MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0101

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

KDD on the CIFAR10 dataset by meta models for binary classification approaches through IP

Madhan Mohan Subramanian[a] | **Karthikeyan Elangovan**[b]

[a] Annamalai University, Chidambaram, Tamil Nadu, India, Research Scholar, Department of Computer and Information Science.
[b] Annamalai University, Annamalai Nagar, Tamil Nadu, India, Research Supervisor, Department of Computer and Information Science, Faculty of Science.

Abstract An ensemble classifiers have been successfully applied to image classification problems. This research finds the efficiency of ensemble model on tiny images for real world problems based on the following metrics like accuracy, precision, recall, receiver operating characteristic curve (ROC), precision recall curve (PRC), kappa, F-Measure, MCC, performance analysis of deviations on like Bagging, Logit Boost (LB), Iterative Classifier Optimizer (ICO), Classification Via Regression (CVR), Multi Class Classifier Updateable (MCCU), and Random Committee (RC) of selected meta models. The RC gives best outcome compare with other models such as 90% of accuracy, 90% of precision, 0.90 of recall, 0.96 of ROC, 0.94 of PRC, 0.80 of kappa, 0.82 of MCC with low deviations. The ICO gives worst outcome compare with other models such as 65% of accuracy, 0.65 of precision, 0.65 of recall, 0.74 of ROC, and 0.74 of PRC with high deviations. This work finds and evolves that the Random Committee model performs well based on comparison of several metrics with other models.

Keywords: random committee, bagging, logit boost, ICO, CVR, MCCU

1. Introduction

For image classification and other computer vision tasks, machine learning has recently surpassed human performance. However, in order for state-of-the-art deep networks to achieve at high levels of accuracy and to generalize well, training them often necessitates thousands to millions of labelled sections, which is a major obstacle to the broad approval of deep neural networks for new submissions.

Manual data labelling takes a lot of time and could not be practical in industries like medical, defense and the sciences where specialized knowledge is needed. Lacking the time or knowledge to label the enormous numbers prerequisite for training, authentication, and analysis but having access to a lot of unlabeled data is a common problematic in deep learning's real-world scenarios. One possible solution is to train a network on just a single manually labelled image per class until its performance is on par with that of a fully supervised network.

This paper presents in section 2 has related works on this work; in segment 3 has materials and methods; in segment 3 encompasses result and discussion and segment 5 presents the conclusion of this work.

2. Literature Survey

A machine learning model is applied through ensemble learning to a fresh issue. AAL produces rare jobs from unlabeled data. We randomly name a portion of the unlabeled images to build a support set. By combining the labels and data from the sustenance set's photographs to generate a target set (Antoniou and Amos et al 2019). Any meta-learning system can be trained using the few-shot challenges. Small real-labeled datasets can be used a trained model directly without any modifications. The learned models perform well in trials in the few-shot learning challenges for Omniglot and Mini-Imagenet. Modern deep learning relies heavily on supervision. Input-output pairs are used in supervised learning techniques to develop a model that associates input data with output labels. (Berthelot et al 2019; Cubuk et al 2019; Doon Raveen et al 2018; Fu et al 2022) The right output for the training set and for previously unobserved data points should be predictable by a successful supervised model. (Hu et al 2021; Manoharan et al 2021) After embedding-based models, gradient-based meta-learning models were found. Inference-time model-state updates use a support set to learn task-specific information that they can then strongly generalize to a task's target set (Kavi et al 2020; La Grassa et al 2021) Performances for CIFAR-10 and CIFAR-100 one-shot semi-supervised learning were reported by (Lucas 2021). By keeping pseudo-labels for samples whose probability exceeds a threshold, pseudo-labeling leverages the model to produce fake labels for unlabeled data. (Karuppusamy 2021; Rasim Caner Calik et al 2018). (Smith and Leslie 2022) suggested accuracy of

up to 95% utilizing CIFAR-10 images with a single labelled sample per class, despite the fact that fully supervised accuracy is only 94.5%. Additionally, SVHN images attain test accuracies of 97.8%, much like fully supervised images. Rigid empirical studies have demonstrated that labelling big datasets is not compulsory for deep neural network training.

MixMatch, ReMixMatch, and FixMatch are only a few of the semi-supervised learning studies that Google Research has published (Sohn et al 2020; Thambu Gladstan and Mohan 2014; Tripathi Milan 2021). ReMixMatch was improved by distribution alignment and augmentation anchoring. Anchors for augmentation mimic pseudo-labeling. The most recent semi-supervised learning system, FixMatch, did well. CIFAR-10 one-shot semi-supervised learning is discussed in FixMatch. Both consistency regularization and pseudo-labeling are employed by the Fix Match algorithm. Consistency regularization employs unlabeled data by making the supposition that the model will produce the same results on perturbed versions of the image as it did on the original. Instead of assuming pseudo-labels for unlabeled data, strong data augmentation should be employed for constancy regularization (Rajesh et al 2014; Thakkar et al 2018). Consistency regularization is increasingly used in unverified, self-supervised, and semi-supervised learning (Xie 2019).

Classification algorithms are classified using semi-supervised learning research. Self-training repetitions) use high-confidence pseudo-labeled data from preceding repetitions and labelled data to train a classifier. In our tests, self-training increased performance to match supervised training using the entire labelled training dataset. This study illustrates how Meta learning performed while using approximately 100 CIFAR-10 images and 10 labels.

3. Materials and Methods

This work considers randomly selected 100 images form the CIFAR-10 dataset which has 60000 color images with of 32x32 dimensions. These 100 images has categorized in 10 classes, each class has10 tiny images. The classes are categorized like truck, ship, horse, frog, dog, deer, cat, bird, auto mobile and airplane images (Table 1).

Table 1 CIFAR 10 dataset.

S. No	Name of the Images	1	2	3	4	5	6	7	8	9	10
1	truck										
2	ship										
3	horse										
4	frog										
5	dog										
6	deer										
7	cat										
8	bird										
9	automobile										
10	airplane										

To achieve the best results, the open source mining tool Weka 3.9.5 has implemented the following techniques.

- **Bagging:** Several different independent models, commonly referred to as base models, are integrated to create an efficient, optimal prediction model using ensemble learning, a popular and favored machine learning technique.
- **Iterative Classifier Optimizer (ICO):** It optimizes the amount of iterations for the specified classifier and applies cross-validation.
- **Logit Boost(LB):** It is a boosting classification algorithm.

- **Classification via Regression (CVR):** It is a classification method that can transform problems into regression functions.
- **Multi Class Classifier Updateable (MCCU):** It uses 2-class classifiers to handle multi-class datasets.
- **Random Committee (RC):** It is a group of base classifiers that can be randomly chosen.

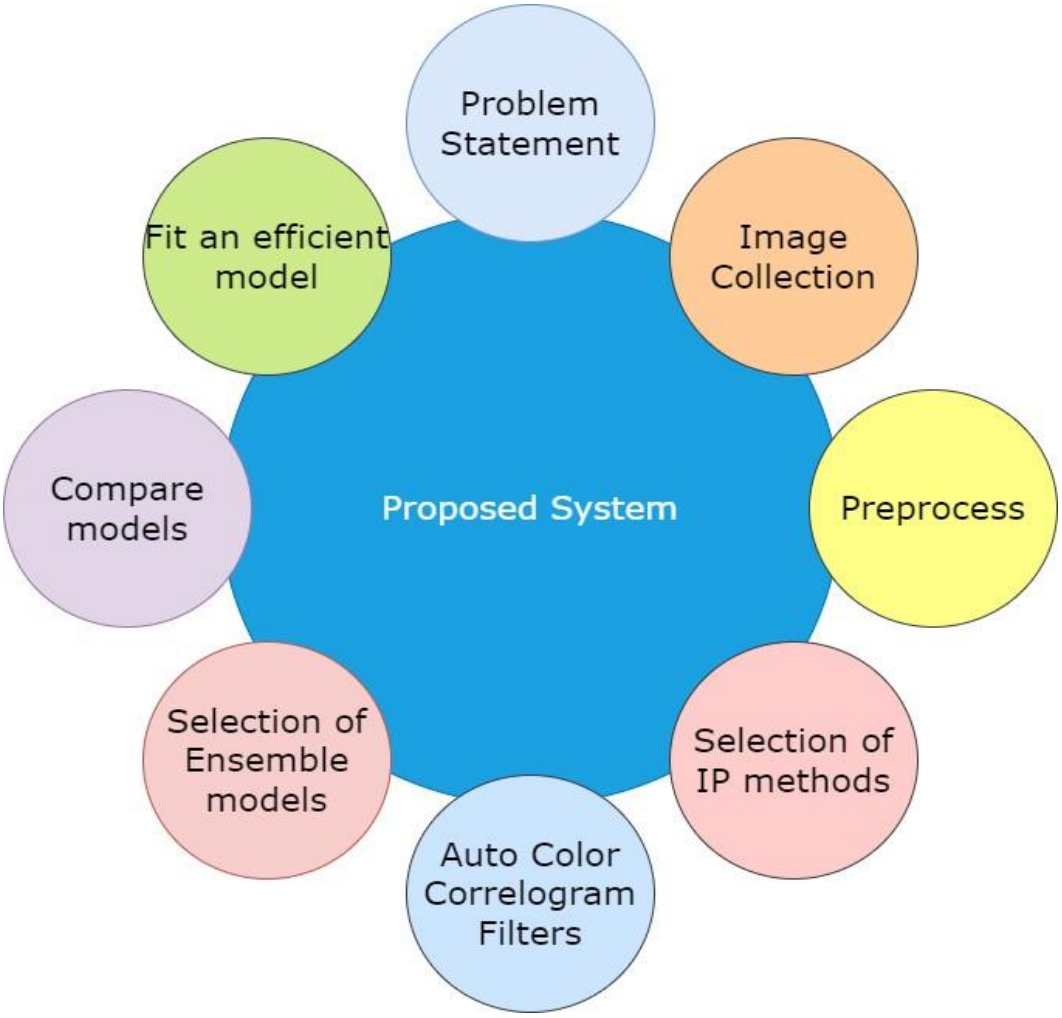


Figure 1 Architecture.

In this study, the aforementioned models were used in image processing with a 90:10 fold cross validation (training and testing ratio) utilizing an auto color correlogram filter.

4. Results and Discussions

This segment focuses the consequence of the selected classifiers like Bagging, Logit Boost (LB), Iterative Classifier Optimizer (ICO), and Classification via Regression (CVR), Multi Class Classifier Updateable (MCCU), and Random Committee (RC) on Meta models.

The Table 2 shows the accuracy, recall, precision, PRC and ROC value on like Bagging, Logit Boost (LB), Iterative Classifier Optimizer (ICO), Classification Via Regression (CVR), Multi Class Classifier Updateable (MCCU), and Random Committee (RC) on meta models.

Table 2 Classifiers vs. Outcomes.

Classifier	Accuracy	Precision	Recall	ROC	PRC
Bagging	85%	0.85	0.85	0.81	0.81
ICO	65%	0.65	0.65	0.74	0.74
LB	75%	0.78	0.75	0.86	0.86
CVR	85%	0.85	0.85	0.81	0.81
MCCU	80%	0.81	0.80	0.80	0.74
RC	90%	0.92	0.90	0.96	0.94

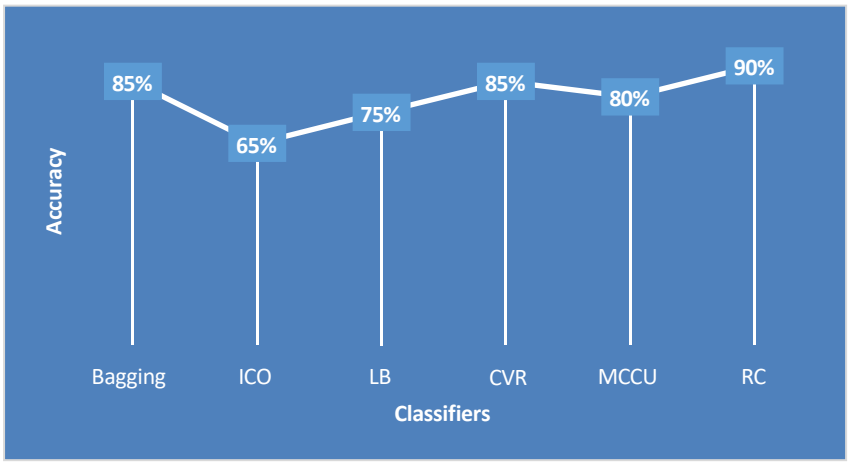


Figure 2 Classifiers vs. Accuracy.

The above figure 2 represents the accuracy level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 90% of accuracy; The ICO gives least outcome compare than other models such as 65% of accuracy; The Bagging and CVR gives same outcome such as 85% of accuracy; The MCCU gives 80% of accuracy; and the LB gives 75% of accuracy.

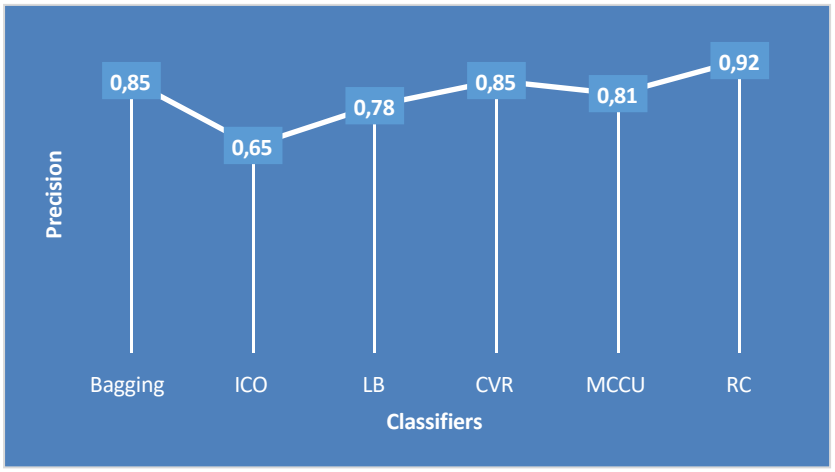


Figure 3 Classifiers vs. Precision.

The above figure 3 represents the numerous precision level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 90% of precision; The ICO gives least outcome compare than other models such as 0.65 of precision; The Bagging and CVR gives same outcome such as 0.85 of precision; The MCCU gives 0.81 of precision; and the LB gives 0.78 of precision.

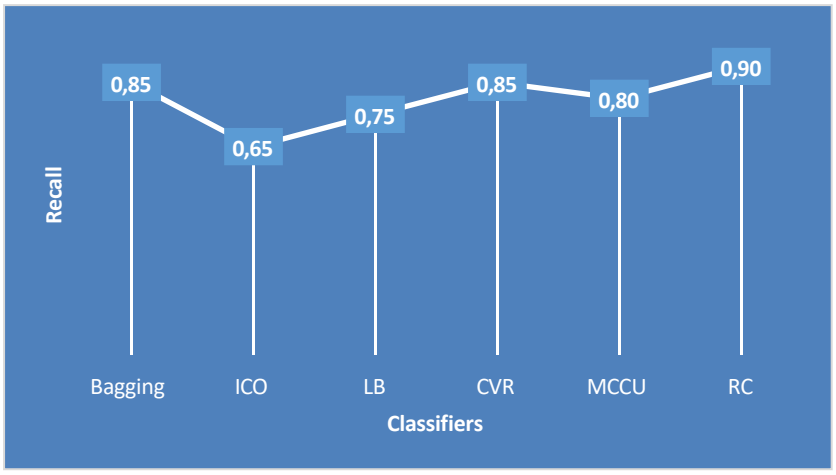


Figure 4 Classifiers vs. Recall.

The above figure 4 represents the numerous recall level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 0.90 of recall; The ICO gives least outcome compare than other models such as 0.65 of recall; The Bagging and CVR gives same outcome such as 0.85 of recall; The MCCU gives 0.80 of recall; and the LB gives 0.75 of recall.

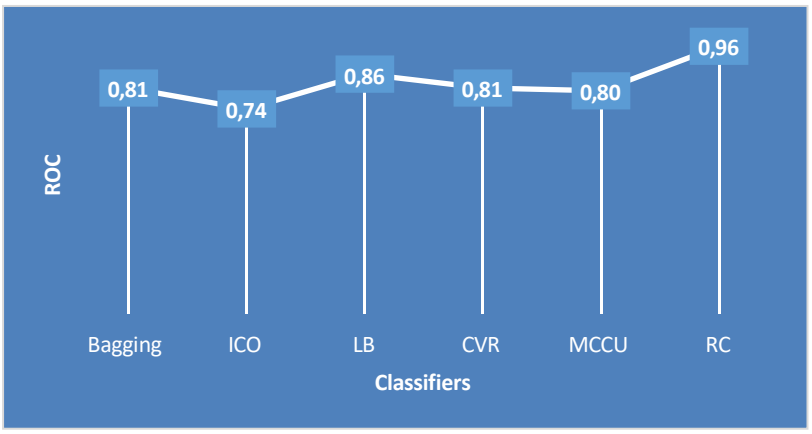


Figure 5 Classifiers vs. ROC.

The above figure 5 represents the numerous ROC level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 0.96 of ROC; the ICO gives least outcome compare than other models such as 0.74 of ROC; the Bagging and CVR gives same outcome such as 0.81 of ROC; the LB gives 0.86 of ROC; and the MCCU gives 0.80 of ROC.

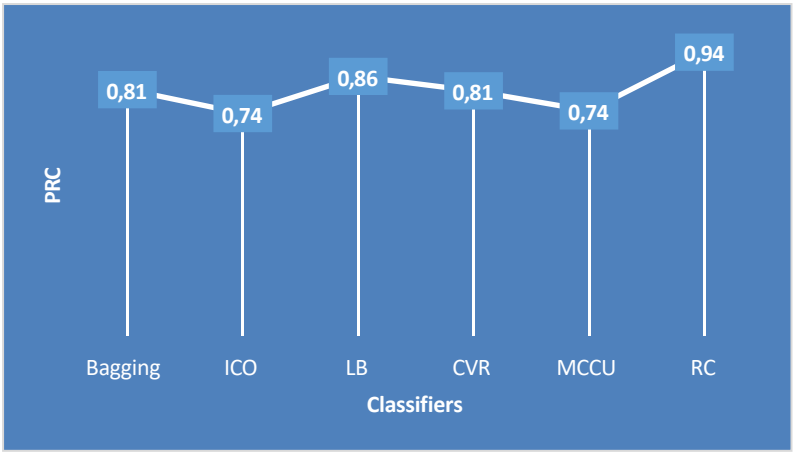


Figure 6 Classifiers vs. PRC.

The above figure 6 represents the numerous PRC level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 0.94 of PRC; the ICO and MCCU gives same as well the least outcome compare than other models such as 0.74 of PRC; the Bagging and CVR gives same outcome such as 0.81 of PRC; and the LB gives 0.86 of PRC.

The above table 3 represents the time ingesting, Kappa, F-Measure, Matthews Correlation Coefficient value (MCC) of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models.

Table 3 Classifiers vs. Outcomes.

