosting) are the various predictive methods used to construct models for foreseeing mechanical properties of steel.

5.1 Random Forest

An augmentation to decision trees and in alike with ensemble method, bagging is Random Forest. A technique which can perform classification as well as regression. Multiple decision trees are generated during the phase of training, and every data-point is sent across all trees to produce the mean predictions that are combined to create its own prediction. This technique shuns the problem of overfitting. Random Forest regression is used in this study to build the model to predict the objective variables. 85% collected data is used for training and 15% is consumed for testing. The performance of the constructed model is assessed using the correlation coefficient of actual values and predicted values of test data.

5.2 Support Vector Machine (SVM)

A supervised technique to handle categorical and continuous variables for classification and regression is Support Vector Machine. A very well known kernel trick method to manage nonlinear input spaces is SVM. Here a plotting of each feature in the n-dimensional space is performed using the SVM algorithm. The different types of SVM kernels are sigmoid, radial basis function (RBF), polynomial and linear [13, 14]. From the composed data, in order to train the model 85% data is used and the remaining data is utilized for testing. The work here concentrates on nu—SVM regression model with radial kernel. By calculating the coefficient of correlation performance assessment of the trained model is accomplished.

5.3 Decision Tree

A supervised method that offers classification and regression techniques is decision tree. It is based on a tree like model for deciding upon decisions. Decision trees cater better results for categorical and continuous data. Attribute identification to decide upon the attribute that acts as the root at different levels is the major demand that decision trees resolve. Thus, attribute selection plays a crucial role in decision tree algorithm. ID3, C4.5, CART, CHAID, MARS, conditional inference trees, etc. are some variants of the decision tree method involved in problem solving. The approach here is based on the split of the data considered as: 85% data for training the model and 15% of data is used for testing. The accuracy of the model is experimented

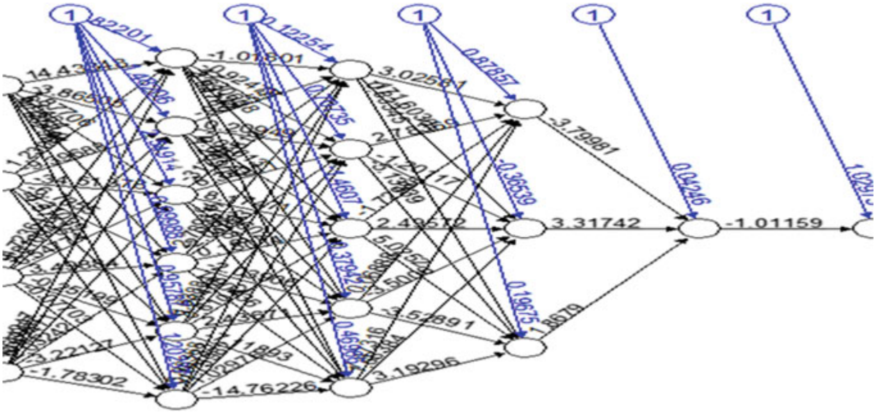


Fig. 2 ANN model designed for predicting tensile strength

further by examining correlation coefficient of test data values and predicted data values.

5.4 Artificial Neural Network (ANN)

Another approach towards classification as well as regression is neural network algorithm [15–22]. The design of neural network here includes 3 layers, the input layer, hidden layer and the output layer. In the proposed method two neural network models were developed, one for tensile strength and one for yield point. The research here involved to design a model to foresee the tensile strength values with 85% of training data and 15% of testing data. The architecture of the neural network used in this study consists of one input layer, 4 hidden layers and an output layer. To start with there are 4 neurons at the input stage, the hidden layers have varying neurons at each stage first hidden layer consists of 6 neurons, the second hidden layer has 5 neurons, third hidden layer consists of 3 neurons, fourth layer has one neuron and the target layer with one neuron. The output layer stage is wherein the model decides upon the value of the target variable tensile strength. The presented ANN model of this study is illustrated below (Figs. 2 and 3).

5.5 Ensemble Method (Boosting)

Ensemble techniques combine multiple models for improving the accuracy of predicting to yield better results. Boosting is one such technique that is used for improving weak models by enhancing their performance and converting them to

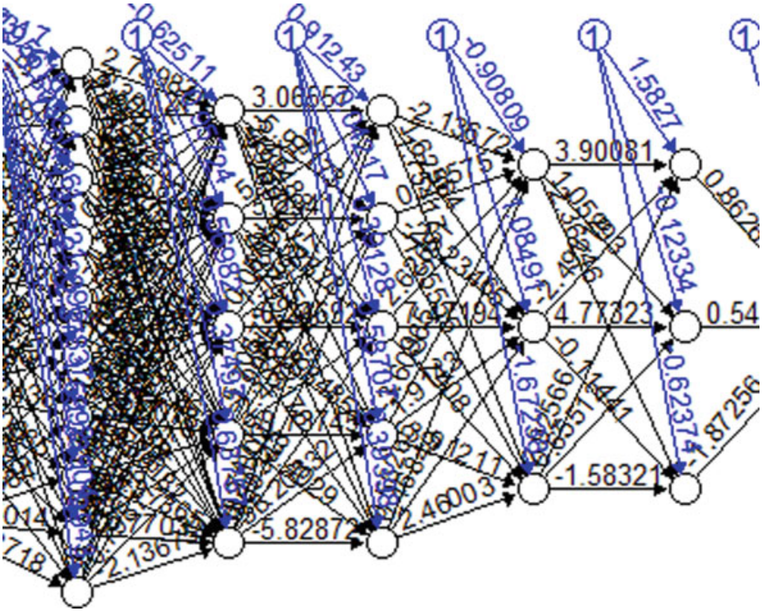


Fig. 3 ANN model designed for predicting yield point

strong models. This method increases bias while decreasing the variance. This research is experimented by treating 85% of data for training purpose and 15% for testing as part of the implemented gradient boosting machine algorithm. 15,000 and 18,000 trees are used for prediction.

6 Experiment and Evaluation

The section here concentrates on the quality evaluation of developed models used in foreseeing tensile strength and yield point. The competence of the models used is estimated by authenticating the predictions on the test data. Correlation Coefficient (R) is the statistical parameter utilized to assess the various built models using regression based prediction techniques. Correlation Coefficient ranges between −1 and +1. Strong negative correlation between the two variables is given as −1, If there is no association between the two variables then R value is 0, strong positive correlation between the two variables is given as 1. The equation (1) below represents the coefficient of correlation:

$$\boldsymbol{R} = \frac{\sum_{i=1}^{i=N}\left(x_{e}^{i} - \underline{x}_{e}\right)\left(x_{p}^{i} - \underline{x}_{p}\right)}{\sqrt{\sum_{i=1}^{i=N}\left(x_{e}^{i} - \underline{x}_{e}\right)^{2}\sum_{i=1}^{i=N}\left(x_{p}^{i} - \underline{x}_{p}\right)^{2}}} \tag{1}$$

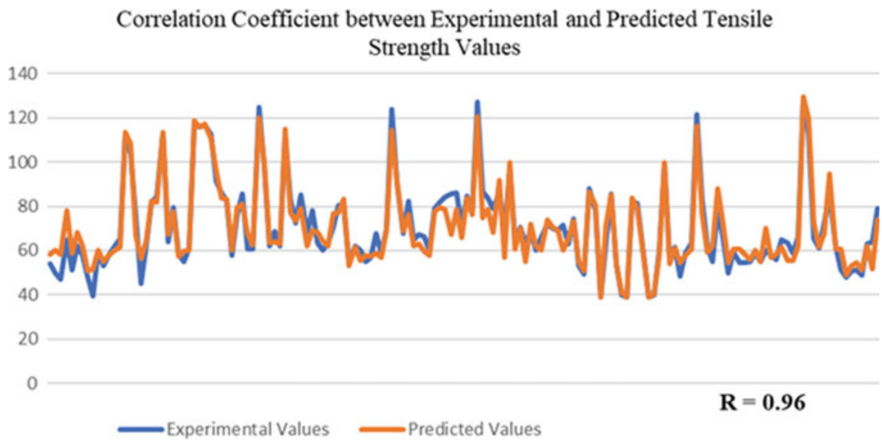


Fig. 4 Correlation for experimental and predicted values of tensile strength

In the above formula xe is experimental value, xp is predicted value, $\underline{x}_e$ is mean of xe values. and $\underline{x}_p$ is mean of xp values. N is the total number of values in the data set.

The correlation coefficient represents how accuracy and efficiency of the implemented model. The regression plots shown below represent the relation between predicted and experimental values.

6.1 Random Forest

Figures 4 and 5 are the regression plots which determines the correlation coefficient of experimental, predicted values.

6.2 Support Vector Machine

A regression plot is demonstrated below that depicted the R value as 0.95 and yield point values as 0.86 (Figs. 6 and 7).

6.3 ANN

The implemented neural network models is corroborated by employing coefficient of correlation which are shown in Figures 8 and 9.

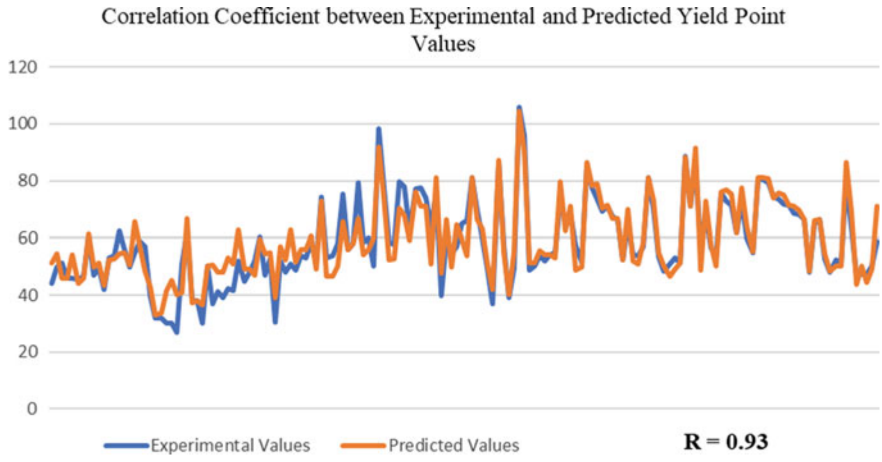


Fig. 5 Correlation for experimental and predicted values of yield point

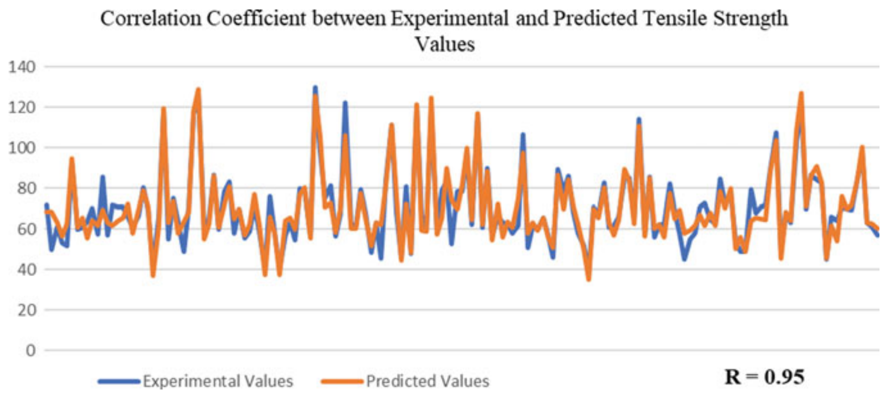


Fig. 6 Correlation for experimental and predicted values of tensile strength

6.4 Ensemble (Boosting)

Implemented boosting techniques for predicting mechanical properties are evaluated by statistical measure coefficient of correlation (Figs. 10 and 11).

6.5 Decision Tree

The below graphs signify decision tree models predicting tensile strength, yield point values relating to the experimental values (Figs. 12 and 13).

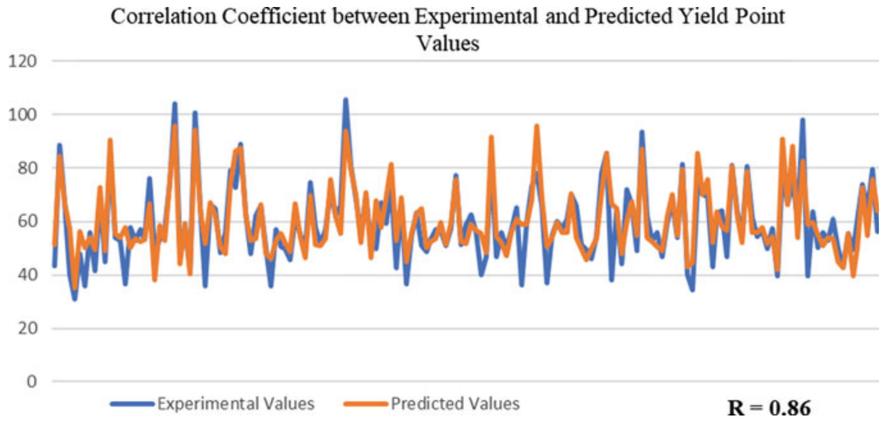


Fig. 7 Correlation for experimental and predicted values of yield point

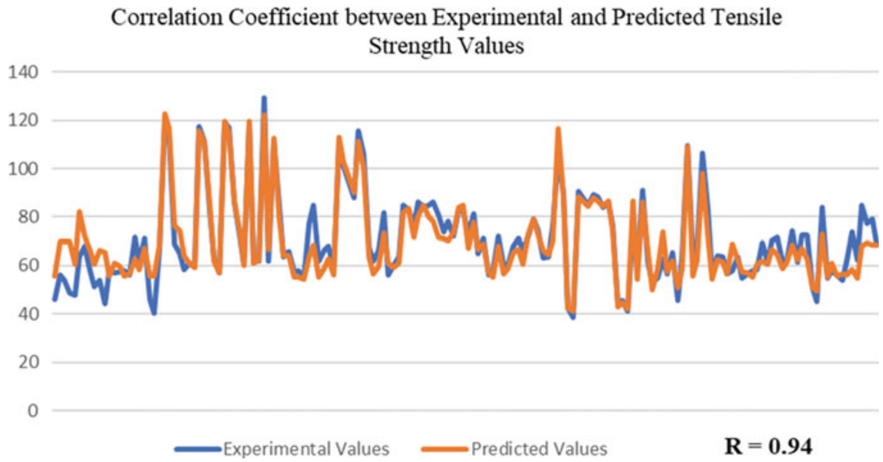


Fig. 8 Correlation for predicted & experimental tensile strength values

6.6 Result Analysis

Figures 14 and 15 show the accuracy of implemented techniques and the algorithm that predicts tensile strength and yield point values accurately. The algorithm that proved efficient in foreknowing the properties of steel is Random forest model. An accuracy of 96.3% for tensile strength and 93% of accuracy for yield point is foreseen using Random Forest.

So, Random Forest algorithm is the accurate algorithm incorporated with the GUI (Graphical User Interface) to make further predictions requested by users.

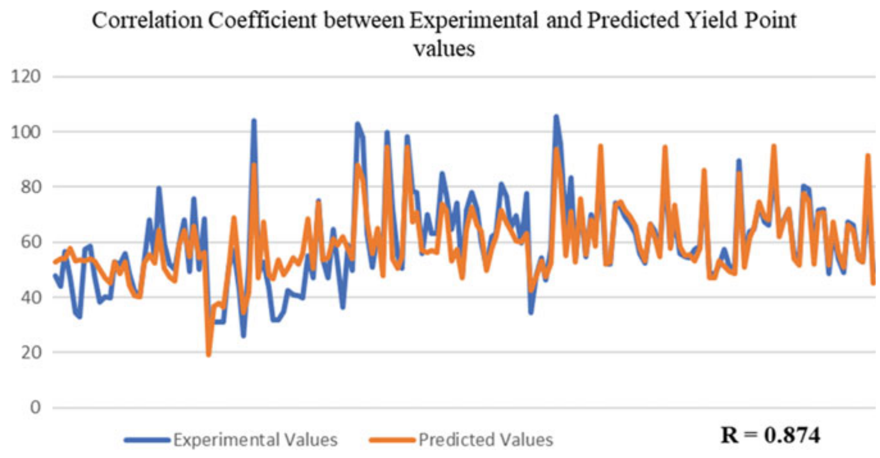


Fig. 9 Correlation for experimental and predicted values of yield point

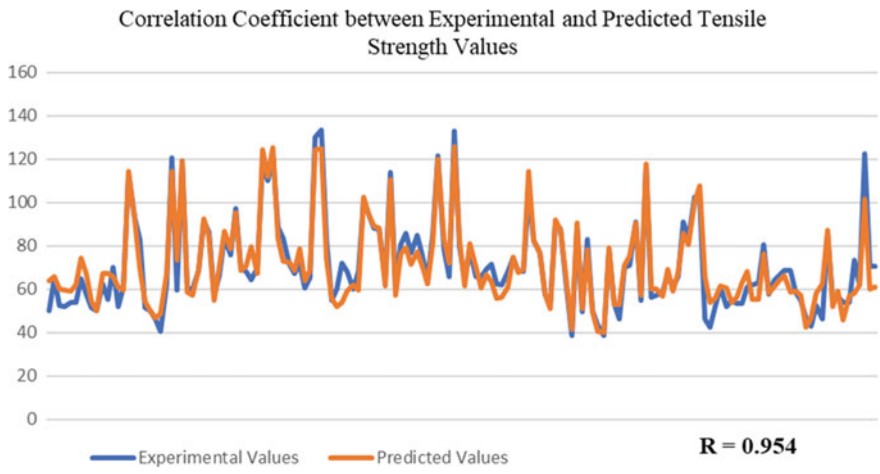


Fig. 10 Correlation for experimental and predicted values of tensile strength

6.7 Graphical User Interface (GUI)

The developed GUI with the name "STEEL PROPERTY PREDICTOR" is shown in Figure 16 with input values. The user needs to select the input values for which the GUI incorporated with developed data science algorithm or technique Random Forest predicts tensile strength and yield point for the input parameters given by the user. There are three manufacturing processes for which mechanical properties of stainless steel can be predicted they are AOH, BOH, BEA. The properties vary depending on the process of manufacture.

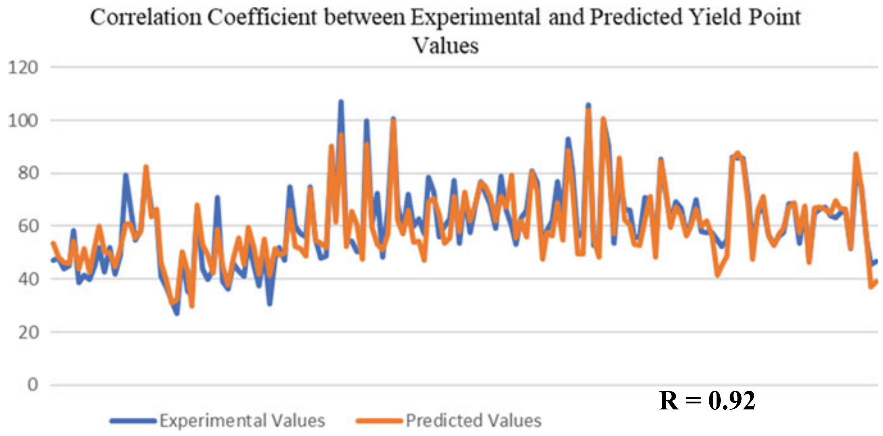


Fig. 11 Correlation for experimental and predicted values of yield point

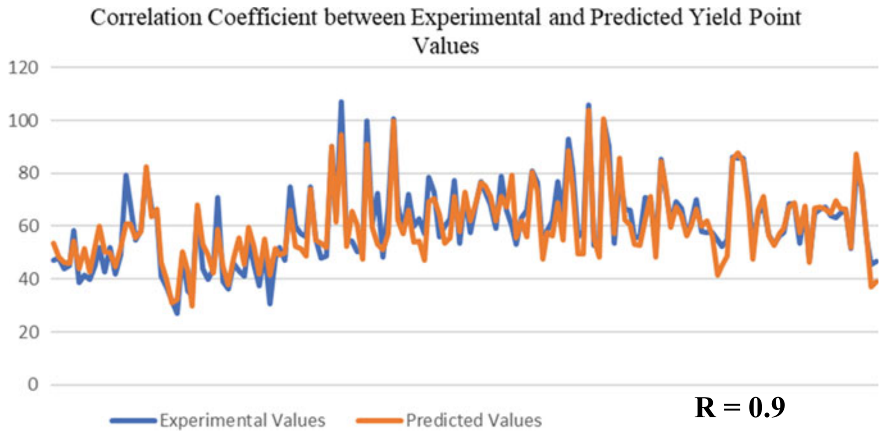


Fig. 12 Correlation for experimental and predicted values of tensile strength

7 Conclusion

The research here assures that soft computing techniques will aid us in benefitting in terms of accuracy and decisiveness in foreseeing the mechanical properties of steel. The dataset here was experimented using 5 different machine learning algorithms, wherein, the models were trained and tested for the performance. Eventually these developed models and the algorithms were integrated to interact through a graphical user interface. The developed accommodating GUI with an inclusion of the implemented ML algorithms yielded better results in predicting the tensile strength and yield point values of stainless steel at varying processing parameters. The target to beat the conventional methodology of experimenting using Universal

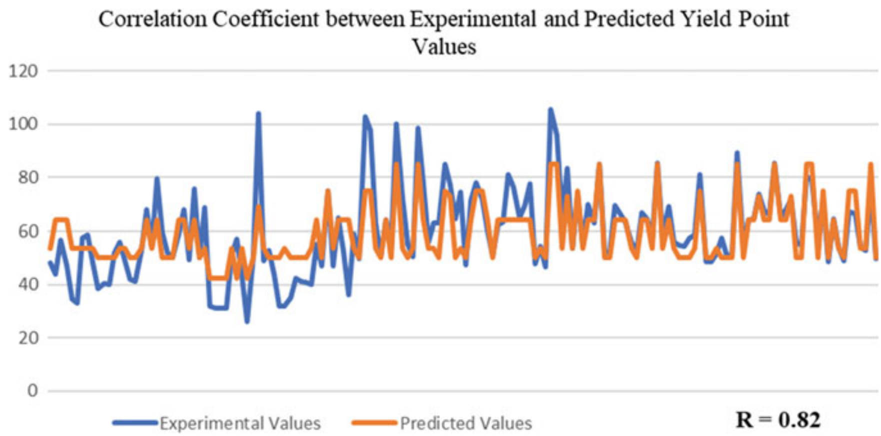


Fig. 13 Correlation for experimental and predicted values of yield point

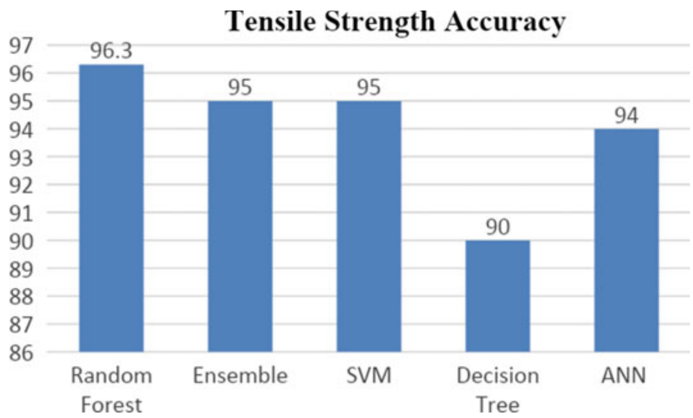


Fig. 14 Prediction accuracy of various algorithms for tensile strength

Tensile Test machine (UTM) to predict mechanical properties of steel is achieved as part of this work. The models developed as part of this study executed good accuracy in predicting the output values that are above 90%. This work can be further extended to study other alloys based on the input processing parameters so as to predict their mechanical properties which is very significant for many applications with an advantage of less time consumption and cost.

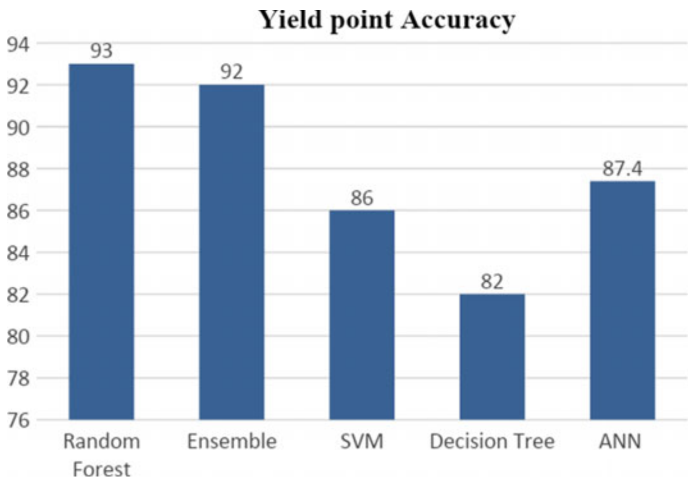


Fig. 15 Prediction accuracy of various algorithms for yield point

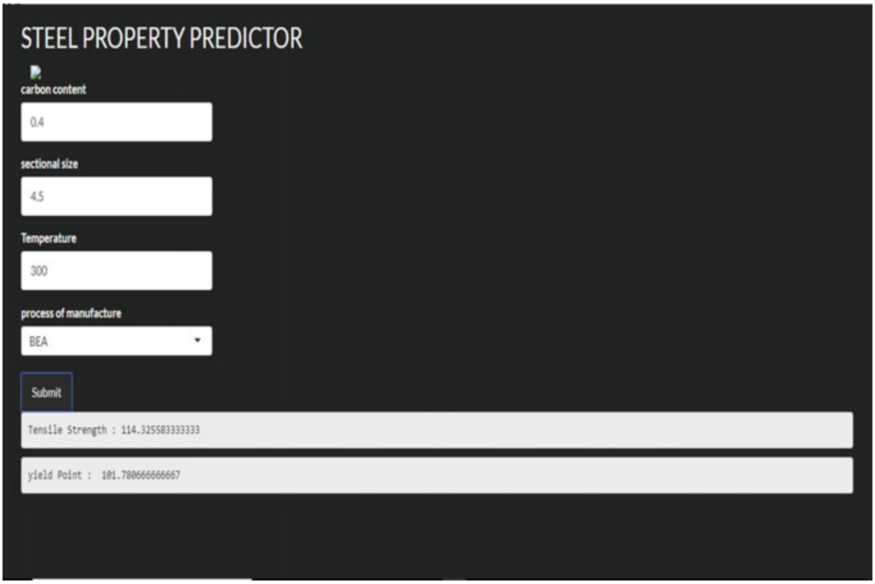


Fig. 16 GUI for steel property prediction

References

1. "Development of a tool to predict mechanical properties of steel by exploring data science techniques" project grant sanctioned by All India Council for Technical Education under Research Promotion Scheme
2. Newton CH (1993) The building of material databases. Race Street, Philadelphia

3. Rajan K (2018) Informatics for materials science and engineering. Woburn, United States.
4. Thankachan T, Soorya Prakash K, Pleass CD, Rammasamy D, Prabhakaran B, Jothi S (2017) Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen. Elsevier
5. Weng Y, Zhao Y, Tang G, Liu Z (2013) Prediction of the mechanical properties of hot-rolled C-Mn steels by single index model. IEEE
6. Somkuwar V (2013) Use of artificial neural network for predicting the mechanical property of low carbon steel. Blue Ocean Res J
7. Xu Z, Liu X, Zhang K (2019) Mechanical properties prediction for hot rolled alloy steel using convolutional neural network. IEEE
8. Fragassa C, Babic M, Bergmann CP, Minak G (2019) Predicting the tensile behaviour of cast alloys by a pattern recognition analysis on experimental data. MDPI
9. Guo S, Yu J, Liu X, Wang C, Jiang Q (2019) A predicting model for properties of steel using the industrial big data based on machine learning. Research Gate
10. Santos I, Nieves J, Penya YK, Bringas PG (2009) Machine-learning-based mechanical properties prediction in foundry production. Research Gate
11. Nagwani NK, Deo SV (2014) Estimating the concrete compressive strength using hard clustering and fuzzy clustering based regression techniques. Research Gate
12. Llewellyn DT (2013) Steels: metallurgy and applications. Woburn, Massachusetts, United States
13. Wen YF, Cai CZ, Liu XH, Pei JF, Zhu XJ, Xiao TT (2009) Corrosion rate prediction of 3c steel under different seawater environment by using support vector regression. Corros Sci 51(2):349–355. https://doi.org/10.1016/j.corsci.2008.10.038
14. Rao BV, Gopalakrishna SJ (2009) Hardgrove grindability index prediction using support vector regression. Int J Miner Process 91(1–2):55
15. Senussi GH (2017) Prediction of mechanical properties of stainless steel using an artificial neural network model. Elsevier.
16. Desu RK, Krishnamurthy HN, Balu A, Gupta AK, Singh SK (2015) Mechanical properties of austenitic stainless steel 304L and 316L at elevated temperatures. Elsevier
17. Hn Bhadeshia HKD (1999) Neural networks in materials science. ISIJ Int 39(10):966–979
18. Schmidhuber (2015) Deep learning in neural networks: an overview 61:85–117
19. Reddy NS, Krishnaiah J, Hong SG, Lee JS (2009) Modelling medium carbon steels by using artificial neural networks. Mater Sci Eng 508(1–2):93–105
20. Fu LM (1990) Building expert systems on neural architecture. 1st ZEE Znr. Co@ on Artificial Neural Networks, London, October 16–18, pp 221–225
21. Gupta AK (2010) Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression. Int J Prod Res 48:763–778
22. Singh SK Mahesh A Gupta prediction of mechanical properties of extra deep drawn steel in blue brittle region using artificial neural network

Comparison of Higher-Order Spectral Features Clubbed with Classification Techniques for the Labeling of Adventitious Sounds in the Breathing Cycle

Rupesh Dubey, Rajesh M. Bodade, and Divya Dubey

Abstract Adventitious sounds appearing in the breathing cycle convey excellent information about the patient's health. In this paper, instead of the conventional method of fixed time interval analysis of the adventitious sound during the breathing cycle, a waveform-based approach is followed. A comparative analysis of accuracy to label adventitious sounds by the higher-order spectral features based on wavelet bi-phase and power spectrums (16 in number) clubbed with classification techniques is caried out. Features are analysed with different classification process namely decision tree, SVM, k-NN, ensemble learners. There are a total of 36 results presented using Matlab© 2019b. Many results having specific feature classifier combinations have achieved accuracy as high as 100% like bi-phase and weighted k-NN, bi-phase and subspace k-NN, power spectrum and fine k-NN, power spectrum and weighted k-NN, power spectrum and bagged tree, power spectrum, and subspace k-NN. Average accuracy achieved with features based on wavelet bi-phase outperforms that achieved by the power spectrum by 0.11%.Please check and confirm if the authors and their respective affiliations have been correctly identified. Amend if necessary.Information is updated

Keywords Higher-order spectral · k-NN · SVM · Classification tree · Wheezes · Crackles

1 Introduction

As per the WHO report, more than 64 million people are suffering from chronic obstructive pulmonary disease COPD. WHO predicts COPD to take up the title of

R. Dubey (✉)
Military College of Telecommunication Engineering (MCTE), MHOW, Ministry of Defence, Government of India, DAVV University, Indore, India

R. Dubey · R. M. Bodade
IPS Academy, Institute of Engineering and Science, Indore, India

D. Dubey
SDBCE, Indore, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_148

the third most death-causing disease by the year 2030 [1]. Adventitious sounds like wheezes and crackles reveal the patient's health; are heard during breathing cycles of patients suffering from obstructive pulmonary diseases [2]. Wheezes sounds last for a length greater than 150 ms, whereas for crackle, the time length is less than 20 ms [2]. The waveform of wheezes is continuous, just like the sinusoidal signal, while for Crackles, the waveform is discontinuous. Physicians, merely with their ears, can recognize adventitious sounds. The traditional stethoscope was invented in 1821 but suffered limitations of lower frequency (<120 Hz). Signal processing techniques at the initial stage to identify wheezes targeted their time-expanded waveforms [3]. Further works focused on peak detection in the spectrum, amplitude, and pitch range [4, 5]. Many experimental types of research clubbed sets of features and classifiers like k-NN, ANN, CNN, and many more [6]. In most cases, the accuracy of algorithms based on the above criteria depends on the acoustics amplitude of the signals. There remains a scope to focus on techniques independent of amplitudes.