Classifier	Time	Kappa	F-Measure	MCC
Bagging	0.23	0.70	0.85	0.70
ICO	0.80	0.39	0.65	0.30
LB	0.05	0.50	0.74	0.52
CVR	0.23	0.70	0.85	0.70
MCCU	0.16	0.60	0.80	0.61
RC	0.08	0.80	0.90	0.82

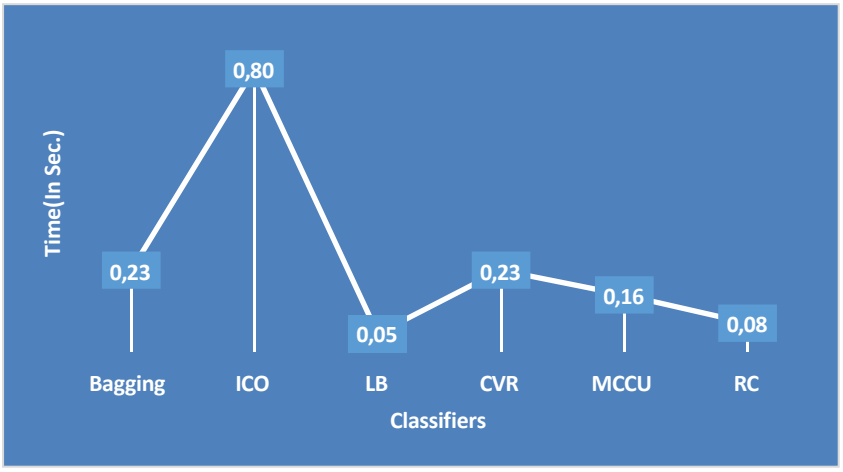
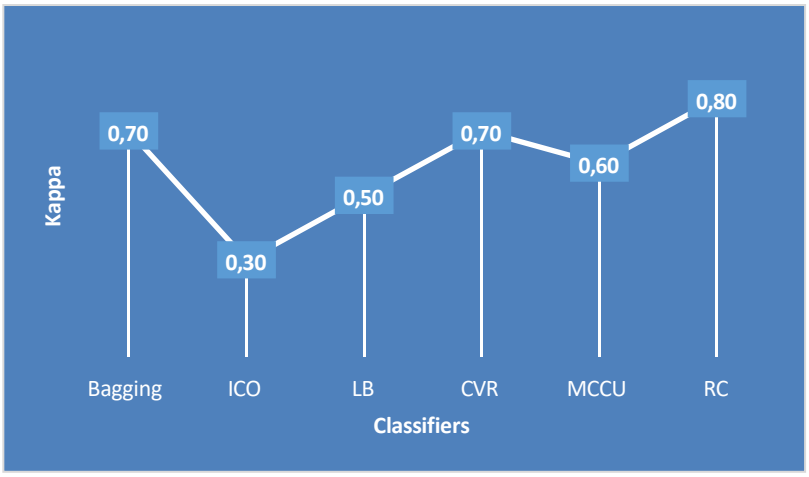


Figure 7 Classifiers vs. Time.

The above figure 7 represents the numerous time consumption in seconds for making model of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The ICO takes 0.80 seconds for creating this model which is highest compare with other models; the LB takes 0.05 seconds for building this model which is the least time for creating a model; The Bagging and CVR takes same time for making their models which is 0.23 seconds; the MCCU takes 0.16 seconds and the RC takes 0.08 seconds for constructing their models.



.

Figure 8 Classifiers vs. Kappa.

The above figure 8 represents the numerous kappa levels of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 0.80 of kappa; The ICO gives least outcome compare than other models such as 0.30 of kappa; The Bagging and CVR gives same outcome such as 0.70 of kappa; The MCCU gives 0.60 of kappa; and the LB gives 0.50 of kappa.

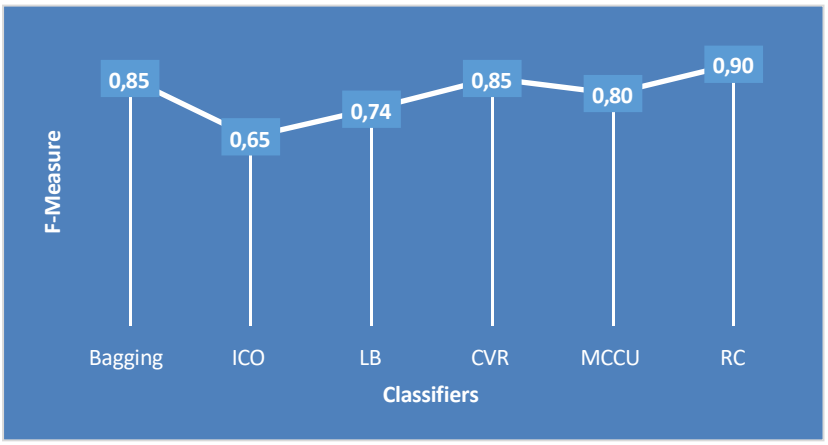


Figure 9 Classifiers vs. F-Measure.

The above figure 9 represents the numerous F-Measure levels of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 0.90 of F-Measure; The ICO gives least outcome compare than other models such as 0.65 of F-Measure; The Bagging and CVR gives same outcome such as 0.85 of F-Measure; The MCCU gives 0.80 of F-Measure; and the LB gives 0.74 of F-Measure.

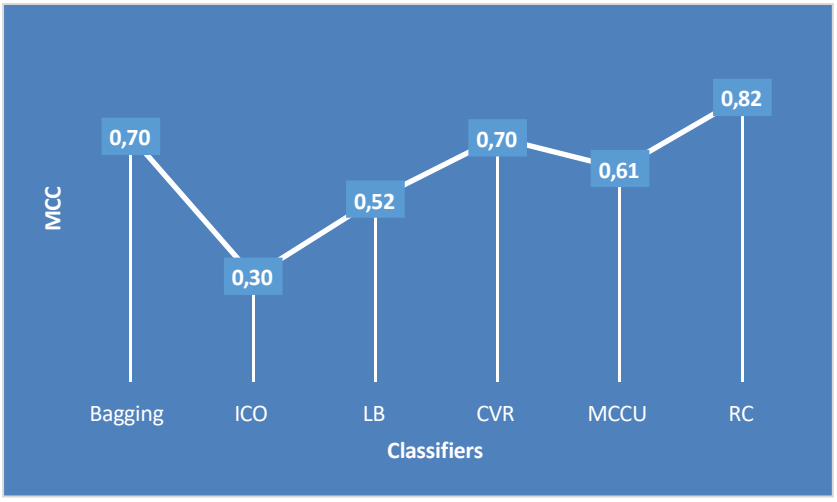


Figure 10 Classifiers vs. MCC.

The above figure 10 represents the numerous MCC levels of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC gives highest outcome compare with other models such as 0.82 of MCC; The ICO gives least outcome compare than other models such as 0.30 of MCC; The Bagging and CVR gives same outcome such as 0.70 of MCC; The MCCU gives 0.61 of F-MCC; and the LB gives 0.52 of MCC.

Table 4 Classifiers vs. Errors.

Classifier	MAE	RMSE	RAE	RRSE
Bagging	0.39	0.42	77.30%	84.69%
ICO	0.39	0.50	78%	99.49%
LB	0.22	0.45	47%	89.84%
CVR	0.39	0.42	77.30%	84.69%
MCCU	0.20	0.45	40.00%	89.44%
RC	0.29	0.33	58%	66.63%

The above table 4 depicts the Mean Absolute Error (MAE), Relative Absolute Error (RAE), Root Measure Squared Error (RMSE), and Relative Root Squared Error (RRSE) of selected classifiers.

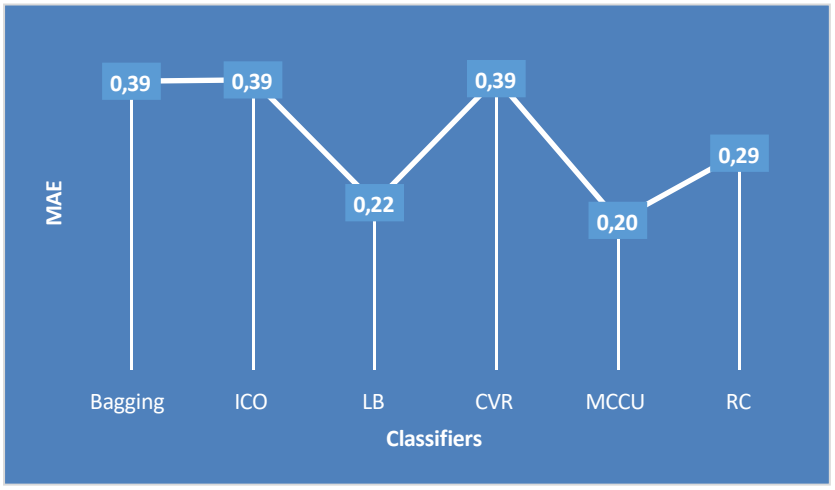


Figure 11 Classifiers vs. MAE.

The above figure 11 represents the numerous MAE levels of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The MCCU performs well compare than other models which is 0.22 of MAE; the Bagging, ICO and CVR performance are same as well worst compare than other models which is 0.39 of MAE; the RC is 0.29 of MAE and the LB is 0.22 of MAE.

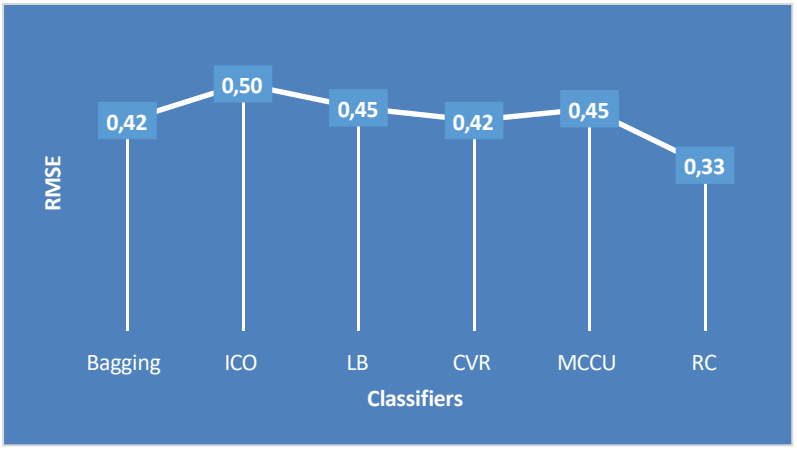


Figure 12 Classifiers vs. RMSE.

The above figure 12 represents the numerous RMSE levels of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC performs well compare than other models which is 0.33 of RMSE; the ICO is having worst enactment which is 0.50 of RMSE; the ICO and MCCU are having same enactment which is 0.45 of RMSE; the Bagging and CVR are having same performance which is 0.42 of RMSE.

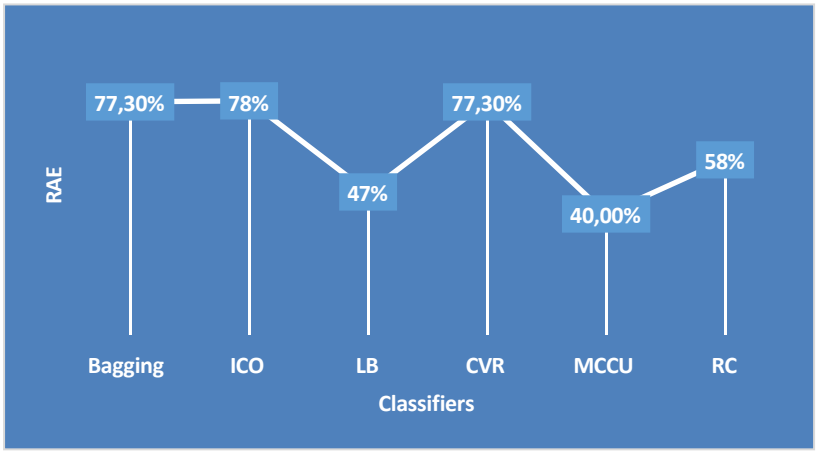


Figure 13 Classifiers vs. RAE.

The above figure 13 represents the numerous RAE level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The MCCU performs well which is 40% of RAE; the LB shows 47% of RAE; the ICO performs worst case which is 78% of RAE; the Bagging and CVR are showing worst performance as well same which is 77.30% of RAE; and the RC performance is 58% of RAE.

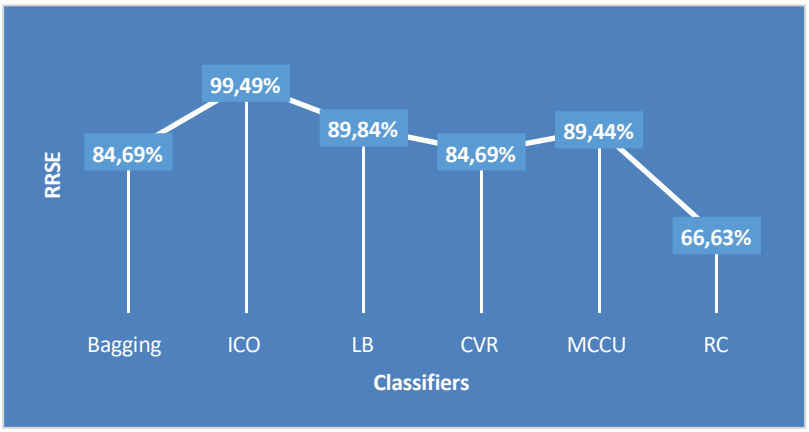


Figure 14 Classifiers vs. RRSE.

The above figure 14 represents the numerous RRSE level of Bagging, LB, ICO, CVR, MCCU, and RC on Meta models. The RC shows best enactment which is 66.63% of RRSE; the ICO shows worst enactment which is 99.49% of RRSE; the LB and MCCU shows same performance which is 89.44% of RRSE; the Bagging and CVR shows 84.69% of RRSE.

5. Conclusion

Based on the results of this study, it can be said that the RC provides the best results compared to other models, such as an accuracy of 90%, while the ICO provides the worst results compared to other models, at 65% accuracy, 90% precision, 65% recall, and 65% F-measure. The RC yields the best result in assessment to other models, such as the ROC (0.96), while the ICO yields the worst results (0.74), the PRC (0.94), and the ICO and MCCU yield the same results (0.74), respectively. In comparison to other models, the ICO's 0.80-second creation time for this one is the longest; the LB's 0.05-second creation time for this one is the shortest. When compared to other models, the RC yields the best results (0.80 kappa), while the ICO yields the worst (0.30 kappa), the RC yields the best (0.90 f-measure), and the ICO yields the worst (0.65 f-measure), and so on. When it comes to the mean correlation coefficient (MCC), the RC yields the best (0.82 kappa), while the ICO yields the worst (0.30 k The best RRSE performance can be found in the RC, at 66.63 percent, while the worst can be found in the ICO, at 99.49 percent. The Random Committee model is recommended in this work because it performs well across multiple metrics.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Antoniou A, Amos S (2019) Assume, Augment and Learn: Unsupervised Few-Shot Meta-Learning via Random Labels and Data Augmentation.

Berthelot D, Carlini N, Cubuk ED, Kurakin A, Sohn K, Zhang H (2019) Remixmatch: semi-supervised learning with distribution alignment and augmentation anchoring. arXiv [Preprint]. arXiv:1911.09785. DOI: 10.48550/arXiv.1911.09785

Cubuk ED, Zoph B, Shlens J, Le QV (2019) Randaugment: practical data augmentation with no separate search. arXiv preprint arXiv:1909.13719. DOI: 10.1109/CVPRW50498.2020.00359

Dharmasiri HML, Goonetillake MDJS (2013) A federated approach on heterogeneous NoSQL data stores. in 2013 International Conference on Advances in ICT for Emerging Regions (ICTer).

Doon R, Rawat YK, Gautam S (2018) Cifar-10 classification using deep convolutional neural network" in 2018 IEEE Punecon, IEEE.

Fu M, Cao YH, Wu J (2022) Worst case matters for few-shot recognition. arXiv [Preprint]. arXiv:2203.06574. DOI: 10.48550/arXiv.2203.06574

Gallinucci E, Golfarelli M, Rizzi S (2018) Schema profiling of document-oriented databases. Information Systems 75:13-25.

Hu Z, Yang Z, Hu X, Nevatia R (2021) Simple: similar pseudo label exploitation for semi-supervised classification in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashvile, TN, 15099–15108. DOI: 10.1109/CVPR46437.2021.01485

Karuppusamy P (2021) Building Detection using Two-Layered Novel Convolutional Neural Networks. Journal of Soft Computing Paradigm (JSCP) 3:29-37.

Kavi BO, Zeebaree SR, Ahmed OM (2020) Deep Learning Models Based on Image Classification: A Review. International Journal of Science and Business 4:75-81.

Klettke M (2016) NoSQL schema evolution and big data migration at scale. in 2016 IEEE International Conference on Big Data (Big Data).

La Grassa R, Gallo I, Landro N (2021) Learn class hierarchy using convolutional neural networks. Applied Intelligence 1-11.

Lucas T, Weinzaepfel P, Rogez G (2021) Barely-supervised learning: semi-supervised learning with very few labeled images. arXiv preprint. arXiv:2112.12004.

Manoharan JS (2021) Study of Variants of Extreme Learning Machine (ELM) Brands and its Performance Measure on Classification Algorithm Journal of Soft Computing Paradigm (JSCP) 3:83-95.

Ozler H (2019) A curated list of JSON/BSON datasets from the web in order to practice/use in MongoDB. Available in: http://www.github.com/ozlerhakan/mongodb-json-files.

Rajesh A, Mohan E (2014) Classification of microcalcification based on wave atom transform", Journal of Computer Science 10:1543-1547.

Rasim CC, Demirci FM (2018) Cifar-10 image classification with convolutional neural networks for embedded systems. 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA).

Saur K, Dumitraş T, Hicks M (2016) Evolving NoSQL Databases without Downtime. in 2016 IEEE International Conference on Software Maintenance and Evolution (ICSME).

Scavuzzo M, Nitto ED, Ceri S (2014) Interoperable Data Migration between NoSQL Columnar Databases. in 2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations.

Scavuzzo M, Tamburri DA, Nitto ED (2016) Providing Big Data Applications with Fault-Tolerant Data Migration across Heterogeneous NoSQL Databases. in 2016 IEEE/ACM 2nd International Workshop on Big Data Software Engineering (BIGDSE).

Smith, Leslie N, (2022) Conovaloff, Adam, Building One-Shot Semi-Supervised (BOSS) Learning Up to Fully Supervised Performance, Frontiers in Artificial Intelligence 5.

Sohn K, Berthelot D, Carlini N, Zhang Z, Zhang H, Raffel CA, (2020) Fixmatch: simplifying semi-supervised learning with consistency and confidence. Adv. Neural Inform Process Syst 33:596–608. DOI: 10.48550/arXiv.2001.07685

Swaroop P (2016) NoSQL Paradigm and Performance Evaluation. SSARSC International Journal of Geo Science and Geo Informatics 3.

Thakkar V, Tewary S, Chakraborty C (2018) Batch Normalization in Convolutional Neural Networks-A comparative study with CIFAR-10 data, 2018 fifth international conference on emerging applications of information technology (EAIT):1-5.

Thambu G, Mohan E (2014) Object Recognition Based Wave Atom Transform. Research Journal of Applied Sciences and Technology 8:1613-1617.

Tripathi M (2021) Sentiment Analysis of Nepali COVID19 Tweets Using NB SVM AND LSTM. Journal of Artificial Intelligence 3:151-168.

Xie Q, Hovy E, Luong MT, Le QV (2019) Self-training with noisy student improves imagenet classification. arXiv preprint arXiv:1911.04252. DOI: 10.1109/CVPR42600.2020.01070

Zafar R, (2016) Big Data: The NoSQL and RDBMS review. In 2016 International Conference on Information and Communication Technology (ICICTM).

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0102

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

A decision-making framework utilizing machine learning techniques, based on analysing the sentiment of tweets

Putta Durga[a] | **T. Sudhakar**[b]

[a]VIT-AP University, Amaravati, Andhra Pradesh, India, Research Scholar, School of Computer Science and Engineering.
[b]VIT-AP University, Amaravati, Andhra Pradesh, India, Associate Professor, School of Computer Science and Engineering.

Abstract Before making any decision, especially a financial one, people use social media to chat to, learn about, and obtain advice from others. How can social media opinions be relied on and utilised economically? Due to social media sites like Facebook and Twitter, where millions of individuals offer real-time thoughts on anything using slang and emoticons, data science researchers have access to more unstructured yet useful information. Social media includes Twitter. 326 million active social media users sent 500 million tweets daily in July 2018, indicating the sector's rapid expansion. Tweets can be viewed, edited, and sent. Knowing how to interpret tweets, which might be facts or views, is helpful. Such a study may predict the stock market, election, response event, news, and subjectivity scales. Tweets are analysed using Sentiment Analysis (SA). On the other hand, Sentiment analysis provides a broad overview of tweet polarity and is not beneficial for decision-making. In this research, we design a natural language processing (NLP) Twitter SA model to overcome the challenges of tweet sentiment recognition and build a hybrid system that can recognise and classify sentiment in Twitter's real-time reactions to any topic. We use a three-classifier machine learning method to analyse tweets from the Sentiment140 dataset (Logistic Regression, Bernoulli Naive Bayes, and SVM). Our findings are supported by the Term Frequency-Inverse Document Frequency analysis (TF-IDF). After that, accuracy and F1 Scores are used to evaluate these classifiers.

Keywords: sentiment analysis, nlp, dss, twitter, machine learning

1. Introduction

The term "sentiment analysis" is used to describe the process of locating and labelling the emotions that are expressed in a piece of writing. When analysed, tweets can yield a wealth of information about public opinion. These statistics help us learn how individuals feel about certain issues. As a result, to calculate customer satisfaction, we need to create an automated ML sentiment analysis model. Model implementation becomes more challenging when both relevant and useless characters (called noise) are present in the data.

The social media platform Twitter has become a popular place for people to express their thoughts and feelings about current events in the digital age. When it comes to sharing information about current events, politics, and public opinion, Twitter has become the most popular online social networking tool. Anyone can use this to find out what people are talking about on Twitter right now (Zhongyu Lu 2015). When it comes to social data, live streaming is a convenient method of obtaining both positive and negative assessments of any current issue. Due to the exponential growth in social media data (Twitter, Facebook, etc.), enterprise data warehouses must develop effective methods of dealing with this new type of data.

People can share their thoughts and build relationships through posts, comments, messages, and the likes section of social media platforms, which serve as virtual barriers between users. A social media site that is rapidly expanding is Twitter. Even in July 2018, Twitter had 326 million active users who sent out 500 million tweets per day (Ruz et al 2020). Tweets are 140-character messages that anybody may write, send, edit, and read. People frequently use Twitter to report on a wide range of events. As a bonus, Twitter also gives a trove of tweets for the public to peruse.

A "tweet" on Twitter is a short text message that can express the user's thoughts and feelings. An enterprise or group could learn a lot from analysing this (Daniati E 2020). The findings of this study can be applied to the forecasting of future events and the identification of emerging trends. Also, the analysis result may be a description of the framework of the entity relationship between Twitter users. Sentiment analysis, a subfield of NLP with its sequence of steps, is necessary for obtaining this research's findings.

Clustering, pre-processing, counting, recording word occurrences and indexing is the cornerstones of most existing event discovery methods (Zhongyu Lu 2015).

(Haifei Huang et al 2015; Avery Ching et al 2015). Existing topic recognition approaches only address issues of global magnitude, while interesting new issues of lesser magnitude receive far less attention (Erich Schubert et al 2014). These new developments will alert consumers to developing trends so they can act swiftly.

Decision-making in a business is crucial. This necessitates the creation of what is essentially a decision support system embedded within a management information system (DSS). It helps businesses efficiently manage their marketing, promotions, and product offerings. The influence of social media platforms is equally crucial to these events. This DSS tracks consumer feedback on advertisements and services across multiple platforms. This helps managers make judgements not based on their gut but on data and modelling criteria. The process of sentiment analysis produces these standards. Because of this, the ensuing decisions are more efficient and in line with the company's aims, and managers have a plethora of options from which to choose.

In this study, we suggest developing and implementing practical methods for mining event information and sentiment from Twitter streams in real-time. The attitudes on Twitter data are classified using a variety of ML models, including Logistic Regression, Bernoulli Naive Bayes, and SVM. Collected information about popular tweets is loaded into the Hive database for analysis. The suggested method maintains its superiority even when the data size and the number of events rise, making it a good candidate for large-scale applications. Experiments conducted on the real dataset demonstrated the effectiveness of the method for recognising viewpoints in real-time tweets.

This paper follows the structure described below. A review of the relevant literature is provided in Section 2, while the system architecture is described in Sections 3 and 4. The results and methodology used are discussed in Section 5. Section 6 contains the conclusion.

2. Related Work

Here, we'll look at the most important studies on sentiment analysis and product rankings as they relate to decision-making assistance systems. To boost sentiment analysis effectiveness, numerous strategies have developed over time. These methods have progressed from lexicon-based algorithms and ML algorithms to more recent and sophisticated DL algorithms.

By using this paradigm, Islam A. et al. (Islam 2021) can transform data mining's unstructured output into labelled data, which can then be used to construct an advisory and decision-making system (DMSS). Data mining techniques have allowed them to amass vast amounts of information, but the real challenge lies in sifting through all of that noise to find the gems of insight. Machine learning and deep learning algorithms were found to significantly improve categorization accuracy across the board. The Covid corpus, processed with a long short-term memory, yielded the greatest accuracy in classifying (99% during training and 96% during testing) (LSTM).

(Kavitha et al 2018) devised a streaming algorithm to retrieve information from Twitter through a keyword search. Data visualisation, report generation, and analysis are all goals of the Twitter data visualisation application. Configuring the system with Hadoop, Hive warehouse, and Apache Flume retrieves real-time data from Twitter. The keyword file is uploaded to the Hadoop cluster where the flume agent will gather the necessary data via the flume channel and then dump the information into the Hadoop Distributed File System. This programme can be used to monitor public opinion in real time on issues such as Elections, Government actions, or any other topic of public interest.

The TSS model suggested by (Guo 2019) uses a novel baseline correlation approach, which not only shows good prediction accuracy but also lessens the computational cost and allows for quick decisions to be made without access to past data. R is used for a variety of statistical purposes, including lexicon-based sentiment analysis, polynomial regression, and classification modelling. Using the suggested baseline model and without consulting historical TSS or market data, the resulting TSS accurately forecasts the future stock market trend by 15-time samples (30 working hours). In particular, TSS's performance in predicting an upward market is found to be significantly better than its performance in predicting a negative market. TSS's prediction of the future market's upward trend is 97.87% accurate when using logistic regression and linear discriminant analysis.

(Daniati E 2021) aimed to provide a decision-making framework using sentiment analysis. This study also represents an advancement over the prior decision-making framework, which relied on sentiment analysis. A switch from the SAW approach to the TOPSIS method allowed for better modelling of the criteria. Also, a comparison is made between TOPSIS and SAW in terms of the final value of the decision alternatives. Scores and rankings at the end of the process are utilised as comparison criteria. When compared to TOPSIS, the final score obtained by the SAW approach is higher. In addition, the rankings produced by TOPSIS and SAW are distinct.

To make decision-making as easy as possible (Pai 2022) created a visual representation of real-time tweets by classifying and determining the polarity score of live tweets according to whether it is positive, negative, or neutral using Vader, which is more efficient than traditional algorithms, and presenting the user with a graph of the subject matter, which is updated by the tweets posted every second.

The work of (Talpada et al 2019) shows the relationship between DL and lexical and semantic-based sentiment classifiers, (ii) the accuracy of sentiment prediction using these methods on a manually annotated sentiment dataset, (iii) the effect of domain-specific knowledge on the accuracy of sentiment prediction methods, and (iv) the use of Twitter-based sentiment to understand the impact of telemedicine on heart attack and epilepsy. They use four different sentiment-predicting strategies in their study. They gathered a dataset of 1.84 million tweets that were posted in the past linked to health. Their research shows that traditional lexical and semantic approaches to sentiment prediction are more accurate than Deep Learning techniques. They found that the target text's prevalence of domain-specific words affected the accuracy with which sentiment could be predicted using only generic information.

SentDesk was an attempt by (Oppong et al 2019) to create a system that may aid businesses by taking user sentiment into account while making decisions. Some businesses in Ghana used this technique, and the results showed that it was superior to human evaluation. Despite this, the authors have not taken into account user discontent or hesitation when making decisions by ignoring the hazy sentiment phrases that were included in the opinions.

To generate lists of recommended purchases (Najmi et al 2015) examined the opinions expressed in online reviews of lexicon-based products such as high-definition televisions and cameras by computing the emotion score of opinions, the brand score of each product, and the helpfulness of each review. Multi-criteria decision-making (MCDM) is a ranking system that takes into account elements of items and customer contentment, discontent, and reluctance for these aspects, whereas the aforementioned methods only evaluated user pleasure for products by assessing sentiment. In addition, a decision-making tool for ranking things is presented in (Sedef Çalı 2019); this strategy combines MCDM and aspect-based SA approaches to evaluate potential options.

(Jawa and Hasija 2015) came up with a model that uses the basics of the Interest graph and Sentiment analysis to figure out how different things relate to each other and make suggestions, like whom to follow on Twitter or what to buy online.

This study makes several key contributions, which can be summarized as follows:

- Investigation of a lexical dictionary's implementation viability and assessment of an ensemble's effectiveness for sentiment classification of tweets.
- Integration and evaluation of three feature engineering approaches: TF-IDF, BOW, and a combination of BOW and TF-IDF.
- Introduction of an ensemble model consisting of LR, support vector classifier (SVC), and Bernoulli naive Bayes (NB) for tweet sentiment classification. The ensemble's combination proved to yield promising results, especially considering LR's proficiency in binary classification tasks.
- Extensive experimentation was conducted to compare the proposed approach's performance against conventional state-of-the-art machine learning models.

3. Sentiment Analysis

Research into the analysis of written expressions of opinion, sentiment, and emotion about entities (such as products, services, organisations, people, events, issues, and themes) is an active subfield of natural language processing known as sentiment analysis. Sentiment analysis aims to recognise, extract, and categorise feelings expressed by users on online forums like social media, blogs, and product review sites. With the proliferation of social media platforms like Twitter and Facebook, as well as review aggregators like Rotten Tomatoes, Amazon, and Yelp, sentiment analysis has attracted the interest of academics and businesspeople alike. Separating subjective and objective phrases, or identifying opinions and facts, is the goal of the categorization process known as "subjectivity." Meanwhile, sentiment categorization is the process of figuring out if a given statement is positively or negatively oriented based on the emotions it conveys. Moreover, some scholars are interested in measuring the intensity of sentiment polarity by calculating its intensity (or strength). The investigation of how people feel about various aspects of things is called "feature-based sentiment analysis," and it's the subject of an in-depth study. The feature-based sentiment analysis of smartphone screens, for instance, examines the emotional expressions of smartphone users and attempts to determine whether they are good or negative. Word-level, phrase- or sentence-level, document-level, and feature-level sentiment analysis are all possible jobs. 280-character snippets of information disseminated via Twitter are called "tweets," and the site is one of the most popular ways to communicate online. The fineness of detail provided by the word level is appropriate here. Several methods have been used to automatically analyse the tone of text based on the predicted emotional tone of individual words, phrases, and entire texts. Such examples are dictionary-based methods and ML (machine learning) approaches. Here, we offer a method for mining Twitter data's sentiment that takes advantage of both supervised and unstructured methods. The next paragraph will show an example of our proposed method in action.

4. Proposed Approach

The suggested structure is comprised of the following four stages:

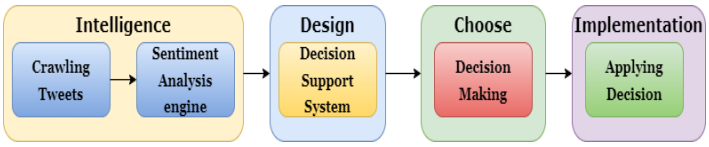


Figure 1 Decision-Making Framework into Sentiment Analysis.

The decision-making procedure is illustrated in Figure 1. Crawling Tweets and a Sentiment Analysis Engine is part of the Intelligence Phase. Twitter content is "crawled" to collect tweets that match certain search criteria. A computer tool called Sentiment Analysis Engine can sort a set of tweets into positive, negative, and neutral categories. During the planning phase, the polarity label serves as a criterion. Thus, the profit criteria are labelled positively and neutrally but have varying degrees of importance. At the same time, the labels for the cost criteria are all negative. As an input for the value of each criterion, we use the total number of tweets that fit the label.

The framework's architecture for generating decisions via sentiment analysis is depicted in Figure 1. It should be noted that this new framework is an update to the old one. It's also possible to return to the beginning of the process, at the intelligence block, by following a route that begins at the Implementation block. When decisions are put into action, there is a corresponding flow of information or feedback. There must be input in the form of feedback if the decision made was inappropriate or if there were roadblocks that ran counter to the decision-making phase, otherwise, the existing processes in the intelligence and design phase would not be revised.

The complete Sentiment Analysis Engine (SAE) model, depicted in Figure 2, comprises a number of phases. Let's break this down step by step.

4.1. Twitter Dataset

We're making an effort to gather all of the labelled tweet data at this point. The tweets in this dataset were retrieved from kaggle.com. This dataset contains information about English tweets that have been assigned positive, negative, or neutral polarity labels. In addition, the Twitter dataset is converted into a corpus for usage in Text Pre-processing.

The Sentiment140 Dataset includes 1,600,000 tweets that were collected with the use of the Twitter API. There are a total of 1,600,000 tweets in the training set for the sentiment-140 dataset, 800,000 of which contain positive emoticons and 800,000 of which contain negative emotions. The test set consists of 177 negative tweets and 182 positive tweets, with only a subset of this data containing emoticons. Any business or individual interested in quickly and easily determining whether or not their brand, product, or issue is generally well-liked or disliked on Twitter can utilise this dataset. There are no non-English tweets in this data collection. The dataset contains the following columns:

- **target:** tweet's divisiveness (positive or negative)
- **ids:** The tweet's one-of-a-kind id
- **date:** the date of the tweet
- **flag:** It is about the question. If such a query does not already exist, then the term "no query" applies.
- **user:** That is the name of the Twitter person who sent out the tweet.
- **text:** This pertains to the actual content of the tweet.

Table 1 Dataset Summary.

Split	Examples
Test	498
Train	1,600,000

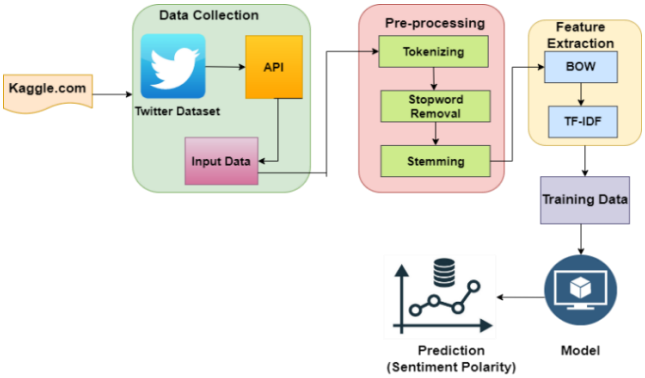


Figure 2 Sentiment Analysis Engine Architecture.

4.2. Text Pre-processing

At this point, we are attempting to gather a pristine collection of words and terms. This is done to ensure that this group of words can be organised structurally and has value. Tokenizing, removing slang, eliminating stop words, and stemming are the components that make up this procedure. The process of tokenizing seeks to produce a list of words that have no punctuation in between them. In addition, slang removal involves changing instances of slang to their regular word counterparts. The method known as Stop Word Removal works to get rid of words like conjunctions that don't have any significant meanings on their own. The stemming procedure is the final step. This procedure attempts to alter the fundamental structure of words.

4.3. Tokenization

Tokenization refers to the process of disassembling complete sentences into their component words as the first step in preparing text data for analysis. By isolating words rather than whole phrases, meaning is lost. This is even though it is a frequent approach to analysing massive text databases. For text data analysis, it is sufficient for computers to look at what words appear in an article and how often these terms appear. This is fast, easy, and accurate, and it yields informative results.

4.4. Stopwords

The news story has been cleaned up after some editing, however, there are still terms like "and," "we," and "etc." that we would prefer not to see. Getting rid of the filler words (also known as "stopwords") is the next step. Words that are used repeatedly but have no real significance in any given context are called stopwords. Words like "I," "the," "a," and "of" are all examples of stopwords. If these terms were removed from articles, it wouldn't change how they were understood. If we import the stopwords from the NLTK library, we can get rid of them.

4.5. Lemmatization

All news articles will be reduced to a list of intelligible words by removing stopwords, symbols, numbers, and punctuation. But to tally up the occurrences of each word, you have to change them back to their base forms, eliminating the tense markers that were previously obscuring the count. To determine the frequency of occurrence of the word "open" in a given news story, for instance, one must count all instances of the word, including its many forms ('open,' 'opens,' and 'opened'). As a result, lemmatization is a crucial stage in text processing. Stemming is another method for reverting a term to its base form. The dissimilarities are as follows:

```
text = 'I ate two apples yesterday'
```

```
words=[token.lemma_ for token in nlp(text) if not token.is_punct]
words
```

```
['-PRON-', 'eat', 'two', 'apple', 'yesterday']
```

```
porter = nltk.PorterStemmer()
[porter.stem(t) for t in text.split()]
```

```
['I', 'ate', 'two', 'appl', 'yesterday']
```

Figure 3 Sample Example of Lemmatization.

One way to break down a word is to use lemmatization, while another is to use stemming, which gets at the linguistic roots of the term. Because some words become unintelligible after being stemmed, lemmatization was chosen instead. The lemma provides a clearer meaning than the linguistic root.

4.6. Feature Extraction

Bag of Words, TF-IDF, word embedding, and Natural Language Processing (NLP) based features like word count, noun count, etc. are just a few of the many feature extraction methods available. In this research, we looked at how the use of TF-IDF word level and Bag of Words features affected the sentiment analysis of a dataset containing tweets.

4.7. Bag of Words

The bag of words model is employed in NLP and IR projects for the purpose of representing texts and extracting features from them. A document is represented as a multiset of its words, with syntax and word order ignored while word frequency is preserved. Text clustering, document similarity, and classification are just some of the applications where this representation is beneficial.

4.8. TF-IDF Vectorizer

Information retrieval and natural language processing applications make use of TF-IDF (Term Frequency-Inverse Document Frequency), a statistical measure. It indicates how significant a document's term is in comparison to the overall corpus. A notion at its most fundamental level is that a term that is common in a single document but scarce across the entire corpus provides more insight than a word that is common in both the document and the corpus.

Term Frequent (TF). The following equation provides a measurement of the frequency with which a given term, t, appears in a given text, d:

$$(tf_{t,d} = \frac{n_{t,d}}{number\ of\ terms\ in\ the\ document}) \quad (1)$$

The term "t" appears in the given document "d" a total of n times, which is denoted here by the numerator value "n." As a result, every document and every term would have its very own unique TF value.

The Inverse Document Frequency (IDF) is a measurement used to determine how significant a term is and is denoted by Equation (2). The value of the IDF is necessary for us because only computing the TF on its own is insufficient to comprehend the significance of words:

$$(idf_t = log\ \frac{number\ of\ documents}{number\ of\ documents\ with\ term\ 't'}) \quad (2)$$

4.9. Training and Testing Data

In this step, we'll be working to train a model using actual data. A classifier of ML is used to train data. Bernoulli Naive Bayes, Support Vector Machine and Logistic Regression were used to perform the classification in this study. To categorise tweets that haven't been labelled, we utilise the model we learned from the labelled data as training.

Our evaluation metrics are applied to the trained model to determine how well it is doing. As a result, we utilise the following metrics to assess the accuracy of the models:

Confusion Matrix and Receiver Operating Characteristic Curve for Accuracy.

In the testing phase, unlabelled tweets (Test Data) are subjected to various classification methods. The polarity of this Tweet will be marked as either positive (+), negative (-), or neutral (=). The model and classifier are put to the test to classify the data. When utilising data for model development via training, it is important that the classifier employed is appropriate for the classifier type.

5. Results and Discussion

```
df.head()
```

	target	ids	date	flag	user	text
0	0	1467810369	Mon Apr 06 22:19:45 PDT 2009	NO_QUERY	_TheSpecialOne_	@switchfoot http://twitpic.com/2y1zl - Awww, t...
1	0	1467810672	Mon Apr 06 22:19:49 PDT 2009	NO_QUERY	scotthamilton	is upset that he can't update his Facebook by ...
2	0	1467810917	Mon Apr 06 22:19:53 PDT 2009	NO_QUERY	mattycus	@Kenichan I dived many times for the ball. Man...
3	0	1467811184	Mon Apr 06 22:19:57 PDT 2009	NO_QUERY	ElleCTF	my whole body feels itchy and like its on fire
4	0	1467811193	Mon Apr 06 22:19:57 PDT 2009	NO_QUERY	Karoli	@nationwideclass no, it's not behaving at all....