Airway obstruction produces changes in non-linear harmonic peak interaction. Some researchers concentrated on analysis based on bi-coherence [7] and phase spacing [8]. Also, works are clubbing continuous wavelet transform with higher-order, i.e., third-order spectra [9, 10]. Some approaches used integrated power to find out the severity of asthma disease [11].

Here for the analysis of nonlinearity in adventitious sounds, wavelet bi-phase, and power spectrum are used. There are 16 unique features based on the bi-phase, and the power spectrums are analyzed for their ability to accurate labeling by using different classifiers. These features classify adventitious sounds like wheezes, crackles, and normal sounds. Features clubbed with other classifiers for the proper labeling have yielded accuracy as high as 100%. Here the comparative analysis is presented between features based on the bi-phase and power spectrum.

The flow of the paper goes as Sect. 2 discusses the mathematics and analysis of the feature set under consideration. Section 3 presents data processing and acquisition techniques. Section 4 provides a view of the data validation technique. Section 5 presents a brief introduction of classification techniques followed in this research; Sect. 6 discusses the summary of experimental results obtained, and the paper concludes with Sect. 7.

2 Feature Analysis and Mathematics

2.1 Wavelet Bi-Spectrum (WBS)

The transient properties of the signal are revealed by convolving with wave-like structures (wavelets). Continuous wavelet transform defined as follows [12]

$$W_x(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\varphi^*(\frac{t-b}{a})dt \tag{1}$$

here, $x(t)$ is the signal in time-domain ($x(t) \in L^2(R)$), * represents complex conjugate, and $\psi(t)$ is the mother wavelet and a and b as scaling and dilation factor respectively, and both are continuous. The combination of bi-spectrum with the wavelet is called wavelet bi-spectrum (WBS). Here time–frequency analysis takes place by using continuous wavelet transform. Wavelet type complex Morlet is selected [9, 12].

$$\boldsymbol{\varphi(t)} = \frac{1}{\sqrt{\boldsymbol{\pi f_b}}} \boldsymbol{e}^{\frac{-t^2}{f_b}} \boldsymbol{e}^{\boldsymbol{j2\pi f_c t}} \tag{2}$$

where f_c is the wavelet central frequency, and f_b is the bandwidth parameter. The WBS is defined as [13].

$$WB_x(a_1,a_2) = \int_T W_x^*(a,\tau)W_x(a_1,\tau)W_x(a_2,\tau)d\tau \tag{3}$$

The calculation of the above integration takes place over a finite interval of time T:$\tau_0 \le \tau \le \tau_1$, and a,a_1, and a_2 are the scale length of wavelet components and signal, respectively.

Instantaneous Wavelet bi-amplitude and bi-phase

The WBS is defined over time interval T, instantaneous WBS (IWBS) is defined as follows:

$$IWB_x(a_1,a_2,t) = |IWB_x(a_1,a_2,t)|e^{j\le IWB_x(a_1,a_2,t)} = A_x e^{j\varphi x} \tag{4}$$

2.2 Power Spectrum

The Welch method is used to calculate the power spectrum with less computational cost [14]. The Welch method provides options to select a type, width, and amount of overlap of the window. The window type here is Kaiser window with a 50% overlap. The window width chosen is such that the frequency resolution of the Welch spectrum equal to the average dB in width of the DFT spectrum.

If $X(j), j = 0, \ldots, N-1$ be the samples from second-order stochastic signal sequence, for the above signal it is pre assumed that $E(x) = 0$ *and* $P(f)$ is spectral density, for $|f| \le \frac{1}{2}$.

The segments overlap of length L with D units apart, the Kth segment is

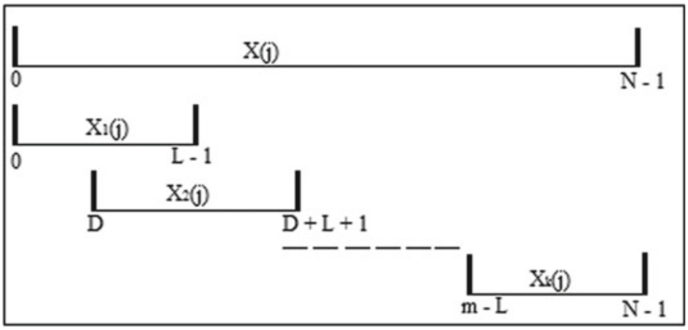


Fig. 1 Division of Segments of the signal under consideration

$$\begin{aligned} X_1(j) &= X(j) \quad & j = 0, \ldots, L-1 \\ X_2(j) &= X(j+D) \quad & j = 0, \ldots, L-1 \end{aligned} \tag{5}$$

Finally

$$X_k(j) = X(j + (K-1)D) \quad j = 0, \ldots, L-1 \tag{6}$$

here N = (K − 1)D + L segments are as shown in Fig. 1.

Modified periodogram calculation takes place for each segment of Length *L*.

$X_1(j)W(j), \ldots, X_k(j)W(j)$, here $W(j)$ is data window, $j = 0, \ldots, L - 1$, now take FFT

$$A_k(n) = \frac{1}{L}\sum_{j=0}^{L-1} X_k(j)W(j)e^{\frac{-2kijn}{L}} \quad \text{and } i = (-1)^{1/2} \tag{7}$$

$$I_k(f_n) = \frac{L}{U}|A_k(n)|^2 \quad k = 1, 2, \ldots, K, \text{ here } f_n = n/L \; n = 0, \ldots, L/2 \tag{8}$$

$$U = \frac{1}{L}\sum_{j=0}^{L-1} W^2(j) \tag{9}$$

Spectral estimation is the average of these periodogram, so the final equation of power spectrum becomes

$$\widehat{P}(f_n) = \frac{1}{K}\sum_{k=1}^{K} I_K(f_n) \tag{10}$$

Table 1 Features based on wavelet bi-phase and power spectrum and their abbreviations

Feature	Feature description	Abbreviation
1	Global max value in the amplitude domain in wavelet bi-phase	$GMax\O_x$
2	Global min value in the amplitude domain in wavelet bi-phase	$GMin\O_x$
3	The distance of the C_i from the contour S of the i_{th} GP at the bi-frequency in wavelet bi-phase	$D^{GPi}\O_x$
4	Amplitude above mean in wavelet bi-phase	$Amean\O_x$
5	Mean wavelet bi-phase related to LPs*	$mean\O_x{}^{LP}$
6	Average instantaneous wavelet bi-phase across the examined total time interval T	$m\O_x(\omega_1, \omega_2)$
7	Maximum wavelet bi-phase across time-related to LPs	$max\O_x{}^{LP}$
8	The standard deviation of the wavelet bi-phase related to LPs*	$std\O_x{}^{LP}$
9	Global max value in the amplitude domain in the power spectrum	$GMaxP_x$
10	Global min value in the amplitude domain in the power spectrum	$GMinP_x$
11	The distance of the C_i from the contour S of the i_{th} GP at the bi-frequency domain in the power spectrum	$D^{GPi}P_x$
12	Amplitude above mean in the power spectrum	$A_{mean}P_x$
13	Mean power spectrum related to LPs	$meanP_x{}^{LP}$
14	Average instantaneous power spectrum across the examined total time interval T	$mP_x\ (\omega_1,\omega_2)$
15	Maximum power spectrum across time-related to LPs	$maxP_x{}^{LP}$
16	The standard deviation of the power spectrum related to LPs	$stdP_x{}^{LP}$

2.3 Feature Set

See Table 1.

The features are selected considering the non-linear nature of the signal under consideration in the time–frequency domain. The characteristics of the signal appear in the quadratic phase coupling of distinct harmonic peaks and their gradual appearance over time. Features 5 and 8 marked with an asterisk are carried forward from research [5].

Figures 2, 3, 4, 5, 6 and 7 gives a view of the wavelet bi-phase and power spectrum obtained for different adventitious sounds. The figures provide the scope of a rich feature base.

2.4 Global Peaks (GPs)

Peaks appearing in the signal for the entire duration of time signal under consideration, i.e., defined over $T = T_{Total}$, are referred to as global peaks (GPs). The different

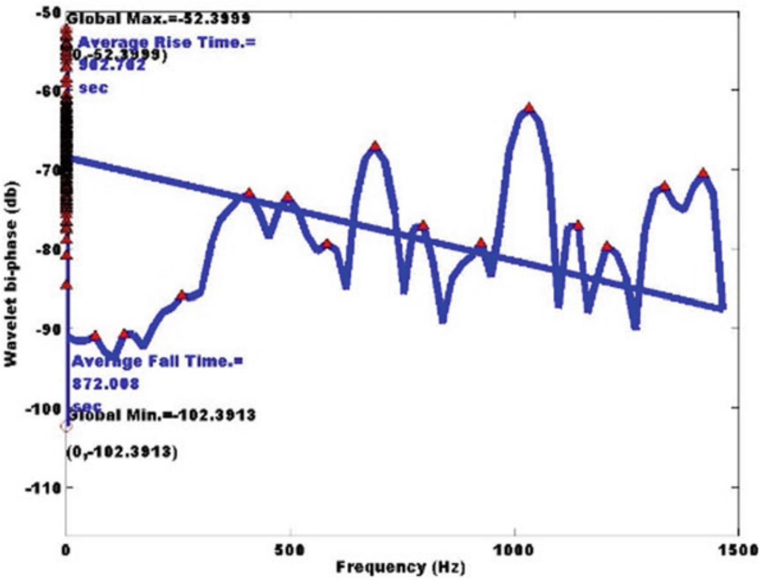


Fig. 2 Higher-order features of the wheezes using bi-phase features (color print)

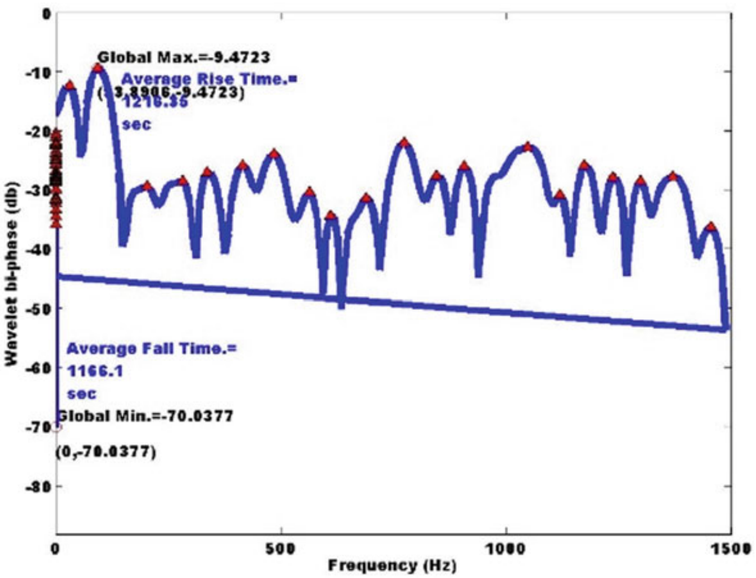


Fig. 3 Higher-order features of the crackles using bi-phase features (color print)

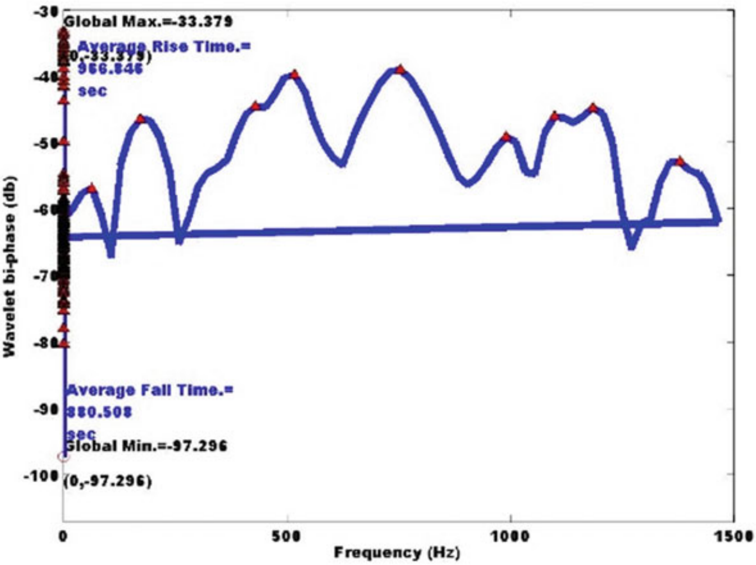


Fig. 4 Higher-order features of the normal sound using bi-phase features (color print)

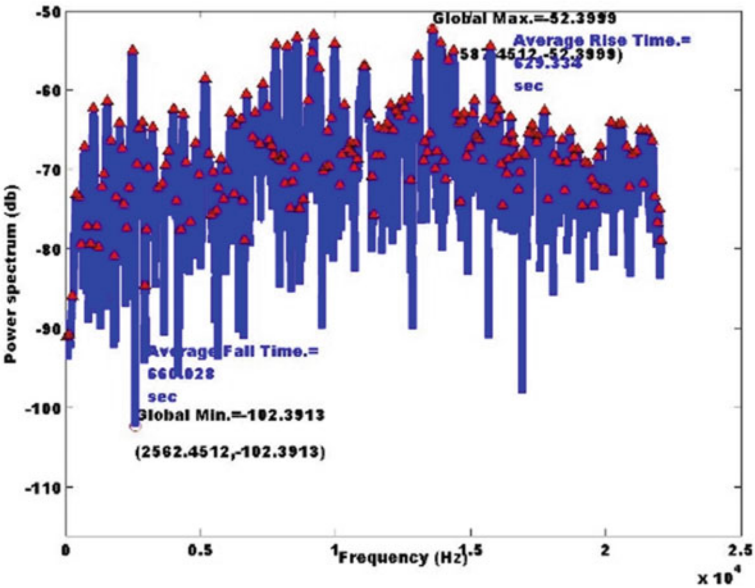


Fig. 5 Higher-order features of the wheezes using power spectrum features (color print)

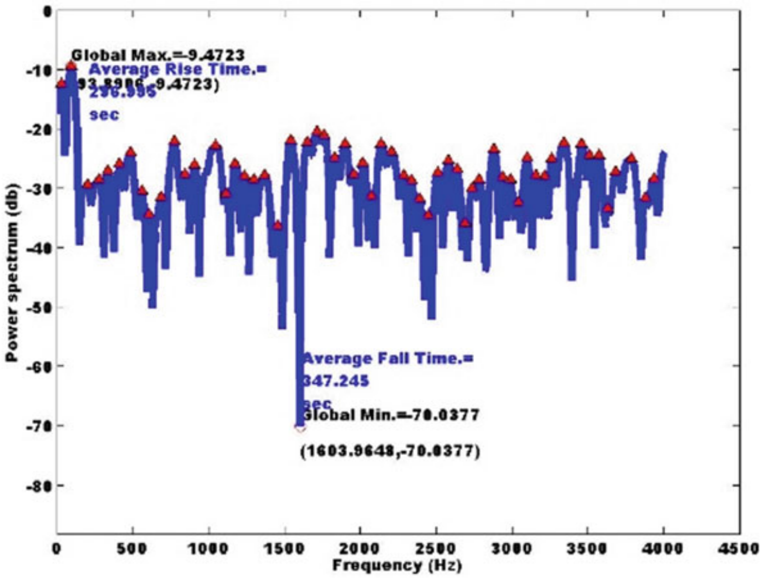


Fig. 6 Higher-order features of the crackle using power spectrum features (color print)

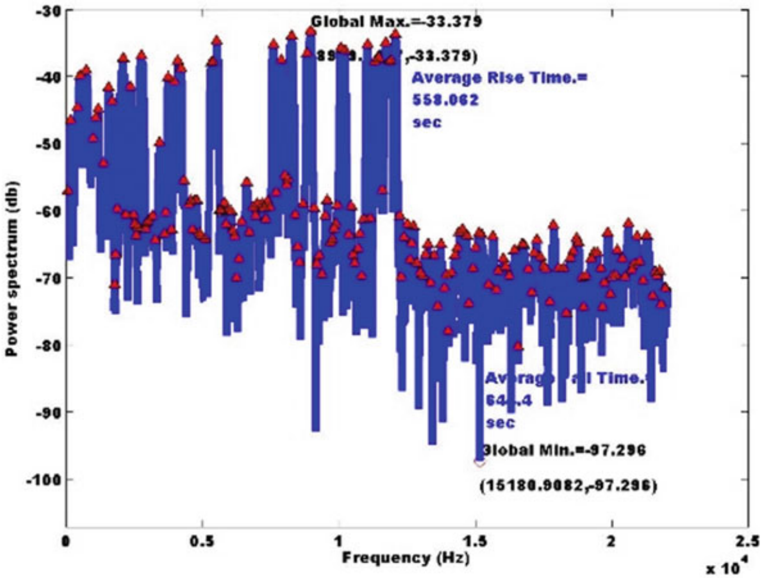


Fig. 7 Higher-order features of the normal sound using power spectrum feature (color print)

characteristics of GPs provide the bi-phase and power spectrum-related characteristics of the proposed feature. The peak is identified as GP if its amplitude is higher than the mean of the average amplitude of peaks considered under duration T. The averaging of instantaneous wavelet bi-amplitude across the whole-time interval T stationary in bi-frequency is performed to reduce the computational burden.

$$mA_x(\omega_1, \omega_2) = \overline{A_x(\omega_1, \omega_2, t)} \tag{11}$$

here $A_x(\omega_1, \omega_2, t)$ is the IWBC amplitude of instantaneous wavelet bi-amplitude over the area Δ exceeding statistical Noise [9, 15]

$$\begin{aligned} mAb_x^{GP_i}(\omega_{c1}, \omega_{c2}) = mAb_x^{GP_i}(\omega_1, \omega_2)\big|mA_x^{GP_i}(\omega_1, \omega_2) = \max \\ C^i = (\omega_{c1}, \omega_{c2}),\ i = 1, 2, \ldots, l \end{aligned} \tag{12}$$

2.5 Local Peaks (LPs)

These are the peaks seen in the detailed perspective span of the signal based on window overlapping sections Δ obtained using IWBS analysis.

$$Ab_x^{LP_i}(\omega_{c1}, \omega_{c2}, t) = Ab_x^{LP_i}(\omega_1, \omega_2, t)\big|mA_x^{LP_i}(\omega_1, \omega_2, t) = \max \tag{13}$$

here l is the number of LPs, and i being the location of the maximum peak.

3 Data Analysis

3.1 Data Acquisition

Data from different means have been used for the analysis of algorithms. The R.A.L.E.® (Respiration Acoustics Laboratory Environment) lung sounds 3.2 [16] and internet resources [17–19]. R.A.L.E.® Lung Sounds are the program for education learning is designed for students, doctors, educators, nurses, and health professionals. The respiratory acoustician laboratory of the University of Manitoba Winnipeg, Canada runs the program. It has more than 50 respiratory sounds recording collections of various age groups and diseases. The quiz section contributes an additional 24 cases of Lung sound. The collection is felicitated by awarded the merit of computer-based materials by the health sciences communications association. The data comprise of wheezes (normal, mono, and polyphonic) 252, crackle (fine and coarse) 70, normal sound (bronchial, tracheal, and bronchovesicular sounds) 50.

3.2 Data Pre-processing

Sound signals are sampled at 4 kHz, 16 bits with 1024 points per segment in the range −5.0 V to +5.0 V (−32,767 to +32,768). As per the Computerized Respiratory Sound Analysis (CORSA) guidelines, a first-order Butterworth filter (HPF) is used for high pass signals at 7.5 Hz to filter out DC offsets. Low pass filtration (LPF) is done at 2.5 kHz by using eighth order Butterworth; also, bandpass filtration (BPF) is done (150 Hz ~ 2 kHz) for heart sound cancellation. Signal was divided into the segments of their waveform using Goldwave® Software. A Pulmonologist performed manual validation of the database in the medical clinic at Indore, India.

4 k-Fold Cross-Validation Approach

Cross-validation is one of the primarily used methods as a robust model for data selection. The method has some prior assumptions and is rarely tied to a particular feature of an algorithm. It aims to estimate the performance of the learned model from data using an algorithm. One of the Cross-validation applications is to assess the algorithm's generalization under test or compare the performance of two or more different algorithms.

The conventional k-fold cross-validation follows:

D = Training set
k = Number of fold (k = 10)
C = Selected Classifier
Divide D into k folds
Model-based on C using k-1 folds
Test the model in step 2 using the k fold
Repeat step 3 for every fold [20] (Fig. 8).

5 Classification Methods

5.1 Decision Tree [21]

The decision tree is one of the non-parametric supervised learning methods. The tree structure has two types of nodes, leaf and internal nodes. The majority vote of training examples reaching that leaf designates the class label. Internal nodes branch out according to answers. Splits are associated with feature tests and are non-leaf nodes. The split value in the case of a coarse tree is 100. The decision tree gives the results when a leaf node is approached after the conduction of the series of feature tests. Here in each step, a data set is identified, and a split is selected, then this split

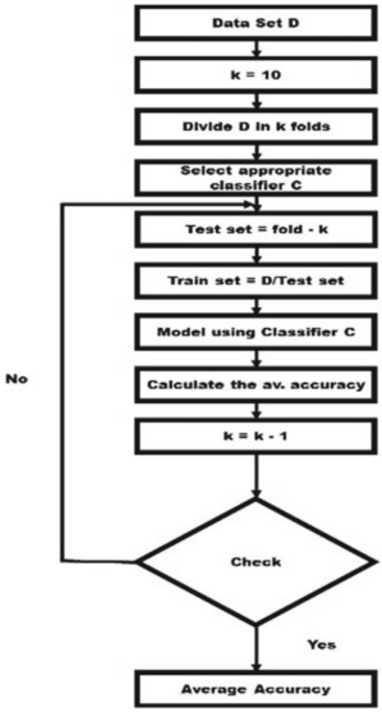


Fig. 8 Flow chart of cross-validation method

is used to divide the data set into subsets, and each subgroup remains the data set for the next upcoming step (Fig. 9).

Types of decision tree

As seen in Fig. 9, there are three categories of trees coarse, medium, and fine tree. The coarse tree with a maximum of 100 nodes does not attain high training accuracy. They are more robust and easier to interpret.

The medium tree uses fewer leaves than the fine tree. In a fine tree, a lesser number of leaves gives coarse distinctions between clauses. A fine tree has many leaves and mostly has higher accuracy on training data.

5.2 Support Vector Machine (SVM) [22]

SVM is one of the most effective supervised classification methods. SVM is preferably used as a non-probabilistic classification method. SVM training builds a model that assigns new labels to one type or the other. Linear SVM uses a linear Kernel similarly cubic and quadratic.

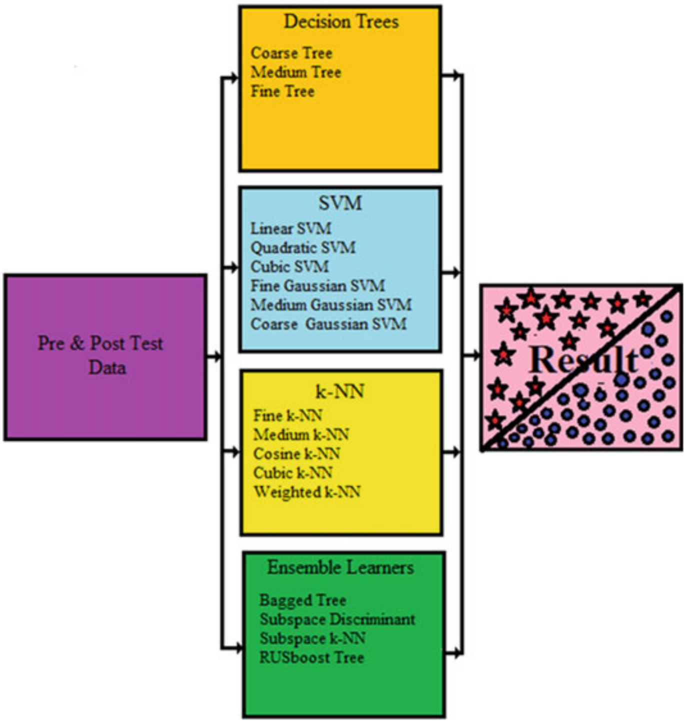


Fig. 9 Supervised classification approaches followed in the research (color print)

Fine SVM use cubic kernel, medium SVM uses quadratic kernel also Gaussian SVM uses Gaussian Kernel with scale. SVM is finely detailed and less finely and mid-class between the two, respectively.

5.3 k-Nearest Neighbor Classifier [23]

In k-NN, specifying the number of nearest neighbors provides the decision of the labeling process. It is a non-parametric learning algorithm in which no underlying data is presumed. Here most of the training data is employed during the testing phase. With number k, it makes a circle, and the class or label of an unknown feature is assigned based on the most significant number of counts of a particular class among all classes in that circle.

5.4 Ensemble Learner [24]

In the Ensemble learner method, the decision to label a class is achieved by combining the decisions of the individual classifiers. Here bagging and boosting classifiers are most widely used.

In the Bagging tree, original data sets provide the additional data sets to decrease prediction variance. Here, combinations with repetitions are performed to produce multi-sets of the same size as actual data. An increment in the size of the training set decreases the variance, which increases the expected outcome. The boosting classifier involves a two-step approach. In the first step, subsets of the original data are obtained to achieve a series of averagely performing models. In the second step, to achieve an increment in performance, a voting scheme is used to club the previous performances.

6 Results

A comparative analysis of accuracy to label adventitious sounds by the features based on Higher-order spectral wavelet bi-phase and power spectrums (16 in number) clubbed with classification techniques is summarized in Table 2. The classification process achieved with classifiers like Decision Tree, SVM, k-NN, and Ensemble learners are analyzed. There are a total of 36 results presented using Matlab© 2019b. Many classifiers clubbed with specific features have achieved accuracy as high as 100%, like bi-phase and weighted k-NN, bi-phase and subspace k-NN, power spectrum and fine k-NN, power spectrum and weighted k-NN, power spectrum and bagged tree, power spectrum, and subspace k-NN. Average accuracy achieved with features based on wavelet bi-phase outperforms that achieved by the power spectrum by 0.11% (Figs. 10, 11, 12 and 13).

Parallel coordinate graphs

Parallel Coordinate graphs are suitable for the visualization of high-dimensional geometry and are used to analyze multi-variable data. N parallel lines draw these graphs, typically erected vertical and spaced equally. The solid lines in the plot mention the median values for each group, and they represent quartile values of the same color. For example, in wave-type crackles, the solid lines represent the median value measured for each variable. The dotted lines above solid lines signify the 75% of measurements for each variable, and the dotted lines below the solid lines show 25 percentage measurements for each variable for the wave-type crackle and similarly for the wave-type normal sound and wheezes.

Table 2 Accuracy chart of feature classifier combination

Data mining technique	Classifier type	Accuracy wavelet bi-phase (%)	Accuracy power spectrum (%)
Decision trees	Coarse tree	93.5	93.3
	Medium tree	99.5	98.9
	Fine tree	99.5	98.9
SVM	Linear SVM	92.2	91.9
	Quadratic SVM	94.9	94.4
	Cubic SVM	94.9	95.2
	Fine Gaussian SVM	98.7	97.8
	Medium Gaussian SVM	95.2	94.9
	Coarse Gaussian SVM	91.1	91.4
k-NN	Fine k-NN	99.5	100
	Medium k-NN	91.9	92.2
	Cosine k-NN	93.0	93.0
	Cubic k-NN	91.9	91.7
	Weighted k-NN	100	100
Ensemble learners	Bagged TREE	99.7	100
	Subspace discriminant	90.1	90.3
	Subspace k-NN	100	100
	RUSboost TREE	99.7	99.5

7 Conclusion

The adventitious sound heard during the breathing cycle in obstructive pulmonary diseases like Asthma and COPD possess non-stationary characteristics. Features based on higher-order spectral analysis bi-phase and power spectrum provide a valuable tool for classifying adventitious sounds like wheezes, crackle, and normal sound. The combination of feature and classifier achieves accuracy as high as 100% to classify adventitious sounds heard during the respiratory cycle. Both features have nearly equally performed in labeling the sounds.

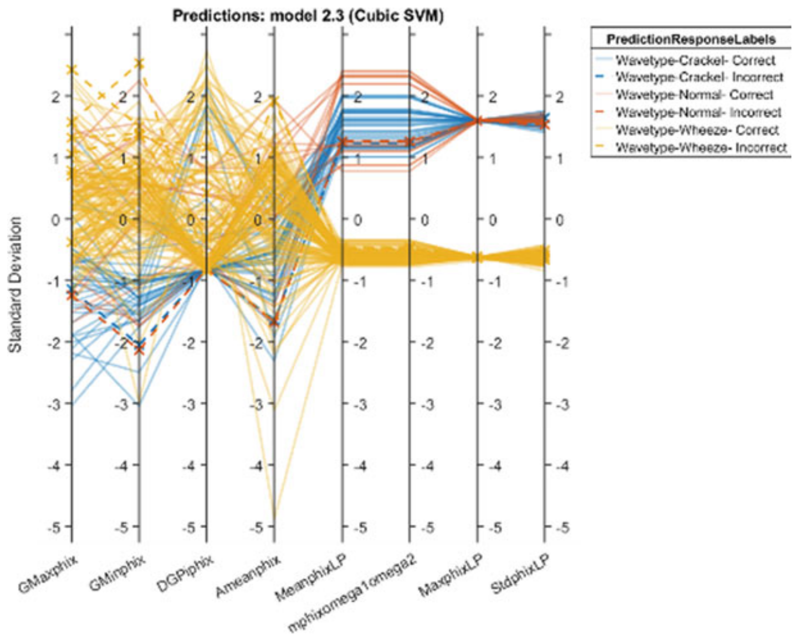


Fig. 10 Parallel coordinates graph of cubic SVM in biphase (color print)

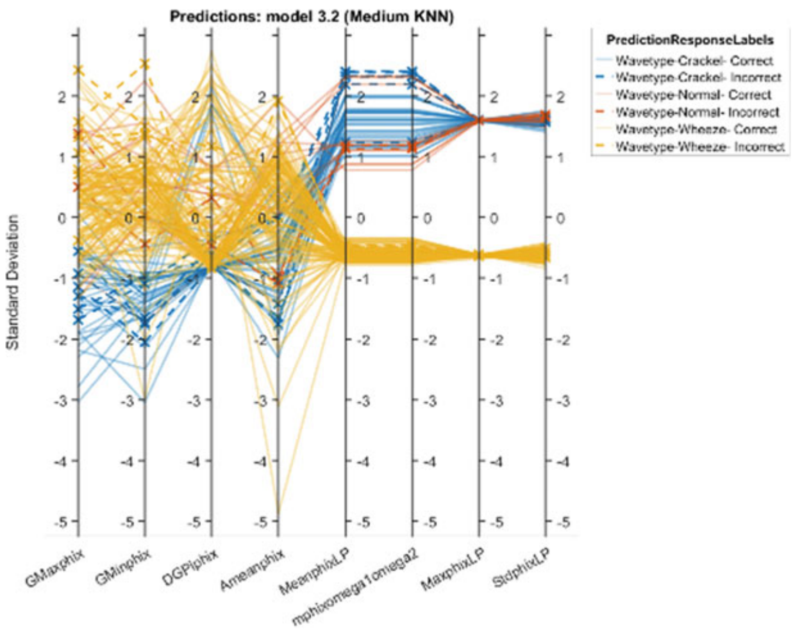


Fig. 11 Parallel coordinates graph of medium k-NN in bi-phase (color print)

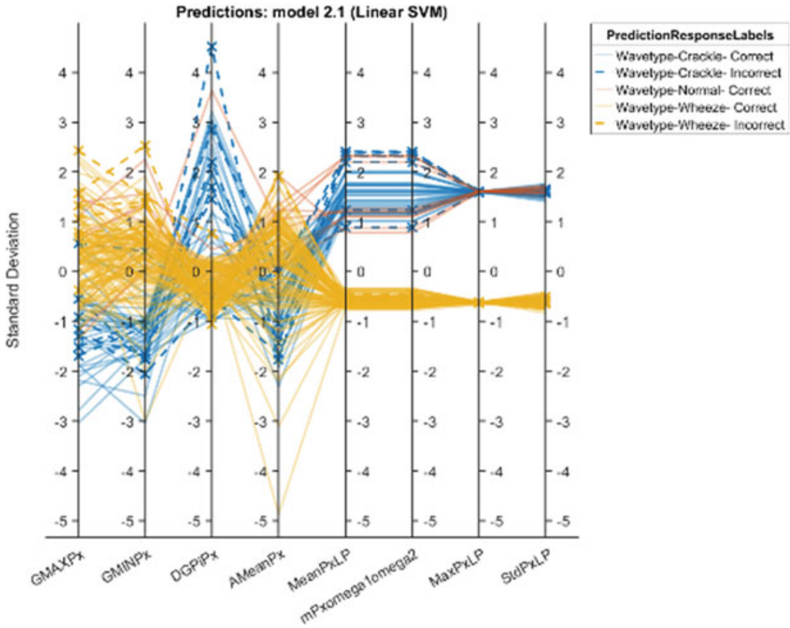


Fig.12 Parallel coordinates graph of Linear SVM in the power spectrum (color print)

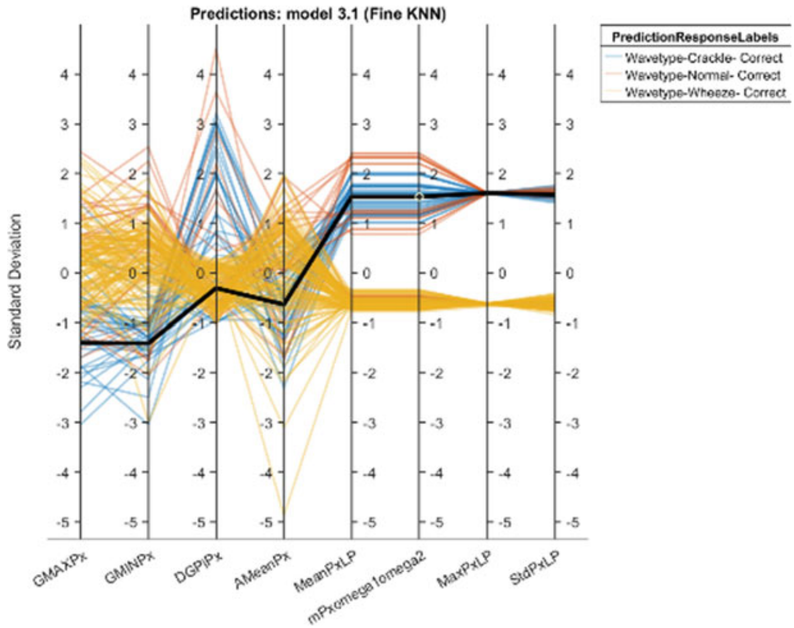


Fig. 13 Parallel coordinates graph of fine k-NN in the power spectrum (color print)

Conflict of Interest No conflict of interest to declare.

Funding No funding received.

References

1. Media WHO (2020) World Health Organisation. [Online]. http://www.who.int/respiratory/asthma/en/
2. Gavriely N (1995) Breath sound methodology. CRC Press, Boca Raton, FL
3. Murphy RL, Holford SK, Knowler WC (1977) Visual lung sound characterization by time expanded waveform analysis. New Engl J Med 296(17):968–971
4. Grotberg JB, Gavriely N, Shabtai-Musih Y (1992) Spectral content of forced expiratory wheezes during Air, He, and SF6 breathing in normal human. J Appl Physiol 72(2):629–635
5. Taplidou SA, Hadjileontiadis LJ (2007) Wheeze detection based on time-frequency analysis of breath sounds. Comput Biol Med 37(8):1073–1083
6. Dubey R, Bodade RM (2019) Classification techniques based on neural networks for pulmonary obstructive diseases: a review. In: International conference recent advances in interdisciplinary trends in engineering & applications (RAITEA-2019), Indore, India, 2019, pp 21–27
7. Hadjileontiadis LJ, Panas SM (1997) Nonlinear analysis of musical lung sound using the bicoherence index. In: Proceedings 19th annual IEEE international conference EMBS, vol 3, Chicago, 1997, pp 1126–1129
8. Ahlström C, Hult P, Ask P Wheezes analysis and detection with nonlinear phase space embedding. In: Proceedings IFMBE 13th Nord. Baltimore Conference (NBC), Umea, Sweden, 2005, pp 305–306
9. Taplidou SA, Hadjileontiadis LJ (2007) Nonlinear Analysis of wheezes using Wavelet bicoherence. Comput Biol Med 37(4):563–570
10. Stefano S, Soo-Kyung S, DE Gennaro M, Nicolino A (2008) Respiratory sound analysis in healthy and pathological subjects: a wavelet approach. Biomed Sig Process Cont 3(3):181–191
11. Ghulam Nabi F, Sundaraj K, Kiang Lam C (2019) Identification of asthma severity levels through wheeze sound characterization and classification using integrated power features. Bio Med Sig Process Cont 52:302–311
12. Addison PS (2002) The illustrated wavelet transform handbook: introductory theory and applications in science. Institute of Physics (IoP) Publishing, Bristol, PA
13. Nikias CL, Athina PP (1993) Higher-order spectra analysis, a non-linear signal processing framework, 2nd edn. Prentice-Hall Inc., USA
14. Welch PD (1967) The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust AU-15(2):70–73
15. Van Milligen BP, Hidalgo C, Anchez ES (1995) Non-linear phenomena and intermittency in plasma turbulence. Phys Rev Lett 74(3):395–398
16. Pasternak H (2008) RALE Lung Sound 3.2. [Online]. http://www.rale.ca/Pricing.htm
17. Littman (2011) [Online]. http://allnurses.com/general-nursing-student/littman-20-examples-525290.html
18. MedEdPortal (2007) The J Teach Learn Resour. [Online]. http://www.mededportal.org/publication/129/
19. Keroes J et al (2018) Medical Simulation and Training, LLC. [Online]. https://www.practicalclinicalskills.com/auscultation-lesson-description?coursecaseorder=5&courseid=201
20. Stone M (1974) Cross-validatory choice and assessment of statistical predictions. J R Stat Soc 2(36):111–147
21. Quinlan JR (1986) Induction of decision trees. Mach Learn 1:81–106

22. Cortes C (1995) Support-Vector Networks. Mach Learn pp 273–297
23. Altman NS (1981) An introduction to Kernel and nearest neighbor nonparametric regression. Ann Math Stat 25:1–31
24. Dietterich TG (1985) web.engr.oregonstate.edu. [Online]. https://web.engr.oregonstate.edu/~tgd/publications/mcs-ensembles.pdf

Smart Water Bottle with Pill Alarm for Cognitively Disabled Geriatrics

Syed Musthak Ahmed, M. Pranay Kumar, Ch. Mohith Sai, A. Ramya Sri, and D.Vineeth

Abstract Untimed medicine administration can always show adverse effects on the health of the person. Many hospitals are restructuring by optimizing medical resources towards home healthcare. The proposed work is carried out to help aged people/ patient to take the required medicine in the right proportion and at the right time at home i.e., transforming health care from hospital centric to home centric. With rapidly growing population, it has become extremely difficult to monitor as well as look after the health of older people who find it difficult to remember of their own medical care. The basic idea is to integrate alarm with light based slot sensing on a normal pill boxes integrated with a water bottle. The design is based on a smart and safe medical box that assists patients in taking their pills on time. A smart system is summed up with a mobile application where one can set multiple times for taking medicines without a caretaker. The nodeMCU module is incorporated in developing the system to coordinate the complete processes of alarming and reminding to take pills from the defined boxes that contains the prescribed medicines for the patient. Each box will have its own timing information which will be continuously compared to real world time. Once the information matches with the real time the buzzer starts blowing thereby indicating the patient to take his medicine.

Keywords Smart water bottle · NodeMCU · ICT · Internet of Things · iHome system

1 Introduction

Now-a-days, taking care of aged people has become difficult especially for the people who go to work or to attend for other domestic needs. Managing work and taking care of aged/diseased people is a big concern. Due to aging there is a prevalence of developing diseases particularly concern to health. Major aging problems include physical disability, cognitive and sensory declination [1, 2]. Also hospital restructuring is taking place to reduce the burden in hospital by take caring and transforming

S. M. Ahmed (✉) · M. P. Kumar · Ch. M. Sai · A. R. Sri · D.Vineeth
S R Engineering College, Telangana State, Warangal, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_149

the health care system from hospitals to home after post hospitalization i.e., transforming health care from hospital to home. With the technical advancements of IoT [3–6] several reforms have been bought in health care industry [7–13] such as 'intelligent medicine boxes (iMedBox), Intelligent pharmaceutical packaging (iMedPack), flexible and wearable bio-medical sensor device (Bio- Patch)'and so on. The new awaited feature is towards intelligent pill boxes with automated alert system [13–17]. By applying information and communication technology many of the health care problems can be solved at home [16, 17].

Leveraging information and communication technology (ICT) has solved many of the social health problems [8] of diseased and elderly people by keeping them healthy with proper monitoring by following few steps like: 1) Real time monitoring the person 2) Checking whether the prescribed pills are consumed 3) Maintaining timely consumption of medicine 4) Being away and at the same time meeting the desirable needs of health care of the person.

Several works have been reported in the past to resolve issues of the diseased, aged and impaired people. The ultimate aim is to see for the well being of the person by providing care and to resolve our day to day works without dedicating for the diseased at home [16].

2 Existing Method

At present, there are many types of medicine kits in the market, which are roughly divided into two categories: In the first category, the patient can only manually fill the medicines in the medicine box, and take the medicine. In this process, there is a possibility of accidently taking medicine from a different pill box [14, 15] due to lack of memory. Also, they are required to physically attend to collect water to take their medicine. In the other category, the pill boxes are filled by care taker and the person will take the assigned medicine indicated by indicator at specified time. In this method, the patient is assured of the right medicine [13]. However, the person has to seek for a water bottle. Such a system, is shown in Fig. 1.In fact, the burden on

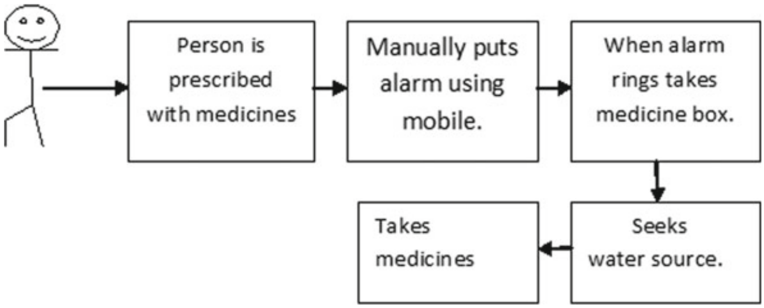


Fig. 1 Block diagram of existing module

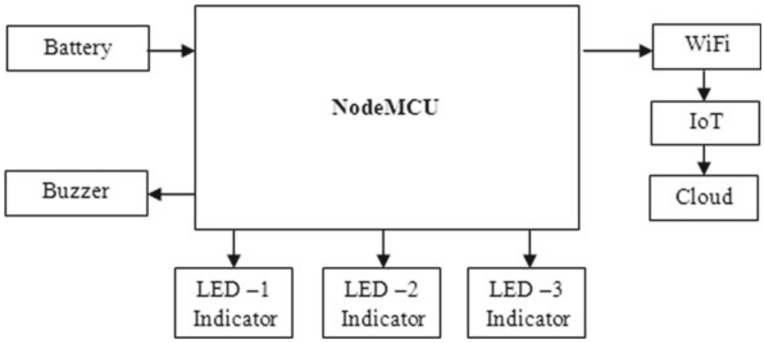


Fig. 2 Block diagram of implemented module

the user is not reduced. In this process, attention is still needed in monitoring them of taking medicines.

3 Proposed Method

The present method reduces the burden on the care taker. Both monitoring and care taking are incorporated by making use of the current technology. The block diagram of the proposed implementation is shown in Fig. 2. It makes use of NodeMCU to control the buzzer for alerting, LED's for locating the particular pill box from which the medicine is to be taken along with control circuitry to monitor the process.

The implemented system will help the users with regular medication to carry a single product which is an integration of water bottle, pill boxes and pill consuming reminding system. This smart bottle reminds the user to take their medicines on time by an alarm and indication of a particular pill box. The developed system is assimilated with a mobile application which controls the complete process. The process of smart bottle utility is explained by a flow chart shown in Fig. 3.

4 Objectives

The objectives behind the development and implementation of the project are to:

1. Assist working category people to attend to their job and at the same time support health care at home.
2. Reduce burden, worry and tension on people attending jobs about the patient's healthcare at home by incorporating a reminder system.
3. Bring transformation in Health care from Hospital Centric to Home Centric Environment.

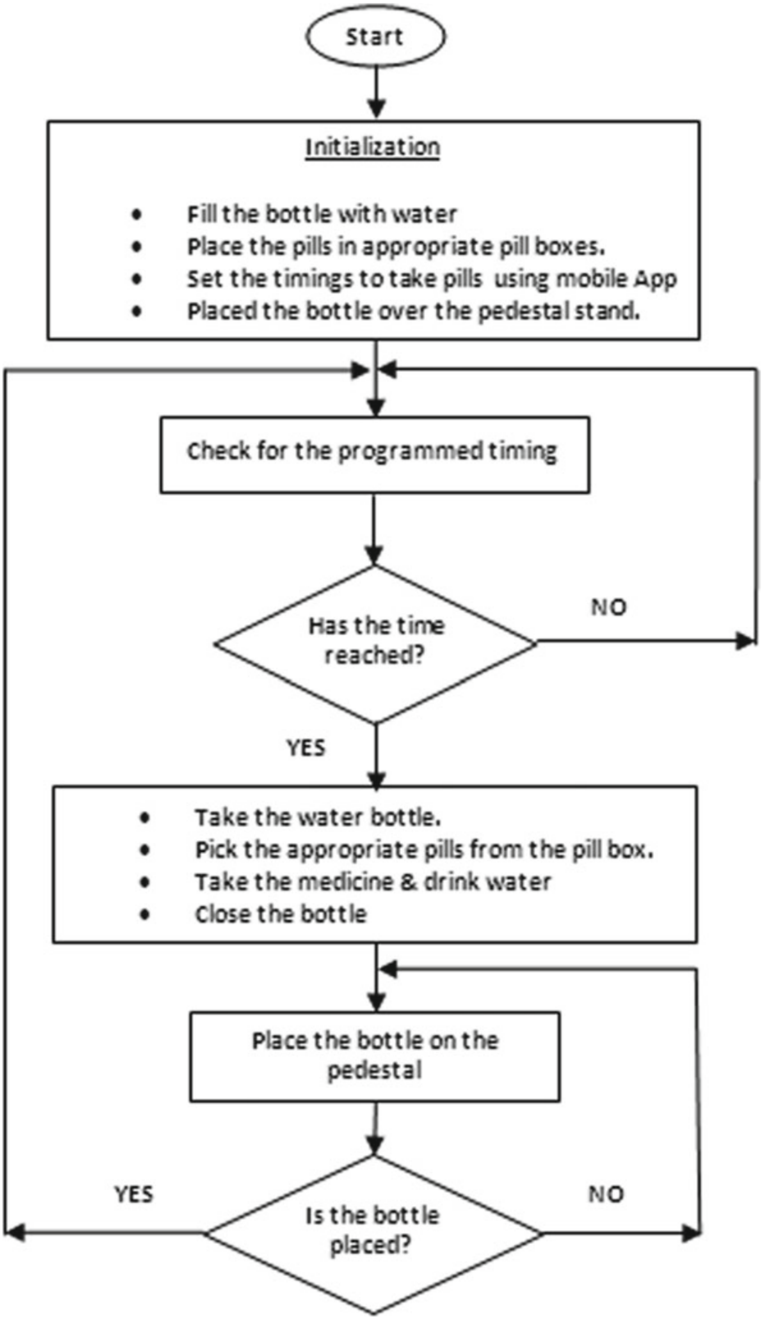


Fig. 3 Flowchart of the proposed system

4. (Post Health Care Service).
5. Develop a product that can be deployed to support medical care under medical Health Research scheme**.**

5 Methodology

In implementing the product the following methodological steps are carried out:

1. A water bottle with a hollow column structure to fix pill box.
2. A pill box compartments to hold pills to be taken at defined timings.
3. An Indicator circuit with LED's pointing the pill compartment.
4. A buzzer as an alarm to remind timing to take medicine.
5. Synchronization of Step 3 and Step 4 incorporating a control circuit
6. Use of Mobile App and IoT for communication and proper functioning of circuit.

6 Flowchart of the Working Module

See Fig. 3.

7 Results and Discussion

The complete system consists of a smart water bottle having a column of pill boxes placed over a pedestal whose function is to hold the water bottle and the complete hardware circuit. The bottle is attached with LED indicators to pick up a particular pill from the pill Box. A buzzer is fitted into the pedestal which sounds at the programmed time as a reminder. If a person doesn't attend or take the medicine then it is programmed such that the module repeats every 2 min repeatedly for 3 times. The various modules in developed system are pill box, water bottle, and pedestal stand carrying water bottle at the top and control circuitry at the bottom. The pill box and pill box attached to the water bottle shown in Fig. 4.

The pedestal stand is fabricated such that it holds the water bottle vertically over it. The indicator LED's are mounted on one side of the water bottle which indicates the particular pill that has to be taken during that time. The buzzer and control circuit is fixed below the pedestal stand. The complete set up is shown in Fig. 5.

Initially the bottle is filled with water and the pill boxes are loaded with the necessary medicine that has to be taken by the person. The bottle is placed on the pedestal stand which now gets connected to the control circuit fitted below the pedestal stand. Now, the timing at which the pills have to be consumed as defined by the medical practitioner is programmed into the system using a mobile App. Whenever the set

Fig. 4 Smart water bottle **a** pill box with compartments **b** water bottle with pill box

Fig. 5 Smart water bottle placed over the Pedestal

time is reached, the buzzer starts blowing as a reminder to the person and the corresponding LED starts glowing at the pill compartment indicating the position from where the medicine has to be taken. In the event, the person has missed to hear alarm due to some reason the system is programmed such that it repeats again with the programmed delay so that the person is again reminded of his medication.

8 Conclusion and Future Scope

In the recent years, the growth of aged adults has increased significantly. To monitor and care with ease by keeping track of their medical doses in time, a water bottle with pills container is designed. In this paper an attempt has been made to make Smart water bottle with pill alarm for cognitively disabled geriatrics with very high efficiency and low cost. The designed smart water bottle would significantly reduce human assistance like nurses or family members. The developed smart water bottle will remind patients about taking their medicines at appropriate times. The developed bottle is provided with suitable grip to hold bottle, pill boxes fixed on bottle to hold pills along with the indicator that tells a particular pill to be consumed from the pill box. The buzzer reminds of the time and the indicator gives the location of the medicine. Thus, it reminds the people with disabilities to interact with the smart bottle to take their medicine in time, thereby ensuring their health and well being. This leads to transformation from hospital centre health care to home centre health care i.e. iHome system. The main benefit of the product is its portability. Here the water bottle is integrated with reminding system as well as support to pick the particular medicine.

The developed work can further be extended by adding few more features. The smart water bottle can be connected to a GSM module so that it reports the take care person whether the patient or user has taken medicines or not. By incorporating suitable sensors we can even record the user's report of taken medicine and also a message to refill the medicines in the pill compartments at the end of the day.

Acknowledgements The authors acknowledge the Management, Principal and Center for Creative Cognition of S R Engineering College, Warangal Urban for their support in providing necessary facilities during the course of project work.

References

1. Ahmed SM, Yasmeen A, Babu BJ (2018) Wheelchair with auto navigation for adults with physio and cognitive impairments. Int J Innov Technol Exploring Eng (IJITEE) 8(252)
2. Ahmed SM, Shireen A, Babu BJ, Shruthi (2020) Powered wheelchair for mobility with features to address physical strength, cognitive response, and motor action development issues. In: International conference on data science, machine learning and applications (ICDSMLA-2019) at CMR Institute of Technology, Kandlakoya, Hyderabad. 29th & 30th March—2019 & chapter no. 121, lecture notes in electrical engineering 601, © Springer Nature Singapore Pte Ltd. 2020. https://doi.org/10.1007/978-981-15-1420-3_121
3. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29(7):1645–1660
4. Sandeep CH, Naresh KS, Pramod KP (2018) Security challenges and issues of the IoT system. Indian J Public Health Res Dev (IJPHRD) 9(11):748–753
5. Ahmed SM, Kovela B, Gunjan VK (2019) IoT based automatic watering system through soil moisture sensing-a technique to support farmers' Cultivation in Rural India. In: International conference On cybernetics, cognition and machine learning applications (ICCCMLA 2019),

16–17, March 2019, & chapter no. 120, Lecture notes in electrical engineering, vol 601. © Springer Nature Singapore Pte Ltd. 2020. https://doi.org/10.1007/978-981-15-1420-3_120.
6. Li X, Lu R, Liang X, Shen X, Chen J, Lin X (2011) Smart community: an internet of things application. IEEE Commun Mag 49(11):68–75
7. Yang G, Xie L, Mantysalo M, Zhou X, Pang Z, Da Xu L, Kao-Walter S, Chen Q, Zheng L-R (2014) A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box. IEEE Trans Ind Inf 10(4):2180–2191
8. Pang Z (2013) Technologies and architectures of the Internet-of-Things (IoT) for health and well-being, PhD Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden
9. Afzal al B, Umair M, Shah GA, Ahmed E (2019) Enabling IoT platforms for social IoT applications: vision, feature mapping, and challenges. ELSEVIER Future Genera Comput Syst 92:712–731
10. Yang G, Xie L, Mntysalo M, Zhou X, Pang Z, Da Xu L, Kao-Walter S, Chen Q, Zheng L-R (2014) A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box. IEEE Trans Ind Inf 10(4)
11. Koop CE, Mosher R, Kun L, Geiling J, Grobb E, Long S, Macedonia C, Marrell RC, Satava R, Rosen JM (2008) Future delivery of health care: cybercare. IEEE Eng Med Biol Mag 27(6):29–38
12. Kortuem G, Kawsar F, Fitton D, Sundramoorthy V (2010) Smart objects as building blocks for the Internet of things. IEEE Internet Comput 14(1):44–51
13. Castillejo P, Martinez J-F, Rodriguez-Molina J, Cuerva A (2013) Integration of wearable devices in a wireless sensor network for an E-health application. IEEE Wires Commun 20(4): 38–49
14. Bhati S, Soni H, Sinh Zala V, Vyas P, Sharma Y (2017) Smart medicine reminder box. Int J Sci Technol Eng 3(10)
15. Zeidan H, Karam K, Daou RAZ, Hayek A, Boercsoek J (2018) Smart medicine box system. In: Ieee international multidisciplinary conference on engineering technology (IMCET)
16. Schuz B, Marx C, Wurm S, Waner LM, Zlegelmann JP, Schwarzer R, Tesch-Romer C (2011) Medication beliefs predict medication adherence in older adults with multiple illnesses. J Psychosom Res 70(2):179–1871
17. Hodges S, Taylor S, Villar N, Scott J, Bial D, Fischer PT (2013) Prototyping connected devices for the Internet of Things. Computer 46(2):26–34

An Intelligent Anti-Theft Vehicle Locking System Using IoT

B. Saritha, CH. S. S. R. Bharadwaja, M. Nikhitha, CH. Nethra Reddy, K. Arun, and Syed Musthak Ahmed

Abstract In the present days, chattel crimes are increasing rapidly. Around 10 million cases are recorded annually. Vehicle crime is a predominant criminal activity observed globally. This is causing a panic to the owners of the vehicles in terms of monetary loss and mental agony. To curb this, many researchers have bought out several preventive measures in terms of security during the fleet. The present work is a step towards providing security to vehicles under the control of the owner. The proposed work is implemented by incorporating IoT platforms for the safety and security of the vehicles. The present work is implemented by incorporating a GPS module in the vehicle. Once the vehicle is missing, the owner gets an alert message on the mobile. The message received the geographical location of the vehicle. On receiving the message, the owner enters into the mobile app and clicks on the "STOP" button which locks the engine instantly. The present work is a step towards incorporating security to vehicles from thefts, assisting the owners from huge damages and fleet management. The developed system is the cheap alternative for vehicle security.

Keywords IoT · Mobile application · Fleet management · NodeMCU · GPS module

1 Introduction

In recent years vehicle theft has become a major issue which should be traced and detected. The safety and security of the vehicle is essential. Even though there are many existing mechanisms, they have few limitations like high cost, complexity etc. So, an efficient security mechanism is needed. With the advent of Information and communication technology (ICT) and applying it to fleet management several problems will be resolved particularly to the owner's owning the vehicles. IoT is such a technology where we can interact with the objects without human intervention as

B. Saritha (✉) · CH. S. S. R. Bharadwaja · M. Nikhitha · CH. Nethra Reddy · K. Arun · S. M. Ahmed
S R Engineering College, Warangal, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_150

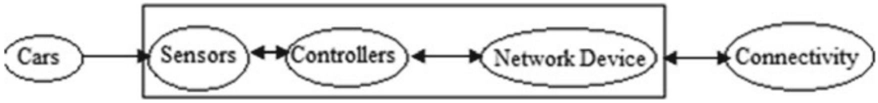


Fig. 1 Architecture of IoT

shown in Fig. 1. This IoT platform can be extended to several applications in the field transportation, medical, home automation [1–8] and so on.

By leveraging IoT into the vehicle, the theft of vehicles could be curbed by incorporating suitable circuitry. Several researchers have produced their works on vehicle security by using microprocessors, ARM processors [9] Arduino board as a control circuitry and communication using GSM/GPS module [10–15]. Even the government authorities will be benefited by incorporating this system to find the location of the vehicles and make announcements to support the public [16].

Thus, by providing dynamic interaction between the vehicle, surrounding and the user enables us to take suitable action to improve fleet management [17]. The monitoring method incorporates the following steps:

1. Tracking of vehicles using GPS.
2. Information about geographical location of the vehicle on the mobile.
3. Misallocation of vehicle assistance.
4. Safety, security and recovery with complete networking of modules.

2 Existing Method

Most of the systems designed for vehicle tracking and security are incorporated with a beeper sensor, PIR sensor and alerting alarms. In spite of it, commercially available products incorporated for security purposes didn't drop down the rate of theft happening around the globe. Basically, there are two categories in recognizing the vehicle presence:

1. Vehicle parked in house.
2. Vehicles parked outdoors.

In the former case, vehicles parked can be detected easily from getting stolen using a buzzer in the house and by sensing the touch. But this system cannot work when the vehicle is far away from the owner. So, the alarming method may not be that beneficial. A list of techniques has been introduced to reduce car theft with the help of pressure tilts and door sensors. These techniques however bear some limitations such as high false-rate, high rate and easy to be disabled. In order to solve these problems recent advancements in computer hardware and software have enabled the automobile industry to develop affordable automated biometrics based identification and verification systems. But, biometrics even has their own disadvantages with its 100% accuracy. This existing system is clearly mentioned in Fig. 2.

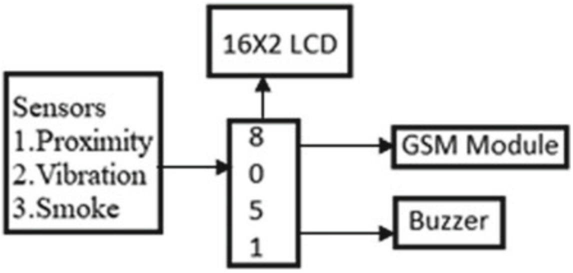


Fig. 2 Block diagram for an existing system

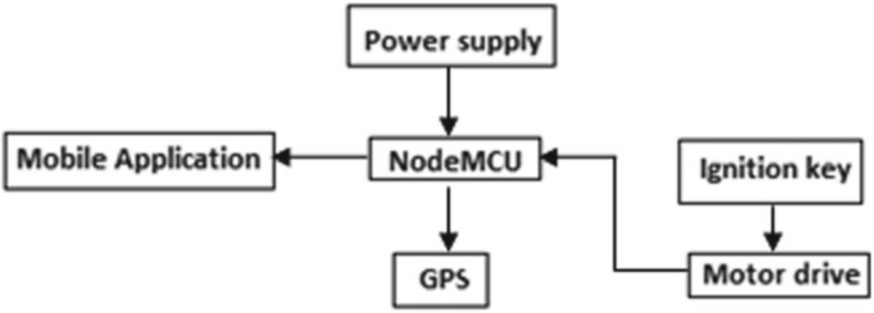


Fig. 3 Block diagram of proposed system

3 Proposed System

The proposed project is developed to prevent vehicle theft. This functionality is achieved by detecting vehicle status during theft mode by a SMS generating automatically to the owner's mobile. Upon receiving the message the owner gets the geographical position of the vehicle. The owner then operates his mobile app by which he can control the locking of the engine. In order to achieve this, the developed system controlled by the mobile App is interfaced with nodeMCU in order to disable the ignition of the vehicle thereby preventing crimes to a great extent. NodeMCU is integrated with Wi-Fi and dual-mode Bluetooth operating at low cost and at low power. The block diagram of the developed system is shown in Fig. 3.

The operation of the system developed is explained by a flow chart given in Fig. 4

4 Objectives

The main objectives of the anti theft system for vehicles is to:

1. Establish a connection between vehicle and the user.
2. Enables the vehicle to notify the user in situations of theft.

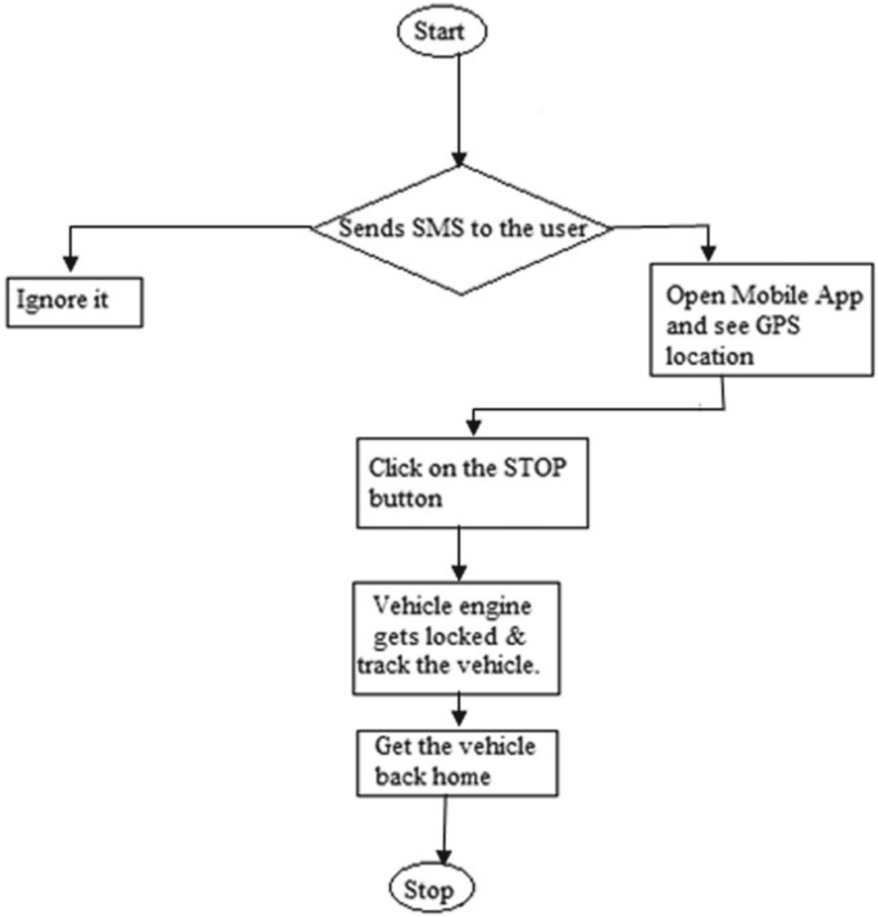


Fig. 4 Flowchart of proposed system

3. Facilitate Locking of engine by installed mobile application.
4. Locating and recovery of the vehicle.
5. Support transportation by assisting fleet management, vehicle safety and security.