Figure 4 Five most notable data records.

Figure 4 is an illustration of the exploratory data analysis of the top five records with the most occurrences, which will be displayed.

```
[ ]  df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1600000 entries, 0 to 1599999
Data columns (total 6 columns):
 #   Column  Non-Null Count    Dtype
---  ------  --------------    -----
 0   target  1600000 non-null  int64
 1   ids     1600000 non-null  int64
 2   date    1600000 non-null  object
 3   flag    1600000 non-null  object
 4   user    1600000 non-null  object
 5   text    1600000 non-null  object
dtypes: int64(2), object(4)
memory usage: 73.2+ MB
```

Figure 5 Data information.

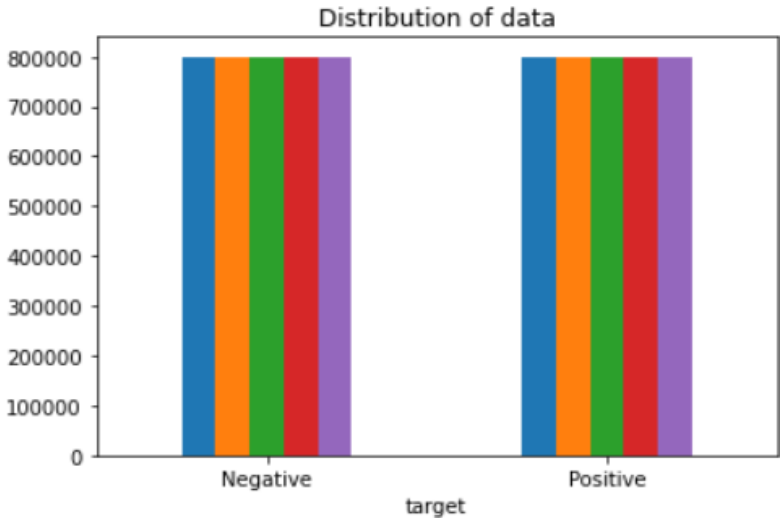


Figure 6 Data Visualization of Target variables.

Figure 6 illustrates the variables that are contained within a Twitter dataset, such as the positive and negative samples that are on the target set.

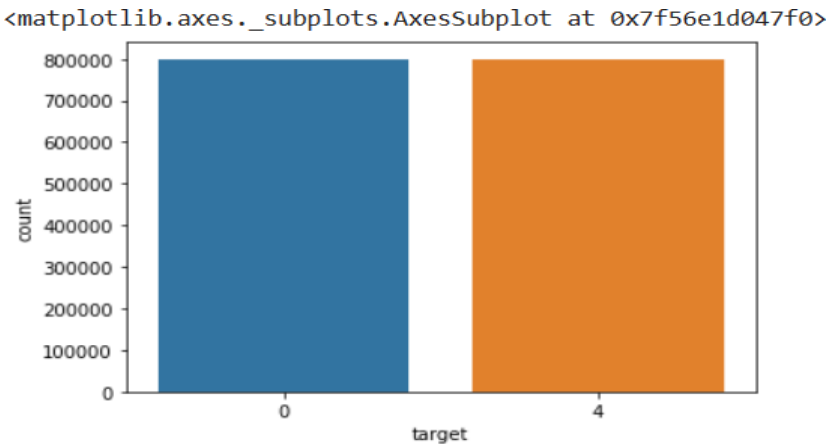


Figure 7 Plot representation of target values.

Upon observation of entries for each class, it becomes apparent that the negative class spans from the 0th to the 799,999th index, while the positive class encompasses entries starting from the 800,000th index until the end of the dataset. Tables 2 and 3 present the counts of positive and negative words, as well as the word count per message.

Table 2 Positive Vs. Negative.

No. of positive tagged sentence is	788435
No. of negative tagged sentence is	790177
Total lengh of the data is	1578612

Table 3 word count per message.

Text	Word count
Is very bad for his school mate	7
I ignored the latest movie AV	6
Omg its already 7:30 :Clk	5

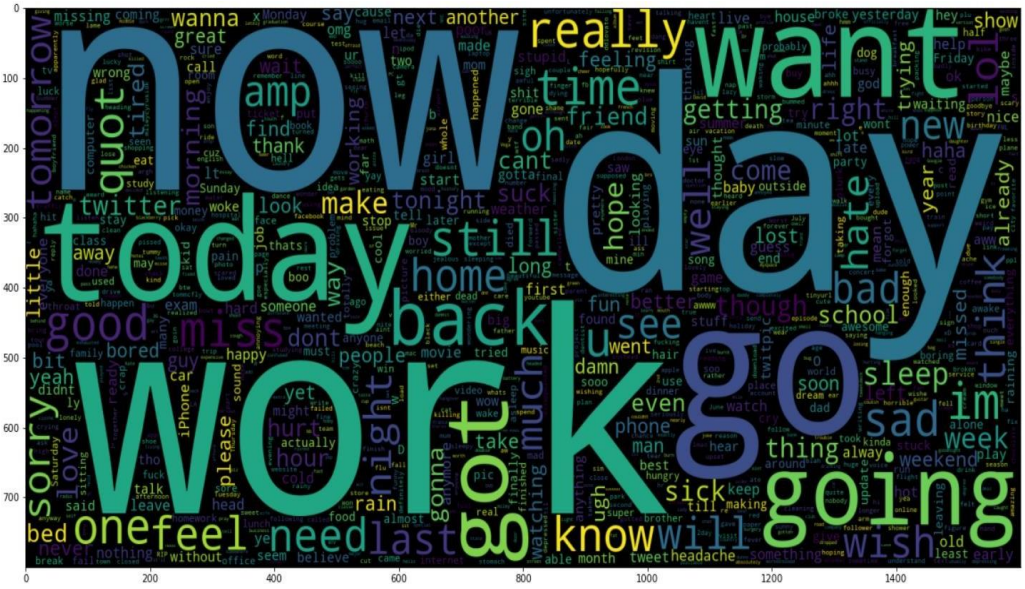


Figure 8 Word cloud for people to use in their unfavourable tweets.

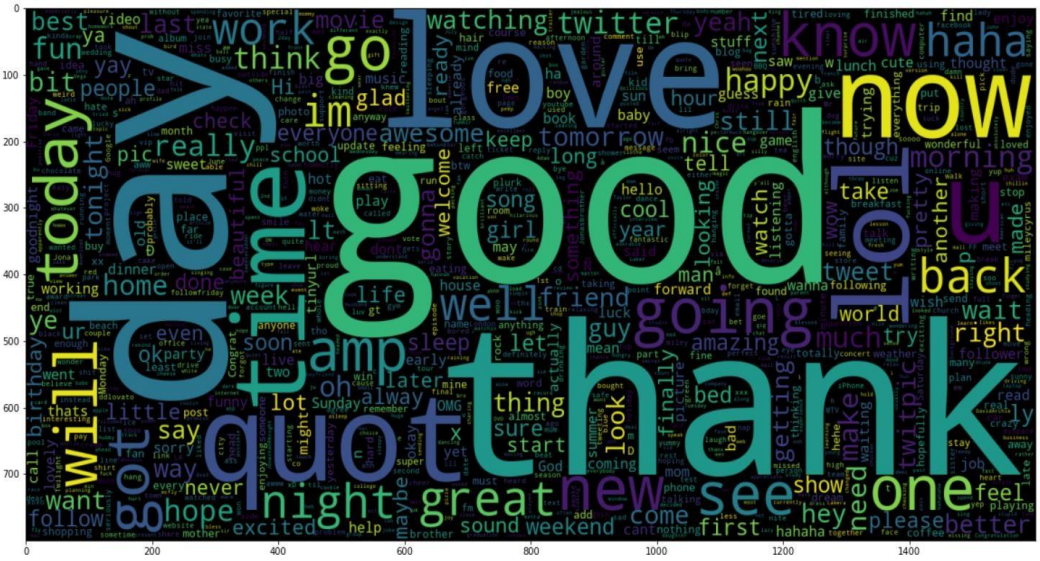


Figure 9 Word cloud for use in tweets that are positive.

The positive and negative tweets that were included in the dataset are displayed in Figs. 8 and 9, respectively.

Specifically, we have applied the following three models to the problem statement:

Several Statistical Methods (including Bernoulli Naive Bayes, SVM, and Logistic Regression) selecting these models was motivated by a desire to evaluate all classifiers, from the most basic to the most complicated, on the dataset in search of the one with the highest performance.

TN **39.77%**	**FP** **10.36%**
FN **9.56%**	**TP** **40.31%**

Figure 10 The Confusion Matrix for the Bernoulli NB Approach.

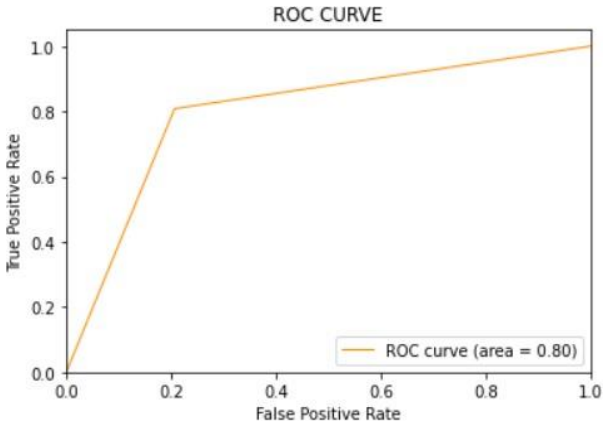


Figure 11 ROC curve for Bernoulli NB.

TN **40.49%**	**FP** **9.63%**
FN **8.77%**	**TP** **41.11%**

Figure 12 Confusion Matrix for Linear SVC Model.

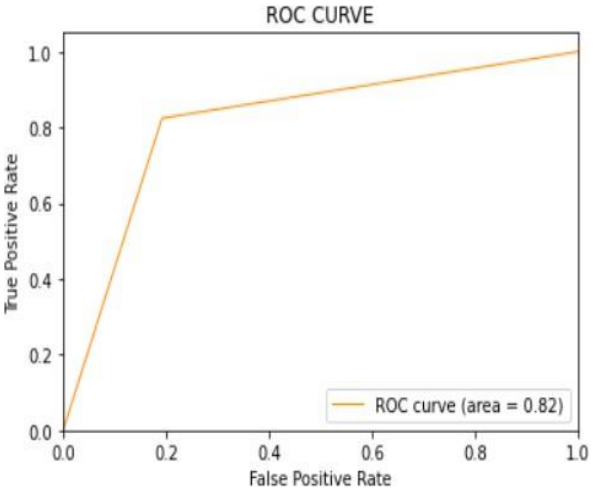


Figure 13 ROC curve for Linear SVC Model.

TN **40.97%**	**FP** **9.16%**
FN **8.21%**	**TP** **41.66%**

Figure 14 Confusion Matrixes for LR Approach.

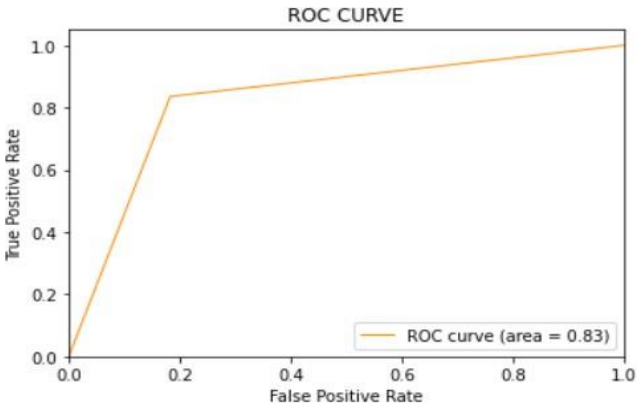


Figure 15 ROC curve for LR.

When compared to the other models we tried, the Logistic Regression Model is the winner. To classify a tweet's sentiment, it obtains nearly 83% accuracy.

It's worth noting that the BernoulliNB Model is also the quickest to train and make predictions with. As a classifier, it is 80% effective. Summary of findings Accuracy after analysing all of the models, we found the following information to be true. With regards to model accuracy, Logistic Regression outperforms SVM, which in turn outperforms Bernoulli NB.

F1 Score: The F1 Scores for class 0 and class 1 are as follows: (a) For class 0: Bernoulli NB (accuracy = 0.90) SVM (accuracy =0.91) LR accuracy = 0.92) (b) For class 1, Bernoulli NB has an accuracy of 0.66, Support Vector Machines have an accuracy of 0.68, and Logistic Regression has an accuracy of 0.69.

There is no difference in the ROC-AUC between any of the three models.

Based on our analysis, we determine that the LR is the optimal model for the aforementioned data.

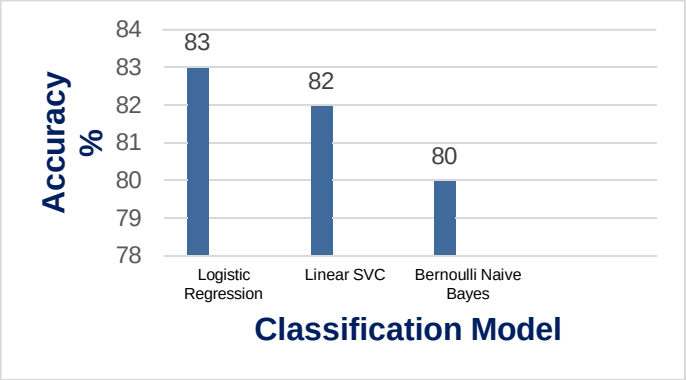


Figure 16 Accuracy of each Model.

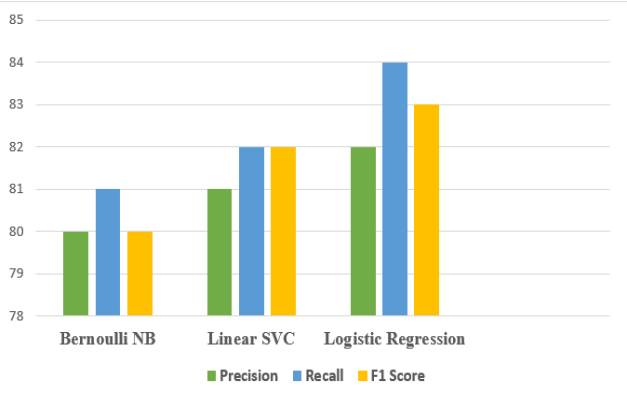


Figure 17 Performance graph of different ML models used in this work.

All of these models did quite well on the Twitter dataset, and their relative accuracy in this suggested strategy is displayed in Fig. 16. Accuracy for LR, Linear SVC, and Bernoulli NB are, respectively, 83%, 82%, and 80%. It turned out that LR classification was the most accurate method.

Table 4 Accuracy and Evaluation parameters for three models.

Model	Accuracy	Precision	Recall	F1-Score
Bernoulli NB	80	80	81	80
Linear SVC	82	81	82	82
Logistic Regression	83	82	84	83

Table 5 Evaluation of the Suggested Research Design against existing Works.

Ref	Year	Classifier	Dataset	Accuracy in (%)
Athindran 2018	2018	NB	Self-collected	77
Vanaja 2018	2018	NB and SVM	Self-collected	81
Rathi 2018	2018	Adaboost, SVM	Sentiment140	67, 80
Wongkar 2019	2019	NB	Twitter	75.58
Prabhakar 2019	2019	AdaBoost (Bagging, Boosting)	Twitter-Airlines	68 F1 score
Hourrane 2019	2022	Ridge classifier	Sentiment 140	76.84
Alzyout 2021	2021	SVM	Twitter US Airline Sentiment	78.25
Proposed	**2023**	**LR + Linear SVC + Bernoulli NB**	**Sentiment140**	**83%, 82%, 80%**

Table 4 provides a detailed overview of each model's performance, while Table 5 evaluates the suggested research in the context of existing works.

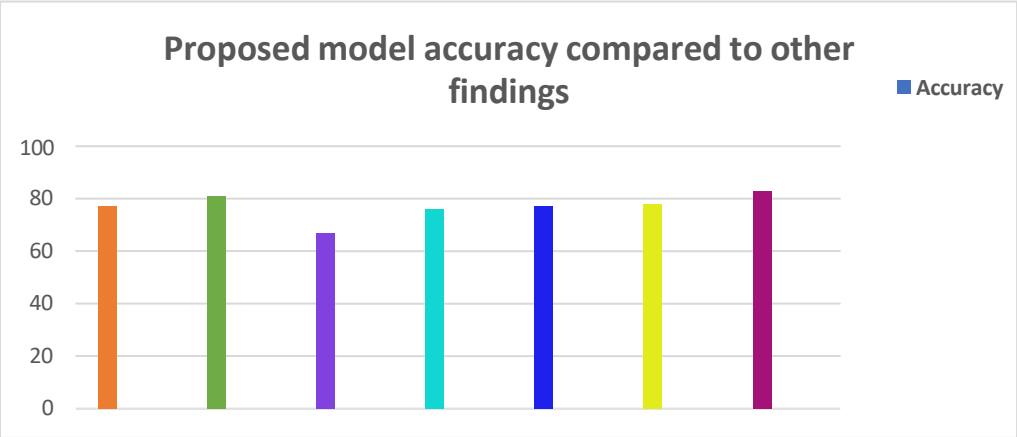


Figure 18 Comparison of the Proposed Model's Accuracy with the Results of Other Studies. (Source: Athindran 2018; Vanaja 2018; Rathi 2018; Wongkar 2019; Hourrane 2019; Alzyout 2021).

6. Conclusion

Sentiment analysis on Twitter is part of the broader field of text and opinion mining. The purpose of this research was to explain why different machine-learning methods yielded different results when applied to Twitter data for sentiment analysis. In this paper, we offer a method for extracting Twitter sentiment analysis by first balancing and scoring a sample of tweets, and then classifying them into different categories using machine learning classifiers. Logistic Regression, Support Vector Machine and Bernoulli Naive Bayes are some of the classifiers utilised here. This method is fine-tuned with a dataset derived from Twitter that can be accessed through the NLTK corpus resources.

Fig. 2 depicts the suggested framework, which consists of four distinct phases: tweet crawling, sentiment analysis engine (SAE), decision support system (DSS), and decision making. From the constructed sentiment analysis, the framework can provide a variety of alternate decisions. There is no need for the decision maker to manually enter the criterion value. The sum of the evaluations obtained by SAE analysis can be used to determine criteria values. Our strategy will be enhanced by exploring the use of bigrams and trigrams shortly. We also plan to look into other forms of deep learning, such as recurrent neural networks, multilayer perceptron's, and deep neural networks.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Alzyout M, Bashabsheh EA, Najadat H, Alaiad A (2021) Sentiment Analysis of Arabic Tweets about Violence Against Women using Machine Learning. In Proceedings of the 2021 IEEE 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain, 24–26, pp. 171–176.

Athindran NS, Manikandaraj S, Kamaleshwar R (2018) Comparative analysis of customer sentiments on competing brands using hybrid model approach. In Proceedings of the 2018 IEEE 3rd International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 15–16, pp. 348–353.

Avery Ching, Sergey Edunov, Maja Kabiljo, Dionysios Logothetis Sambavi Muthukrishnan (2015) “One Trillion Edges: Graph Processing at FacebookScale”, Proceedings of the VLDB Endowment 8:1804-1815.

Daniati E, Utama H (2020) Decision Making Framework Based on Sentiment Analysis in Twitter Using SAW and Machine Learning Approach, in 2020 3rd International Conference on Information and Communications Technology (ICOIACT), pp. 218–222.

Daniati E, Utama H (2021) TOPSIS in Decision-Making Framework Based on Twitter Sentiment Analysis, 2021 4th International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, pp. 268-273. DOI: 10.1109/ICOIACT53268.2021.9564015.

Erich Schubert, Michael Weiler, Hans-Peter Kriegel (2014) SigniTrend: Scalable Detection of Emerging Topics in Textual Streams by Hashed Significance Thresholds, Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp: 871-880.