5 Methodology

In implementation the following steps are followed:

1. Making use of the NodeMCU and a mobile application to control vehicles remotely.
2. In cooperating IoT for communication for proper functioning of circuits.

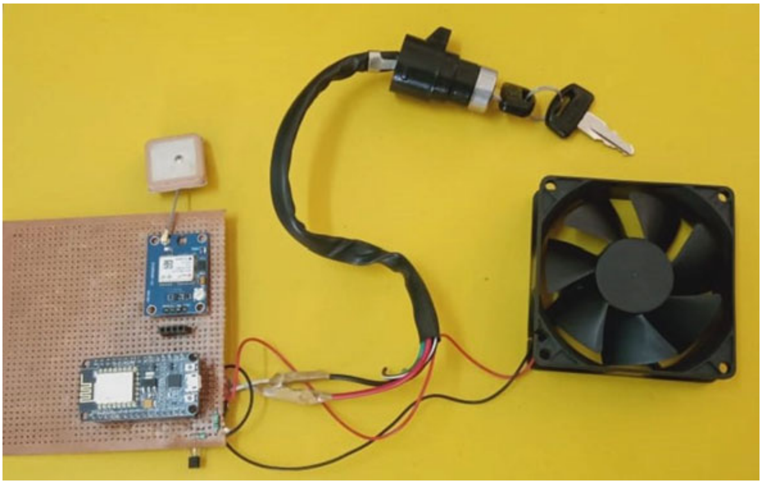


Fig. 5 Vehicle theft detection with engine locking using IoT

3. Installation of the module at an appropriate place in the vehicles.
4. A message signal during the mode of vehicle theft.
5. Implementing the mobile application with a STOP button and GPS to lock the vehicle.

6 Results

The complete system installation and net working includes Tracking system using GPS Module with a mobile application assimilated using IoT. If the vehicle is theft or moved without the knowledge of the owner, then a notification is given to the user. Then the user gets alert and opens an application in mobile where we can track the location of the car i.e., the GPS location. There is also feasibility where we can see the longitude and latitude values respectively. Then the owner of the vehicle can lock the engine by clicking the STOP button shown in below Fig. 5.

7 Conclusion

Vehicle theft, although not as intrusive as violent crimes, causes greater loss to the victims in terms of monetary values. The proposed system provides vehicle safety and detects theft efficiently and effectively at very low cost. This project "An intelligent anti-theft vehicle locking system using IoT" has been successfully designed and tested. Integrating feature of all the hardware components used has developed it. Presence of every module has been responses out and placed carefully thus, contributing to the best working of the unit.

8 Future Scope

As a future scope of work the project can be extended:

1. By adding a few more functions for information to vehicle security systems by sending information to police forces if the vehicle is stolen, by sending GPS location and an alert message for a quick action.
2. By incorporating artificial intelligence (AI) a voice message can be sent to the owner and even intimation to the thief sitting in the car.

Acknowledgements The authors acknowledge the Management, Principal and centre for IoT and EMBEDDED systems of S R Engineering College, Warangal Urban for their support in providing necessary facilities during the course of the project work.

References

1. Ahmed SM, Anusha A (2017) Vehicle tracking and monitoring system to enhance the safety and security driving using IoT. In: IEEE international conference on recent trends in electrical, electronics and computing technologies, © 2017 IEEE, pp 49–53. 10.1109-ICRTEECT.2017.35
2. Ahmed SM, B. Kovela, Gunjan VK (2020) IoT based automatic watering system through soil moisture sensing-a technique to support farmers' cultivation in rural India. In: International conference on cybernetics, cognition and machine learning applications (ICCCMLA 2019), 16–17, March 2019, & chapter no. 120, Lecture notes in electrical engineering vol 601, © Springer Nature Singapore Pte Ltd. 2020. https://doi.org/10.1007/978-981-15-1420-3_120
3. Saritha B, Chitti S (2018) Constructing a system to monitor and control indoor environment. Int J Eng Tehnol 7(1.8):247–350
4. Sandeep CH, Naresh KS, Pramod KP (2018) Security challenges and issues of the IoT system. Indian J Public Health Res Dev (IJPHRD) 9(11):748–753
5. Ramani R, Selvaraju S, Valarmathy S, Thangam R, Rajasekaran B (2012) Water-level monitor for bore well and water tank based on GSM. Int J Eng Sci Technol (IJEST) 4(10). ISSN 0975-5462
6. Song K-T, Yang C-C (2005) Front vehicle tracking using scene analysis. In: Proceedings of the IEEE international conference on mechatronics & automation
7. Ramya V, Palaniappan B, Karthick K (2012) Embedded controller for vehicle in-front obstacle detection and cabin safety alert system. Int J Comput Sci Inf Technol (IJCSIT) 4(2)
8. Kulkarni V, Babu V (2015) Embedded smart car security system on face detection. Special issue of IJCCT 3(1):112–116
9. Joshiand MS, Mahajan DV (2014) Arm 7 based theft control, accident detection and vehicle positioning system. Int J Innov Technol Explor Eng 4(2):29–31
10. Khedher MAA (2011) Hybrid GPS-GSM localization of automobile tracking system. Int J Comput Sci Inf Technol 3(6):75–85
11. Maurya K, Singh M, Jain N (2012) Real time vehicle tracking system using GSM and GPS technology-an anti-theft tracking system. Int J Electron Comput Sci Eng pp 1103–1107

12. Kaushik N, Veralkar M, Parab P, Nadkarny K (2014) Anti-theft vehicle security system. Int J Sci Res Dev 1(12):2845–2848
13. Chen H, Chiang Y, Chang F, Wang H (2010) Toward real-time precise point positioning: differential GPS based on IGS ultra rapid product. SICE annual conference, IEEE Xplore 14 Oct 2010
14. Peijiang C, Xuehua J (2008) Design and implementation of remote monitoring system based on GSM. vol 42, pp 167–175
15. Al-Hindawi AMJ, Talib I (2012) Experimentally evaluation of GPS/GSM based system design. J Electron Syst 2(2)
16. Rashed MAA, Oumar OA, Singh D (2014) A real time GSM/GPS based tracking system based on GSM mobile phone. IEEE J Signals Telecommun 3(1):65–68
17. Bosi I, Ferrera E, Brevi D, Pastrone C (2019) In-vehicle IoT platform enabling the virtual sensor concept: a pothole detection use-case for cooperative safety. In: 4th international conference on Internet of Things, big data and security, pp 232–240

Diabetic Retinopathy Classification Techniques in Medical Analysis Using Deep Representations

Morarjee Kolla and T. Venugopal

Abstract The development of Image Processing techniques in classification using conventional features has not yet achieved the expected results. The medical analysis is still beginning with these traditional imaging functions. The use of Deep representations can attain improvements in medical analysis instead of conventional features. Researchers explored several techniques using Deep Classification in the Medical field. Dealing with the leading cause of vision impairment and blindness of Diabetic Retinopathy using the existing methods suffers with memory bottleneck. The purpose of this paper is to describe the various deep classification techniques used in Diabetic Retinopathy and discuss novel ideas that may be useful for advanced medical analysis and research.

Keywords Deep classification · Deep representations · Diabetic retinopathy · Image processing · Medical analysis · Traditional features

1 Introduction

Most of the data used currently for analysis is available in an unstructured format and is useful for medical research and analysis. Storing and analyzing this data is one of the most predominant areas of research. It is now challenging to develop efficient classification techniques suitable for a wide variety of health applications. Searching for user queries in an extensive database is a current challenging research area. Understanding user requirements and defining them is a difficult task. Many images are uploaded and stored online at any given time to develop high-speed technologies quickly. High-visibility effects and end-user interface communication opens up opportunities to transform information and services through an image content. Now, finding the right and effective results for the given query is a challenging task.

M. Kolla (✉)
Department of CSE, CMR Institute of Technology, Hyderabad, Telangana, India

T. Venugopal
Department of CSE, JNTUH College of Engineering Jagtial, Nachpally, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_151

With the development of visual processing, imaging plays a vital role in many applications and real-time fields. The maintenance and recovery of active audio-visual techniques are of particular interest. Finding the right image for an organized image repository is a daunting task in response to the demand for latency and retention of images found on the internet at high speed and accuracy with user queries and applications.

Emergency groups did not organize the classification of images as images based on the known image label's nature. There is plenty of space to relax with the machine's small space and the elderly's views. Modern image processing from different images is a daunting challenge for Content-Based Image Retrieval (CBIR) [1]. CBIR is still struggling to find useful articles from major news outlets. Two problems with finding these elegant images are the place of exploration of the elegant space's elegance. Recent advances in deep classification technologies have shown to integrate images with high-quality information from centre space. Furthermore, established algorithms such as these integrators use a lot of time and space to process memory and data. This study reveals how automation helps to create classified images and complex files for visibility.

1.1 Traditional Image Classification

The idea of image classification is to break the larger groupings of images as a model. So that images with the same tag are collected and use as a comparison. It provides unwanted physical tension with a large collection of images and thus will find many useful applications. This approach can greatly improve image classification performance in a large image area by regrouping tagged images instead of confusing individual images. The image collection also aims to organize a large number of similar tagged images into the groups. We use this approach to solve image representation, graphical rendering, and visualization issues. The primary purpose of taking an image is to obtain information about the content. Users are interested in the content of the group tag associated with images. The primary purpose of the image classification is to get a detailed map of labeled images. So classified groups can provide reliable information to store in an image collection. We use image classification to assign a label to an image subtle to computer vision and pattern recognition [2] and have different application forms. Typically, the process of image classification depends on feature extraction. The handcrafted classification methods are facing difficulties in classifying the images using low-level features. We introduce high-level feature-based deep classification methods to overcome the difficulties and gain more popularity recently.

1.2 Deep Image Classification

With advances in computational computing, previously managed methods have been successfully taught in large-scale data sets [3]. Researchers in this field have identified millions of already labeled samples and more massive data sets such as ImageNet. Also, video-based composite data acts as a driver of deep neural networks for event-detected events. The marking remains a mandatory task in all cases to exploit controlled learning algorithms' capabilities, but the task of tagging data is tedious, difficult, slow, and costly [4]. In order to overcome this, much work has focused on examining non-descriptive image data representations. Exercises managed with a predefined set of elements require expanding and generalizing the model for real-time data. Introducing many semi-managed and uncontrolled learning methods helps to reduce data labeling efforts. The effectiveness and stability of the classification algorithm depend on the type of input provided by the user. The deep learning approach integrates the process of feature extraction and classifying online image activities. Advanced features of deep learning are superior to hand-made features. Multiple layers transform data input into a classified result while analyzing high-level features based on the deep learning model concept [5].

1.3 Diabetic Retinopathy Classification Using Deep Representations

It is one of India's highest diabetes rates by 2030, with 79 million people suffering from diabetes. However, the lack of trained specialists in India limits asymptomatic patients' effective screening and leads to advanced diabetic eye disease patients. Significant progress is made in Artificial Intelligence (AI) since the 1950s, with the advent of machine learning in the 1980s and deep learning by 2010 [6]. Deep learning (DL) is a new field of machine learning and impacts neural networks and mimics the human brain in decision-making. It requires a large amount of data for training. Diabetic Retinopathy (DR) is a complication of diabetes that causes the retina's blood vessels to damage, and sometimes it may lead to vision loss [7]. Deep Convolutional Neural Network (DCNN) can learn from in-depth image data, which is the essential deep representations to classify retinal images. Figure 1 shows the five different DR classification images of the Kaggle dataset [8].

2 Related Work

This chapter contains previous articles that illustrate the concept of the proposed scale. Related work articles talk about discussing and presenting past and present

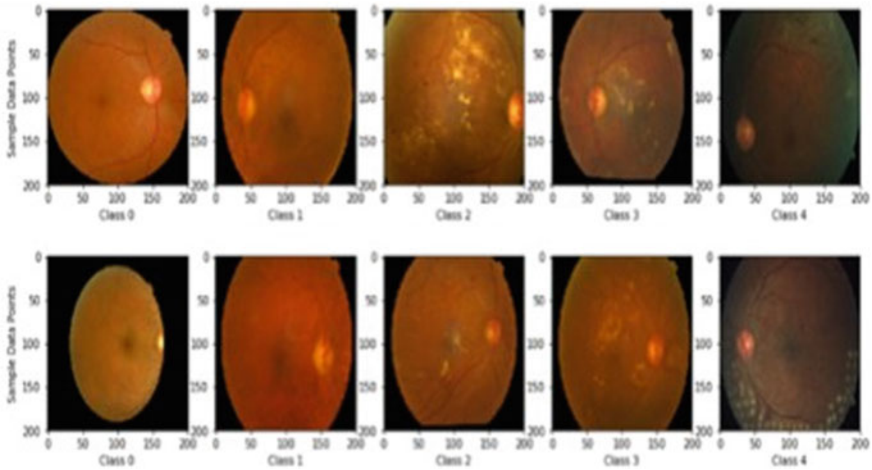


Fig. 1 Representative retinal images of DR at various stages of the disease

works on DR classification. This research includes shortcomings in this area related to the proposed studies.

2.1 DR Classification in Medical Analysis

Raman et al. [9] designed a system to detect outlier features in retinal images. Zhao et al. [10] proposed a new technology based on intelligence in detecting leaks in fluorescent angiography. Usman Akram et al. [11] approach to the hybrid classifier is on diagnosing retinal damage. Roychowdhury et al. [12] helped reduce AdaBoost claims classification's potential according to classification characteristics.

2.2 DR Classification Using Deep Representations

DR Classification with Inception-v3 architecture used is to classify retinal images [13]. ResNet34 pretrained CNN architecture used is to identify DR affected and Normal images [14]. Jiang H et al. [15] combined existing pretrained architectures with data augmentation and produced DR into two classes. Some researchers classified DR into five-stage classification from the Kaggle using various pretrained networks [16–18]. Khan SH an et al. [19] achieved the best accuracy than pretrained models with their custom CNN. This chapter aims to discuss various deep classification approaches using in the medical analysis along with their performance. A comparison with some of the main DR classification methods using different datasets is shown in Table 1.

Table 1 Comparison of various DR classification methods in medical analysis

S. No.	Reference and dataset	Methodology	Result	Advantages
1	Wang et al. [20] Kaggle	Detection using regression activation map (RAM)	Parameters reduced by 21.8%, speed is increased by 11.8% compared to the baseline method	Predict ROI on different Clinical level
2	Smailagic et al. [21] Messidor	Trains the model incrementally by using labeled data	O-MedAL test accuracy (92.02%), MedAL (91.18%) ResNet18 Baseline (85.71%)	Create an optimized and labeled training set from unlabeled data
3	Li et al. [22] ISBI, IDRiD and Messidor	Disease-specific and dependent attention modules to learn useful features for diseases	Joint Accuracy 84.1%, the accuracy of DR and Diabetic Macular Edema (DME) are 91.7%	Joint modeling the disease and its complication for fundus images
4	Li et al. [23] Kaggle	Multiclass SVM optimized using TLBO	Achieve a recognition rate up to 86.17%	Developed an app called “Deep Retina”
5	Zeng et al. [24] Kaggle	Siamese-like binary classification CNN	AUC of 0.951 is obtained	0.011 higher ROC obtained
6	Hashim et al. [25] DIARETDB0 and HRF	The Haralick features are used to classify the normal and DR images	Classification accuracy of 81.92% is achieved	Haar-DWT and GLCM features are combined
7	Bhaskaranand et al. [26] EyePACS	The MA turnover estimation tool aligns retinal images	Sensitivity was 91.7% and Specificity was 91.5%	Retinal images were taken with desktop fundus cameras
8	Prabhu et al. [27] Kaggle	Random forest algorithm and ANN are explored	Classification accuracy of random forest 87.5% and ANN 78.5%	Detection of different stages of the disease
9	Sahlsten et al. [28] Messidor	The ensemble of six models trained with 512 × 512 sized retinal images	AUC 0.953 sensitivity 0.945 and specificity 0.992 achieved	Accurately classifying five-grade DR and four-grade DMA scales
10	Qureshi et al. [29] RIMONE, DIARETDB0	Developed by various computational intelligence techniques	Classification accuracy of 98.2% is achieved	CAD systems for the diagnosis of DR

3 Diabetic Retinopathy Classification Techniques in Medical Analysis

Nowadays, DR classification techniques help find disease-causing medical analysis patterns to detect and control DR's progression early.

3.1 Discussion and Observation

The current study reviewed several articles on DR classification methods. Most of them use deep learning approaches than handcrafted approaches with data augmentation and reduce overfitting on training data. There is a need for an efficient DR early detection system due to the rapid increase in the number of DR affected people in India. The result of DR mostly depends on the size of the training data. Hence, there is a need for quality data set with balanced classes. Customized CNN architectures are producing noticeable results than existing pretrained networks even though they are time-consuming. By finding a reliable DR system model, we can identify different lesions and DR steps, leading to a better screening system to avoid visually impaired DR patients.

3.2 Need for DR Classification Using Compact Deep Representations

This paper examines and discusses the deep classification methods used in the clinical management of the DR level. Many studies develop and list compact deep representations that can reduce memory consumption and speed the execution [30, 31]. Deep learning models are now taking up vast amounts of storage space. The development of models using compact deep representations without compromising performance is an active research area nowadays to overcome this. Deep classification models installed on mobile devices in the future are widely used for real-time business applications.

4 Future Directions and Challenges

The automatic diagnosis of DR based on digital fundus images in medical images has long been a dynamic study. The foundation of research interests is a significant reduction in healthcare costs and a high potential for new medical industry products. More efficient algorithms are needed to identify the structural and retinal changes associated with DR detection. Early insertion and treatment of DR are essential public health interventions that can reduce the likelihood of vision loss. Also, testing is not done on existing screening programs with 2D background photography to DME's level. Existing selection systems cannot provide an additional description of the peripheral retina and require the pupils' pharmacological enlargement. The incredibly expansive field of imaging and confocal examination laser ophthalmoscopy has a unique potential to eliminate these deficiencies.

Today, advanced imaging techniques have played an essential role in computer-aided systems for detecting abnormalities in DR. some guidelines can help us make full use of the deep learning methods more effectively. Most of the existing studies use CNN techniques to develop a wide range of DR designs using retina's digital images. On the other hand, the diagnosis and interpretation of retinal photographs asking ophthalmologists require much time and expensive tasks. Therefore, it is vital to develop the right deep learning strategies from a limited database collection. Also, we improve the performance of existing deep learning methods by combining DL-based frameworks that adapt to flow conditions. Therefore, it can reduce computing costs and training needs for a variety of independent deep learning models.

5 Conclusion

This paper explores different DR classification methods and discusses the advantages and disadvantages. We use different DR classification methods to deal with useful models and data from these various sources. Data selection and classification methods are essential in medical diagnostics, and information in this field is imperative. The primary objective of this detailed DR classification work is how to implement policy change in treatment. Each set is unique for different treatment plans. Using deep DR classification, we need to choose the right fit for the right business from the available DR classification methods. Class imbalance and annotation problems need to solve with advanced methodologies in the future. Future research directions for effective utilization of deep DR classification models in medical analysis with a scope of research discussed above will help find existing medical applications.

Acknowledgements This research work was supported in part by a Collaborative Research Project grant from TEQIP-III, JNTU Hyderabad, India (**JNTUH/TEQIP-III/CRS/2019/CSE/14**).

References

1. Eakins JP, Graham ME (1999) Content-based image retrieval, a report to the JISC Technology Applications programme
2. Huang Y et al (2014) Feature coding in image classification: a comprehensive study. IEEE Trans Pattern Anal Mach Intel 36(3):493–506

3. Yang J, Devi P, Dhruv B (2016) Joint unsupervised learning of deep representations and image clusters. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp 5147–5156
4. Yang B, Xiao F, Sidiropoulos ND, Mingyi H (2017) Towards k-means-friendly spaces: Simultaneous deep learning and clustering. In: International Conference on Machine Learning, pp 3861–3870. PMLR
5. Litjens G et al (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88
6. Shah P, Mishra DK, Shanmugam MP, Doshi B, Jayaraj H, Ramanjulu R (2020) Validation of deep convolutional neural network-based algorithm for detection of diabetic retinopathy–artificial intelligence versus clinician for screening. Indian J Ophthalmol 68(2):398
7. Vo HH, Abhishek V (2016) Discriminant color texture descriptors for diabetic retinopathy recognition. In: 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP), pp 309–315. IEEE
8. Kaggle dataset [Online]. Available https://kaggle.com/c/diabetic-retinopathy-detection
9. Raman V, Patrick T, Putra S (2016) Proposed retinal abnormality detection and classification approach: computer aided detection for diabetic retinopathy by machine learning approaches. In: 2016 8th IEEE International Conference on Communication Software and Networks (ICCSN), pp 636–641. IEEE
10. Zhao Y, Zheng Y, Liu Y, Yang J, Zhao Y, Chen D, Wang Y (2016) Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy. IEEE Trans Med Imaging 36(1):51–63
11. Akram MU, Khalid S, Tariq A, Khan SA, Azam F (2014) Detection and classification of retinal lesions for grading of diabetic retinopathy. Comput Biol Med 1(45):161–171
12. Roychowdhury S, Koozekanani DD, Parhi KK (2013) DREAM: diabetic retinopathy analysis using machine learning. IEEE J Biomed Health Inform 18(5):1717–1728
13. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316(22):2402–2410
14. Esfahani MT, Ghaderi M, Kafiyeh R (2018) Classification of diabetic and normal fundus images using new deep learning method. Leonardo Electron J Pract Technol 17(32):233–248
15. Jiang H, Kang Y Mengdi G, Dongdong Z, He M, Wei Q (2019) An interpretable ensemble deep learning model for diabetic retinopathy disease classification. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp 2045–2048. IEEE
16. Dutta S, Manideep BC, Basha SM, Caytiles RD, Iyengar NC (2018) Classification of diabetic retinopathy images by using deep learning models. Int J Grid Distrib Comput 11(1):89–106
17. Wang X, Yongjin L, Yujuan W, Wei-Bang C (2018) Diabetic retinopathy stage classification using convolutional neural networks. In: 2018 IEEE International Conference on Information Reuse and Integration (IRI), pp 465–471. IEEE
18. Wan S, Liang Y, Zhang Y (2018) Deep convolutional neural networks for diabetic retinopathy detection by image classification. Comput Electr Eng 1(72):274–282
19. Khan SH, Zeeshan A, Danish Rizvi SM (2019) Classification of diabetic retinopathy images based on customised CNN architecture. In: 2019 Amity International Conference on Artificial Intelligence (AICAI), pp 244–248. IEEE
20. Wang Z, Jianbo Y (2018) Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation. In: Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence
21. Smailagic A, Pedro C, Alex G, Kartik K, Mostafa M, Jonathon F, Devesh W, et al (2020) O-MedAL: online active deep learning for medical image analysis. Wiley Interdiscip Rev: Data Mining Know Dis 10(4):e1353
22. Li X, Hu X, Yu L, Zhu L, Fu CW, Heng PA (2019) CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular Edema grading. IEEE Trans Med Imaging 39(5):1483–1493

23. Li YH, Yeh NN, Chen SJ, Chung YC (2019) Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network. Mob Inf Syst 1:2019
24. Zeng X, Chen H, Luo Y, Ye W (2019) Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network. IEEE Access 5(7):30744–30753
25. Hashim MF, Mohd Hashim SZ (2014) Diabetic retinopathy lesion detection using region-based approach. In: 2014 8th. Malaysian Software Engineering Conference (MySEC), pp 306–310. IEEE, 2014
26. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda S, Solanki K (2016) Automated diabetic retinopathy screening and monitoring using retinal fundus image analysis. J Diabetes Sci Technol 10(2):254–261
27. Prabhu N, Deepak B, Nita S (2019) Diabetic retinopathy screening using machine learning for hierarchical classification. Int J Inno Technol Explore Eng 8(10):1943–1948
28. Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, Kaski K (2019) Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. Sci Rep 9(1):1–1
29. Qureshi I, Ma J, Abbas Q (2019) Recent development on detection methods for the diagnosis of diabetic retinopathy. Symmetry 11(6):749
30. Courbariaux M, Hubara I, Soudry D, El-Yaniv R, Bengio Y (2016) Binarized neural networks: training deep neural networks with weights and activations constrained to + 1 or -1. arXiv preprint arXiv:1602.02830
31. Courbariaux M, Bengio Y, David J-P (2015) Binaryconnect: training deep neural networks with binary weights during propagations. In: Advances in neural information processing systems, pp 3123–3131

A Systematic Exploration of Image Fusion: A Review

Tushar and Abhijit Nayak

Abstract Image fusion in the world of image analysis is a very trendy area of study. It has a broad variety of uses, such as monitoring, diagnosis and imaging, in different fields. All the methodologies have been addressed in this survey paper and different methods are presented to address the vulnerabilities of various issues. In this article, thorough review is carried out. Several attempts have been made to recognize the obstacles and indeed the milestones gained thus far. This paper subject to numerous study papers and have sought to cover any part of image fusion. On the basis of their domain, i.e., in the spatial and frequency domain, multi-scale transform image fusion methods were addressed. Next, with its main implementations and methods interconnected, image fusion is described in depth. Second, in tabulation, analysis of different fusion approaches is performed. Finally, the latest informative discussions of image fusion and potential aspects have been ended. This survey would be seen as a guideline for a domain connected to image fusion. The goal of the survey was to collectively examine the fundamental image fusion process with its categorization based on its modal systems and algorithms. The study shows that tremendous work has been done in the field of image fusion, but in this sense, there is still enormous space for fresh and creative work. The study also shows that any algorithm has its own strengths and disadvantages in this sense.

Keywords Image fusion · Frequency domain · Multi-scale transform · Fusion process

1 Introduction

A frame reflecting a sample surface can be represented in detail as an image. The pixels have spatial details and each pixel has a numerical meaning in the simplest form [18, 19]. Second, their signals come from multiple modalities, presenting scenario

Tushar
USICT, GGSIPU, Sector 16 C, Dwarka, New Delhi, India

A. Nayak (✉)
Bhagwan Parshuram Institute of Technology, New Delhi, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_152

details from numerous aspects i.e., reflected light is collected by visible photographs, whereas thermal emission is captured by infrared images. This mix is therefore more insightful than that of signals of a single-modality. Second, ultraviolet and transparent photos have properties inherent in virtually all objects. The knowledge given by a single picture or by the individual analysis of many pictures is often enhanced by image fusion. The basic purpose of the study can be incredibly diverse if fused photos come from the same platform. It may include examining images obtained at various depths on the same sample, it may rely on defining the nature of the process observed by images [14, 15], it may be geared towards quantitative analysis using model comparison and test sets, or it may be possible to use multiple images from separate individuals in biological samples to describe the general population trends [16, 17]. The mixture of images taken from the same sample in a very different way on the same platform is often intended to improve the spatial resolution [9, 10] of the original measurement. This is the principle of so-called super-resolution techniques, which concentrate on matching data of multiple pictures that are low resolution, which are slightly altered by a motion process of sub-pixels, from each other. Super-resolution algorithms go past basic metrics and often provide a final image of even more spatial information than the parent's average measurements.

Method of Image Fusion
Methods for image fusion may be narrowly divided into two categories—spatial domain fusion and domain fusion transformation. Spatial domain approaches come under fusion strategies such as averaging, Brovey process, principal component analysis (PCA) and IHS-based methods. The high pass filtering dependent technique is another significant spatial domain fusion tool. The high frequency knowledge is injected into an upsampled edition of MS photos here. The downside of spatial domain approaches is that in the fused picture, they create spatial distortion. Spectral distortion [19, 20] becomes a detrimental influence when we begin sorting, such as the issue of grouping. Spatial distortion may be treated very well by methods to image fusion in the frequency domain. Multiresolution processing [12] has been a very valuable method for remote sensing pictures to be examined. A very valuable method for fusion has been the discrete wavelet transform. There are also several other ways of fusion, such as Laplacian pyramid centred, curvelet transform based, etc.

2 Literature Review

Zhang et al. (2020), suggested a general system for image fusion based on the convolutionary neural network, called IFCNN. They use two convolutionary layers to remove the salient image features from several input images in the experiments. The algorithm used is very focused on the theory of transform-domain image fusion algorithms. The convolutionary characteristics of several input images are then combined by a suitable fusion rule. The proposed model is entirely convolutionary, but without any post-processing procedures it could be trained from first of this experiment to last. In addition, the findings also check that equal or even better results have been obtained [1].

Ma et al. (2020), proposed a method which helps the fused picture to maintain the thermal radiation in the infrared image with sharpening infrared target limits and the abundant textural information in the visible image. Owing to their substantially discriminatory thermal emission, this research aims to easily observed from the backdrop of infrared photographs, whereas visible images provide large spatial resolution textural information that are helpful to improving target detection. Their findings improved infrared photographs with distinctly highlighted and edge-sharpened objectives as well as ample details in depth [2].

Zhang et al. (2020), proposed a multi-focus image fusion technique which is a form of enhancement to produce full-clear photographs that can overcome the constraint of field-depth in optical lens imaging. This paper deals subjective perceptual effects and objective measurements for multi-focus pictures. They suggested a GAN model as unsupervised learning algorithms for multi-focus image fusion [3].

Tan et al. (2020) documented a broad range of medical images of patient's organ. Each of these pictures represents a process which makes the organ unique and contributes to several observations of specific phenomena (such as stroke). A detailed analysis of all plans allows recognizing more acceptable treatment options. Multimodal medical imagery is a research field which involves the creation of robust algorithms that allow the combination of picture information from a variety of methods. A modern multimodal image fusion algorithm is introduced in this article for a number of medical diagnostic issues. They also tested the algorithm in a data collection including more than 100 pairs of images like glioma, bronchogenic alzheimer's and metastatic carcinoma. Quantitative and qualitative review guarantees that the suggested algorithm may not surpass all of the algorithms that exist and offers useful ideas for medical diagnostic [4].

De Juan et al. (2019) studied all data processing methods which intended to combine multifaceted image data from the same network or related spectroscopic systems. The fusion of images can imply the development, by regression models of each associated picture, of a single multiset or multiway structure with all involved images. Any situation also overcomes the product of individual image processing with the data analysis performed on fused image systems. They found many diverse scenarios and applications for image fusion, ranging from 3D hyperspectral images to other related 2D image simulation of systems, quantitative investigation or the easy usage of spectroscopic know-how from various spectroscopic platforms [5].

Trinidad et al. (2019) suggested three cases for image fusion which are colour conversion for the colour inferences on the monochrome display, HDR fusion to merge malaligned brake exposure and data transfer from a cost-beneficial camera in order to reproject a high-definition image [6].

Amin-Naji et al. (2019) focused on Convolutionary Neural Networks (CNNs). Compared to the previous state of the art approaches that were performed in the spatial and transform realms, they substantially increased the built judgement map. Nevertheless, these approaches have not achieved a satisfactory initial decision chart, and in order to obtain a satisfactory decision map, they must undergo rigorous post-processing algorithms. In this article, a novel approach focused on CNNs is suggested with the aid of ensemble learning. Using multiple templates and datasets rather than just one is quite reasonable. In addition, the proposed approach incorporates a modern simple form of dataset of multi-focus images [7].

Xia et al. (2019), suggested a modern multi-modal medical image fusion scheme that incorporates both the multi-scale transformation characteristics and the deep convolutionary neural network. Experimental findings reveal that our proposed approach not only achieves better results by fusing the multiple images effectively, but also guarantees that the different quantitative parameters are enhanced relative to other current approaches [8].

Li et al. (2018), proposed multi-scale transformations are expected by most current multi-modal picture fusion approaches. However, this condition does not always contribute to the fusion outcome containing the original source picture strength, and a strong computational sophistication is needed for multi-scale transformations. They discussed the issue of multi-modal image fusion with a low computational complexity in the spatial domain in this article [9].

Shao and Cai (2018) proposed a residual learning which used to extensively research the relationship between high- and low-resolution pictures. The system suggested comprises primarily of two methods viz Spatial and spectral characteristics which are derived by convolutional layers of separate depths of both pictures, respectively. The proposed approach offers improved outcomes relative to other classical approaches by analyzing the output of the QuickBird and Gaofen-1 photos which consider as universe data set for image fusion [10].

Zhu et al. (2018), proposed a multi-modality image fusion which is an efficient approach for fusing complementary details into an integrated image from multi-modality pictures. This approach not only increases the clarity of human eyes, but also complements each image's limitations [11].

Liu et al. (2018), proposed the fusion of the same scene's infrared and visible images which aims to creating a combined picture that can offer a more accurate explanation of the scene. In this paper, they recommend a system of infrared and visible image fusion based on convolutionary neural networks (CNNs). In particular, to obtain a weight map that combines the pixel behavior data from two source pictures, a Siamese convolutionary network is implemented. In image fusion as a whole this CNN-based method coped with two critical problems, as calculation of activity level and weight assignment [12].

Song et al. (2018) suggest a new method in spatiotemporal fusion based upon deep neural networks (CNNs). We are designing a fusion model in the prediction point that includes high-passe modulation and weighing technique so that information can be used as far as possible in the previous images, rather than merely taking CNN findings as a fusion result. In specific, they first map the MODIS input pictures to temporary pictures by way of a non-linear Mapping method (CNN) and further boost photos migrating to Landsat by using the fusion model. Then the LSR Landsat images are converted into transitory pictures by the learning SR CNN, which are further improved by the fusion model for Landsat images [13].

3 Compares and Findings

The decomposition systems, decomposition and fusion technique instruments used in the techniques mentioned in the literature are shown in Table 1. Multi-scale decomposition is used in all common techniques. The efficacy of the final production depends primarily on this method of decomposition and the fusion technique used.

4 Conclusion

Image fusion is an approach used to amalgamate the corresponding characteristics into a single composite image in a series of input images that retains all the essential features of the input images. Pansharpening is often referred to as image fusion. It is a process used for combining and incorporating the geometric data of a panchromatic (Pan) high-resolution picture and the colour details of a multispectral (MS) low-resolution image for the creation of a high-resolution MS image. This paper also explained the different image/picture strategies in a detailed and comparable way in this article. As mentioned here, numerous different kinds of techniques are very effective in producing fused images that are more appropriate for identification, tracking and visual awareness. Table above is the summary version of all the surveyed techniques, and their strengths and demerits are mentioned.

Table 1 Comparison of different image fusion methods

Author (year) [ref]	Proposed method	Image type	Fusion strategy	Fusion type
Zhang et al. (2020) [1]	Convolutional neural network	Multi-focus image	Transform-domain image fusion	Transform-domain image fusion
Ma et al. (2020) [2]	GANs for IR/VIS fusion	Infrared and visible image	Edge-enhancement loss to improve the quality	Infrared and visible image fusion
Zhang et al. (2020) [3]	Unsupervised GAN model	Multi-focus image	Adaptive decision block	Multi-focus image fusion
Tan et al. (2020) [4]	Multimodal image fusion algorithm	Medical Images	Pulse-coupled neural network fusion	Multimodal medical image fusion algorithm
de Juan et al. (2019) [5]	Incomplete multisets or structures	3D hyperspectral	Subpixel motion step in Time	Multimodal image fusion
Trinidad et al. (2019) [6]	Novel cascaded Feature extraction method	Multiple misaligned image	HDR fusion and transfer for reprojecting a high-definition image	Multi-view image fusion
Amin-Naji et al. (2019) [7]	CNNs architecture for multi-focus image fusion	Multi-focus test images	Fused image among the other state	Multi-focus image fusion
Xia et al. (2019) [8]	Multi-scale transformation and deep convolutional neural network	Multi-modal medical images	Low frequency images fusion process	Medical Image Fusion
Li et al. (2018) [9]	Spatial domain with a low computational complexity	Medical images	Structure-preserving filter	Multi-modal image fusion
Shao and Cai (2018) [10]	Deep convolutional neural network	Remote sensing images, panchromatic (PAN) images	Feature fusion procedure	Remote sensing image fusion
Zhu et al. (2018) [11]	Source multi-modality images	Multi-modality images	Multi-modality	Multi-modality image fusion
Liu et al., (2018) [12]	Convolutional neural networks (CNNs)	Infrared and visible images	Fusion method based on convolutional neural networks (CNNs	Infrared and visible image fusion
Song et al. (2018) [13]	Deep convolutional neural networks	MODIS and LSR Landsat images	CNNs-based spatiotemporal method	Spatiotemporal satellite image fusion

References

1. Zhang Y, Liu Y, Sun P, Yan H, Zhao X, Zhang L (2020) IFCNN: A general image fusion framework based on convolutional neural network. Inf Fusion 54:99–118
2. Ma J, Liang P, Yu W, Chen C, Guo X, Wu J, Jiang J (2020) Infrared and visible image fusion via detail preserving adversarial learning. Inf Fusion 54:85–98
3. Zhang H, Le Z, Shao Z, Xu H, Ma J (2020) MFF-GAN: an unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion. Inf Fusion 66:40–53
4. Tan W, Tiwari P, Pandey HM, Moreira C, Jaiswal AK (2020) Multimodal medical image fusion algorithm in the era of big data. Neural Comput Appl 1–21
5. de Juan A, Gowen A, Duponchel L, Ruckebusch C (2019) Image fusion. In: Data handling in science and technology. Elsevier, vol 31, pp 311–344
6. Trinidad MC, Brualla RM, Kainz F, Kontkanen J (2019) Multi-view image fusion. In: Proceedings of the IEEE international conference on computer vision, pp 4101–4110
7. Amin-Naji M, Aghagolzadeh A, Ezoji M (2019) Ensemble of CNN for multi-focus image fusion. Inf Fusion 51:201–214
8. Xia KJ, Yin HS, Wang JQ (2019) A novel improved deep convolutional neural network model for medical image fusion. Clust Comput 22(1):1515–1527
9. Li W, Xie Y, Zhou H, Han Y, Zhan K (2018) Structure-aware image fusion. Optik 172:1–11
10. Shao Z, Cai J (2018) Remote sensing image fusion with deep convolutional neural network. IEEE J Sel Top Appl Earth Observations Remote Sens 11(5):1656–1669
11. Zhu Z, Yin H, Chai Y, Li Y, Qi G (2018) A novel multi-modality image fusion method based on image decomposition and sparse representation. Inf Sci 432:516–529
12. Liu Y, Chen X, Cheng J, Peng H, Wang Z (2018) Infrared and visible image fusion with convolutional neural networks. Int J Wavelets Multiresolut Inf Process 16(03):1850018
13. Song H, Liu Q, Wang G, Hang R, Huang B (2018) Spatiotemporal satellite image fusion using deep convolutional neural networks. IEEE J Sel Top Appl Earth Observations Remote Sens 11(3):821–829
14. Yan H, Li Z (2020) Infrared and visual image fusion based on multi-scale feature decomposition. Optik 203:163900
15. Liu X, Liu Q, Wang Y (2020) Remote sensing image fusion based on two-stream fusion network. Inf Fusion 55:1–15
16. Farid MS, Mahmood A, Al-Maadeed SA (2019) Multi-focus image fusion using content adaptive blurring. Inf Fusion 45:96–112
17. Wang H, Li S, Song L, Cui L (2019) A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals. Comput Ind 105:182–190
18. Liu Y, Chen X, Wang Z, Wang ZJ, Ward RK, Wang X (2018) Deep learning for pixel-level image fusion: recent advances and future prospects. Inf Fusion 42:158–173
19. Tang H, Xiao B, Li W, Wang G (2018) Pixel convolutional neural network for multi-focus image fusion. Inf Sci 433:125–141
20. Vargas E, Espitia O, Arguello H, Tourneret JY (2018) Spectral image fusion from compressive measurements. IEEE Trans Image Process 28(5):2271–2282

Designing Framework for Intrusion Detection in IoT Based on Spotted Hyena-Based ANN

Archana Bathula, Samya Muhuri, Suresh Merugu, and Suneet K. Gupta

Abstract Internet of Things (IoT) is the platform for resource sharing amidst the computing platforms, mainly memory, power, large data, and storage. Securing the data in the IoT has become a prime concern as the hackers can access through the invaluable information from the IoT database. Once the attacks are incurred on the IoT environment, the loss of data is inevitable, and it can have a major effect on the progress of the IoT platform. In this paper, an Intrusion Detection System (IDS) with improved artificial intelligence is proposed. This paper makes use of standard benchmark datasets from diverse sources for performing the experiment. A well-performing machine learning algorithm called Artificial Neural Network (ANN) is developed with improved network architecture. The usage of a renowned meta-heuristic algorithm called Spotted Hyena Optimization (SHO) is used for selecting the optimal hidden neurons for ANN. The main objective of the improved training is to reduce the error difference between the target and the measured outputs to enhance the detection accuracy. Finally, the experimental outcomes and simulations prove the stability and robustness of the proposed model in terms of a variety of performance metrics over other machine learning models.

Keywords Internet of Things · Intrusion detection system · Artificial neural network · Security · Spotted hyena optimization

A. Bathula (✉)
Computer Science and Engineering, CMR College of Engineering & Technology, Hyderabad, India

S. Muhuri · S. K. Gupta
Computer Science Engineering, Bennett University, Greater Noida, India
e-mail: samya.muhuri@bennett.edu.in

S. K. Gupta
e-mail: suneet.gupta@bennett.edu.in

S. Merugu
Head Research and Development, CMR College of Engineering & Technology, Hyderabad, India
e-mail: msuresh@cmrcet.org

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_153

1 Introduction

Internet of Things (IoT) is the latest evolutionary technology on Internet, which provides the connectivity to each device on the concerned applications [1]. Detection of intruders in the system is considered a prime concern for developing secured IoT environments. The existing intrusion detection systems (IDS) can be divided into different groups like misuse/signature-based detection systems, specification-based intrusion detection, behavioural or anomaly-based detection, and hybrid intrusion detection systems [2]. Moreover, the intrusion detection is done over the host level or the network level, and this is particularly selected based on the security policy driven by the IDS monitoring system. Most research on the IDS is primarily focused on the rule-based detection techniques [3]. However, the rule-based detection does not deal effectively with the zero-day threats. For this case, it is necessary to build the anomaly-based detection techniques for the IoT.

The machine learners can do the automated data processing and analysis, thus can easily interpret the data in the IoT. Thus, the machine learners can interpret the actions of the IoT access, and can easily predict the presence of the hacker during the data access [4]. The anomaly-based detection techniques along with the machine learners can come in handy for detection the known or unknown attacks in IoT, for the practical cases. The various machine learners like cluster analysis [5], statistical analysis [3], ANN [6] and deep learning [7] are effectively employed in intrusion detection. Moreover, the use of machine learning algorithms for the intrusion detection has proved more effective, since it is successful in dealing with a large amount of data [8]. The main contribution of this work towards establishing a secured environment for the IDS is given below:

To establish a secure model with the use of the machine learner like ANN and the hidden neurons in it are optimally selected by the SHO heuristic algorithm and developed a framework namely, Spotted Hyena-Artificial neural network (SH-ANN) and improved the detection accuracy. The datasets from standard modules are considered for implementing this model and compared it with the state-of-the-art models.

The structure of the paper is arranged in the following manner: Sect. 2 briefs some literary works with the description of their features and challenges in the field of the intrusion detection. Section 3 explains the proposed SH-ANN architecture and the description of the dataset used for the experimentation. A brief description of the proposed SH-ANN model is explained here and its results are discussed in Sect. 4. Finally, the conclusion of this work is depicted in Sect. 5.

2 Literature Survey

2.1 Related Works

In 2019, Chaabouni et al. [9] This paper has evaluated the Network Intrusion Detection Systems (NIDSs) for various network threats. For the implementation, various free and open-source tools were used and here the results were compared with the existing models. The efficiency of the NIDS was evaluated for the success against the security and privacy. In 2019, Zhang et al. [10] The improved version of the genetic algorithm was employed for learning the parameters for the deep belief network (DBN). The improved GA has selected the optimal number of hidden layers, and hidden neurons, and the training of DBN was focused on various network attacks. High detection rate was achieved through the DBN as the optimal layers were used for the training. In 2019, Anthi et al. [11] have developed a three layer IDS, and the supervised approach was employed for the detection purpose. The learning task has analyzed the presence of various cyber-attacks on the IDS platform. The system effectiveness was tested against the complex events. In 2020, Almiani et al. [12]. Here, a multilayered Recurrent Neural Network (RNN) was employed for creating a fog computing security, and the layer was designed in such a way that was near to the end users. Various metrics were used and the robustness of the scheme was proved against other models. In 2020, Pacheco et al. [13] An adaptive IDS model was designed on the basis of the ANN and it has proven that the model had the capability for the characterization of the abnormal behavior in the fog nodes. The presence of various abnormalities like misuses, cyber-attacks, and system glitches were identified for any sources. Further, the detection rate was achieved to be high, and the false alarm rate was low.

2.2 Problem Statement

Security of IoT has become a critical concern. Some of the features as well as the challenges of the existing IDS are depicted in Table 1. Deep RNN [12, 23] offers security against the cyber-attacks and it further effectively works in the real time environments. But, it results in high false alarm rates. ANN [13, 21] provides a low false alarm rate and high detection rates. Still, the major limitations are the high cost of signature matching and network packet overload. ML [9, 24] examines the free as well as the open source network sniffing software and also offers a good success rate in privacy and security. Yet, it cannot function in a real-world IoT-dedicated dataset. GA and DBN [10, 22] finds the low frequency attacks happening in the intrusion detection systems and it can further be employed in the recognition and classification. But, the accuracy is not enhanced and the training time is too minimized. Supervised technique [11, 25] provides various resources that help the research into the automating IoT-oriented cyber-attack detection and the malicious

Table 1 Features and challenges of detection state-of-the-art techniques for the intrusion in IoT

Author [citation]	Methodology	Features	Challenges
Chaabouni et al. [9]	ML	• It provides good success rate in privacy and security	• It does not deeply explore the Big Data, fog, and edge computing methods
Zhang et al. [10]	GA and DBN	• It can be used in the recognition as well as classification. And also, to find the low frequency attacks occurring in the IDS	• It does not optimize various parameters of the deep network and does not enhance the accuracy and minimize the training time
Anthi et al. [11]	Supervised technique	• It recognizes the malicious packets on the network when an attack occurs • It offers various resources for helping the research into the automating IoT-based cyber-attack detection	• It cannot be implemented in industrial IoT, heterogeneous IoT, and real environment • It does not use the deep learning techniques for identifying the malicious activities occurring inside the IoT environment
Almiani et al. [12]	Deep RNN	• It can work effectively in the real time environments	• It returns high false alarm rates
Pacheco et al. [13]	ANN	• It offers low false alarm rate and high detection rates	• The high cost of signature matching and network packet overload are the major drawbacks

packets on the network are recognized when an attack occurs. Still, the deep learning techniques are not used to recognize the malicious activities that occur inside the IoT environment.

3 Architectural View of Intrusion Detection System for Internet of Things

3.1 Proposed Model and Description

In this work, a novel intrusion detection framework has been developed for detecting the intruders in the IoT platform and its architecture is presented in Fig. 1. The IoT

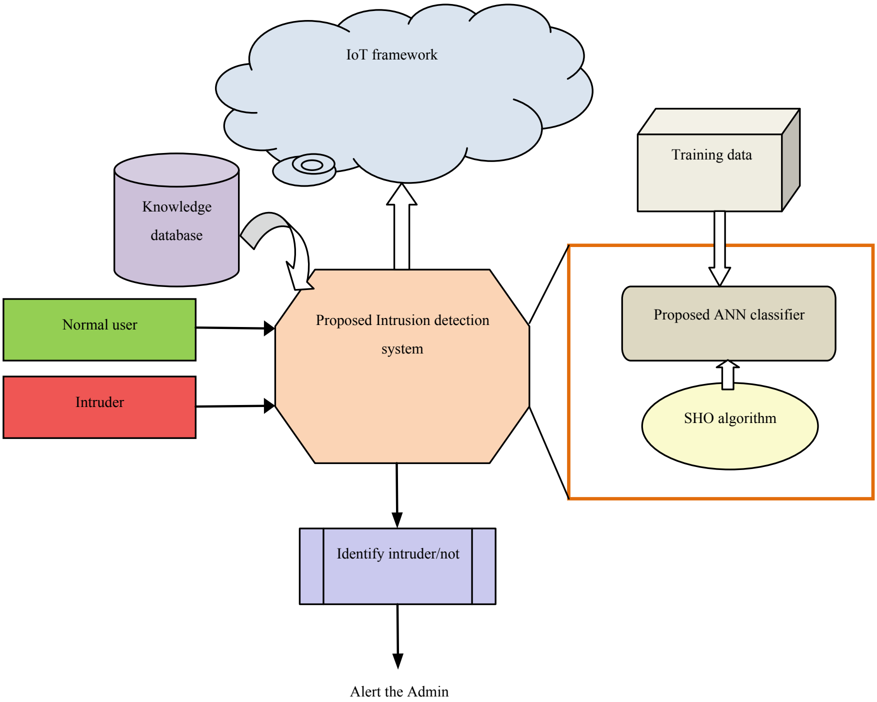


Fig. 1 Block diagram of the proposed Intrusion detection system with the SHO-ANN classification

devices are connected virtually through the wired/wireless medium. Once a user enters the system to access the devices in the IoT cloud, they are checked by the proposed IDS model. The proposed IDS are trained with the sufficient data through the training database with the proposed ANN classifier to determine the user to be an intruder normal. Here, in the proposed work, the ANN follows an optimal structure to enhance the detection accuracy and structure of the ANN is influenced by the Spotted Hyena Optimization (SHO), and it allows the ANN classifier to minimize the error difference between the target and the measured outputs. After the training of ANN, a knowledge database is created by ANN. In the testing phase, when the user enters the IDS, the ANN checks the knowledge base, and declares the user to be normal or an intruder. If the intruder is detected, the IDS directly alert the admin, and block the access of the intruder into the IoT, such a way the proposed system secures the communication in the IoT.

3.2 DataSet Description

For this work, three standard databases like, KDDcup99 dataset [14], IoTID20 [15], IoT Botnet dataset [16] are considered.

3.3 Data Normalization

Data normalization allows regularizing the data contents in the database. Data normalization eliminates the presence of redundant information in the database. The procedure for the data normalization is as follows: Consider the database P from the IoT, and it is subjected for the data normalization. For the data normalization, it is necessary to consider two data variables $g = 0$ and $h = 0$. The data variables g and h refers to the minimum and the maximum normalized values. Equation (1) represents the formula for the mathematical computation of the normalized data from the database P.

$$P^{\text{norm}} = (g - h) * \frac{\left(P_b - P^{\min}\right)}{\left(P^{\max} - P^{\min}\right)} + g \tag{1}$$

Here, P^{norm} refer to the normalized database, and the term P_b indicates the bth data element in the database to be normalized. The terms $P^{\max}$ and $P^{\min}$ indicates the minimal and the maximal value in each data record of the database P.

3.4 Proposed SH-ANN

After the data normalization, the normalized data input is provided to the proposed SH-ANN module for the detection of the intruders or the normal users in the IDS module. The normalized data are fed to the ANN for the training, and the architecture of the proposed SH-ANN for the intrusion detection is depicted in Fig. 2. In the proposed SH-ANN, the hidden neurons present in the ANN architecture are selected. The basic architecture of the ANN module is explained here. ANN classifier is a feed forward neural network with many inputs and the hidden layers. The feed forward ANN module consists of three layers, input, output and the hidden units. The training of ANN is done through the back-propagation algorithm, and the intrusion detection is carried out. The functioning of the ANN layer is given as follows: The normalized data input is feed to each layer of the input units, and then they are multiplied by the weight provided by the hidden units. Equation (2) depicts the data unit received by each hidden units of ANN.

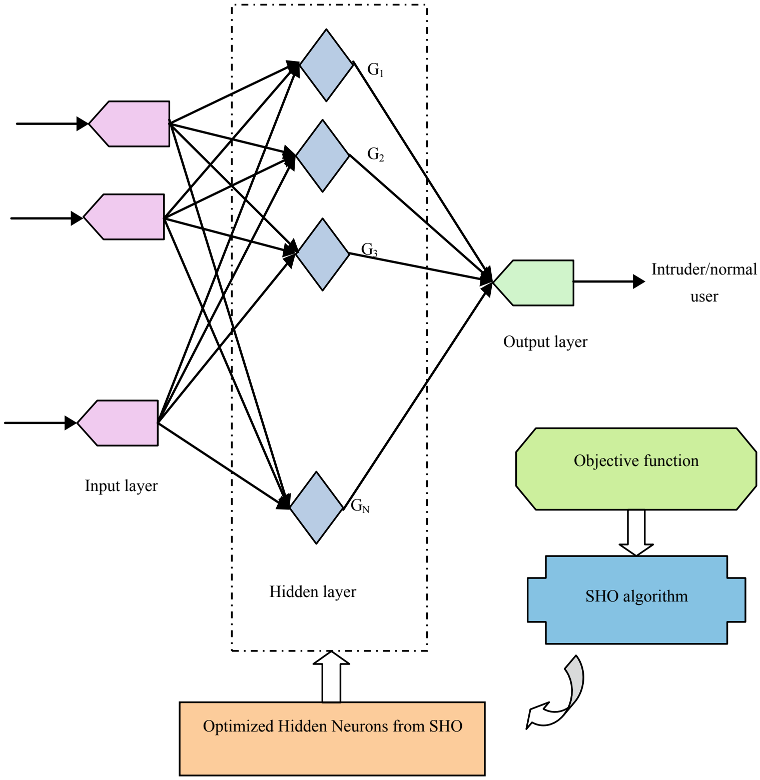


Fig. 2 Architecture of the proposed spotted hyena based artificial neural network

$$G(k) = \beta_k + \sum_{m=1}^{D} X_m Y_{mk} \tag{2}$$

Here, β_k indicate the bias provided for the kth hidden unit, and X_m indicate the mth data input. The weight between the kth hidden unit and the mth input layer is represented as Y_{mk}. After passing through the hidden unit, the data is passed over the bipolar sigmoid activation function, and it is represented in Eq. (3).

$$B(x) = \frac{2}{(1 + \exp(-x))} - 1 \tag{3}$$

The hidden units are passed over the output layers through the activation function, and Eq. (4) represents the final output from the ANN layer.

$$O = \alpha + \sum_{k=1}^{N} Y_k B(G(k)) \tag{4}$$

Here, α indicates the bias for the output layer. During the ANN training, the response of the classifier is compared with the true value to identify and rectify the error deviation, and it is represented in Eq. (5).

$$Q = \frac{1}{2\Phi}\sqrt{Tr - O} \tag{5}$$

Here, Tr refer to the target value of the ANN classifier, and O indicates the predicted response. The weights of the ANN are selected using the gradient descent algorithm during the training process. After the training, the ANN classifier predicts the presence of intruder in the IDS platform. It is necessary to optimize the total hidden neurons for the training of the ANN classifier. For this procedure, the SHO is used in this work. The SHO algorithm is inspired from the way of hunting of spotted hyenas. The SHO algorithm [20] performs the optimization procedure through three conventional steps (1) piercing the prey, (2) encompassing the prey, and (3) striking the prey. After the piercing is done, the prey is encompassed, and the final best solution is identified. The solution space for the SHO algorithm is initially randomly set, and based on the distance parameter of the search agent, the solution is update. Equation (6) expresses the distance measure between the prey and the search agent.

$$T_{ij} = \left|W.R_{prey}(s) - R(s)\right| \tag{6}$$

Here, T_{ij} indicates the distance interval between the ith search agent and the jth prey. The terms $R_{prey}(s)$ and $R(s)$ indicates the distance vector for the prey and the search agent respectively. The term W indicates the position vector calculated through the Eq. (7).

$$W = 2 \cdot w_1 \tag{7}$$

Here, the value of the constant w_1 varies in range of [0,1]. After piercing the location of the prey, the position of the search agent gets updated as in Eq. (8).

$$R(s+1) = R_{prey}(s) - U.T_{ij} \tag{8}$$

Here, U is the coefficient for the update, and it is expressed as in Eq. (9).

$$U = 2a \cdot w_2 - a \tag{9}$$

Here, a is the notation for the minimization, and its value gets reduced from 5 to 0, and the term w_2 can have value similar to w_1. For determining the value of a during the optimization procedure, Eq. (10) is used.

$$a = 5 - \left(s * \left(\frac{5}{\max_s}\right)\right) \tag{10}$$

The solution update mainly depends on the position of the best pierce agent that directs the solution to move faster towards the optimal solution. The hunting characteristics of the search agent are inspired from the Eqs. (11), (12), and (13), respectively.

$$T_{ij} = \left| W \cdot R_{ij} - R_{sf} \right| \tag{11}$$

$$L_{sf} = L_{ij} - U \cdot T_{ij} \tag{12}$$

$$V_{ij} = R_{sf} + R_{sf+1} + \cdots + R_{sf+E} \tag{13}$$

Here, V_{ij} indicates the total number of optimal solutions, and E refer to the total count of the spotted hyenas considered for the solutions, and it can be represented in Eq. (14).

$$E = C_{num}\left(R_{ij}, R_{ij+1}, \ldots, (R_{iv} + b)\right) \tag{14}$$

Here, the term C_{num} indicates the total number of the candidate solutions, and the random vector b ranges in the value of [0,1]. Finally, the prey attack is symbolized using the Eq. (15).

$$R(s+1) = \frac{V_{ij}}{E} \tag{15}$$

In the final expression, the optimal solution is pierced by the optimal location of the pierce prey agent, and thus the position resides on the V_{ij} vector. The final optimal solution is retained at the end of the iteration, and the optimal hidden neurons are selected for the ANN network.

3.5 Objective Evaluation

For the objective evaluation, fitness is derived based on the accuracy and precision metrics. Thus, the fitness for the SHO is selected as the arg max function, with the hidden neurons providing the maximal accuracy and precision. The expression for the objective function of SHO is given as in Eq. (16).

$$Obj = \underset{\{G_N\}}{\arg\max}\{A + P\} \tag{16}$$

Here, A and P indicates the accuracy and precision respectively. The SHO algorithm selects N hidden neurons represented as G_N having the maximal accuracy and precision. Equation (17) and Eq. (18) provide the mathematical formulation of the accuracy and precision.

$$A = \frac{tr_P + tr_N}{tr_P + tr_N + fs_P + fs_N} \tag{17}$$

$$P = \frac{tr_P}{tr_P + fs_P} \tag{18}$$

Here, tr_P, tr_N, fs_P, and fs_N indicates the response of the classifier as true positive, true negative, false positive, and false negative respectively.

4 Results and Discussion

4.1 Simulation Setup

The entire simulation for the proposed SH-ANN based IDS environment was implemented on the PYTHON tool. The simulation results were compared with the other state-of-the-art techniques such as ANN [17], PSO-ANN [18], and GWO-ANN [19].

4.2 Evaluation Metrics

Here, the evaluation is done by considering the metrics like accuracy, precision, sensitivity, specificity, False Positive Rate (FPR), False Negative Rate (FNR), False Discovery Rate (FDR), Negative Predictive Value (NPV), F1-score, and Matthews Correlation Coefficient (MCC).

4.3 Analysis for the Dataset 1

Figure 3 presents the analysis of the proposed SH-ANN based on accuracy on dataset 1. While the training % of 35 is chosen from the dataset, the proposed SH-ANN achieved accuracy greater than 4.25% to ANN, greater than 2% to PSO-ANN, and greater than 2% to GWO-ANN. Similarly, for the training % as 85, the proposed SH-ANN achieved accuracy superior than 4.8% to ANN, superior than 2.8% to PSO-ANN, and superior than 2.1% to GWO-ANN.

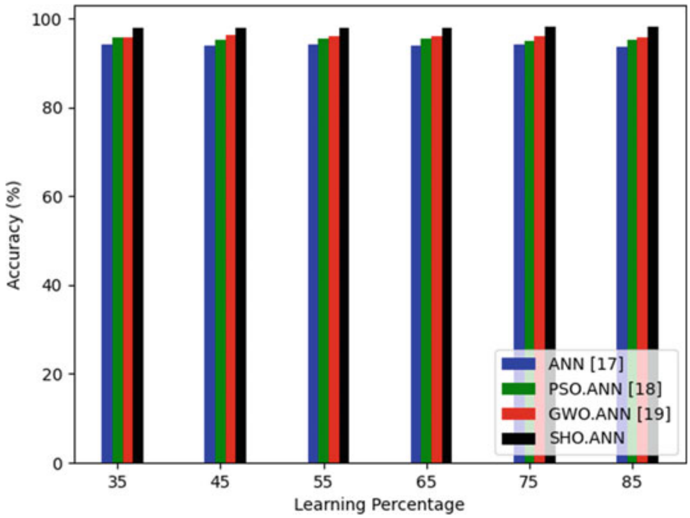


Fig. 3 Analysis of proposed SH-ANN based on the dataset 1

4.4 Analysis for the Dataset 2

Figure 4 depicts the analysis of the proposed SH-ANN against the existing models based on accuracy on the dataset 2.

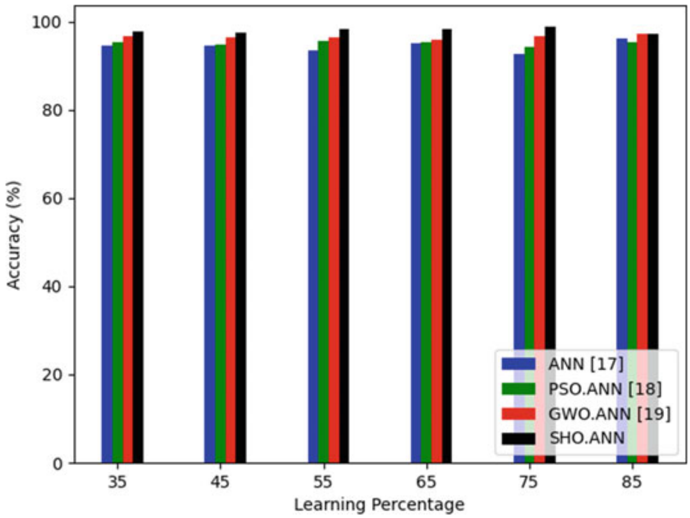


Fig. 4 Analysis of proposed SH-ANN based on the dataset 2

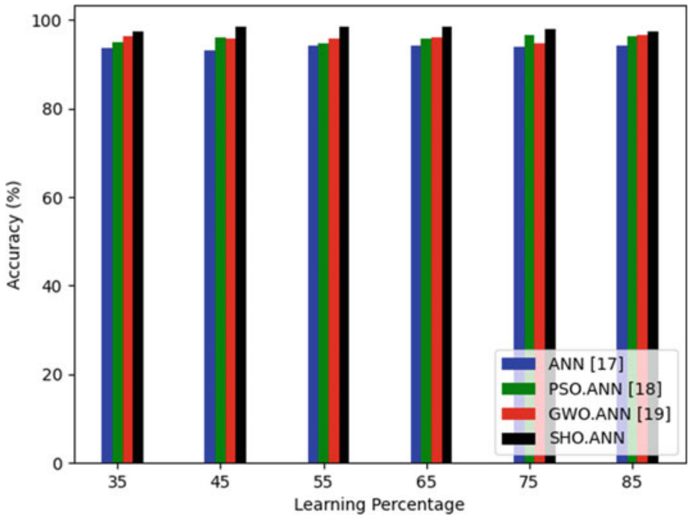


Fig. 5 Analysis of proposed SH-ANN based on the dataset 3

For the training % of 35 from the dataset 2, the proposed SH-ANN achieved accuracy greater than 4.5% to ANN, greater than 2.6% to PSO-ANN, and greater than 2.1% to GWO-ANN. Similarly, for the training % as 65, the proposed SH-ANN achieved accuracy superior than 3.1% to ANN, superior than 2.6% to PSO-ANN, and superior than 2.1% to GWO-ANN.

4.5 Analysis for the Dataset 3

Figure 5 presents the analysis of the proposed SH-ANN against the comparative models based on accuracy on the dataset 3. While the training % of 35 is chosen from the dataset 3, the proposed SH-ANN achieved accuracy greater than 4% to ANN, greater than 2.5% to PSO-ANN, and greater than 2.1% to GWO-ANN. Also, for the training % as 85, the proposed SH-ANN achieved accuracy improved than 3.7% to ANN, improved than 1.5% to PSO-ANN, and improved than 1.5% to GWO-ANN.