Guo X, J Li (2019) A Novel Twitter Sentiment Analysis Model with Baseline Correlation for Financial Market Prediction with Improved Efficiency, 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, 2019, pp. 472-477. DOI: 10.1109/SNAMS.2019.8931720.

Haifei Huang, Jianxin Li, Richong Zhang, Weiren Yu, Wuyang Ju (2015) LiveIndex: A distributed online index system for temporal microblog data, IEEE 17th International Conference on High-Performance Computing and Communications (HPCC), pp. 884-887.

Hourrane O, Idrissi N (2019) Sentiment Classification on Movie Reviews and Twitter: An Experimental Study of Supervised Learning Models. In Proceedings of the 2019 IEEE 1st International Conference on Smart Systems and Data Science (ICSSD), Rabat, Morocco, 3–4:1–6.

Islam A, Chang K (2021) Real-Time AI-Based Informational Decision-Making Support System Utilizing Dynamic Text Sources. Appl Sci 11:6237. DOI: 10.3390/app11136237.

Jawa V, Hasija V (2015) A Sentiment and Interest Based Approach for Product Recommendation, 2015 17th UKSim-AMSS International Conference on Modelling and Simulation (UKSim), Cambridge, UK, 2015, pp. 75-80. DOI: 10.1109/UKSim.26

Kavitha G, Saveen B, Imtiaz N (2018) Discovering Public Opinions by Performing Sentimental Analysis on Real Time Twitter Data, 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), Kottayam, India, 2018, pp. 1-4. DOI: 10.1109/ICCSDET.2018.8821105

Najmi E, Hashmi K, Malik Z (2015) CAPRA: a comprehensive approach to product ranking using customer reviews. Computing 97:843–867. DOI: 10.1007/s00607-015-0439-8.

Oppong, Stephen, Asamoah, Dominic, Oppong, Emmanuel, Lamptey, Derrick (2019). Business Decision Support System based on Sentiment Analysis. International Journal of Information Engineering and Electronic Business 11:36-49. DOI: 10.5815/ijieeb.2019.01.05.

Pai AR, Prince M, Prasannakumar CV (2022) Real-Time Twitter Sentiment Analytics and Visualization Using Vader, 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1-4. DOI: 10.1109/CONIT55038.2022.9848043.

Prabhakar E, Santhosh M, Krishnan AH, Kumar T, Sudhakar R (2019) Sentiment analysis of US Airline Twitter data using new AdaBoost approach. Int J Eng Res Technol (IJERT) 7:1–6.

Rathi M, Malik A, Varshney D, Sharma R (2018) Mendiratta, S. Sentiment analysis of tweets using machine learning approach. In Proceedings of the 2018 IEEE Eleventh International Conference on Contemporary Computing (IC3), Noida, India, 2–4:1–3

Ruz GA, Henríquez PA, Mascareño A (2020) Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers, Futur Gener Comput Syst 106:92–104.

Sedef Çalı, Şebnem Yılmaz Balaman (2019) Improved decisions for marketing, supply and purchasing: Mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment, Computers & Industrial Engineering 129:315-332. DOI: 10.1016/j.cie.2019.01.051

Talpada H, Halgamuge MN, Tran Quoc Vinh N (2019) An Analysis on Use of Deep Learning and Lexical-Semantic Based Sentiment Analysis Method on Twitter Data to Understand the Demographic Trend of Telemedicine, 2019 11th International Conference on Knowledge and Systems Engineering (KSE), Da Nang, Vietnam, pp. 1-9. DOI: 10.1109/KSE.2019.8919363.

Vanaja S, Belwal M (2018) Aspect-level sentiment analysis on e-commerce data. In Proceedings of the 2018 IEEE International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 11–12:1275–1279.

Wongkar M, Angdresey A (2019) Sentiment analysis using Naive Bayes Algorithm of the data crawler: Twitter. In Proceedings of the 2019 IEEE Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, 16–17:1–5.

Zhongyu Lu, Weiren Yu, Richong Zhang, Jianxin Li, Hua Wei (2015) Discovering Event Evolution Chain in Microblog, IEEE 17th International Conference on High-Performance Computing and Communications (HPCC), pp. 635-640.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0103

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

IoT system architecture for monitoring and analyzing public transport data

Ihor Zakutynskyi[a] | **Ihor Rabodzei**[b]

[a] National Aviation University, Kyiv, Ukraine, Radio Electronic Devices and Systems Department, Faculty of Air Navigation, Electronics and Telecommunications.
[b] National Aviation University, Kyiv, Ukraine, Department of Information Technology Security, Faculty of Cyber Security, Computer and Software Engineering.

Abstract The crucial role of public transport in the economy and modern city development is widely acknowledged. However, contemporary public transport systems are confronted with several challenges that require resolution. These challenges include real-time monitoring, data management, passengers flow optimization, and road accident prediction. The Internet of Things (IoT) presents a promising avenue for the development of modern public transport management systems. By leveraging a diverse range of technologies, such as sensors, edge devices, cloud computing, and various communication infrastructures, IoT systems can enable the creation of robust and automated public transport infrastructure. In this research, we propose an architecture for a public transport monitoring and management system. The proposed system can effectively collect, store, and monitor public transport data, thereby facilitating efficient management and timely response to risks that arise in the transport system. Implementation of such a system would enable public transport administrators to monitor and swiftly respond to potential problems such as overloaded passenger flows, emergencies, and technical issues with transport systems.

Keywords: IoT, ITS, system architecture, monitoring, NB-IoT

1. Introduction

Every year, the urban population is growing, both in Ukraine and around the world (World economic forum 2022), which leads to a significant increase in the number of vehicles on the roads. The rapid growth in the number of private vehicles causes traffic jams, deteriorating air quality, an increase in the number of road accidents, etc. The quality of public transportation often leaves much to be desired, due to problems with suboptimal routes, non-compliance with schedules, and overcrowding. The introduction of modern information technologies, in particular the Internet of Things (Kyriazis and Varvarigou 2013), can positively affect the development of public transport systems and make them more modern and attractive to the urban population.

The concept of the Internet of Things was proposed by Kevin Ashton in 1999 to describe a network of physical objects that are connected to the Internet (Kevin Ashton 2015). Since then, physical devices connected to the Internet have become part of everyday life.

The application of the IoT concept to transportation systems, including public transportation systems, is particularly promising (Sherly and Somasundareswari 2015).

Modern vehicles are equipped with thousands of sensors that allow recording information on movement, technical conditions, fuel consumption, and the number of passengers, and detecting anomalies in real-time. IoT platforms allow receiving and transmitting data to Cloud systems, which in turn allows further analysing and building predictive models based on the so-called Big Data.

The analysis of the accumulated data can be applied (Aceves and Aceves 2002; Ryley et al 2022) to solve such problems as:

1) Vehicle monitoring/tracking.
2) Passenger flow forecasting.
3) Reducing the number of road accidents.
4) Prediction and prevention of potential technical malfunctions.

The purpose of this study is to build optimal system architecture for collecting and analysing public transportation data. The system should be easily integrated into existing systems and should not require additional infrastructure. The proposed architecture should implement the following functionality:

- Provide passengers with real-time information on arrival times and routes.
- Warn drivers of potentially dangerous situations in time to avoid accidents.

- Control traffic according to a road or weather conditions.
- Monitor the number of passengers.
- Controlling vehicle speed and compliance with traffic rules.
- Monitoring the technical condition of the vehicle.
- Possibility of third-party integrations (Providing a public API).

2. Literature Review

The Internet of Things (IoT) is a system of interconnected devices that share data over the Internet. IoT has become an important technology in various industries including manufacturing, healthcare, agriculture, and smart cities. IoT is used to monitor and analyse various aspects of transportation, including vehicles, drivers, and traffic. IoT systems rely on a complex architecture that includes various components such as sensors, drives, gateways, and cloud platforms. This literature review focuses on IoT architectures for monitoring and analysing the transport sector.

The authors of this article (Radonjić 2022) propose a complex IoT architecture for determining the status of rotating machines by sound signals. Based on the information provided by this system, preventive maintenance can be more reliably planned for a specific rotary machine. This system can be used in any industrial plant with fixed rotational frequency machines.

The authors (Majid Moazzami et al 2021) propose an IoT-based smart transport system architecture that includes various components. The system is designed to monitor and analyse traffic flow, vehicle location, and driver behaviour. The proposed architecture provides real-time traffic management and reduces congestion.

In this paper (Surachet Sangkhapan 2021), the researchers studied the problem of road accidents caused by public buses and propose a smart bus management system based on NB-IoT.

The authors (Amara Aditya et al 2023) propose IoT architecture for a smart parking system that includes various components including sensors, actuators, and cloud platforms. The system is designed to monitor and analyse the availability of parking spaces and provide parking instructions to drivers in real-time. The proposed architecture reduces congestion and improves parking efficiency.

This paper (Kavitha 2022) presents an architecture of network-based congestion control in a WSN-based IoT system and a mechanism for data transmission in an intelligent transportation system.

In general, IoT systems are becoming increasingly important in the transportation sector, IoT architectures are used to monitor and analyse various aspects of transportation. The proposed architectures provide real-time traffic management, reduce congestion, improve parking efficiency, and reduce fuel consumption.

In this article, we will propose the architecture of the public transport monitoring and management system.

3. Materials and Methods

This study uses experimental and analytical approaches.

Before designing the systems, a technical task was formulated, which described the functionality that the system should implement, as well as its technical characteristics.

The main requirements were:

A. Connectivity

The architecture should have a reliable and secure communication infrastructure that enables seamless connectivity between the public transport vehicles, infrastructure, and passengers.

B. Data Management

The architecture should be capable of collecting, processing, and analysing large amounts of data from various sources such as sensors, GPS, and passenger information systems. This data can then be used to optimize routes, schedules, and maintenance schedules.

C. Real-time monitoring and control

The architecture should support real-time monitoring and control of public transport vehicles, enabling operators to identify and respond to issues quickly. This helps to improve safety, efficiency, and passenger experience.

D. Interoperability

The architecture should support interoperability with existing and future systems, enabling integration with other transportation systems such as ride-sharing and bike-sharing services.

E. Security

The architecture should be designed with security in mind, incorporating measures to prevent unauthorized access to data and systems.

F. Scalability

The architecture should be scalable to accommodate future growth and changes in technology.

In the process of system development, the initial step involved performing a design of the system architecture. In the next step, researchers selected basic methods of connection and communication. The fundamental technologies and communication protocols between components and services were analysed, and performance tests were conducted to facilitate decision-making. Also, the input data types and data structures were analysed, which facilitated the selection of optimal storage and databases. Various solutions and methods such as linear methods and neural networks were evaluated for data processing.

Also, best practices for microservices-based IoT software architecture were analysed using next practices:

1) Performance testing to see if the software can handle large amounts of data.
2) Security testing to check whether the software meets the requirements for data security and protection against malicious attacks.

4. System Architecture

In this paper, we propose a system for monitoring and managing public transport data based on a classic IoT architecture. The system can be divided into two parts – Hardware (Vehicle module) and Cloud. The Hardware part is an IoT device that is installed in the vehicle and performs the task of collecting and transmitting data. The cloud part is responsible for storing, aggregating, and analysing the accumulated data.

The above model (Figure 1) allows collect large amounts of well-structured chronological data, which is used to train neural network models to predict the number of passengers on transport routes, as well as models to predict the likelihood of an accident. The application of neural networks represents a potent predictive technique that facilitates the generation of forecasts grounded on intricate interdependencies.

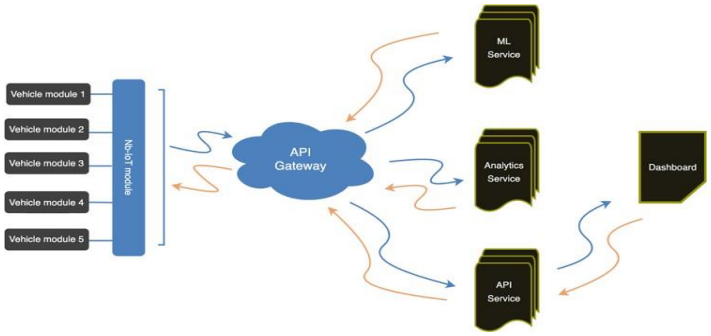


Figure 1 General system structure.

4.1. Evaluation

4.1.1. Hardware module

A hardware module is installed directly in the vehicle and collects and transmits data via the cellular network and consists of the next components:

A. GPS module

This module would allow the transport system to track the location of each vehicle in real-time, which could help optimize routes and schedules, and provide passengers with accurate information about arrival times.

B. Communication module

A communication module would allow the transport system to transmit data between vehicles, transit stations, and other systems, such as traffic management systems or emergency services. This could help improve coordination and response times.

C. Sensor module

A sensor module could include various sensors, such as temperature, humidity, and air quality sensors, which would help monitor the conditions inside the vehicle and at transit stations. This could help improve passenger comfort and safety, as well as identify areas for improvement in the system.

D. Payment module

A payment module could allow passengers to pay for their fares using contactless payment methods, such as credit cards or mobile payment apps. This could help reduce lines and wait times at ticketing kiosks and provide a more convenient and seamless experience for passengers.

E. Driver assistance module

A driver assistance module could include features such as collision detection, lane departure warning, and adaptive cruise control, which would help improve driver safety and reduce the risk of accidents.

Information from physical devices is transmitted in real-time via cellular networks using the MQTT protocol via NB-IoT connection (Figure 2).

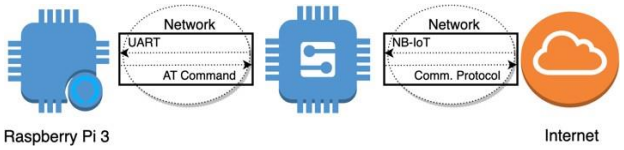


Figure 2 Data transmitting diagram.

There are various reasons why MQTT (Message Queuing Telemetry Transport) can be an appropriate protocol for smart public transport systems using IoT technology.

F. Lightweight

MQTT is a lightweight protocol that requires minimal bandwidth and is well-suited for low-power, resource-constrained IoT devices. This makes it an ideal choice for IoT smart public transport systems that rely on battery-powered devices.

G. Low latency

MQTT supports low latency, enabling real-time monitoring and tracking of public transport vehicles. This is essential for ensuring that public transport systems are running efficiently and on time.

H. Scalable

MQTT is highly scalable and can handle a large number of IoT devices, making it suitable for smart public transport systems that may have thousands of vehicles that need to be monitored and managed.

I. Reliable

MQTT has built-in mechanisms to ensure reliable message delivery, even in unreliable network conditions. This is critical for public transport systems where reliable communication is essential for the safety and efficiency of the system.

The utilization of MQTT in an IoT-based smart public transport system can enhance the system's dependability, efficiency, and safety by enabling real-time monitoring and tracking of public transport vehicles while minimizing bandwidth and power consumption.

Later, after the data is saved, microservice communication takes place via the HTTP protocol.

In this study, the Vehicle module is implemented based on a Raspberry Pi 2 model B with SIM7020E NB-IoT HAT module (Figure 3).

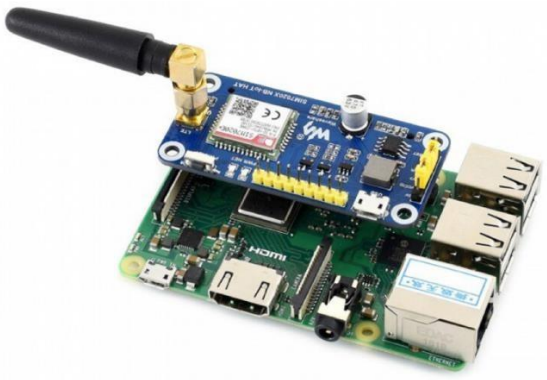


Figure 3 Raspberry Pi 2 model B with SIM7020E NB-IoT HAT module.

The Narrow Band-Internet of Things (NB-IoT) HAT is compatible with Raspberry Pi and can be operated using serial AT commands. It facilitates communication protocols such as LWM2M/COAP/MQTT, among others (Table 1).

Table 1 SIM7020E NB-IoT Specification.

SIM card	NB-IoT specific card
Protocols	LWM2M/COAP/MQTT/TCP/UDP/HTTP/ HTTPS
Control	AT commands (V.25TER, 3GPP TS 27.007, and SIMCOM AT Commands)
Band	FDD-LTE B1/B3/B5/B8/B20/B28
Data rate	Uplink≤62.5Kbps Downlink≤26.15Kbps
Baudrate	300bps~921600bps (115200bps by default)
Voltage	5V/3.3V
Current	Single module current (VBAT=3.3V): Idle mode: 5.6mA Sleep mode: 0.4mA PSM mode: 5uA eDRX mode: 70uA (eDRX=655.36s)
Dimension	30.5mm x 65.0mm

NB-IoT can be a suitable technology for a smart public transport system for several reasons:

A. Wide coverage

NB-IoT has better coverage than traditional cellular networks and can provide connectivity even in areas with weak network coverage. This is particularly useful for public transport systems that operate in remote or rural areas where cellular coverage may be weak.

B. Low power consumption

NB-IoT devices require low power, which means that they can operate on battery for extended periods, making them ideal for IoT applications like public transport systems where devices may not always be connected to a power source.

C. Low latency

NB-IoT has a low latency, meaning that data can be transmitted quickly, which is crucial for real-time tracking and monitoring of public transport vehicles.

D. Cost-effective

NB-IoT technology is cost-effective, which makes it an affordable solution for public transport systems that may have limited budgets.

Overall, the use of NB-IoT technology in a smart public transport system can improve the efficiency, safety, and convenience of the system while reducing costs and increasing reliability.

4.1.2. MQTT Broker

The MQTT Broker acts as an intermediary entity, receiving messages from Vehicle modules (i.e., publishers) and delivering them to clients who have subscribed to receive such information (i.e., subscribers).

4.1.3. API Gateway

The API Gateway service functions as a means for storing and retrieving data. Serving as a central point of access, it facilitates entry to stored data, analytics, and neural network-derived forecasting outcomes. Additionally, the service offers a Public API, which enables integration with third-party systems.

4.1.4. ML (Machine learning) Service

ML service used for building neural network models based on stored data. Machine learning techniques can enhance the efficiency and efficacy of smart public transport systems via several mechanisms.

Predictive maintenance constitutes a pertinent approach, wherein machine learning algorithms predict maintenance requirements for public transport vehicles before they fail, reducing downtime and enhancing system reliability.

Real-time routing and scheduling is another mechanism, wherein machine learning algorithms analyse real-time traffic data, weather conditions, and passenger demand to optimize the routing and scheduling of public transport vehicles. Such optimization can reduce travel time, enhancing the overall passenger experience.

Demand prediction is another application, where machine learning algorithms predict public transport demand based on historical data, weather conditions, and city events, enabling transport authorities to adjust service frequency and avoid overcrowding or empty vehicles.

Machine learning can also be used to optimize fares by analysing fare data from public transport systems, the adjusting pricing based on demand and supply, time of day, and other factors to increase revenue. Finally, passenger behavior analysis using machine learning can enable transport authorities to analyse travel patterns and preferences of passengers, facilitating the design of improved services that cater to passenger needs and enhancing the overall experience.

In this study, researchers used AWS services to build, train and deploy models. The general architecture is shown in Figure 4 (Amazon Web Services 2022).

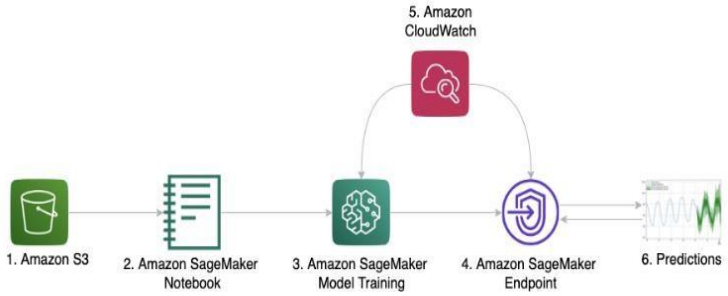


Figure 4 ML Service training process.