4.6 Overall Analysis

The overall analysis of the proposed SH-ANN model is explained in this section. The proposed SH-ANN is analyzed for the data contents from the dataset 1, dataset 2, and dataset 3 over 10 different metrics. Table 2 presents the overall analysis of the proposed SH-ANN model for the dataset 1. Here, the proposed SH-ANN had

Table 2 Overall analysis of the proposed SH-ANN based IDS for the dataset 1

Performance metrics/methods	ANN	PSO-ANN	GWO-ANN	Proposed SH-ANN
Accuracy	0.9410	0.9511	0.9621	0.9816
Sensitivity	0.9392	0.9507	0.9597	0.9803
Specificity	0.9427	0.9514	0.9642	0.9827
Precision	0.9357	0.9456	0.9597	0.9806
FPR	0.0573	0.0486	0.0358	0.0173
FNR	0.0608	0.0493	0.0403	0.0197
NPV	0.9427	0.9514	0.9642	0.9827
FDR	0.0643	0.0544	0.0403	0.0194
F1_score	0.9374	0.9482	0.9597	0.9804
MCC	0.8816	0.9019	0.9239	0.9630

achieved high accuracy over the state of art techniques with superior values of 4.31% to ANN, 3.2% to PSO-ANN, and 2% to GWO-ANN respectively. Table 3 presents the overall analysis of the proposed SH-ANN for the dataset 2. Here, the proposed SH-ANN had achieved high MCC over the state of art techniques with superior values of 53% to ANN, 38% to PSO-ANN, and 13.5% to GWO-ANN respectively. Table 4 explains the overall analysis of the proposed SH-ANN over the dataset 3. The proposed SH-ANN had achieved high sensitivity over the state of art techniques with superior values of 4% to ANN, 1% to PSO-ANN, and 3.1% to GWO-ANN respectively.

Table 3 Overall analysis of the proposed SH-ANN based IDS for the dataset 2

Performance metrics/methods	ANN	PSO-ANN	GWO-ANN	Proposed SH-ANN
Accuracy	0.9277	0.9434	0.9688	0.9883
Sensitivity	0.9315	0.9461	0.9668	0.9896
Specificity	0.8667	0.9000	1.0000	0.9667
Precision	0.9912	0.9935	1.0000	0.9979
FPR	0.1333	0.1000	0.0000	0.0333
FNR	0.0685	0.0539	0.0332	0.0104
NPV	0.8667	0.9000	1.0000	0.9667
FDR	0.0088	0.0065	0.0000	0.0021
F1_score	0.9604	0.9692	0.9831	0.9938
MCC	0.5871	0.6523	0.7941	0.9020

Table 4 Overall analysis of the proposed SH-ANN based IDS for the dataset 3

Performance metrics/methods	ANN	PSO-ANN	GWO-ANN	Proposed SH-ANN
Accuracy	0.9400	0.9667	0.9480	0.9787
Sensitivity	0.9394	0.9648	0.9479	0.9775
Specificity	0.9500	1.0000	0.9500	1.0000
Precision	0.9970	1.0000	0.9970	1.0000
FPR	0.0500	0.0000	0.0500	0.0000
FNR	0.0606	0.0352	0.0521	0.0225
NPV	0.9500	1.0000	0.9500	1.0000
FDR	0.0030	0.0000	0.0030	0.0000
F1_score	0.9674	0.9821	0.9718	0.9886
MCC	0.8816	0.9019	0.9239	0.9630

5 Conclusion

This paper has presented an IDS framework for detecting the presence of intruders in IoT framework. A machine learning approach was integrated with the optimization model for the effective handling of various intruders. This work has developed the SH-ANN framework specifically for IoT model, such that the timely detection of intrusion can be done. The proposed SH-ANN model was developed by using the ANN as the machine learner, and the SHO optimization algorithm for the structure refinement of machine learning. The SHO algorithm has selected the optimal hidden neurons for the ANN, such that the efficacy of the intrusion detection was enhanced. The proposed SH-ANN had achieved high accuracy over the state of art techniques with superior values of 4.31% to ANN, 3.2% to PSO-ANN, and 2% to GWO-ANN respectively. This proves the efficiency of the proposed model in the intrusion detection in the IoT framework.

References

1. Prabavathy S, Sundarakantham K, Shalinie SM (2018) Design of cognitive fog computing for intrusion detection in Internet of Things. J Commun Netw 20(3):291–298
2. Chaabouni N, Mosbah M, Zemmari A, Sauvignac C, Faruki P (2019) Network intrusion detection for IoT security based on learning techniques. IEEE Commun Surv Tutorials 21(3): 2671–2701
3. Lin F, Zhou Y, An X, You I, Choo K-K-R (2018) Fair resource allocation in an intrusion-detection system for edge computing: ensuring the security of Internet of Things devices. IEEE Consum Electron Mag 7(6):45–50
4. Fahim M, Sillitti A (2019) Anomaly detection analysis and prediction techniques in IoT environment: a systematic literature review. IEEE Access 7:81664–81681
5. Yang A, Zhuansun Y, Liu C, Li J, Zhang C (2019) Design of intrusion detection system for Internet of Things based on improved BP neural network. IEEE Access 7:106043–106052

6. Benkhelifa E, Welsh T, Hamouda W (2018) A critical review of practices and challenges in intrusion detection systems for IoT: towards universal and resilient systems. IEEE Commun Surv Tutorials 1–1
7. Manimurugan S, Al-Mutairi S, Aborokbah MM, Chilamkurti N, Ganesan S, Patan R (2020) Effective attack detection in internet of medical things smart environment using a deep belief neural network. IEEE Access 8:77396–77404
8. Moustafa N, Turnbull B, Choo KR (2019) An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of Internet of Things. IEEE Internet Things J 6(3):4815–4830
9. Chaabouni N, Mosbah M, Zemmari A, Sauvignac C, Faruki P (2019) Network intrusion detection for IoT security based on learning techniques. IEEE Commun Surv Tutorials 21(3):2671–2701
10. Zhang Y, Li P, Wang X (2019) Intrusion detection for IoT based on improved genetic algorithm and deep belief network. IEEE Access 7:31711–31722
11. Anthi E, Williams L, Słowińska M, Theodorakopoulos G, Burnap P (2019) A Supervised intrusion detection system for smart home IoT devices. IEEE Internet Things J 6(5):9042–9053
12. Almiani M, AbuGhazleh A, Al-Rahayfeh A, Atiewi S, Razaque A (2020) Deep recurrent neural network for IoT intrusion detection system. Simul Model Pract Theor 101
13. Pacheco J, Benitez VH, Félix-Herrán LC, Satam P (2020) Artificial neural networks-based intrusion detection system for Internet of Things Fog nodes. IEEE Access 8:73907–73918
14. KDDcup99dataset: http://kdd.ics.uci.edu/databases/kddcup99/task.html
15. IOT intrusion detection dataset: https://sites.google.com/view/iot-network-intrusion-dataset/home
16. IOTBOTNETdataset: https://sites.google.com/view/iotbotnetdatset/home
17. Wang G, Hao J, Ma J, Huang L (2010) A new approach to intrusion detection using artificial neural networks and fuzzy clustering. Expert Syst Appl 37(9):6225–6232
18. Marin, F, Walczak B (2015) Particle swarm optimization (PSO). A tutorial. Chemomet Intel Lab Syst 149:153–165
19. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Adv Eng Softw 69:46–61
20. Dhiman G, Kumar V (2017) Spotted hyena optimizer: A novel bio-inspired based metaheuristic technique for engineering applications. Adv Eng Softw 114:48–70
21. Merugu S, Jain K, Mittal A, Raman B (2019) Sub-scene target detection and recognition using deep learning convolution neural networks. In: ICDSMLA 2019. Lecture notes in electrical engineering, pp 1082–1101
22. Yadav BC, Merugu S, Jain K (2019) Error assessment of fundamental matrix parameters. In: ICCCE 2018. Lecture notes in electrical engineering, vol 500. ISSN 1876-1100
23. Syed A, Merugu S, Kumar V (2020) Augmented reality on sudoku puzzle using computer vision and deep learning. Adv Cybern Cogn Mach Lect Notes Electr Eng 567–578
24. Koppula VK, Soumya DS, Merugu S (2020) Nurse alarming device for bedridden patients using hand gesture recognition system. In: Advances in cybernetics, cognition, and machine. Lecture notes in electrical engineering, pp 377–385
25. Merugu S, Reddy MCS, Goyal E, Piplani L (2019) Text message classification using supervised machine learning algorithms. In: ICCCE 2018. Lecture notes in electrical engineering, vol 500. ISSN 1876-1100

Object Detection System for Visually Impaired Persons Using Smartphone

D. Ravi Kumar, Hiren Kumar Thakkar, Suresh Merugu, Vinit Kumar Gunjan, and Suneet K. Gupta

Abstract Portable assistive technology systems are developed to enhance the capability of persons with disabilities. One of the most important senses for humans is vision. Vision is the most important sense which helps us in understanding the perception of the surrounding environment. Visually impaired persons face a lot of difficulty in understanding the perception around them, particularly in outdoor environment where objects are continuously changing and moving from one place to another. Object detection solutions would greatly assist visually impaired persons in avoiding from the barriers which they face in their daily routine life. The aim of the object detection system is to provide a simple, user-friendly, handy, economical and efficient solution for the visually impaired persons. Motive of this system is to develop a solution that detects the objects present using camera, as the input device in real time and communicate the same to the user using smartphone through headphones. The system would be using an audio device such as speakers or headphones in providing the information about objects to assist the visually impaired persons. The proposed system helps in identifying and avoiding the objects both in outdoor and indoor environments that affect day to day life activities and occupational performance of the visually impaired persons. The information about the objects in surrounding environment would be very much helpful to the visually impaired persons in their daily life.

Keywords Object detection system · Visually impaired · Blindness · Assistive system · Computer vision · Image recognition

D. R. Kumar (✉) · S. Merugu
Department of Research & Development, CMR College of Engineering & Technology, Hyderabad, India
e-mail: dravikumar@cmrcet.org

H. K. Thakkar · S. K. Gupta
Department of Computer Science Engineering, Bennett University, Greater Noida, India

V. K. Gunjan
Department of Research & Development, CMR Institute of Technology, Hyderabad, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_154

1 Introduction

There are several organizations that take cognizance of the welfare of the blind and visually impaired. [1] The World Health Organization has reported that there are 2.2 billion people who are either visually impaired or completely blind through a recent study conducted by them. [2] In India, the Ministry of Health and Family Welfare of the Government of India has conducted a survey called "The National Blindness & Visual Impairment Survey India 2015–2019". It was conducted for approximately 85,000 people. The results of the survey showed the estimation of the number of people with vision problems. The results are presented for the age group of 50 and above. The prevalence for blindness is 1.99% and the prevalence for visual impairment is 13.76%. The survey data supports that there is an astonishingly large number of people with vision related problems. The numbers of people that need support are huge and a simple, efficient system is needed in place for their assistance. A successful execution of a supporting system will have a large impact on the ways and livelihood of the people with visual impairment.

The challenges that the completely blind and visually impaired people face directly impact the quality of their lives and also how do they perform their daily activities. The traditional method is that of a supportive guide. A person is needed who can efficiently guide them in travelling from one place to another. The white cane is a helpful aid in navigating with obstacles in near proximity of the person. It can only help identify obstacles that are very close to the person. This brings up the issue of safety of the blind and visually impaired. Independent mobility also is matter of concern. An object detection system is the need of the hour to combat the safety issues and also enable them to have mobility that is not dependent on anyone. The object detection system has to be a simple and most importantly user friendly. It also should be budget friendly so that it can be afforded by them.

We are fortunate to have technologies that can make a huge difference in the life of the blind and visually impaired. Some advanced technologies like computer vision, object detection and object identification are available in today's world. The camera of a smart phone can be used as the vision apparatus which helps in the object detection system. It then detects the objects that are in front of the camera. Yolo V3 is an object detection model that can be used for detection of objects in a real time video. The detected object can be identified and then communicated as an audio message to the user. It can be done either on the speaker or using a head phones. This application is designed to be used on an Android based smartphone. Smart phones not only help in the regular telephonic conversations but also help the blind and visually impaired if the smart phones have a camera. The primary goal of the application is to enable the user identify the object in front of the camera. The computer vision techniques supported by the object detection systems can enhance the quality of lives of the blind and visually impaired by many fold. We are proposing a system that is simple and user-friendly. It is not only handy but also economical and efficient. This overcomes the issues of an expensive, additional sophisticated system. The use of smartphones is not just for the super advanced applications. A common man is able to use a smartphone with a camera in a budget friendly manner. The proposed system will not need any additional hardware other than those are already available in a smartphone. Smartphone comes with an integrated camera. It can be used to capture the real time video. This in turn is used to detect the objects. The identified object is communicated to the user as an audio instruction. This paper highlights the computer vision techniques and presents the overview of the object detection system, which can enhance the quality of lives of the visually impaired persons.

2 Literature Survey

Various aids have been designed to develop a system that provide visual assistance and which help the persons facing vision problems in their daily routine life by communicating about the obstacles present in front of them. The following literature review presents the outcome and the limitations of some of the existing assistive solutions for blind and visually impaired.

In [3], authors proposed an android based system which captures the images through the smart phone camera and communicates the user about the object detected, the distance and direction of the object to the user via audio instructions using speakers or headphones. The proposed android based system provides the perception of the surrounding environment and enables them to navigate independently. The system has three modules: Object detection module, Distance calculation module and Direction. The object is classified and module isolate different objects present in the image is done in the object detection module. Object detection is done using TensorFlow's object detection API. Open source computer vision library (OpenCV) and Triangle similarity law are used to determine the distance of the object. Using 'x' coordinates of the bounding box of the object, the direction 'Left', 'Center' and 'Right' is determined and communicated the same to the visually impaired persons. The proposed system was able to detect with an accuracy of 87%. Some of the limitations identified are the distance calculation functionality is yet to be included in the system. The proposed system can be retrained to detect more objects.

The aim of the system [4] is to minimize the need of dedicated and wearable devices by providing an android smartphone based application system that assist visually impaired persons in recognizing objects and in moving around. The R-CNN algorithm is used for classification of real time objects. Image processing algorithms are used to process the video captured by the inbuilt camera of the smartphone. After objects are scanned and segmented, Fast R-CNN algorithm is used to identify multiple regions. Later, using Google API text to speech conversion is done to communicate the same to the user. The proposed system revealed that better accuracy is achieved in detecting with low latency by SSDMobileNetV2 model and performed 35% better in speed than MobileNetV1 when paired with TensorFlow Lite. The system has a limitation to focus on better performance in detecting multiple objects with better accuracy.

The goal of the proposed systems in [5–7] is to develop an Android application which provides necessary information about the objects and environment to the visually impaired persons. The proposed system minimizes the usage of the conventional dedicated devices and wearable devices. The paper also provides an overview of the object detection and significance of ORB over SIFTS and SURFS in different cases. The video is captured by the camera then the video is divided into frames. Based on the key matching in the database the image edge detection is processed then text to speech conversion is done using synthesizer which instructs the user.

The objective of the papers [8, 9] is to review the different navigation solutions available for visually impaired persons that work both in indoor and outdoor environment. A thorough review of the existing navigation solutions in both indoor and outdoor is made. In Assistive Travel Aids, the author classified available aids into three major categories: electronic orientation aids (EOAs), position locator devices (PLDs) and electronic travel aids (ETAs). Electronic orientation aids (EOAs)—used in finding the navigation path to the visually impaired persons. Position locator devices (PLDs)—both Global positioning system (GPS) and Geographic information system technologies are used to guide user navigation. Electronic travel aids (ETAs)—help visually impaired persons to avoid obstacles. Some of the limitations observed are in EOAs, it is difficult to add complex computing devices. Using PLDs, the device cannot detect objects and ETAs can be improved in the detection range of the obstacles. The study found lacking of some core features like portability, comfort ability, adaptation time and user learning which are very helpful and important for independent mobility in the existing navigation systems for visually impaired persons.

To develop an intelligent assistance system [10] composed of wearable smart glasses, mobile device app, an intelligent walking stick and an online information platform that assist in object detection. The components wearable smart glasses, mobile device app, an intelligent walking stick and an online information platform are composed in the system. Whenever the smart glasses detect objects it reminds user by smart walking stick. If the user falls down then the related information is stored online and shares the same information with their family members. Apart from object detection assistance, the major advantage in using this device is that whenever user falls down or any collision event happens then related information is sent to family members. A few limitations observed are features like recognizing front images like traffic signs and to develop intelligent walking guide are yet to be included.

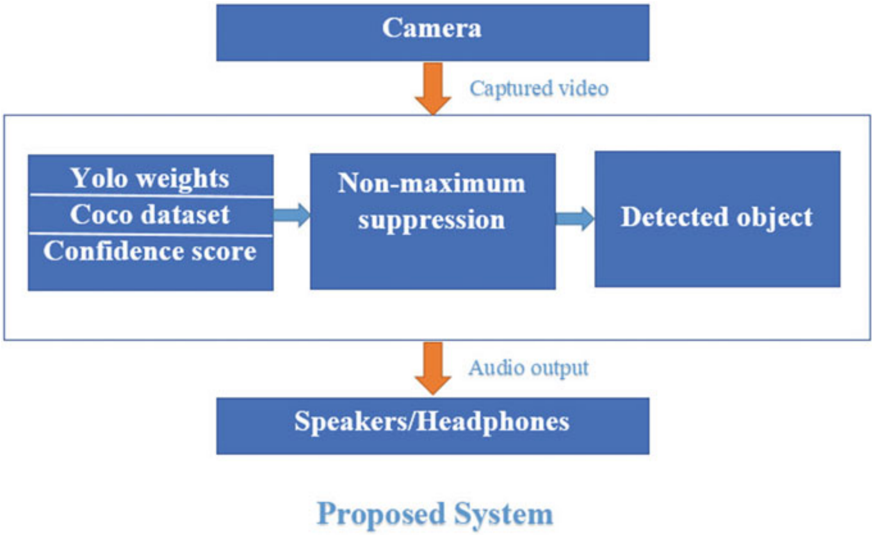
The aim of the device in [11] is to provide the safe navigation assistance and perception of the environment in the real time. The obstacle avoidance system in the device learns from the RGBD, pilots choice-of-action input and semantic map, it ensures to provide safe and appropriate feedback about the surrounding environment to the visually impaired. A few limitations of the device are the sonar and bump sensors are yet to be incorporated to ensure the safety of the user in exception cases.

Matusiak et al. [12, 13] explained about the main features of the smartphone Android application software developed to assist visually impaired persons. The software has three modules, where the main module in the system recognizes and matches the scanned objects with the objects stored in the system database. The other two modules are capable of identifying direction of the major colors regions in the captured images. The SIFT (Scale-Invariant Feature Transform), a computer vision algorithm is used in the proposed application to describe image features.

Devi et al. [14], Bhatnagar et al. [15] explains about determining obstacles and potholes and notify the user through audio output and to share active location of the user with the care taker. The obstacles detection part is carried out by ultrasonic sensors attached to the device interacts with the Microcontroller then the respective voice commands are given to the user. Potholes are identified by the infrared sensors and vibrates the aid to let the user know. GSM gathers the active location of the user [16, 17]. Whenever user press panic button then emergency assistance request is made. Apart from travelling assistance, emergency request is also included. The limitation observed in the system is since all components are placed on stick, weight may be a constraint and repeated hitting of stick on ground may cause malfunction of the device.

3 Proposed System

Visually impaired persons face many challenges in identifying objects and performing the day to day activities. Most of the time they are dependent on others in moving from one place to other. They face a lot of difficulty especially when moving in outdoor environment, where objects are continuously moving. This object detection system would assist the visually impaired persons by providing the perception of the surrounding environment or position of the objects. This device can assist visually impaired persons in avoiding the obstacles in both indoor and outdoor environment. The proposed system also helps the user in identifying the objects around them. It would minimize the visually impaired person's difficulties and help them lead a quality life.

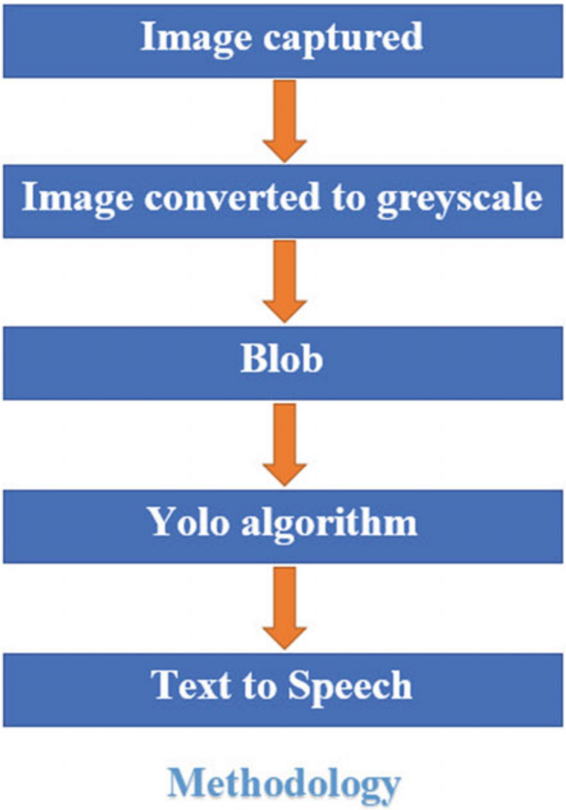


The proposed system aims to be simple, user-friendly, handy, economical and efficient solution. Most of the existing assistive systems are highly sophisticated and expensive. Smartphones are widely used these days and the usage of smartphones by all persons has become very usual in the recent past. Therefore this system uses all the advanced built-in features of a smartphone. Smartphone's integrated camera is used to capture the real time video to detect the objects present and headphones or smartphone's speaker is used to communicate to the user through audio instructions.

The proposed application can easily be accessed by the visually impaired persons. The Yolo V3 (You only look once) algorithm is used to detect the real time objects captured from the continuous streaming by the smartphone camera. It is considered as one of the most powerful real time object detector algorithms. Unlike other algorithms like R-CNN and Faster R-CNN which examine several regions of the image to identify objects, Yolo passes image or video only one time through its network and uses a unique neural network using the characteristics of the complete image to predict multiple boxes containing an object. This is a significant feature in Yolo, which reduces the processing speed when compared with other algorithms. Processing speed plays a key role particularly while detecting objects in real time video stream. OpenCV provides a function that facilitates image pre-processing for deep-learning classification. Group of connected pixels in an image that share common properties (Blob ex: Grayscale value) of each frame captured is identified. Bounding boxes are created for each object identified and based on the Yolo V3 pre-trained weights, confidence score and coco dataset each object is processed and labelled. The coco dataset contains all the objects or class names on which the model is trained. Non-maximum suppression (NMS) uses function called "Intersection over Union (IoU)" which is used to determine the best boxes. In order to select the best box, NMS follows three steps. 1. It selects the box with highest score. 2. Computes the overlap with other boxes, removes the overlap which has more than a certain threshold. 3. Iterates the process until there are no more boxes with a lower score than the currently selected box. The above NMS steps will remove all the boxes that have a large overlap with the selected boxes. Only the best boxes remain. After detecting the object by the proposed model, it produces the text format output. Then using text to speech converter, the label of the object detected is converted to speech and communicated the same to the user through headphones.

4 Methodology

The object detection, recognition and communicating speed of the object to the user should be very fast. Since the proposed system is intended for visually impaired persons, the response time of the system plays a key role. Delay in communicating about the obstacles to the user will not meet the purpose. To overcome processing speed and delay issues, Yolo V3 algorithm is used in the proposed system. Yolo V3 algorithm is faster than many other real time object detecting algorithms.



1. Captures the objects in the surrounding environment in real time and this is done by using the camera of the smartphone. The video will get divided into frames
2. The RGB color image is converted into greyscale image
3. Group of connected pixels in an image that share common properties of each frame captured is identified (Blob)
4. Based on the Yolo V3 pre-trained weights, confidence score and coco dataset each object is processed and labeled (Yolo V3 algorithm)
5. Using text to speech converter, the label of the object detected is converted to speech

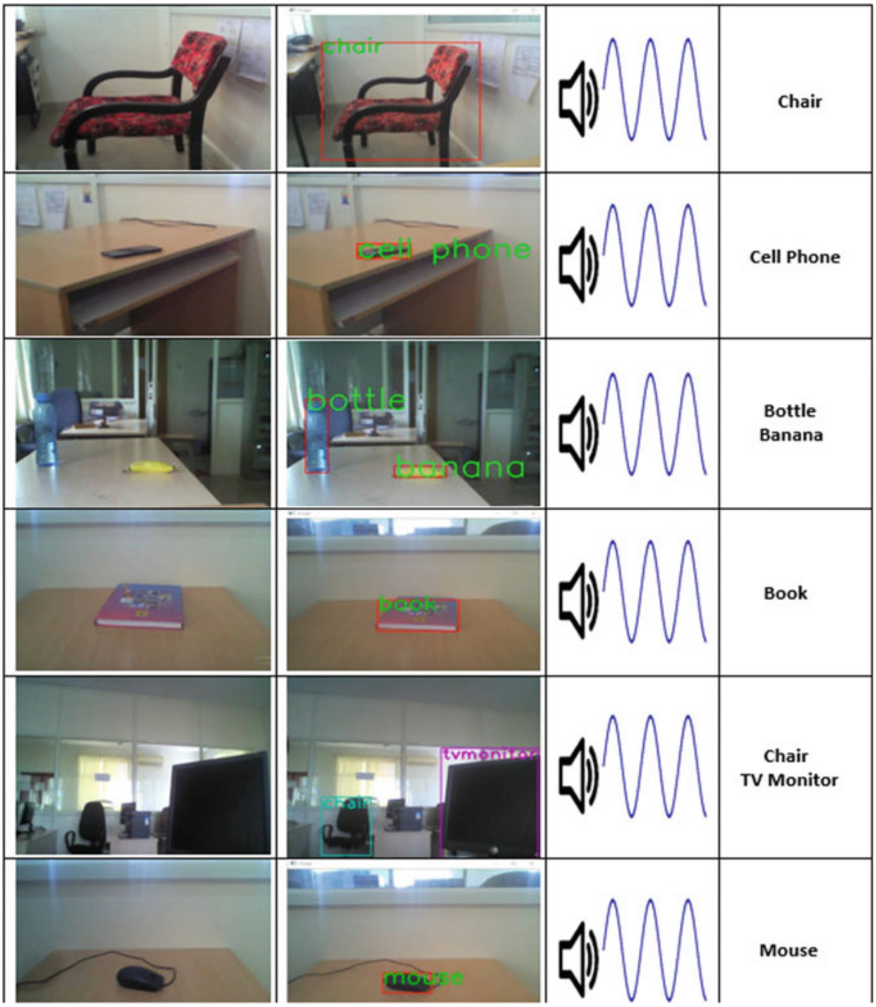
5 Results

The main motive of the proposed system is to assist visually impaired persons by providing the perception of the environment, which helps them in avoiding obstacles or barriers and in moving from one place to another. The goal to provide a simple, user-friendly and handy solution is achieved. The proposed system is capable of detecting the objects present in the surrounding environment with good speed and accuracy. It detects objects effectively in both outdoor and indoor environment. The system is able to successfully detect the multiple objects present in the surrounding environment and communicate the same to the user in audio through headphones or speaker.

The proposed system is tested in detecting objects in indoor environment, outdoor environment and objects which are more than 10 m from the camera. The system is capable of detecting the objects in the surrounding environment and provide audio output to the user. The performance of the application in the above mentioned three categories is satisfactory. The results captured are presented below.

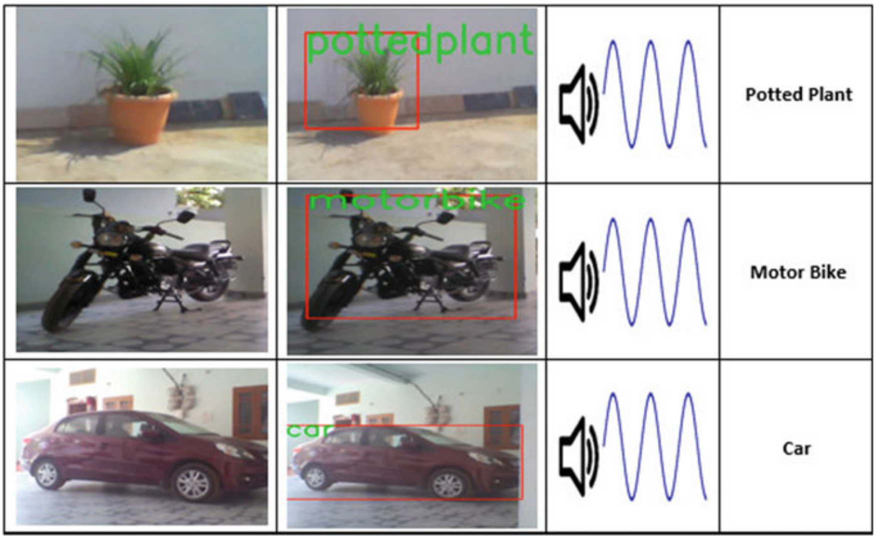
5.1 Indoor Environment

The following figure show the objects captured by the built in camera and the detected objects with the labels and the respective audio output of the proposed system.

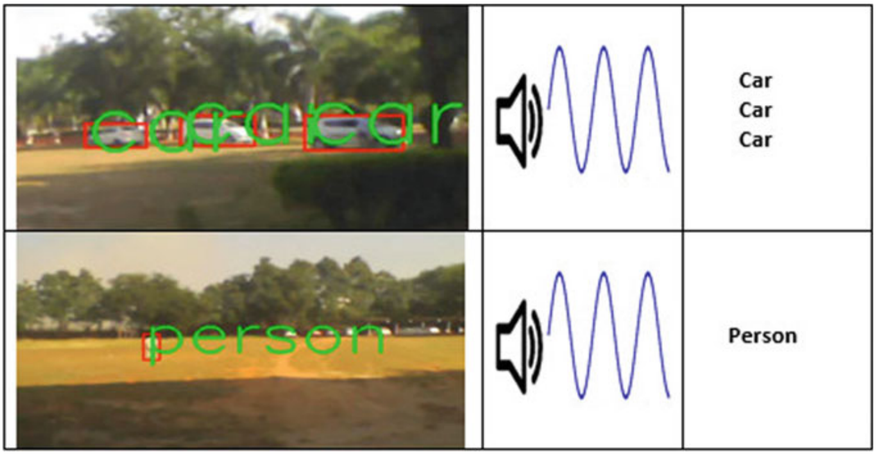


5.2 *Outdoor Environment*

The following figure shows the objects captured by the built in camera and the detected objects with the labels and the respective audio output of the proposed system.



The proposed system successfully detected the objects which are far away in the outdoor environment with good accuracy. Following are the objects that are detected in the outdoor environment. The system was successful in detecting the objects which were more than ten meters away in the outdoor environment.



6 Future Scope

The performance of the system can be further improved by re-training on larger datasets. Providing information about direction of the object will help the user to understand the perception better. The system can be improved in detection time and better accuracy.

7 Conclusion

In recent years, there were many assistive solutions developed for the visually impaired persons to provide assistance in detecting objects present in their surrounding environment and in moving from one place to another. But most of the existing solutions are expensive, highly sophisticated, difficult to handle, designed as a dedicated aid, require training etc. Our primary goal is to provide an assistive solution to the visually impaired persons which is simple, user-friendly, affordable and handy and assist the visually impaired persons in understanding the surrounding environment and help them in moving from one place to another place independently. The real time video stream is captured by the smartphone camera and then the label of the object detected using object detection algorithm is communicated to the user in speech through the speakers or headphones. The audio output communicated to the user would help them in performing their day to day activities and lead a quality life.

Acknowledgement & Funding Statement: Authors would like to thank the Deanship of Scientific Research at CMR College of Engineering & Technology, Hyderabad for supporting this work under TIDEDST Project No. SEED/TIDE/2018/55.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. World Health Organization, "World Report on Vision" https://www.who.int/publications-detail/world-report-on-vision
2. National Programme for Control of Blindness & Visual Impairment (NPCBVI), Government of India "The National Blindness & Visual Impairment Survey India 2015–2019". https://npcbvi.gov.in/writeReadData/mainlinkFile/File341.pdf
3. Badave A, Jagtap R, Kaovasia R et al (2020) Android based object detection system for visually impaired. In: International conference on industry 4.0 technology 2020, pp 34–38. https://doi.org/10.1109/I4Tech48345.2020.9102694
4. Anitha J, Subalaxmi A, Vijayalakshmi G (2019) Real time object detection for visually challenged persons. Int. J. Innov. Technol. Explor. Eng. 8:312–314
5. Jadhav PSP, Tomy S, Jayswal SS et al (2016) Object detection in android smartphone for visually impaired users. 5:332–334. https://doi.org/10.17148/IJARCCE.2016.51171
6. Khairnar DP, Karad RB, Kapse A et al (2020) PARTHA: A visually impaired assistance system. In: Proceedings of 2020 3rd international conference on communication system computing IT application CSCITA 2020, pp 32–37. https://doi.org/10.1109/CSCITA47329.2020.9137791

7. Sait U, Ravishankar V, Kumar T et al (2020) Design and development of an assistive device for the visually impaired. Procedia Comput. Sci. 167:2244–2252. https://doi.org/10.1016/j.procs.2020.03.277
8. Kuriakose B, Shrestha R, Sandnes FE (2020) Tools and technologies for blind and visually impaired navigation support: a review. IETE Tech Rev (Institution Electron Telecommun Eng India) 1–16. https://doi.org/10.1080/02564602.2020.1819893
9. Manjari K, Verma M, Singal G (2020) A survey on assistive technology for visually impaired. Internet of Things 11:100188. https://doi.org/10.1016/j.iot.2020.100188
10. Chen LB, Su JP, Chen MC, et al (2019) An Implementation of an intelligent assistance system for visually impaired/blind people. In: 2019 IEEE International conference on consumer electronics ICCE 2019, pp 4–5. https://doi.org/10.1109/ICCE.2019.8661943
11. Lin Y, Wang K, Yi W, Lian S (2019) Deep learning based wearable assistive system for visually impaired people. In: Proceedings of 2019 International conference on computer vision workshop ICCVW 2019, pp 2549–2557. https://doi.org/10.1109/ICCVW.2019.00312
12. Matusiak K, Skulimowski P, Strurnillo P (2013) Object recognition in a mobile phone application for visually impaired users. In: 2013 6th international conference on human system interaction (HSI 2013), vol 17, pp 479–484. https://doi.org/10.1109/HSI.2013.6577868
13. Matusiak K, Skulimowski P, Strurnillo P (2013) Object recognition in a mobile phone application for visually impaired users. In: 2013 6th International conference on human system interaction (HIS), pp 479–484. https://doi.org/10.1109/HSI.2013.6577868
14. Devi RA, Uthaman I, Shanthanam RR (2020) An IOT Security based electronic aid for visually impaired detection with navigation assistance system, p 6
15. Bhatnagar V, Chandra R, Jain V (2019) IoT Based alert system for visually impaired persons. Springer, Singapore
16. Anand S, Kumar A, Tripathi M, Gaur MS (2019) Human face detection enabled smart stick for visually impaired people. Springer, Singapore
17. Jain S, Sushanth D, Varsha V et al (2019) Design and implementation of the smart glove to aid the visually impaired. In: Proceedings of 2019 IEEE international conference on communication and signal process (ICCSP 2019), pp 662–666. https://doi.org/10.1109/ICCSP.2019.8698009

Agricultural Crowdfunding Through Blockchain

Naga Venkata Mohit Desabathina, Suresh Merugu, Vinit Kumar Gunjan, and Bandreddi Sunil Kumar

Abstract Farmers are the backbone of India's economy but every year there is an increase in the number of suicides in the agriculture sector. The main reason behind their suicides is the financial problems like debt, loan repayment, no security for the loan and the crowdfunding scams and frauds. In the financial sector of agriculture, there are many actors/participants involved and their transactions have to be managed. Blockchain-based agricultural crowdfunding will eradicate all the middlemen and connect the consumer to the farmer directly. Blockchain technology creates a platform between all the participants and provides a shared ledger/database to deliver an immutable and unique version of the truth among all the actors/participants in the network who don't trust each other. Blockchain-based crowdfunding in agriculture creates a market for community members where they can financially back agriculture operations in exchange for groceries and establish a peer to peer network, with donations directly and uninterruptedly going to farmers.

Keywords Crowdfunding · Blockchain · Agriculture · Technological interventions · Smart contracts

1 Introduction

Crowdfunding for agriculture can be defined as raising funds for the farmers, from a group of people instead of an individual centralized entity like banks, insurance

N. V. M. Desabathina
Department of Information Technology, Fr. Conceicao Rodrigues Institute of Technology, Vashi, Navi Mumbai, Maharashtra, India

S. Merugu (✉)
CMR College of Engineering & Technology, Hyderabad, Telangana, India
e-mail: msuresh@cmrcet.org

V. K. Gunjan
CMR Institute of Technology, Kandlakoya, Hyderabad, Telangana, India

B. S. Kumar
Cluster IT Solutions Private Limited, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_155

companies and big farmers. It's an execution of collecting many minute contributions to finance or capitalize on agricultural activities. The blockchain-based crowdfunding process mainly involved two parties, which are the contributors and farmers [1–6]. Despite the efforts of the Government and NGOs, millions of smallholder farmers remain unprotected against financial and agricultural risk as they have limited awareness and access to the crowdfunding platforms. There are multiple problems faced by the traditional crowdfunding platforms such as trust, platform, managing compliance and accounting Issues, centralized architecture and many more. Blockchain technology remains on top for solving all those issues [7, 8].

In short, the blockchain network is composed of data units called blocks, which are digitally linked to each other. In general, each block consists of three elements; a hash of that particular block, a pointer to the hash of the previous block and data (this includes the list of transactions, state of the network header of the block, etc.). Blockchain provides both transparency and security. One of the major features of blockchain technology is that every transaction's (successful/unsuccessful) entry is reflected on each node's ledger in the blockchain network [9, 10]. Blockchain is immutable, that means it's almost impossible to change that transaction's entry, once the transaction is committed then it can't be altered, making blockchain technology the easiest and safest method to transfer funds from consumer to farmer in a completely automated and secure way [11, 12]. Blockchain-based crowdfunding allows us to create a digital contract that will allow the consumer to send funds to the farmers only when a certain amount of funds are raised, otherwise, the funds may return to the consumer in a transparent and safe manner. Hence the Blockchain-based agricultural crowdfunding system protects the consumer's funds and farmer's.

2 Types of Blockchain

There are primarily four types of blockchain:

1. Public blockchain—Public blockchain is also called a permission less blockchain as anyone can join or leave the blockchain network and can verify and append transactions. Private blockchain network is decentralized, visible to the public and allows the dynamic collection of participants who may not even know each other.
   Examples: Ethereum, Bitcoin, Litecoin, etc.
2. Private blockchain—Permissioned or Private blockchain operates only at an organizational-level or at a closed network. Nodes with the specific Identities are given access for joining or leaving the blockchain network and thereby can verify, add records and append transactions to the blockchain network. Like a public blockchain, even the private blockchain could be publically visible but the participants are known to each other and trust each other.
   Examples: Hyperledger projects (Fabric, Sawtooth), Multichain, Corda, etc.

3. Hybrid blockchain—It is the union of the private blockchain and public blockchain, which means it is the combination of the transparency benefits of a public blockchain and privacy benefits of a permissioned blockchain. With such a hybrid network, users/actors in the network have complete authority over his data and decide who gets access to which part of the data in the network. Simple hybrid blockchain can be made by joining a private blockchain with multiple public blockchains [13–15].
   Example: Dragonchain.
4. Consortium blockchain—When a group of organizations or nodes forms a consortium to govern the blockchain network is called as a consortium blockchain network. It is also known as a semi-decentralized blockchain. The process of verification and adding new records is done on the bases of the consensus algorithms made by those preselected nodes.
   Examples: Quorum, R3, etc.

3 Research Method

As the blockchain-based agricultural funding involves the personal detail of farmers and the fund contributor, so using the private blockchain is the best option. Crowdfunding requires direct communication with the contributor and the farmer in real-time when the time arises, that can be achieved by using Hyperledger Fabric. Hyperledger Fabric framework provides distributed ledger solutions on permissioned networks for all ranges of industrial applications. The architecture of hyperledger fabric increases the flexibility of blockchain-based solutions, transparency, user compliance, confidentiality and resilience.

The Hyperledger Fabric network is comprised of:

1. Smart contract: Smart contract is also known as chaincode and contains the network's business logic. It can be written in languages like golang, java or javascript (Nodejs). The chaincode is invoked whenever the decentralized application needs to interact with the ledger.
2. Asset: An asset is a real-world object that has a value and has both state and ownership.
3. Peer nodes: They are a basic entity in the blockchain network as they are responsible for hosting the distributed ledger. It executes chaincode, interfaces with applications, access ledger's data and endorses transactions. The chaincode also contains an endorsement policy, which comprises mandatory and necessary conditions for a valid transaction to be endorsed or committed. Depending upon the roles of the peer nodes, they can act as committing peers or endorsing peers.
4. Shared ledger: It records the ownership and state of the asset. It mainly comprises of two parts:

   (i) The history of the transaction log that contains the transaction's records.

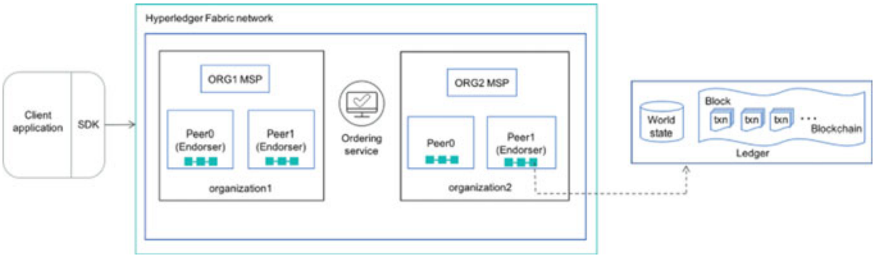


Fig. 1 Hyperledger fabric architecture

(ii) World state database tells the present state of the ledger. It acts as the database of that particular ledger.

5. Organizations: The organization is a conceptual entity that accesses channels (ledgers) and consists of peers. For example, when we develop a hospital network on the blockchain then the doctors or patients are peers and the hospital is the organization. Every peer in the organization has an identity assigned by the Membership Services Provider (MSP) stating that the peer belongs to that particular organization.
6. Channel: Channels are sensible structures built by the collection of peers. There may be internal channels in multiple organizations or a single organization. One single channel can accommodate peers or nodes belonging to the different organization, allowing those groups of peers/nodes to form a separate individual ledger, consisting of transactions.
7. Ordering services: It wraps the transactions into the chunks/blocks that should be sent to the respective peers/nodes in the channel and communicates with endorsing peers and committed peers. It ensures the delivery of transactions between peers on the network. Kafka and Solo are the only two configuration mechanisms for the Ordering service that are supported.
8. Membership Services Provider (MSP): It acts as a Certification Authority that manages the certification used for authentication of member's identity and their roles. Membership Services Provider is responsible for managing the member's identity and authentication of the actors in the blockchain network, making Hyperledger Fabric a permissioned and private blockchain network. Unknown members/identity cannot be a part of the hyperledger network or make a transaction in the network (Fig. 1).

4 Implementation

Initially, the network configuration has to be set up and then develop the application for interaction with the blockchain network by setting up the Nodejs server for processing the requests to the network. The client (angular) is used for bringing up

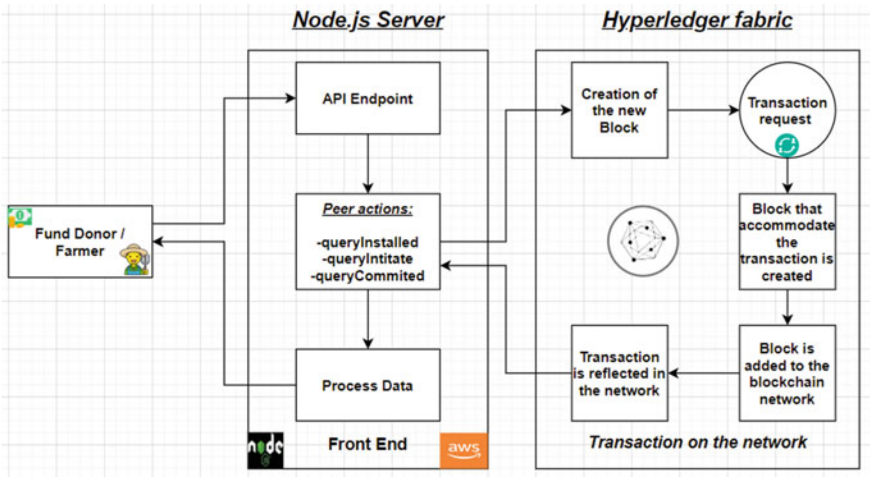


Fig. 2 System architecturere

the client's web interface. The following steps have to be followed while developing the application from scratch:

1. Generating the crypto/certificate using cryptogen tool.
2. Generating the configuration transaction using configtxgen tool.
3. Bring up the nodes in the network based on what is defined in the docker-compose file.
4. Using the command-line interface to set up the network.
5. Using the command line interface CLI to install and instantiate the chaincode.
6. Using the command line interface to invoke and query the chaincode.
7. Develop the Nodejs server, which will communicate with the deployed network.
8. Create the frontend (Angular) for the decentralized application to act as an interface for the blockchain network (Fig. 2).

The complete project is made on Hyperledger framework. The client-interface is browser-based, which means it is implemented on the bootstrap framework using Javascript, HTML 5, andCSS. The actors in the decentralized application are Fund Contributor and Farmer.

The functionality of the Fund Contributor is given Fig. 3.

The functionality of the Farmer is given Fig. 4.

5 Conclusion

There are many social projects with great ideas but abandoned due to lack of investment or due to unavailability of the proper crowdfunding platforms. Blockchain-based crowdfunding not only enhances the present crowdfunding process but also

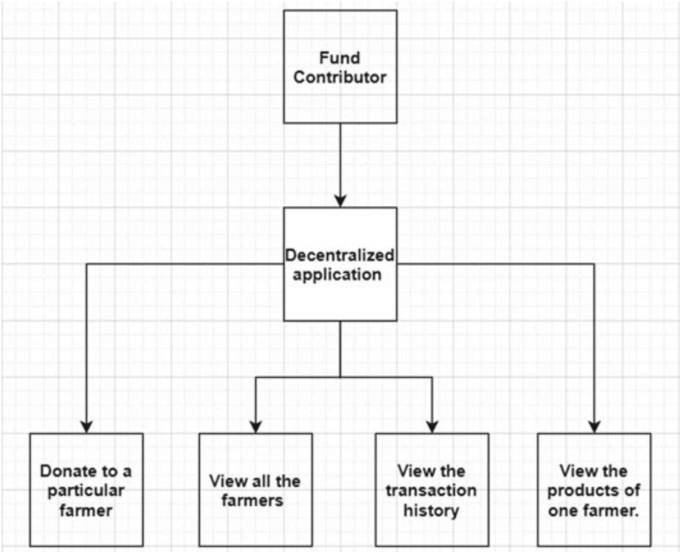


Fig. 3 Fund contributor functionality

improves the current online transactions system by reducing frauds and removing the third parties. Blockchain-based agriculture crowdfunding sees that every penny from the consumer reaches directly to the farmer and the transaction once committed, cannot be altered (immutability). Hence, there is absolutely no chance of showing fake donations and third party taxes. The further advantages to the industry using Blockchain technology include accessible equity, transparency, universal availability, peer to peer exchanges, user compliance and accountability to keep a check.

6 Future Scope

This decentralized application is completely based on crowdfunding only. In Future, the blockchain can be integrated with other emerging technologies like Machine learning, Drone technology and Information of things and complete the whole cycle of agriculture from scratch to home delivery. Thus, using the blockchain to its full potential can completely eradicate the middle-men and make a digital platform between the farmers and consumers with complete security and transparency.

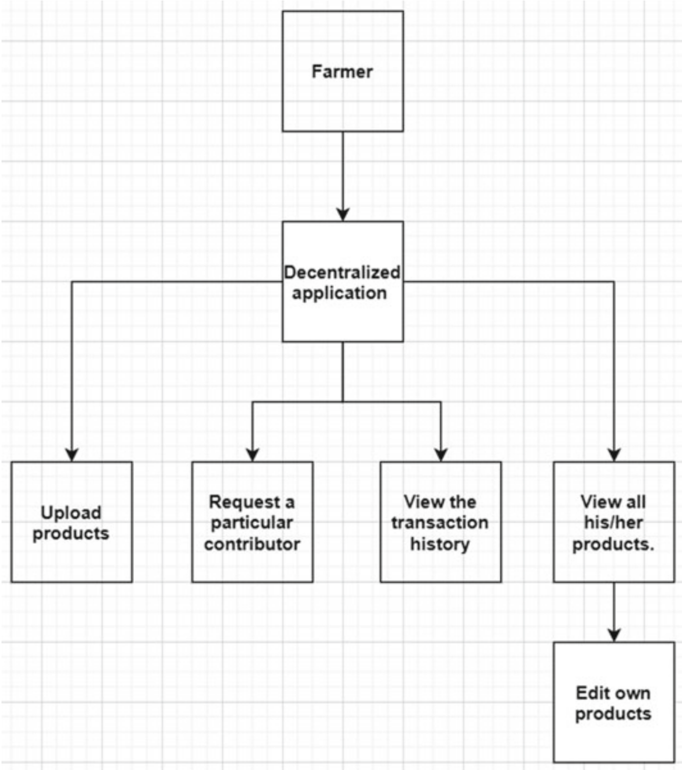


Fig. 4 Farmer functionality

Acknowledgements The project was implemented by the Naga Venkata Mohit Desabathina and the research paper was guided by Dr. Suresh Merugu (Ph.D from Indian Institute Technology Roorkee, Uttarakhand) currently working as Dean-R&D at CMR College of Engineering & Technology Hyderabad. Special Thanks to Mr. Sunil Kumar for proving the guidance regarding the corporate blockchain application.

References

1. developer.ibm.com (Internet Source)
2. Submitted to CSU, San Jose State University (Student paper)
3. www.fao.org (Internet source)
4. pdfs.semanticscholar.org (Internet Source)
5. Submitted to Indian Institute of Management Rohtak(Student Paper)
6. www.gs1nz.org (Internet Source)
7. www.ijaict.com (Internet Source)
8. Submitted to Cummins College of Engineering for Women, Pune (Student Paper)

9. Blockchain-based crowdfunding: a ‘PayitForward’ model of WHIRL. Int J Recent Technol Eng (2019)
10. creditsmate.com (Internet Source)
11. Submitted to National Institute of Technology, Raipur (Student Paper)
12. shaih.github.io (Internet Source)
13. Submitted to Harper Adams University College (Student Paper)
14. “Blockchain–ICBC 2019”, Springer Science and Business Media LLC, 2019
15. https://www.hyperledger.org/ (Internet source)

LWT-DCT Based Image Hashing for Tampering Localization via Blind Geometric Correction

Abdul Subhani Shaik, Ram Kumar Karsh, Merugu Suresh, and Vinit Kumar Gunjan

Abstract There is a tradeoff between hash length and small area tampered localization based on the block-based analysis. In this paper, we have mitigated this issue using only global features. A compact image hash is constructed using LL band of Lifting Wavelet Transform (LWT) followed by Discrete Cosine Transform (DCT). An image map is reconstructed from the hash. Based on the difference of image map obtained from the received hash and received image, tampering regions have been localized. Experimented the blind geometric correction removes the influences of geometric distortion while tampering with localization. While decreasing the hash length, a very small area of tampered regions has been localized using the proposed approach. Due to the blind geometric correction, the proposed method can localize tampering, even if tampering and composite RST geometric transformations occur simultaneously. The experimental results reveal that the proposed technique can upgrade the tampering localization compared to both the block-based and existing global features based state-of-the-art methods.

Keywords Image hashing · LWT · DCT · Geometric transformation · Tamper detection

1 Introduction

THE growing of advanced image editing tools forced to think about sophisticated authentication mechanism [1]. Distinguishing original images from fake ones and localizing the tampered area is a challenging issue for industries and academia. Recently, an image hashing based approach has been widely used for tampered

A. S. Shaik (✉) · R. K. Karsh
Department of ECE, National Institute of Technology Silchar, Assam 788010, India

A. S. Shaik · M. Suresh
Department of ECE, CMR College of Engineering & Technology, Hyderabad, Telangana 501401, India

V. K. Gunjan
CMR Institute of Technology, Kandlakoya, Hyderabad 501401, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_156

area localization. In this technique, an image is represented by an image hash or a digital signature. For content-preserving practices, an image hash should be stable and open to changes in content manipulation [2]. The technique of image hashing was first developed by Venkatesan et al. [3]. In this method, wavelet coefficients are extracted to form an image signature, which is robust in compression, geometric distortions, but sensitive to some CPOs. To localize the area of tampering, Roy and Sun [4] A block-based image hashing technique was first adopted. This technique is stable for compression and rotation, although certain CPOs have not been investigated. In [5] authors consider the LL band of DWT from the 16 × 16 sizes of image blocks to construct hash. This method can detect tampering and robust to filtering operation and compression, but the hash length is very high. To achieve better tampering localization, Yan et al. [6] introduced a method to construct a hash with combined global and local features. This approach is versatile for compression, filtering, and rotation up to 5 degrees, but constrained in the case of tampering in small areas. There are some image alignment techniques [7–10] to eliminate the impact of geometric transformation, which has been employed to authenticate image and to localize tampering.

These image alignment methods require too long hash to reconstruct the distorted image. Yan et al. [11] represented an image hashing based on multi-scale, in which the image is divided into multiple rings. This method is robust to most CPOs, but sensitive to composite RST attack. In another study, Yan et al. [12] Introduced quaternion-Fourier Mellin geometric correction moments and quaternion Fourier tamper detection coefficients. The performance of geometric distortion correction has been degraded in the case of tampering. Recently, Yan et al. [13] Tamper detection based on multi-scale deviation map fusion provided. This method is robust against some CPOs but fails to locate tampering happens simultaneously.

Based on the constraints of modern techniques, the main contribution of the proposed work can be encapsulated as First, a combination of LWT and DCT is used to construct a hash that is robust to most common signal processing operations (CSPOs). An image map is constructed using inverse DCT of hash. The difference of image map from the received hash and received image yields accurate tamper localization. Second, a blind geometric distortion correction approach based on geometric characteristics is analyzed to obtain the tamper location, in the case where tampering and geometric transformation occur simultaneously.

The arrangement of the paper is as follows. Section II represents a detailed description of the proposed methodology. Section III demonstrates the proposed geometric distortion correction approach, image authentication, and proposed the localization process of tampering. The experimental results and analysis have been discussed in section IV. Finally, the future scope and conclusions are mentioned in section V.

2 Proposed Method for Image Hashing

In the image hashing method, a digital signature or image hash is the visual representation of an image, that has been used in different applications like image authentication [14, 15, 28–31], image retrieval [16, 17, 32], and tampering detection [6, 7, 18–21].

The proposed method for image hashing is shown in Fig. 1. An arbitrary size of source image, I(x, y), is re-sized into P × P (i.e., I´(x, y)) using bilinear interpolation. It is necessary to keep the same length of hash, which also maintain robust to scaling operation. The pre-processed image (i.e., I´(x, y)) is factorized into subsequent bands (i.e., LL, HL, LH, and HH) using LWT up to four level decomposition [23] as shown in Fig. 2. Only LL band (i.e., I″ (x, y), 8 × 8) has been used for further processing. Furthermore, DCT (Discrete Cosine Transform) discussed below is applied on I″(x, y) to get compressed hash (i.e., length of n digits).

DCT (Discrete Cosine Transform)
The DCT is the transform of sinusoidal base functions aligned with the DFT. For an N × N image, the DCT is given by

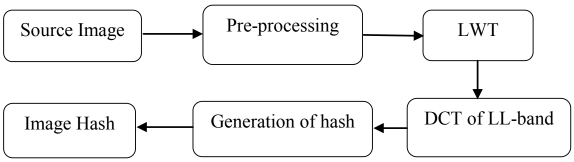


Fig. 1 The proposed method for image hashing

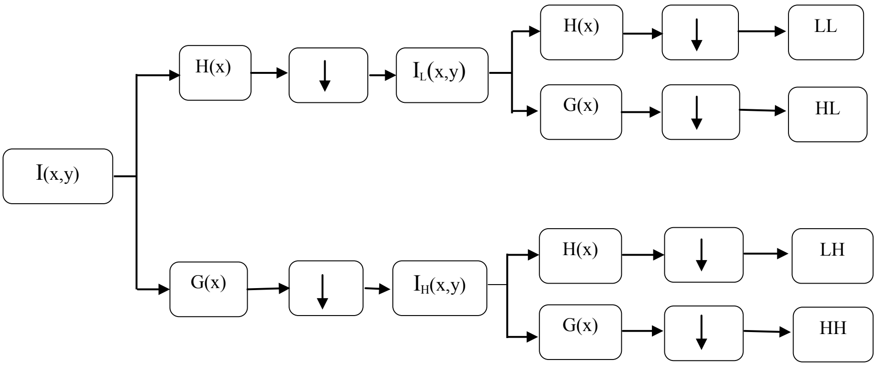


Fig. 2 Lifting wavelet transform

$$F(p,q) = \frac{1}{4}\psi(p)\xi(q)\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f(m,n)\cos\left(\frac{(2m+1)p\pi}{2N}\right)\cos\left(\frac{(2n+1)q\pi}{2N}\right) \tag{1}$$

where f(m,n) is the original image with indices m and n in the spatial domain, the transformed image with indices p and q is F(p, q), and the transformation coefficient multipliers are given by y and x, which are defined as

$$\psi(p) = \begin{cases} 1/\sqrt{2} & \text{for } p=0 \\ 1 & otherwise \end{cases} \tag{2}$$

$$\xi(q) = \begin{cases} 1/\sqrt{2} & \text{for } q=0 \\ 1 & otherwise \end{cases} \tag{3}$$

The fundamental favorable circumstances of the DCT are that it yields a genuine esteemed image and it is a fast transform. After applying DCT, it is conceivable to discard the coefficients that encode those frequency components that the human eye is not extremely sensitive to. In this manner, the information can be decreased, without truly influencing how an image looks to the human eye. The inverse transform for Eq. 1 is given by

$$f(m,n) = \frac{1}{4}\sum_{p=0}^{N-1}\sum_{q=0}^{N-1} \psi(p)\xi(q)F(p,q)\cos\left(\frac{(2m+1)p\pi}{2N}\right)\cos\left(\frac{(2n+1)q\pi}{2N}\right) \tag{4}$$

The discrete cosine transform (DCT) is an efficient way to calculate a hash of frequency data and the calculation of distances is relatively simple. Although image similarity cannot be taken into consideration semantically, it provides a hash as an image ID and is robust against minor distortions like small rotations, rinsing, and compression. The process of application of DCT on an image is shown in Fig. 3.

2.1 Metric for Performance

The metric of performance comparison is the L2 norm. Let, q and q´ are the hash for transmitted and received images, respectively. The L2 norm (or Hash Distance) is given by:

$$HashDistance\ (D) = \sqrt{\sum_{r=1}^{n} |q(r) - q'(r)|^2} \tag{5}$$

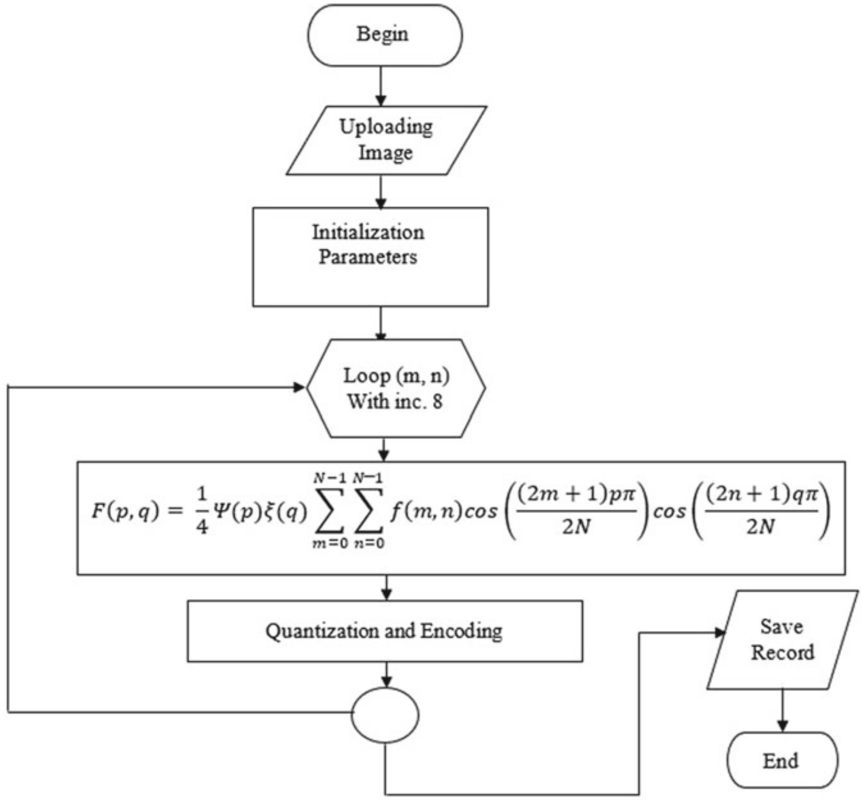


Fig. 3 Flowchart for the application of DCT on an image

where, q(r) and q´ (r) shows that the rth elements of q and q´, respectively. When threshold T1 < D, the comparing image pair is considered to be a different form. Otherwise, nearly similar or images with small tampered area pairs. FPR (False Positive Rate) and TPR (True Positive Rate) are two other performance comparison metrics, discussed as follows.

$$FPR = \left|\frac{\zeta_1}{\eta_1}\right. \tag{6}$$

$$TPR = \left|\frac{\zeta_2}{\eta_2}\right. \tag{7}$$

where ζ_1 reveals the total number of the different content image pairs considered as similar ones, and ζ_2 reveals the total number of similar content image pairs considered as similar ones. η_1 and η_2 represent entire visually different and similar contents image pairs, respectively.

3 Proposed Tampering Localization Method

The proposed tampering localization method is discussed in this section. Sects. 3.1 and 3.2 respectively discussed the blind geometric correction and tampering localization in detail.

3.1 Blind Geometric Correction Approach

The received image may tamper as well as geometrically transformed. For accurate tampering detection, it is required to eliminate geometric distortions in the received image. Hence, the proposed method uses a blind geometric correction, which has been discussed in Algorithm 1. The process of geometric correction on a geometric transformed image is demonstrated in Fig. 4.