4.1.5. Analytics Service

Service involves utilizing data-driven techniques and tools to gather, process, and analyse data from various sources, such as vehicles, ticketing systems, and passenger feedback. The primary goal is to extract valuable insights and intelligence to optimize the system's performance, improve passenger experience, and enhance operational efficiency.

4.1.6. Management Dashboard

Control panel and data visualization. Allows real-time tracking of the location of vehicles, number of passengers, technical condition, etc. This dashboard consists of the following parts:

Login page: This is the first page users see when they want to access the dashboard. It has a form for entering user credentials such as username and password and a submit button to submit the form (Figure 5).

Dashboard Home Page (Figure 6): This page allows select the desired transport unit and gets the necessary information on it. It contains such information as the current number of passengers, aggregated number of passengers (by hours), geolocation, etc.

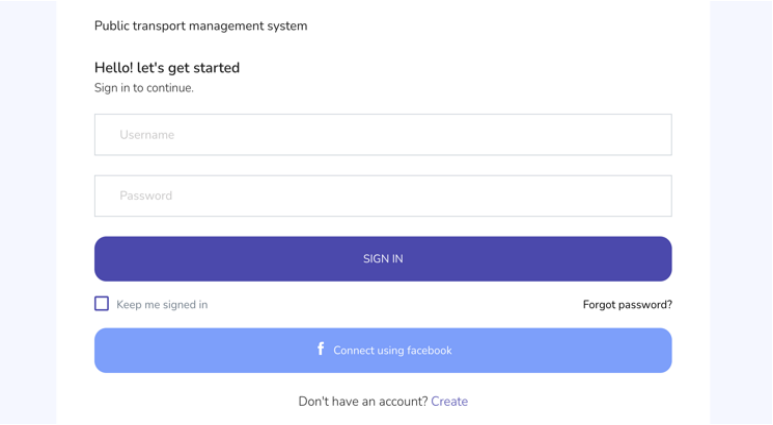


Figure 5 Public transport management system - Login page form.

Settings Page: This page allows users to manage buses in the system. It will include features such as adding new buses to the system and routes, editing bus information, and deleting buses. Users can also view the status of each bus, such as its location and whether it is running or not.

Analytics Page: This page provide for users performance reports of the system, including critical metrics like bus utilization, ridership, and revenue. Users can apply filters and export the reports to diverse formats like PDF or CSV.

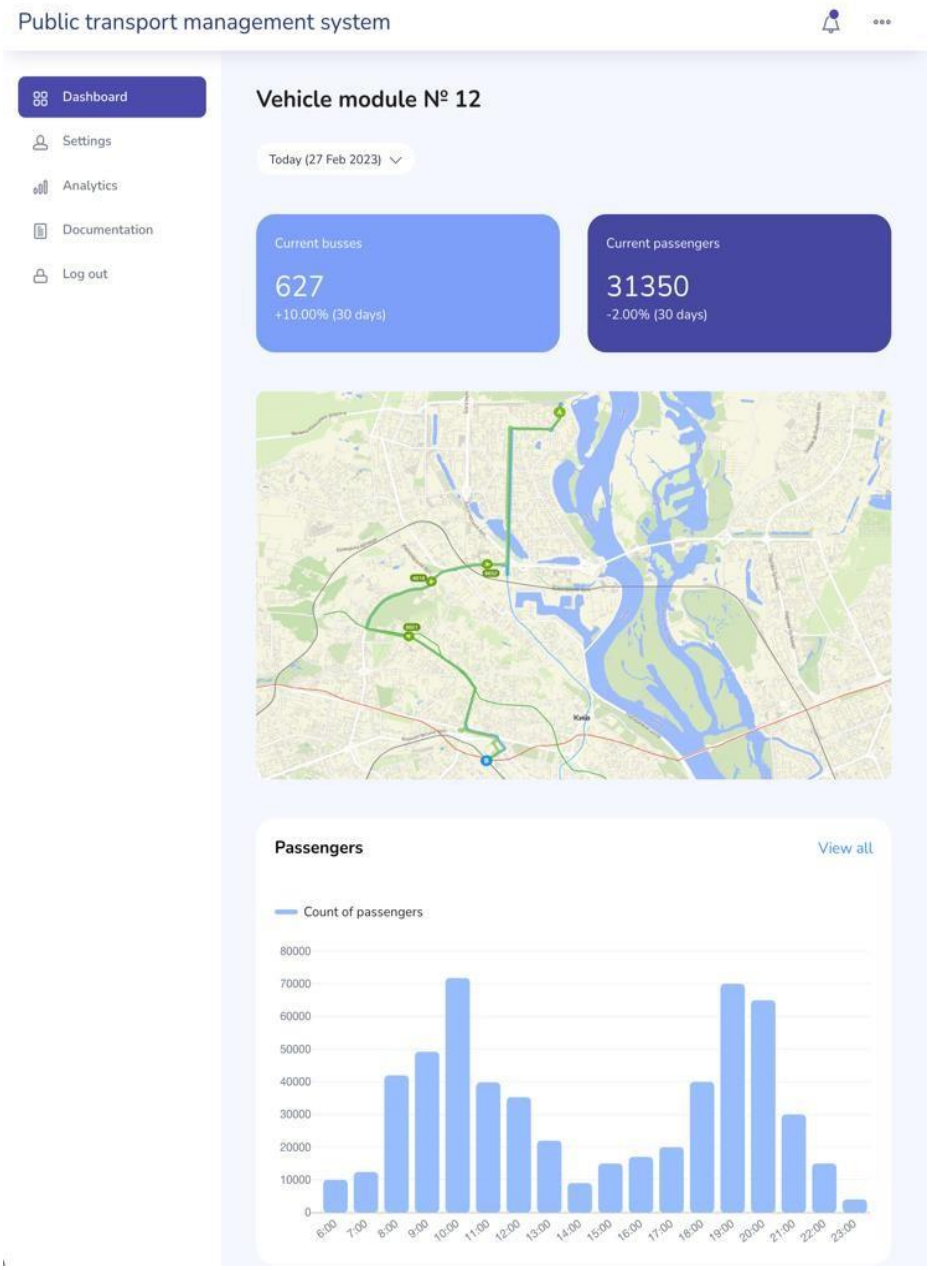


Figure 6 Public transport management system - Dashboard page.

The deployment of the server infrastructure is fast and can be realized with the help of such systems as Terraform or CloudFormation.

5. Discussion

The proposed system architecture has the potential to provide numerous benefits for transportation operators, passengers, and city planners. A data management system can provide real-time information on the location and timing of vehicles. This can help passengers plan their trips more efficiently, reducing wait times and increasing overall satisfaction. Also such systems can improve efficiency by collecting and analysing data, transportation operators can identify patterns and optimize routes to reduce congestion and improve service quality. With data on passenger demand, transportation operators can allocate resources more efficiently, ensuring that there are enough vehicles on the road to meet demand without wasting resources. Also data management system can monitor vehicle conditions and identify potential safety hazards before they become a problem, ensuring the safety of passengers and drivers alike. By optimizing routes and resource allocation, transportation operators can reduce costs and increase profitability.

Overall, proposed system can help improve the efficiency, safety, and sustainability of public transportation, benefiting both transportation operators and passengers.

6. Conclusion

The proposed architecture allows for the rapid realization of efficient collection, monitoring, and analysis of data generated in public transport systems.

The implementation of such a system will allow public transport administrations to track and therefore respond in a timely manner to potential problems, such as overloaded passenger flows, emergencies, problems with the technical condition of transport systems, etc. It also makes it possible to provide passengers with information about the time of arrival and vehicle load in real-time, which will allow them to plan their time more accurately and thus make public transport more attractive. By collecting and analysing traffic data, smart public transport systems can optimize traffic flow and reduce congestion in urban areas. This reduction in traffic congestion also reduces greenhouse gas emissions, improving air quality and contributing to sustainable development.

The system uses publicly available cellular networks to transmit information, which makes it easy to integrate into the city's existing infrastructure.

Acknowledgment

This paper and the research behind it would not have been possible without the exceptional support of the Faculty of Air-navigation, Electronics, and Telecommunications and the Faculty of Cyber Security, Computer and Software Engineering of National Aviation University.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Aceves SM, Paddack EE (2002) Developing intelligent transportation systems in an integrated systems analysis environmet. World Congress on Intelligent Transportation Systems, pp. 1– 10.

Amara Aditya, Shahina Anwarul, Rohit Tanwar, Sri Krishna Vamsi Koneru (2023) An IoT assisted Intelligent Parking System (IPS) for Smart Cities.

Amazon Web Services (2022) Deep demand forecasting with Amazon SageMaker. Amazon Web Services. Available in: https://aws.amazon.com/blogs/machine-learning/deep-demand-forecasting-with-amazon-sagemaker/

Kavitha T, Pandeeswari N, Shobana R, Vinothini VR, Sakthisudhan K, Jeyam A, Jasmine Gnana Mala A (2022) Data congestion control framework in Wireless Sensor Network in IoT enabled intelligent transportation system

Kevin Ashton (2015). That 'Internet of Things' Thing, Kevin Ashton, RFID Journal.

Kyriazis D, Varvarigou T (2013) Smart, autonomous and reliable Internet of Things. Procedia Comput Sci 21:442–448.

Majid Moazzami, Niloufar Sheini-Shahvand, Ersan Kabalci (2021) Internet of Things Architecture for Intelligent Transportation Systems in a Smart City.

Radonjić, Milutin, Sanja Vujnović, Aleksandra Krstić, Žarko Zečević (2022) IoT System for Detecting the Condition of Rotating Machines Based on Acoustic Signals. Applied Sciences 12.

Ryley, Tim, Jonathan Burchell, Lisa Davison (2013) Research in Transportation Business & Management.

Sherly J, Somasundareswari D (2015) Internet of things based smart transportation systems. Int Res J Eng Technol 2:1207–1210.

Surachet Sangkhapan, Panita Wannapiroon, Prachayanun Nilsook (2021) Development of Smart Bus Management System Using NB-IoT.

World Economic Forum (2022) How has the world's urban population changed from 1950 to 2020?

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August xx, 2023
https://doi.org/10.31893/multiscience.2023ss0104

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

A Comprehensive Approach for Assessing the Reliability of Complex Networks using OANN Approach

Anuradha Taluja[a] | **A. K. Solanki**[b] | **Harish Kumar**[c]

[a]Ajay Kumar Garg Engineering College, Ghaziabad, India, department of Computer Science.
[b]Bundelkhand Institute of Engineering Technology, Jhansi India, department of Computer Science.
[c]HR Institute of Engineering and Technology, Ghaziabad, India, Department of Computer Science.

Abstract The optimistic model of network reliability assumes the nodes to be trustworthy while assigning a failure probability to the network's connections. For calculating a network's reliability, the persistent condition is used to guarantee that no two edges have the same failure probability. Accurately estimating the network's reliability across all of its endpoints is an NP-hard problem. The study of network reliability centers on problems arising from topological isolation of individual data nodes. Analysis of network reliability spans the phases of building, deploying, and testing a network. Issues with connection, capacity, and trip time are all metrics used in analyses of computer network reliability. In this paper, an Optimal ANN strategy is proposed as a means of assessing the reliability of the network. In this research, we analyze the methods used in and findings from recent studies on trustworthiness. The term "Optimized Artificial Neural Network" (OANN) is used to describe a strategy that takes into account aspects of both neural networks and another way for measuring trustworthiness. Results are well tested on network of 2^8nodes (mesh network) and hyper-tree network of n nodes. A network's performance may be evaluated with the help of the suggested method, which also contributes to the calculation of its cost and reliability metrics.

Keywords: ANN, ANOVA, BDD, CNN, PDF, min-batch gradient descent

1. Introduction

Computer networks have become the primary means of communication hence it is crucial that secure networks should be developed. In order to determine whether or not the computer network lives up to the expectations set out in its design, a reliability study is performed on the infrastructure and key components that make up the network. The capacity to find a design that achieves maximum reliability within cost restrictions is at the heart of the reliability optimization challenge. Reliability study of huge, complex structural systems, flow networks, computer and communication systems, coherent systems, series. Due to reliable communication networks, calls and messages should always arrive complete, securely, and in order. Combining also requires weighing constraints like price and efficiency. Another aspect of dependability is the ability to quickly recover from infrastructure or service disturbances. Network dependability may be measured in a number of ways, including the rarity of failures, the speed with which the network recovers from outages, and its overall resilience. The potential for longer-than-expected delays due to unreliable networks must be considered. Methods to improve dependability include expanding access to backup paths, strengthening individual network nodes, and improving response times to disruptions. Communication networks, sensor technologies, social networks, and other real-world systems that may be modelled as stochastic networks rely heavily on network reliability theory (Hesham et al 2021; Baes et al 2022). There are five mainstays of system dependability: security, reliability, availability, safety, and resilience. Indicators of a network's dependability or availability, which measure its capacity to carry out a certain task, may be modelled and calculated using a wide range of approaches. The reliability of a system is a numerical rating of how well it is expected to perform over a certain time frame. Most tools are constructed using algorithms that are minimum path set algorithms or minimal cut set algorithms (Negi et al 2022; Clark and Verwoerd 2012). Determining network reliance via a collection of minimum pathways or cuts is NP-hard, unfortunately. It is illogical to judge the trustworthiness of a communications network solely on the performance of hypothetically perfect nodes (Al-Zu'bi et al 2021; Sriram 2022). Real-world applications of complicated network design structures have provided a major impetus towards solving the maximum dependability challenge in computer networks. The component is an object present in the network. The component performs certain operation in complex network. The

decision making of a network designing depends upon certain measures in which calculation of network reliability is one important measure. For calculating complex and huge designing of a network, the system reliability is required. From the last four decennary reliability optimization area of research received much substantial contribution. The problem of finding network reliability of layered or structured is defined as the communication between the network nodes through some specified path. All-terminal and source-sink network reliability calculations are most common (Khanna and Chaturvedi 2018; Zhifu Luan 2017). Calculating network reliability uses several approaches. Most of these approaches are simulation or analytical, requiring substantial computer work. For smaller networks, these strategies are straightforward and successful (Chakraborty and Goyal 2016; Jaime Silva 2015; Kaushik 2013). Because communicating with every network node is the primary goal of reliability, calculating precise reliability in networks of various sizes is fairly difficult. In order to build a successful network, several elements must be in place to assure reliability, availability, and serviceability (Bhola and Soni 2021; Gupta et al 2015). An important study area has been reliability calculation, and in recent years, significant advances have been made in network architecture to decrease cost and defect. Evaluating reliability is essential to tackling many difficult real-world problems. To reduce costs and errors, this challenge requires a cost-effective, error-free design. Reliability is crucial to the design of many systems; hence reliability research measurement has made substantial contributions. Here we provide a more in-depth analysis of the review work that has been done and is being done at the moment in relation to reliability measurement. In the context of reliability evaluation of flow networks employing a subset cut and enumeration approach, has provided a number of methods which are presented for accounting for faulty nodes. Using subset cuts (with perfect nodes) to determine the legitimate additional subset cuts that imperfect nodes make is the basis of three independent articles' methods. Using the suggested decomposition technique, he also introduces a new approach to enumerating subset cuts. With this method, we hope to forestall the emergence of duplicate, higher-order subsets in the network. The number of internal linkages is what is meant by "order."

W. C. Yeh presented a low-cost MCS-RSM method for ensuring network resilience. Stochastic network system dependability evaluation is important in system planning, design, and control. Minimizing resource usage and overall cost while maintaining network stability are examples of real-world network concerns. Altiparmak (2014) presented a method for efficient optimization of all-terminal reliability of networks using an evolutionary approach. Leite da silva et al (2007) proposed a method for composite reliability estimation based on MCS & ANN. Xiao (2008) proposed an Enhanced Factoring Algorithm to calculate the system reliability. Mangey (Bisht et al 2021) suggests a productive algorithm for calculating the network's reliability indices. The authors investigate how the exponentially distributed failure rate can be used to tackle network-related issues utilising the universal generating function. Shuai Lin (2018) in a fault-propagation perspective, focuses on the assessment of complex electromechanical systems' (CEMSs) system reliability. For the evaluation of system reliability, a failure propagation model based on network theory and an enhanced polychromatic set is first suggested. The system effectiveness index is built from the perspective of the node to explore the variation in network efficiency. Gaur (2021) explains the many measures and their mathematical foundations used to assess network connectivity. In this overview research, a number of methods are listed, including State Enumeration, Sum of Disjoint Products, Minimal Cut Set, Factoring Theorem, Cellular Automata, Subset Simulation, Percolation Theory, Binary Decision Diagrams, and Universal Generating Functions. A new encoding technique for employing neural network models to evaluate the reliability of telecommunication networks with similar link reliabilities was given by Altiparmak (2009). Iterative algorithms like tabu search or simulated annealing can be utilized to optimize network designs while taking use of neural estimation's quick processing speed. The large vector length of the inputs required to express the network link architecture and the sensitivity of the neural network model to a particular system size are two major issues they had identified with earlier methods of employing neural networks to model systems.

Anuradha, Clarke R N, Khanna G states that finding reliability in diverse anecdotal networks is an NP-Hard problem (Anuradha 2020; Clarke 2014; Khanna and Chaturvedi et al., 2019). The modeling of wireless reliability differs from that of conventional networks. In compared to a wired network, reliability modelling is different (Gupta 2013; Rahman et al 2021). Most of these methods involve repeated simulations, which is computationally expensive for highly rising and variable-sized growing networks. An enhanced neural network technique improves accuracy and efficiency. The goal is to construct a cost-effective, dependable layered or structured network. The study uses iterative min-batch gradient descent neural technique highly growing network.

2. Statement of the problem

The Calculation of the network reliability or the probability of failure, of a dynamic system under a terrain condition is most important and challenging problems in reliability engineering. The primary goal of the paper is to develop an algorithm that utilizes a highly reliable layered network so that the overall system can have a low cost and fault rate. The primary goal of the paper is to develop a system design that satisfies high reliability requirements while minimizing implementation costs.

A network graph$G = (N, L, P)$ shows interconnection between nodes through links with link reliability. Network reliability is defined by:

$$Minimize\ C(X) and\ R(X) \geq R_0 \quad (1)$$

Where,

$$X = (X_{12}, \dots, X_{ij}, \dots, X_{N-1})$$
$$R_0 = (R_{01}, \dots R_{02}, \dots R_{0N})$$

Assume that g be a subgraph ($g \subseteq G$) of the network. The model DM_x is considered as Damage model assigns the probability that this subgraph occurs. The link failure model is considered with the probability that all the links (l) comprising the subgraph are in functional state and all remaining links ($L - l$) in the network are not. Now whether subgraph g, function $\emptyset(g)$which returns 1 or 0 means g is operational or not. Aim is to design a computer network by selecting a subgraph from growing networks with high reliability and minimum cost. Reliability is defined for a big size network is as follows

$$R(X) = \sum_{i=1}^{N-1} \quad \sum_{j=i+1}^{N} \quad C_{ij}X_{ij} \quad (2)$$

Hence, reliability and cost will be crucial when designing a highly scalable, efficient, and fault-tolerant network.

The challenge is to prepare a stag of complex network using ANN by comparing the results with MCS, MCS-PSO and GA methods with random and heuristic link (Nasim Nezamoddini 2015; Chakraborty and Goyal 2015; Ioannis 2015). ANNs are powerful data-driven models that have been developed that are often used for the dynamic modelling and characterization of nonlinear systems due to their flexible structure and ability to mimic any sophisticated nonlinear behavior (Shyamala SRD 2022; Sajja 2021).

Compared to previous algorithms, the proposed technique enhances reliability and reduces C(X) for mesh networks of 2^8 nodes and hyper-tree networks of n nodes. The proposed algorithm is formerly stated as:

Input used for training algorithm is gathered by improving upper bound and lower bound value of network with a special numerical value called weights which can be positive or negative. The network reliability significantly gets improved by upgrading existing method. There are circumstances that the literature work does not attempt to explore, despite the fact that analyzing the many reliability estimation techniques has discovered that certain approach may affect reliability estimate. No conclusions can be drawn about how other reliability estimators will perform under similar settings from our findings. The findings we've gotten so far are the product of an extensive set of parameter tests. Consequently, the performance network reliability estimate in terrain situations should be tested in further research. We investigate and use a number of known approaches for estimating dependability, including the Monte Carlo System, Particle Swarm Optimization, Multi-Criterion Search, and Artificial Neural Networks.

3. Methods for the evaluation of reliability:

3.1. Multiple Probability Simulation method for Reliability Estimation

Multiple Probability Simulation or Monte Carlo Simulation is a mathematical technique worked with the principle of random variable (RM_v)for modelling risk or uncertainty of a system. The RM_v or inputs are modelled on the basis of probability distributions. These Networks are from the probabilistic network. MCS work on the following principles

- MCS uses BFS technique to determine that there is a direct path available between the nodes.
- The fraction of sampled networks in which there exists a path between the two nodes is an approximation to the exact network reliability.

The exact estimation of network reliability for a stochastic network is NP-hard. Following are the recommendation while using MCS technique

All MPs/MCs of the system must be known in advance (Chakraborty and Goyal 2015).

- To reduce exact calculation, the simulation technique was introduced to estimate the system reliability of a stochastic network.
- Prior knowledge of network series is must either series or parallel.
- Various distribution methods are applied on the network components.

MCS algorithm is one of the more efficient, optimal approaches for estimating network reliability.

Assumptions

1. Network arc is perfectly reliable.
2. Network node, and the system is either operative or have failed.
3. Failure is totally stochastically independent.
4. Maintenance of node and arc is not considered.

Algorithm 1: MCS Algorithm for Reliability Estimation.

MCS Algorithm for Reliability Estimation for variable size network

Step 1: Start and Prepare the data Set for Simulation.
Step 2: Simulate all the element based on to its lifetime distribution
Step 3: Set Life time for link *l* and for node n with

$$LF_l(t) and\ LF_n(t)$$

Step 3: Calculate the life time of each variable and apply in an array LTA[N],
Step 4: Sort the Array LTA [] in ascending order.
Step 5: Record each failure instance (τ) and arrange in chronological order.
Step 5: Calculate the network lifetime reliability $R(LF_N)$ by the formula:

$$R(LF_N) = \sum \quad \tau\ at\ Time\ inerval\ (t)$$

Calculate it for *i=0 to n* instances
Step 6: Calculate the Variance of network lifetime reliability by the formula

$$V(LF_N) = \frac{LF_N(t) * (1 - LF_{N)}}{N}$$

Step7: Store the results in individual array of size N.
Step 8: Stop Algorithm.

The implemented algorithm1 determines the lifetime reliability (as shown in Figure 1) distribution on mesh and hyper tree network. The simulation results will be calculated after K trials times. The results show in Table 1, significant decrease in l $R(LF_N)$ for increasing and variable networks.

Table 1 $R(LF_N)$of MESH $Rand\ (LF_N)$of Hyper Tree.

t	$R(LF_N)$ **of MESH**	$R(LF_N)$ **of Hyper Tree**
0.10	0.9998	0.9989
0.30	0.9997	0.9997
0.40	0.9992	0.9994
0.57	0.9992	0.9992
0.72	0.9988	0.9987
0.87	0.9990	0.9990
1.02	0.9981	0.9979
1.17	0.9877	0.9975
1.32	0.9875	0.9974
1.47	0.9866	0.9966
1.62	0.9854	0.9965
1.77	0.9847	0.9963
1.92	0.9834	0.9960
2.07	0.9828	0.9958
2.22	0.2456	0.9951
2.37	0.2412	0.9947
2.52	0.2402	0.9876
2.67	0.2401	0.8901
2.82	0.1678	0.8345
2.97	0.1672	0.6789

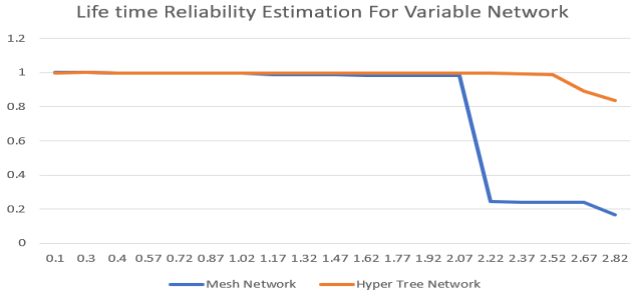


Figure 1 Life time Reliability Estimation for Variable Network.

1. *Particle Swarm Optimization (PSO) method for reliability estimation:*

PSO is an optimization technique proposed by W. Yeh, Y. Lin (Kaushik 2015).It is based on population and now it is most commonly used optimization techniques. It makes use of a number of potential solutions to the optimization issue. Each remedy is referred to as a particle, and the aggregation of these particles at each iteration step creates the swarm. It works on both deterministic and stochastic update rules along with fitness functions. PSO search for optima by updating generations (iterations). Its work on the principle of population of moving particle. The first particles are drawn stochastically toward the positions of their own previous best performance and the best previous performance of their neighbors. Table 2 shows Adjacency matrix for the graph.

$$P(V_{ik}) = +a[LB_{ik} - X_{ik}] + b[GB_{ik} - X_{ik}] \quad (6)$$

Denotes the velocity of the particles and X_{ik} denoted the position vector of the particles. LB_{ik} Denotes the local best value and GB_{ik} denotes the global best value.

$$X_{ik} = + P(V_{ik}) \quad (7)$$

How PSO solve the problem of Reliability? PSO solved problem by having a population of candidate solutions and each particles have their local known best position search space of algorithm means the range in which algorithm compute the optimum control values and if the optimum control value of any particle exceeds in the searching space, then we will re-initialize that value. Algorithm 2 gives the Algorithm for Reliability Estimation. Figure 2 shows the Arpanet Graph of 20 Nodes. The complex equations are a result of having many unknown parameters, complex equations take longer to solve, and numerical methods are no longer sufficient to solve such problems and traditional approaches were inadequate for tackling complicated issues so reliability estimation turned to the particle swarm optimization technique. PSO approach does not require prior knowledge of the precise or approximate reliability functions.

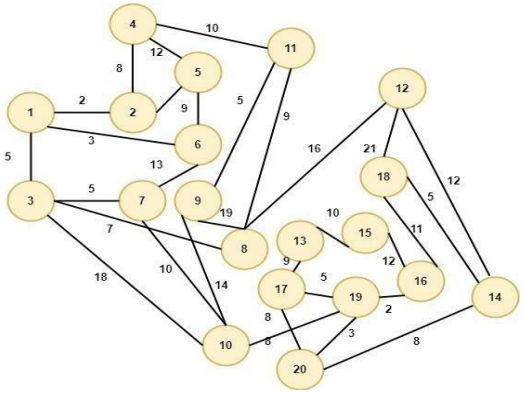


Figure 2 Arpanet Graph of 20 Nodes.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
4	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
5	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
8	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
10	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0
11	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1
15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0
17	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1
18	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0
19	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	1
20	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0
	3	3	4	3	3	3	3	4	3	4	3	3	2	3	2	3	3	3	4	3

Figure 3 Adjacency matrix for the graph shown in Figure 4.

Algorithm 2: Algorithm for Reliability Estimation

PSO Algorithm for Reliability Estimation

Step1: Select a network of N nodes.

Step 2: Assume a weight associate for the link. Weight means number of data packets delivered by the source node over the link.

Step3. Assume a transmission bandwidth and utilization of data has been used as the main parameter for calculating the reliability of the links.

Step 4: Nodes are assumed to be failure free.

Step 5: Reliability of graph is calculated by

$$P_r(Sub_i(s,t)) = \propto_s * \prod_{n} P_r(Sub_{in}(s,t))$$

$$Sub_{in} is\ the\ subset\ of\ Sub_i\ and\ n\ is\ the\ number\ of\ nodes$$

Step 6: Calculate the reliability matrix using the node-pair reliability values.

Step 7: Find the most reliable path between any ordered pair and record it,

Step 8: Record the link throughput of the network.

Step 9: Record the fuzzified values based on the relation.

Step 10: Update the pbest, gbest, velocity, and position of each particle based on function (6) and (7).

Iteration Number=iteration Number+1.

Step 11: If the maximum Iteration Number> threshold, then stop; else go to Step 2.

Computation Results:

Although it shows a lower computing cost, the best solution might not be found. When the confidence level is high (around 1), it is more difficult to reject the hypothesis testing and less likely that the best option will be abandoned. Table 3 shows deviations and no. of success and no. of failure.

Number of particles 20
Max position 1
Min position 0.6
Iteration 100
Cognitive factor 0.8
Social factor 0.8
Simulation 1000
Hypothesis testing simulation 50

Table 2 Five sets of deviations and no. of success.

	LB	UB	Deviation	Time	No of Success	No of Failure
Set1	0.921	0.931	0.004	4.5	160	12
Set2	0.920	0.930	0.005	23.5	159	450
Set3	0.921	0.930	0.004	18.0	429	70
Set4	0.908	0.930	0.010	23.8	-	-
Set5	0.929	0.930	0.000	2.2	-	-

MCS-PSO method for reliability estimation: We have proposed to combo algorithm consist of MCS and PSO to minimize the cost under reliability constraints as shown in Algorithm 3. Figure 5 shows the Reliability estimation using MCS-PSO. Combo algorithm is an integration algorithm for solving the problem of reliability estimation.

Algorithm 3: step of MCS- PSO Algorithm

MCS-PSO Combo Algorithm

Step 1: Select a Network as per step of PSO Algorithm step 1 and 2.

Step 2: Each Particle is assigned with lower bound value on reliability of component.

Step 3: Apply basic Monte Carlo Simulation used for finding system reliability $R_K(G)$.

Step 4:

If $R_K(G) \leq R(X)$

Implement for each particle.

Else

Use Initial value of component's reliability.

Step 5: The i[th] dimension value will be based on hash function $H_f = 0.35 \times (1 - X_i) + X_i$.
Step 6: Apply Monte Carlo Simulation is applied to obtain $R_K(G)$.
Step 7:Calculate the cost $C(X)$ of link which sold be minimum

$$R_K(G) \geq R(X)$$

Step 8: The value of link dimension must be equal to H_f, and go to step 2.

Strength:

1. The reliability functions are not need to be known before using the combo approach.
2. The MCS algorithm is the sole one used by this Combo algorithm to calculate the reliability for a certain particle.
3. The proposed Combo solves the challenging network reliability optimization problem while overcoming the limitations of conventional algorithms.

Weakness:

1. Required more computational time for simulation.
2. Little bit high in complexity evaluation.

Computation Results:

Table 3 Random and Heuristic value calculation at reliability value R(x).

	R(X)=0.93		R(X)=0.97		R(X)=0.99	
	R_V	H_V	R_V	H_V	R_V	H_V
1	1199.67	1171.29	1209.96	1175.76	1201.25	1180.27
2	1199.74	1171.29	1204.65	1178.65	1201.65	1183.34
3	1206.66	1173.96	1202.27	1179.96	1205.65	1185.54
4	1203.56	1171.95	1201.68	1186.45	1203.45	1186.65
5	1214.78	1171.68	1213.65	1188.54	1202.45	1181.45

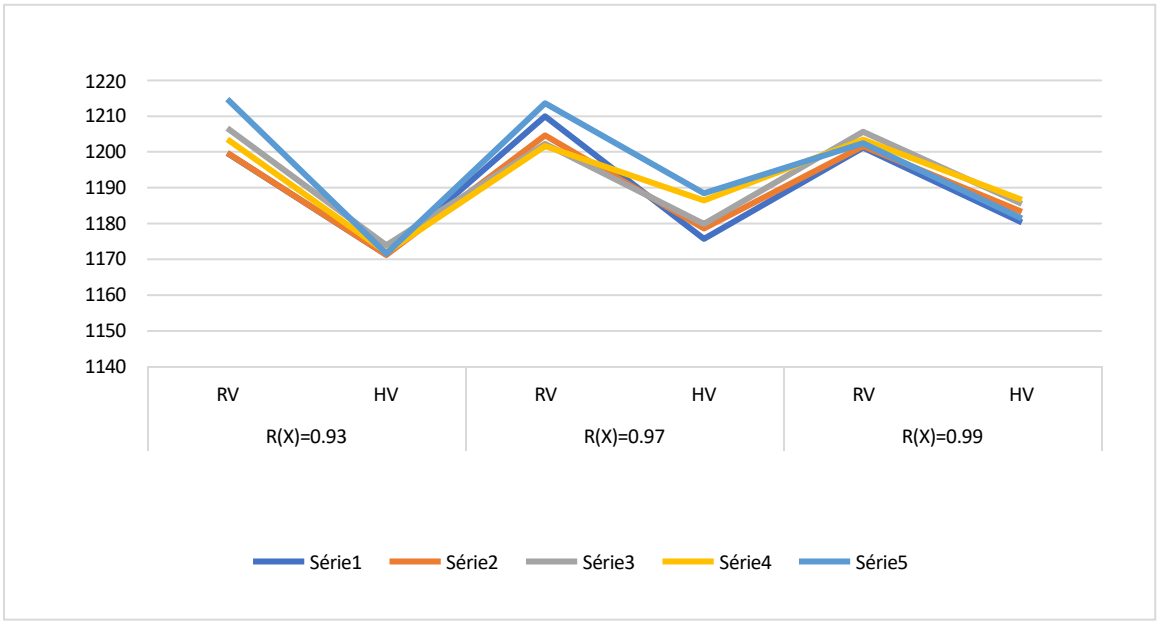


Figure 4 Reliability estimation using MCS-PSO.

2. *Meticulous Method Based ANN Approach OR OANN Approach method for reliability estimation:*

Exact evaluation of growing sized networks can be done with the help of Neural Network. Network designing is done by examine the upper-bound value and this designing is proposed in our paper. Optimization design of artificial neural network training is proposed in our previous paper [32]. Artificial Neural networks technique is inspired by the functioning of biological neuron. If the system (series/parallel) has a failure rate of each module at time t=1 of 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5, the failure rate's graph is as follows:

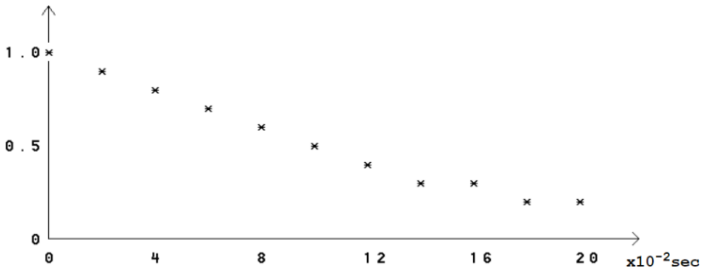


Figure 5 Shows Failure rate at time interval t.

To build an ideal design of a network, which is a variable and increasing size network, and to execute costly calculation in order to evaluate network reliability, the two approaches shown in Figure 6 are the meticulous technique and the extended meticulous method-based ANN approach, respectively. The idea of using ANN training to check the upper bound is also offered. Algorithm 4 demonstrates a method for defining models. The method is broken down into two distinct phases: configuring the experiment and determining where to make adjustments. When training an ANN, the process is repeated until the criterion is met.

Algorithm 4: Defining models:

#Defining the Model
Step 1: Design a best fitted network.
Step2: Network Input to the ANN
Step3: Input no of hidden layer to ANN

$$H_L = \gamma$$

Step 4: Input learning rate(α) to ANN

Step 2: Prepare the Orthogonal Array.
#Set up of the model and initial value for the following must be assigned
connection weights(w),
learning rate(α)
Tolerance Parameter
Calculating the loss function between predicted Actual Values
Optimizing the predicted values
Calculating the Accuracy metrics
Design an Orthogonal Array.
Set up of the neural network Parameter along parallelly with orthogonal array.
 # Setup the statistical Method technique using ANOVA method.
Some more significant parameters such as adaptive learning rate, convergence rate, network topology with high upper-bound for reliability are chosen.

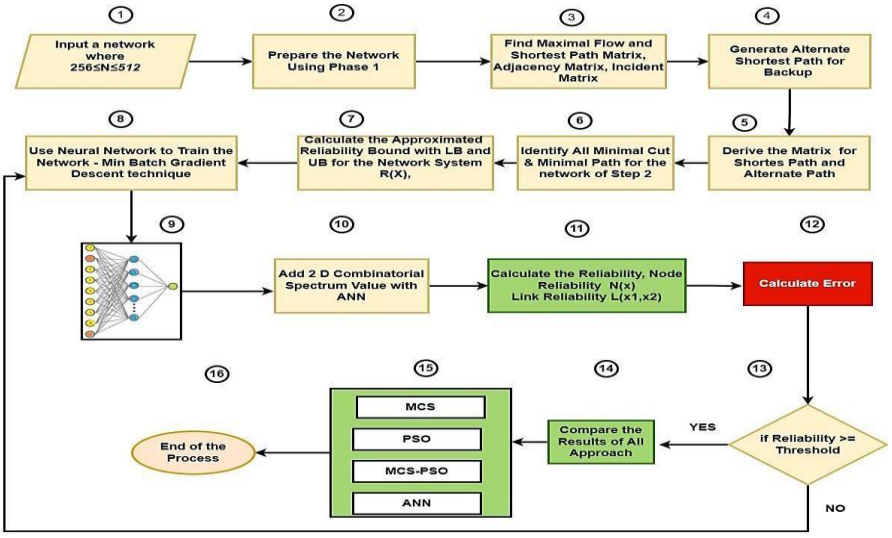


Figure 6 Extended Meticulous Method Based ANN Approach.

Over one thousand simulations are used to validate the algorithm's output. Links factors and reliability are displayed in Figures 7, 8, 9, 10, 11, and 12. Mean absolute error (MAE) and root-mean-squared error (RMSE) are measures of how much of a gap there is between Monte Carlo and ANN estimates of the network's reliability. Table 6 provides reliability estimation utilizing existing one and OANN, while table 5 shows the reliability for data sets.

Figure 4 Reliability for Data Set 1 and Set 2.

Reliability for Data Set 1					
Network	0.80	0.85	0.90	0.95	0.99
	850	850	850	850	850
Reliability for Data Set 2					
Network	0.80	0.85	0.90	0.95	0.99
	850	850	850	850	850

Fixed Link Reliability _CV Outcomes			
Data Set 1		Node = 50 & Edges = 1225	
		RMSE	
	Initial fit data	Hold out data	Upper-bound
I	0.0326	0.04201	0.07272
Ii	0.03337	0.03937	0.07034
Iii	0.03279	0.03995	0.06863
Iv	0.03465	0.03935	0.07301
V	0.03377	0.03603	0.06307
Average	**0.033436**	**0.039342**	**0.069554**
	Fixed Link Reliability Estimation		
Data Set 2		G = (50, 1225)	
		RMSE	
I	0.04505	0.05279	0.07661
Ii	0.03722	0.04815	0.08575
Iii	0.03876	0.04036	0.07011
Iv	0.03858	0.04001	0.07169
V	0.03852	0.04776	0.07796
Average	**0.039626**	**0.045814**	**0.076424**

Five-fold Cross Validation Results for Fixed Link Reliability			
	Fixed Link Reliability Estimation		
Data Set 1		G = (50, 1225)	
		RMSE	
	Initial fit data	Hold out data	Upper-bound
I	0.04774	0.06066	0.09277
II	0.04944	0.05041	0.08848
III	0.04966	0.05425	0.08439
IV	0.05056	0.05248	0.08362
V	0.04857	0.05725	0.08916
Average	**0.049194**	**0.05501**	**0.087684**
	Fixed Link Reliability Estimation		
Data Set 2		G = (50, 1225)	
		RMSE	
I	0.05152	0.06059	0.09947
Ii	0.05247	0.05788	0.09095
Iii	0.05132	0.05777	0.09979
Iv	0.05123	0.05858	0.09379
V	0.05144	0.05959	0.09813
Average	**0.051596**	**0.058882**	**0.096426**

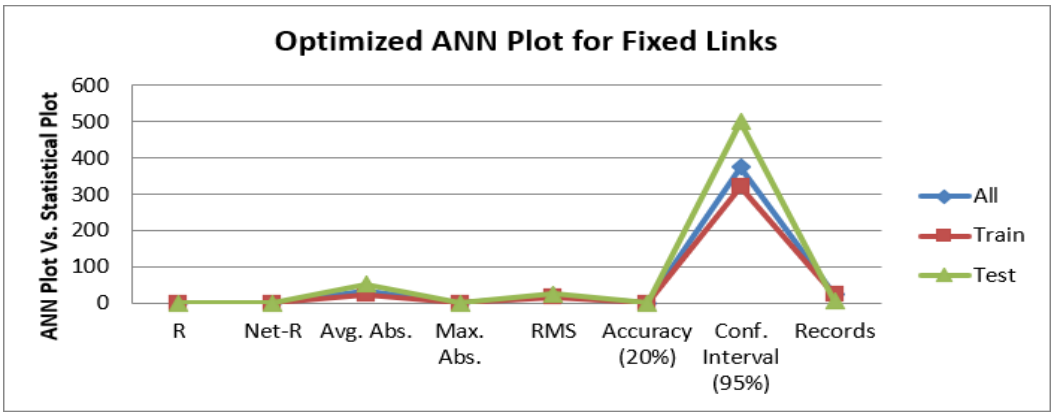


Figure 7 Optimized ANN plot for Fixed Links.