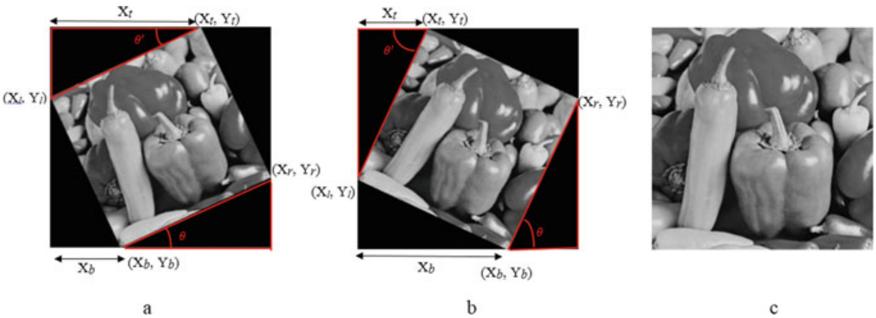


Fig. 4 Correction in geometric transformation **a** rotated in anticlockwise direction, 25° **b** rotated in clockwise direction, − 25° **c** reconstructed image

Algorithm 1: Pseudocode for the geometric transformation correction algorithm by exploiting the inherent characteristics of the geometric transform

Input: F (a query image gone through the geometric transformation, size $M' \times N'$)
Output: Restored image (geometric transformation corrected image)
Compute the indexes of the rightmost, leftmost, and the top and the bottom points ie (X_l, Y_l), (X_r, Y_r), (X_t, Y_t), and (X_b, Y_b) respectively.
Compute $\theta = \arctan(\Delta Y/\Delta X)$, $\theta' = \arctan(\Delta Y'/\Delta X')$ where $\Delta Y = Y_b - Y_r$, $\Delta X = X_r - X_b$, $\Delta Y' = Y_l - Y_t$, and $\Delta X' = X_t - X_l$.
if If $\theta \cong \theta'$ **do**
    if $X_t > X_b$ then
        $\theta_r = \theta$ // image is rotated in anticlockwise direction
    else
        $\theta_r = (90 - \theta)$ // image is rotated clockwise direction
    end if
Rotate the image by $-\theta_r$ to get the rotation corrected image F2'
else
There is no rotation
end if
Compute bounding box over F2' as $(Xmin, Ymin)$ and (width, height)
Crop F2' where rows go from Xmin to Xmin + width and columns go from Ymin to Ymin+ height
Resize F2' into the original dimensions of the query image
Return F2'

3.2 Tampering Localization

Firstly, the received image has been passed through blind geometric correction. Next, generate a hash ('h') of the received image using the proposed image hashing as discussed in section II. An image map is generated for the received hash and the generated hash of the received image. The image map is created using IDCT and ILWT. The two image map is subtracted and normalized. Thereafter, it has been converted into a binary image, which is multiplied with the reconstructed image yields the tampered regions. The flow chart for the proposed tampering localization method is presented in Fig. 5.

4 Analysis of Experimental Results

In this section, a large number of image pairs have been used to test the proposed system and the optimal value of the proposed model parameters is as follows: re-sized image size $P \times P = 128 \times 128$, image authentication threshold and tamper localization, $T1 = 10$, and $T2 = 0.6$ respectively, and hash length $n = 64$ digits. The

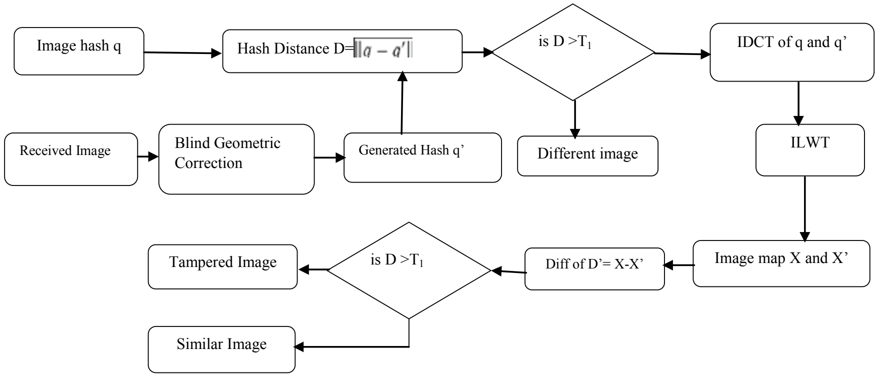


Fig. 5 Flow chart for the proposed tampering localization method

proposed model has been analyzed in two categories: performance of hashing technique for robustness and discriminative capability, and tamper localization, shown in Sects. 4.1, and 4.2.

4.1 *Performance of Hashing Technique for Robustness and Discriminative Capability*

The suggested model is used in this section to segregate the image obtained from "visually identical images" and multiple images. The experiment was conducted on 4,032 visually similar pairs, created as shown in Table 1. 42 source images from the USC-SIPI database[24] are chosen here, such as 37 and 5 from the categories "Aerial" and "Miscellaneous" respectively, with sizes ranging from 512 × 512 to 2250 × 2250. From various combinations of 200 source images, 19,900 different image pairs are also produced. The source images are as follows: 75 from the Ground Truth Database [25], size 756 × 504, 75 from the Nikon D3200, size 3008 × 2000 to 4512 × 3000, and 50 from the Internet, size 256 × 256 to 1024 × 7688. Just in Fig. 6. Redline displays the distribution of hash distances between identical image pairs (exponential curve on the left side), while the blue line shows different image pairs (approximately Gaussian on the right side).

It can be observed from Fig. 6. that red lines are near to zero and the blue line centered around 35, which show the efficacy of the proposed method in case of

Table 1 Optimal threshold T_1 of the proposed technique

Threshold T1	TPR	FPR
9	0.9834	5.0251×10^{-5}
10	0.9901	2.512×10^{-4}
11	0.9920	4.522×10^{-4}

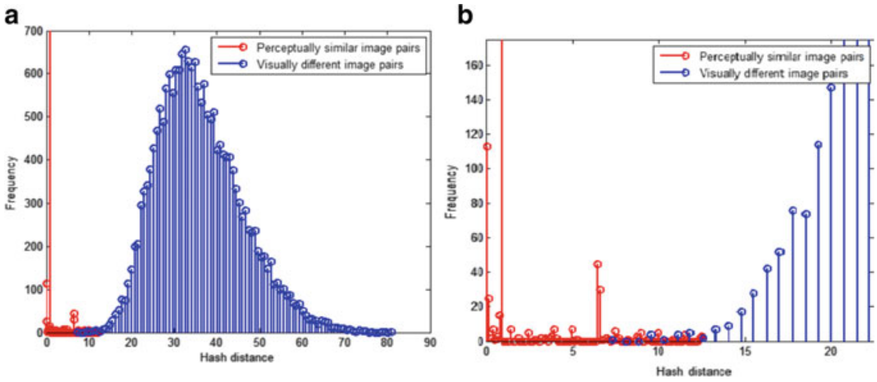


Fig. 6 **a** Distribution of hash distance for perceptually identical and distinct pairs of images. **b** Expanded version

robustness against CPOs and discrimination. The range of threshold (T1) may be from 7.5 to 12.5. Here, based on the trade-off between robustness and discrimination, the optimal threshold is 10, as shown in Table 1. If we vary the threshold (T1) and set it to a suitable value then there may be zero percent false acceptance.

4.2 Tamper Localization

The experiment on 800 tampered image pairs was conducted in this section. In which, from the CASIA V2.0 archive, 400 gross impacted images were taken [22] and 400 minimum influenced images from the NITS database [23] sizes vary from 240 $\times$ 160 to 900 $\times$ 600. It has been observed that 33.17% of affected images are falsely accepted as original ones using only global features (i.e., without image map). The tamper localization is shown in Fig. 7. for some sample images using the proposed method. This method has been recognized in all tampered image pairs, but in some pairs, the small size of not tampered regions are detected as tampered ones. However, the proposed method can locate the arbitrary size of tampering.

5 Performance Comparison with State-Of-The-Art Methods

The proposed method is compared with state-of-the-art methods such as Radon Transform and Discrete Fourier Transform (RT-DFT) based hashing [26], SIFT based hashing [27], Zernike Moments (ZM) based hashing [7], and Ring invariant vector distance. (Ring-IVD) based hashing [2]. All compared methods are evaluated with

Fig. 7 Localization of tampered regions

the same database, discussed in section IV. TPR and FPR reflect robustness and discrimination performance. A better method should have high TPR (i.e., close to one) and low FPR (i.e., close to zero). Overall performance is represented using the receiver operating characteristic (ROC) curve, as shown in Fig. 8. A ROC curve that follows towards the upper left corner is a better technique. The performance comparison is shown in Table 2 and Fig. 7. It can be observed from Table 2 and Fig. 7 that the performance of the proposed method is overall better than the compared ones. The hash length and computational cost are slightly larger than some of the methods. However, the proposed method can locate the tampering region, even if tampering and composite RST occur simultaneously, which is the main focus of the proposed method. All compared methods are implemented using a desktop computer with Intel i5 processor of 12 GB RAM having windows 10 operating system using MATLAB 2018a.

6 Conclusions

In this work, an image hashing technique based on LWT and DCT is proposed. In comparison to the blind geometric correction method, the proposed technique is applied in tampering localization. The experiment was carried out on a large database and the findings revealed that the proposed strategy against CPOs was successful. Besides, strong discriminative capacity and localization of tampered zones. The ROC curve suggests that the approach suggested is equivalent to state-of-the-art techniques. The accuracy of the tamper position can be enhanced in future work.

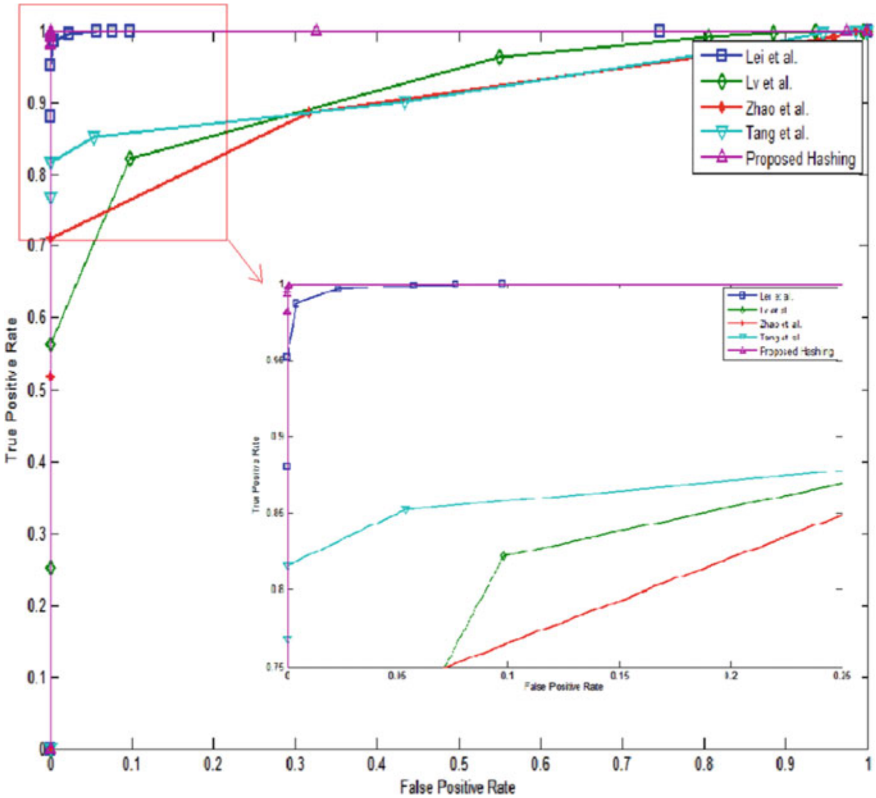


Fig. 8 Performance comparison of the proposed technique with some existing techniques

Table 2 Performance comparison for discrimination with different existing techniques

Methods/operations	Lei et al. [26]	ZV et al. [27]	Zhao et al. [7]	Tang et al. [2]	Proposed hashing
Forgery detection in the case of copy-move/splicing	Yes	Yes	Yes	No	Yes
Robust against arbitrary	No	No	No	Yes	Yes
Rotation				(with information loss)	
Robust against composite RST	No	No	No	No	Yes
Tampered area localization along with rotation	No	No	No	No	Yes
The length of a hash	15 digits	20 digits	70 digits	40 digits	64 digits
TPR at an optimal threshold	0.9871	0.8023	0.7234	0.8236	0.9901
FPR at an optimal threshold	0.0054	0.0453	0.0617	0.0015	2.512×10^{-4}
Average time(s)	0.7016	0.6651	0.5642	0.0749	1.7181

Acknowledgements The authors would like to thank all the people of Speech and Image Processing Laboratory, National Institute of Technology Silchar, India, for offering help and vital facilities for doing this work.

References

1. Mishra M, Adhikary MC (2013) Digital image tamper detection techniques: a comprehensive study. Int J Comput Sci Bus Inf 2(1):112
2. Tang Z, Zhang X, Li X, Zhang S (2016) Robust image hashing with ring partition and invariant vector distance. IEEE Trans Inf Forensics Secur 11(1):200–214
3. Venkatesan R, Koon SM, Jakubowski MH, Moulin P (2000) Robust image hashing. In: International conference on image processing, vol 3, pp 664666
4. Roy S, Sun Q (2007) Robust hash for detecting and localizing image tampering. In: International conference on image processing (ICIP), pp VI-117–VI-120
5. Ahmed F, Siyal MY, Abbas VU (2010) A secure and robust hash based scheme for image authentication. Signal Process 90(5):14561470
6. Pun CM, Yan CP, Yuan (2016) Image alignment-based multiregion matching for object-level tampering detection. IEEE Trans Inf Forensics Secur 12(2):377391
7. Zhao Y, Wang S, Zhang X, Yao H (2013) Robust hashing for image authentication using Zernike moments and local features. IEEE Trans Inf Forensics Secur Inf Forensics Secur 8(1): 5563
8. Wang X, Pang K, Zhou X, Zhou Y, Li L, Xue J (2015) A visual model-based perceptual image hash for content authentication. IEEE Trans Inf Forensics Secur 10(7):13361349

9. Lu W, Varna AL, Wu M (2010) Forensic hash for multimedia information. In: Proceedings of SPIE media forensics and security, San Jose, CA, pp 75410Y
10. Lu W, Wu M (2010) Multimedia forensic hash based on visual words. In 2010 IEEE international conference on image processing (ICIP), Hong Kong, pp 989992
11. Yan CP, Pun CM, Yuan XC (2016) Multi-scale image hashing using adaptive local feature extraction for robust tampering detection. Signal Process 121:116
12. Yan CP, Pun CM, Yuan XC (2016) Quaternion-based image hashing for adaptive tampering localization. IEEE Trans Inf Forensics Secur 11(12): 2664–2677
13. Yan CP, Pun CM (2017) Multi-scale difference map fusion for tamper localization using binary ranking hashing. IEEE Trans Inf Forensics Secur 12(9):21442158
14. Battiato S, Farinella GM, Messina E, Puglisi G (2012) Robust image alignment for tampering detection. IEEE Trans Inf Forensics Secur 7(4):11051117
15. Krawczyk H, Bellare M, Canetti R (1997) HMAC: Keyed hashing for message authentication. IETF RFC-2104
16. Gong Y, Lazebnik S, Gordo A, Perronnin F (2013) Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval. IEEE Trans Pattern Anal Mach Intell 35(12):2916–2929
17. Kulis B, Grauman K (2009) Kernelized locality-sensitive hashing for scalable image search. In: IEEE 12th international conference on computer vision, pp 2130–2137. IEEE
18. Karsh RK, Saikia A, Laskar RH (2018) Image authentication based on robust image hashing with geometric correction. Multimedia Tools Appl 1–21
19. Karsh RK, Laskar RH, Richhariya BB (2016) Robust image hashing using ring partition-PGNMF and local features. Springerplus 5:1995
20. Karsh RK, Laskar RH (2017) Robust image hashing through DWTSVD and spectral residual method. EURASIP J Image Video Process 31:1–17
21. Daubechies I, Sweldens W (1998) Factoring wavelet transforms into lifting steps. J Fourier Anal Appl 4(3):247–269
22. CASIA Tampered image detection evaluation database [Online]. Available: http://forensics.idealtest.org/
23. NITS Image hashing database [Online]. Available https://rishabhphukan.wixsite.com/rkkarsh
24. USC-SIPI Image database. http://sipi.usc.edu/database/. Accessed February 2007.
25. Ground Truth Database. http://www.cs.washington.edu/research/imagedatabase/groundtruth/. Accessed 8 May 2008
26. Lei Y, Wang Y, Huang J (2011) Robust image hash in radon transform domain for authentication, signal processing: Image Commun 26:280–288
27. Lv X, Wang ZV (2012) Perceptual image hashing based on shape contexts and local feature points. IEEE Trans Inf Forensics Secur 7(3):10811093
28. Syed AT, Merugu S, Kumar V (2020) Augmented reality on sudoku puzzle using computer vision and deep learning. In: Advances in cybernetics, cognition, and machine. Lecture notes in electrical engineering, pp 567–578
29. Koppula VK, Soumya DS, Merugu S (2020) Nurse alarming device for bedridden patients using hand gesture recognition system. In: Advances in cybernetics, cognition, and machine, 2020. Lecture notes in electrical engineering, pp 377–385
30. Merugu S, Jain K, Mittal A, Raman B (2019) Sub-scene target detection and recognition using deep learning convolution neural networks. In: ICDSMLA 2019. Lecture notes in electrical engineering, pp 1082–1101
31. Merugu S, Reddy MCS, Goyal E, Piplani L (2019) Text message classification using supervised machine learning algorithms. In: ICCCE 2018. Lecture notes in electrical engineering, vol 500. ISSN 1876-1100
32. Yadav BC, Merugu S, Jain K (2019) Error assessment of fundamental matrix parameters. In: ICCCE 2018. Lecture notes in electrical engineering, vol 500. ISSN 1876-1100

Evaluation of Dyke Rocks as Building Material, Accessing the Properties Using Mat Lab for Quality

Musini Venkateshwarlu, Suresh Merugu, Vinit Kumar Gunjan, K. Suresh, and A. P. Ravichandra

Abstract Environmental change on fauna and flora, affects water quality and quantity due to indiscriminate mining for societal usage. Illegal sand mining is one of the activities posing serious threat to river systems. With the increase of urban population and non availability of sand for construct on of residential complexes leads too many fold increase in illegal mining. Statistics show that requirement of housing increase by one hundred and thirty three present in the next decade. Construction companies, one of the largest users of river sand, need sand with specific parameters. Some of parameters are high compressive strength, durability, economically viable, and availability of large quantity of sand. The dyke rocks chosen for this study lies in the outskirts of Hyderabad making it available. Detailed geological mapping was carried out and field samples and core samples were collected. Laboratory studies to assess the compressive strength were done. The study proved the dyke rock to be suitable to generate Robo sand and create an eco—friendly environment for a better future.

Keywords Dyke · Rob sand and river sand

1 Introduction

Many metropolitan cities have witnessed tremendous growth in construction activity and infrastructural request in India. There are four large Metropolitan cities which are consuming maximum building material over few decades, Urbanization and industrialization rapid growth of Hyderabad and surrounding area led to air pollution also

M. Venkateshwarlu (✉)
Department of Civil Engineering, CMR College of Engineering & Technology, Kandlakoya (V), Medchal Road, Medchal, Hyderabad, Telangana 501401, India

S. Merugu
Research and Development (R&D) Centre, CMR College of Engineering & Technology, Kandlakoya, Medchal Road, Hyderabad, Telangana 501401, India

V. K. Gunjan · K. Suresh · A. P. Ravichandra
CMR Institute of Technology, Kandlakoya, Medchal Road, Hyderabad, Telangana 501401, India

A. Kumar et al. (eds.), *ICDSMLA 2020*, Lecture Notes in Electrical Engineering 783,
https://doi.org/10.1007/978-981-16-3690-5_157

know present aspect [1]. In recent times Hyderabad within twin cities has occupied the first position in consumption of cement and other building materials for construction activities. Construction activity consume large amount of sand either for simple brick construction or for concrete mix. The per capita consumption of cement in Hyderabad is almost ranging between 100 and 150 kg indicating indirect consumption of sand for this purpose. The "WALTA" act 2002 promulgates a ban on the exploitation of river bed sand. It has been reported that the consumption of River sand creates environmental imbalance and the Govt. of India is in process of imposing ban, regulation on norms on digging river beds and also highlighted the environmental impact and greater awareness and protection of nature [2]. In view of this the river sand will became a scare commodity in the construction material. In the back drop of this there is in urgent need for such of those materials which can find place as an alternative to river sand. Number of construction companies has made exercises to identify the material alternatively generated. However, the evidence supporting this assertion is invalid in the studied area as no clastic dykes have been seen cutting the Precambrian basement formation and the funnel structures are absent [3]. The origin of granite is contentious and has led to varied schemes of classification. A classification based on origin of the "parental" magma from which the granite was formed is the most accepted one. The study area, Narsapur, Narsapur mandal in Medak district, Telangana State, is at a distance of 60 km NW of Hyderabad [4]. There is an alteration of mineral constituents of the rock such as apatite, fluorite, biotite, and muscovite and from use of fertilizers in agriculture. Portable drinking water is not available in the region there by using only groundwater. Information System Used to Study on Effect of Groundwater Quality based on influence of fertilizer is agriculture add further contamination of ground water made unsafe to drink groundwater using organic forming reduce the contamination levels [5].

Topography

The area is slightly elevated cut across by streams and observed as a linear elongated mound. The ground level slowly rises up to 20 meters and ground level is at 500R.L. There are open lands and cultivated lands in close proximity to this area. A Granite hill under excavation is about 1k.m. west of this area [6].

Vegetation

The area exhibit spares vegetation. Only few shrubs and herbs are observed along the slops of Paddy cultivation is noticed.

Drainage

The area adjoining locations exhibit dendritic drainage however this linear intrusive rock form water divide in this area [7]. In close proximity to this area there is pond Rayaraopet pond which locally provides the needs of cultivation. It provides material to meet the demand of construction material particularly the river sand in Hyderabad.

Structure

The area exhibits prominent liniment along North South Direction evidently indicated by the emplacement of Dolerite Dyke [8, 9]. The major liniment is along North South along extending more than 40 km has is seen in 56k/11 There are East West Joints and North Northeast, South Southwest Joints presents in this area the intrusive Dyke rock disturbed and shear along East West Joints North Northeast South Southwest Joints and partly by absorbed along the structural trends

General lithological characters:

Granites: Within the vicinity of this intrusive rock peninsular Granite is well exposed. Quartz and feldspar are essential minerals. Quartz has been calculated 22% from this area [10].

Pink Granites: The feldspars represented by both Albite and oligoclase [11] and Horn blende producing Grey colorization to rock. It essentially contain Quartz Feldspars where Feldspars is Albite with >20% of Silica in the form of Quartz and the accessory Hornblende Biotite.

Dolerite: Two types of Dolerites observed as per structure the contain medium to coarse grain mostly Labradorite feldspar with dark Ferro Magnesium minerals coarse medium grained Dolerite is seen as contact of Granite they are intruded by vein of Quartz–Feldspars [12, 13]. Field Investigation has been carried out on the basic intrusive rock trending N-S direction near the Jamilapet. The area has been mapped on 1:330 scale bringing out contacts between the enclosing granite and the intrusive basic rock. A base line along the strike of intrusive body has been laid over a distance of 600 m (0.6 km). The base line has been deviated at North that is at Station N6 and similarly N6 parallel towards East for 15 m and is again extend towards North. For another 60 m and it was close at N8 station. Towards South the base line deviated at S6 station towards West for 40 m and again extends towards south and closed at S12. N-O station has been taken as a reference point which is almost at the centre of the intrusive body. There are 8 Cross lines laid towards North and seven along South. The table depicting him station points etc. are given below.

2 Laboratory Investigations

Representative samples have been further scrutinized and 6 samples select for preparation of Thin Sections [14, 15]. The Same samples have been Cored have been selected with a diameter and length ratio is 1:2–2.5 and the Core sample that can to Compressive strength (Table 1).

3 Petrographic Nature of Rock

Thin Sections were prepared from representative samples. Petrography was interpreted. In general the rock exhibit medium to fine grained texture with domination of plagioclase (basic) feldspar and accessory Quartz can be identified [16]. Ophitic to Subophitic texture is common. Augite is enclosed by Plagioclase. Majority of the

Table 1 Representative samples for investigations

S. No.	Sample No.	Nature of samples
1	2	1. It is melanocratic rock feldspar survivable available but difficult to recognize 2. Semi coarse grain (or) Medium grain
2	7	1. It is melanocratic rock feldspar survivable available but difficult to recognize 2. Fine grained homogenous with less disturbed material
3	10	1. It is melanocratic rock feldspar survivable available but difficult to recognize 2. Semi coarse grain (or) Medium grain
4	15	1. It is melanocratic rock feldspar survivable available but difficult to recognize 2. Intruded by Quartz
5	16	1. It is melanocratic rock feldspar survivable available but difficult to recognize 2. Semi Coarse grain (or) medium grain
6	17	1. It is melanocratic rock feldspar survivable available but difficult to recognize 2. Semi coarse grain (or) medium grain

sections exhibit altered zones as the samples are picked up from the top of intrusive body [17, 18]. The Felspar is sericitised occasionally. The Augite is abundant and.generally represents ferro magnesium minerals [19, 20]. Released Iron oxides are common the figure of rock representative of micro photo section has been given in Fig. 1. To summarize petrography of intrusive body it exhibits true mineralogy of medium grained Dolerite rock with occasional emplacement of epidote and quartz [21].

(a) (b)

Fig. 1 Photograph of a dykes at Jamilpet. **a** Series of fractures of milli metric sizes seen at both sides of the dykes. **b** The rocks at the contact with the massive dyke is exposed

Table 2 Samples selected for compressive strengths

S. No	Sample No.	Diameter (mm)	Length	Remarks
1	2	29.3	62	_
2	7	29.3	61.5	_
3	10	29.3	60	_
4	15	29.3	61.4	Cut by Vein
5	16	29.3	62	_
6	17	29.3	60.4	_

Table 3 Result of compressive strength and conversion

Sample No.	Diameter (d) [mm]	Load (p) [kN]	Area (A) [mm^2]	Compressive strength (ρ) N/mm^2	Conversion (ρ) = p/A kg/cm^2
2	29.3	240	674.26	356	3629.79
7	29.3	265	674.26	393.03	4007.90
10	29.3	220	674.26	326.3	3327.31
15	29.3	80	674.26	118.7	1209.93
16	29.3	180	674.26	267	2722.34
17	29.3	160	674.26	237.3	2419.86

4 Characters Physico Mechanical and Compressive Strength

Six representative samples have been drilled with 29.3 mm core bit (ex: the samples were prepared with core length of 6 cm for testing Compressive Strength. The test was conducted in the Compressive strength. The results have been given in Table 2. Distinct Laboratory photographs indicating loading of the sample, fracturing upon compressive strength and the total failure of the samples calculating in both N/mm^2 and kg/cm^2. The reserves obtained or given table of compressive strength Table 3.

5 Compressive Strength of Meta Dolerite

Compressive strength (***ρ***)

$$\rho = p/A \quad \begin{cases} \text{p - load tending break the sample} \\ \text{A - cross section of the sample} \\ A = \Pi r^2 \text{ or } \Pi d^2/4,\ \ d = \text{diameter} = 29.3\ \text{mm} \end{cases}$$

6 Result and Discussion

The area situated North East of Ghatkesar Mandal near jamilapet has been evaluated both with field investigations and laboratory investigations [13]. A portion of intrusive body is chosen for detailed mapping. The weathering condition of deposit reveals the existence of small boundary out crops and below which the massive dyke is exposed. The deposit is evaluated for its suitability as manufactured sand for construction material and also suitable for decorative tiles. In this process, the evaluation physico mechanical properties and microscopic characters of representative samples for petrographic properties are studied.

The intrusive Dolerite dyke is extending over a strike length of more than 40 km. Along N-S direction and appears to be a major lineament in this area. As the dyke is weathered into small boulders its utility for manufactured artificial sand, called Robo sand, has been studied. The rock is medium to coarse grained (Semi gabbroic) and offer oolitic sub-oolitic textures as the size of boulders of for suitable crusher. The rock can be crushed to 1 mm fraction and slightly more than that however, chocked crushing can lead to form 1 mm fines. The petrographic nature also suggest the presence of hard minerals viz, Feldspar, Augite with very less amount of Quartz. This appeared to be a favorable mineralogy for use. This Dyke rock in the form of boulders for manufacture of artificial sand as the rock exhibit typical physico mechanical characters particularly compressive strength which is quite favorable for the manufacture of rob sand.

7 Conclusion

The location of the intrusive body, the Dolerite Dyke, in close proximity to the city limits enhances its potentiality to be utilize as a construction material Petrografic nature of the rock indicates semi coarse Dolerite which will favorably account for the manufacture of artificial sand. The physico-mechanical properties such as compressive strength indicates the hard nature of the rock type which make produce prohibitive cost on crushers however they are still economical. The top boulders can be used for manufacture of Robo sand and the exposed rock hills blocks ranging from 0.4 to 0.6 cubic meters can be used in the manufacture of Granite blocks.

Acknowledgements We express our thanks to Ch. Gopal Reddy, Secretary, CMRGI, Principal and HOD, CMR College of Engineering & Technology (A), Hyderabad, for help and encouragement to publish this paper.

References

1. Venkateshwarlu M, Merugu S (2020) Fuzzy inference system based assessment of pollution aspects in Hyderabad. In: Advances in cybernetics, cognition, and machine learning for communication technologies, pp 579–587. https://doi.org/10.1007/978-981-15-3125-5_56. Online ISBN 978-981-15-3125-5
2. Barker F (1979) Trondhjemite: definition, environment and hypothesis of origin. In: Barker F (ed) Trondhjemites, dacites and related rocks. Elsevier, Amsterdam, pp 1–12
3. Scholz H, Frieling D, Obst K (2009) Funnel structures and clastic dykes in Cambrian sandstones of southern Sweden—indications for tensional tectonics and seismic events in a shallow marine environment. Neues Jahrbuch fur Geologie und Palaontologie Abh 251(3):355–380
4. Venkateshwarlu M, Prasad B, Ravichandra AP, Prashanthi M (2020) Petrography of granitic rocks of, Medak district, Telangana state. Int J Anal Exp Modal Anal XII(I):570–573. ISSN NO: 0886-9367
5. Musini V, Merugu S, Reddy TR (2020) Information system used to study on effect of groundwater quality. In: Advances in cybernetics, cognition, and machine learning for communication technologies, pp 531–542. http://webvpn.fjmu.edu.cn/https://doi.org/10.1007/978-981-15-3125-5_53. Online ISBN 978-981-15-3125-5
6. Batchelor RA, Bowden P (1985) Petrogenetic interpretation of granitoid rocks series using multicationic parameters. Chem Geol 48:43–55
7. Cox KG, Bell JD, Pankhurst RJ (1979) The interpretation of igneous rocks. Allen and Unwin, London, pp 450
8. Eskola P (1952) On the granulites of Lapland. Am J Sci (Bowen V) 1:133–171
9. Chappell BW, White AJR (1974) Two contrasting granite types. Pac Geol 8:173–174
10. Sitaramayya S, Gnaneshwar P (1981) Geochemical prospecting for groundwater in the granitic terrain around Ghatkesar, A.P. (India). J Geochem Explor 16(1):13–20
11. De La Roche H, Leterrier J, Grande CP, Marcha LM (1980) A classification of volcanic and plutonic rocks using R_1–R_2 diagrams and major element analyses-its relationships and current nomenclature. Chem Geol 29:183-210
12. Frost CD, Barnes CG, Collins WJ, Arculus RJ, Ellis DJ, Frost BR (2001) A geochemical classification for granitic rocks. J Petrol 12:2033–2048
13. Angela LC, in Geological Field Techniques, pp. 11–16, Blackwell Science Publishing, Oxford, 1st edition, 2010.
14. Green J, Poldervaart A (1958) Petrochemical fields and trends. Geochim Cosmochim Acta 13:87–122
15. Divakara Rao V, Rama Rao P, Subba Rao MV (1999) The Ghingee grannite, Tamil Nadu, South India: geochemistry and petrogenesis. Gondwana Res 2(1):117–126
16. Janardan Rao, Y. 1961. The sequence Biotite-Plagioclase-K-feldspars in metasomatism. Mahadevan Volume, Osmania University Press, Hyderabad. pp. 217-225
17. Jayananda M, Martin H, Mahabaleshwar B (1992) The mechanisms of recycling of Archean continental crust: example of the Closepet granite, southern India. In: Proceedings of 3rd international Archean Symposium, Perth. vol 2, pp 13–222. Irvine TN, Baragar WRA (1971) A guide to the Chemical classification of the common volcanic rocks. Can J Earth Sci 8:523–548
18. Santhosh K, Rino V (2006) Mineralogy and geochemistry of micro granular enclaves in Palaeoproterozoic Malanjkhand granitoids, Central India: evidences of magma mixing, mingling and chemical equilibration. Contrib Mineral Petrol 152:591–609
19. Mackenzie WS, Smith JV (1959) Change balance and the stability of alkali feldspars. Acta Cryst 12:73–74
20. Maniar PD, Piccoli PM (1989) Tectonic discrimination of granitoids. Geol Soc Am Bull 101:635–643
21. Nockolds SR, Allen R (1953) The geochemistry of some igneous series. Geochim Cosmochim Acta 4:105–142