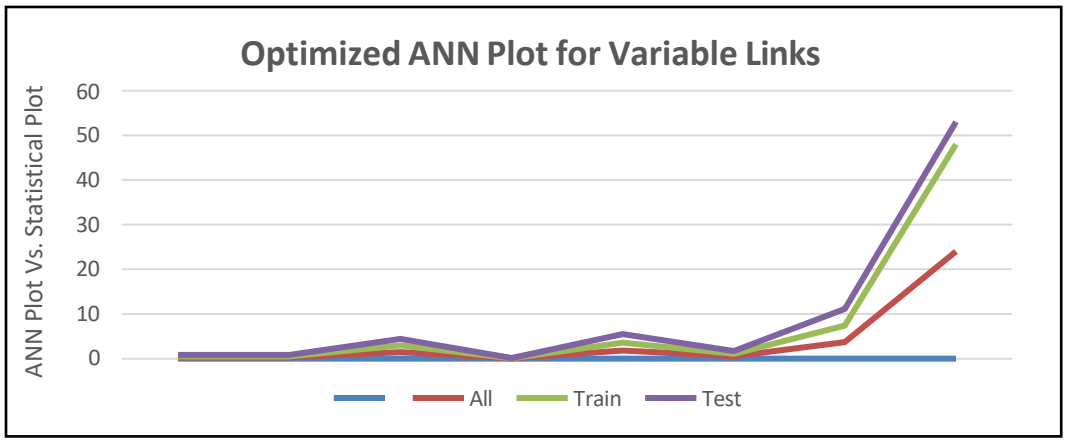


Figure 8 Optimized ANN plot for Variable Links.

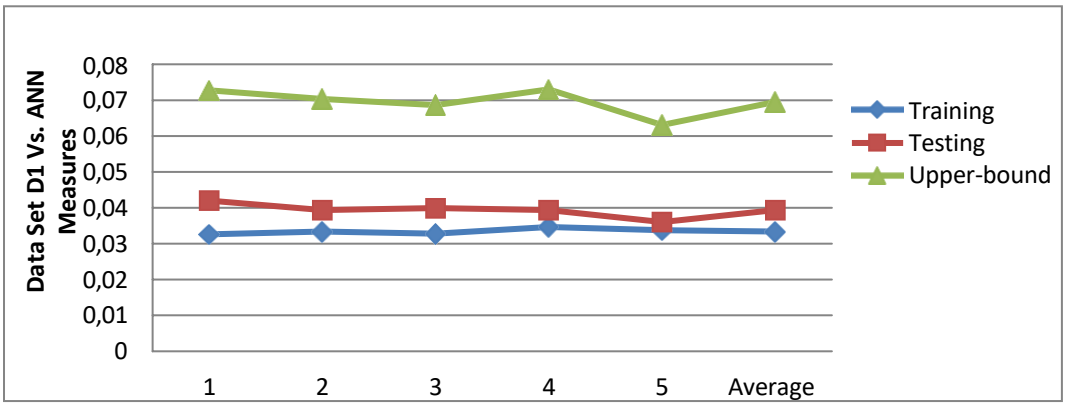


Figure 9 Fixed Link Reliability for Data Set D1.

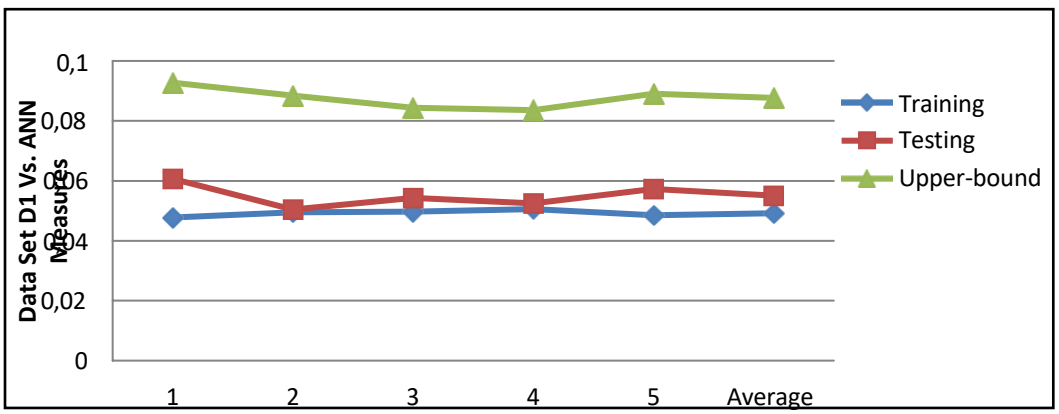


Figure 10 Extended Meticulous Method Based ANN Approach.

Table 5 Reliability Estimation using existing Algorithm (MCS, PSO).

Assessing Reliability Using a Proven Methodology (MCS, PSO)		OANN	
Reliability of MESH	**Reliability of Hyper Tree**	**Improved Reliability of Hyper Tree**	**Improved Reliability of MESH**
0.9998	0.9979	0.9999	0.9999
0.9997	0.9998	0.9998	0.9998
0.9995	0.9996	0.9997	0.9997
0.999	0.9988	0.9989	0.9989
0.9988	0.9987	0.9985	0.9985
0.9983	0.998	0.9974	0.9974
0.998	0.9975	0.9957	0.9957
0.9877	0.9972	0.9974	0.9974
0.9872	0.9971	0.9972	0.9972
0.9862	0.9969	0.9971	0.9971
0.9855	0.9966	0.9970	0.9970
0.9844	0.9963	0.9965	0.9965
0.9833	0.996	0.9960	0.9960
0.9827	0.9958	0.9955	0.9955
0.2455	0.9951	0.9964	0.9964

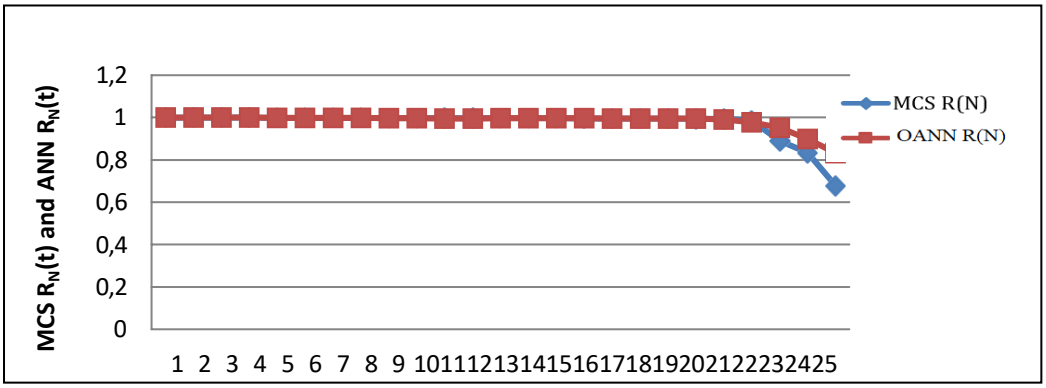


Figure 11 Reliability measures for RN (t) of Hyper-Tree Vs OANN approach for RN(t).

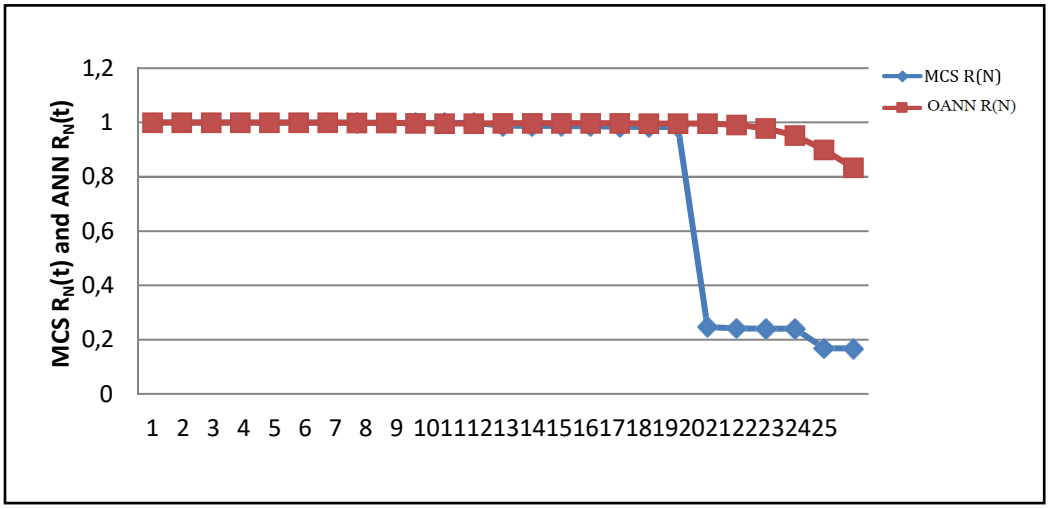


Figure 12 Reliability measures for RN (t) of Mesh Network Vs OANN approach for RN (t).

4. Conclusion

This paper presents alternative approaches for calculating reliability in growing computer networks, so that a tolerant of failure system with maximized reliability with minimum cost is constructed. This paper presents some of the alternative methods for estimating reliability computer networks such as MCS method and MCS-PSO method and results are compared with optimized neural network approach. Following conclusions states effectiveness of optimized neural network approach when it is compared with MCS-PSO and MCS methods are as follows:

1. A new Monte Carlo method known as MCS-PSO for estimating reliability in a complex network system without knowing the Minimal path and Minimal Cuts in a given graph by performing Activity on Arc and Activity on Nodes.
2. The MCS-PSO method computes length and dimensions of the links in a network system and in repeated calculations it finds expected reliability and overall reliability in a given system.
3. The MCS-PSO system model works perfectly for simple network model for evaluating reliability, but, evaluating reliability in a complex growing network requires lot of effort and time which is equivalent to NP-hard.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Altiparmak F, Dengiz B (2014) Design of reliable communication networks: A hybrid ant colony optimization algorithm IIE Transactions 42:273-287, DOI: 10.1080/07408170903039836

Altiparmak Fulya, Dengiz Berna, Smith, Alice A (2009) General Neural Network Model for Estimating Telecommunications Network Reliability. Reliability, IEEE Transactions 58:2-9.

Al-Zu'bi H, Al-Khaleel O, Shatnawi A (2021) FPGA implementation of data flow graphs for digital signal processing applications. International Journal of Communication Networks and Information Security 13:92-114.

Anuradha (2020) Calculation and Evaluation of Network Reliability using ANN Approach, Procedia Computer Science, Volume 167, 2020, pp 2153-2163, ISSN 1877-0509. DOI: 10.1016/j.procs.2020.03.265

Baes AMM, Adoptante AJM, Catilo JCA, Lucero PKL, Peralta JF, de Ocampo ALP (2022) A novel screening tool system for depressive disorders using social media and artificial neural network. International Journal of Intelligent Systems and Applications in Engineering 10:116-121.

Bhola J, Soni S (2021) Information Theory-Based Defense Mechanism Against DDOS Attacks for WSAN. Advances in VLSI, Communication, and Signal Processing. Lecture Notes in Electrical Engineering 683. DOI: 10.1007/978-981-15-6840-4_55

Bisht S, Kumar A, Goyal N, Ram M, Klochkov Y (2021) Analysis of Network Reliability Characteristics and Importance of Components in a Communication Network. Mathematics 2021:1347. DOI: 10.3390/math9121347

Chakraborty S, Goyal NK (2015) Irredundant subset cut enumeration for reliability evaluation of flow networks. IEEE Transactions on Reliability 64:1194–1202. DOI: 10.1016/j.eswa.2015.05.019

Chakraborty S, Goyal NK (2016) An Efficient Reliability Evaluation Approach for Networks with Simultaneous Multiple-Node- Pair Flow Requirements, Quality and Reliability International Engineering 33:1067-1082DOI: 10.1002/qre.2097

Clark ST, Verwoerd WS (2012) Minimal Cut Sets and the Use of Failure Modes in Metabolic Networks. Metabolites 2:567–595.

Clarke RN (2014) Expanding mobile wireless capacity: The challenges presented by technology and economics. Telecommunications Policy 38:693–708.

DOI: 10.1109/TR.2008.2011854

Gaur V (2021) A literature review on network reliability analysis and its engineering applications. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2021:167-181. doi:10.1177/1748006X20962258

Gupta P, Goyal MK, Gupta N (2015) Reliability aware load balancing algorithm for content delivery network. Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India (CSI) Volume 1 pp 427-434

Gupta P, Goyal MK, Kumar P (2013) Trust and reliability based load balancing algorithm for cloud IaaS. In 2013 3rd IEEE International Advance Computing Conference (IACC) pp 65-69.

Hesham E, Gadelrab MS, Elsayed K, Sallam AR (2021) Modelling multilayer communication channel in terahertz band for medical applications. International Journal of Communication Networks and Information Security 13:358-365.

Ioannis S. Triantafyllou (2015) Consecutive-Type Reliability Systems: An Overview and Some Applications, Journal of Quality and Reliability Engineering 2015:1-20. DOI: 10.1155/2015/212303.

Jaime Silva (2015) An Effective Algorithm for Computing All-Terminal Reliability Bounds, Networks, Volume 6 Issue 4, pp 282-295, DOI: 10.1002/net.21634

Kaushik B (2013) Achieving maximum reliability in fault tolerant network design for variable networks, Applied Soft Computing 13:3211-3224, DOI: 10.1016/j.asoc.2013.02.017

Kaushik B (2015) Performance evaluation of approximated artificial neural network (AANN) algorithm for reliability improvement, Applied Soft Computing 26:303–314. DOI: 10.1016/j.asoc.2014.10.002

Khanna G, Chaturvedi SK (2018) A comprehensive survey on multihoop wireless networks: milestones, changing trends and concomitant challenges. Wireless PersCommun 101:677–722 DOI: 10. 1007/s11277-018-5711-8

Khanna G, Chaturvedi SK, Soh S (2019) On computing the reliability of opportunistic multihop networks with mobile relays, Qual Reliability Eng Int 0748-8017. DOI: 10.1002/qre.2433

Leite da Silva AM, Chaves de Resende L, Manso LA, Miranda V (2007) Composite reliability assessment based on Monte Carlo simulation and artificial neural networks, IEEE Trans. Power Systems 22:1202–1209. DOI:10.1109/TPWRS.2007.901302

Nasim Nezamoddini (2015) Reliability and topology-based network design using pattern mining guided genetic algorithm, Expert Systems with Applications 42:7483-7492.

Negi SK, Rajkumari Y, Rana M (2022) A deep dive into metacognition: Insightful tool for moral reasoning and emotional maturity. Neuroscience Informatics 2:100096.

Rahman ASA, Masrom S, Rahman RA, Ibrahim R (2021) Rapid software framework for the implementation of machine learning classification models (2021) International Journal of Emerging Technology and Advanced Engineering 11:8-18. DOI: 10.46338/IJETAE0821_02

Sajja GS (2021) Machine Learning based Detection of Depression and Anxiety. In International Journal of Computer Applications. Foundation of Computer Science 183:20–23. DOI: 10.5120/ijca2021921856

Shuai Lin (2018) System Reliability Assessment Based on Failure Propagation Processes, Complexity 2018:1-19. Available in: DOI: 10.1155/2018/9502953

Sriram GS (2022) Challenges Of Cloud Compute Load Balancing Algorithms. International Research Journal of Modernization in Engineering Technology and Science 4:1186-1190.

Xiao Y, Chen S, Li X, Li Y (2008) An Enhanced Factoring Algorithm for Reliability Evaluation of Wireless Sensor Networks, The 9th International Conference for Young Computer Scientists 2175-2179. DOI: 10.1109/ICYCS.2008.240.

Zhifu Luan (2017) Calculation and Simulation of Transmission Reliability in Wireless Sensor Network Based on Network Coding, International Journal of Online Engineering (IJOE) 13. DOI: 10.3991/ijoe.v13i12.7883

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August xx, 2023
https://doi.org/10.31893/multiscience.2023ss0105

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Migrating data from document-oriented database to graph-oriented database

Lim Fung Ji[a] | **Nurulhuda Firdaus Mohd Azmi**[b]

[a]Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur, Malaysia, department of Software Engineering and Technology.
[b]University Teknologi Malaysia, Malaysia, department of Advanced Informatics, Razak Faculty of Technology and Informatics.

Abstract In data migration between different types of NoSQL database, data may not be directly transferred to the targeted database in compare to migration of data between the same types of database. This is due to the heterogeneity of storage paradigm of the NoSQL databases. For example, migrating data from a document-oriented database such as MongoDB, which stores data in Json (Java Object Notation) format to Neo4j, a graph-oriented database stores data in node, the differences among these databases' storage paradigm requires different representation of data model in the targeted graph-oriented database. This paper proposed a sequential approach to migrate data from MongoDB to Neo4j. The approach migrates MongoDB data to Neo4j and verifies the migrated data using a comparative method. The paper discusses on the migration algorithm and how complex field in MongoDB such as nested document is presented in Neo4j.

Keywords: data migration, NoSQL database, MongoDB, Neo4j, sequential comparison

1. Introduction

Graph-based database is one of the NoSQL type databases that stores data in nodes. Another type of NoSQL, which is the document-based database stores data in Json form. The transfer of data between the two types of databases may be a straight-forward, direct record to record transferring due to the heterogeneities of storage paradigm among the databases. This paper discusses the algorithm on data migration from MongoDB; a document-oriented database, to Neo4j; a graph-oriented database.

2. Document Oriented Databases and Graph Oriented Databases.

Document oriented NoSQL saves data using JSON or BSON. A key value set is used to identify documents. For document-oriented database, searches may be conducted using both the key and the value. The graph database stores data as nodes connected by edges. The edge illustrates the connection between nodes. There is a pointer in the nodes that directs to the subsequent nodes (Swaroop 2016). Table 1 shows some summary on storage architecture for the two types of database (Zafar 2016)

Table 1 Summary on data model of document and graph databases (Zafar 2016).

Types of NoSQL	Data Model	Strength	Weakness	Example
Document oriented	Groups of relationships between key values	Tolerate on incomplete data.	Do not have standard query syntax.	CouchDB, MongoDB
Graph oriented	Data Nodes – "Property Graph"	Apply graph algorithm – identify shortest path, connected ness.	Difficult in clustering, traverse whole graph.	InfoGrid, InfiniteGraph

3. Data Migrations

The process of transferring data from one data source to another, such as from one database to another or within the same database, is known as data migration. For NoSQL databases, there are differences (Dharmasiri and Goonetillake 2013) among these databases and the differences are:

- Each type of NoSQL database has its own data model, i.e. data model is different among types of NoSQL databases.
- Implementation is different even for the NoSQL databases which has the same data model.
- Difference query language
- Supporting different features of CAP

- Apply different model of consistency

In the migration of data between NOSQL databases may requires the consideration of the heterogeneities of the databases' data model. Inappropriate design of data model in the target database may not be able to support the data schema from the original data source and will affect the integrity of data. This paper discusses on the migration of data between different types of NoSQL with the objectives:

Obj1: Examine the sequential algorithm that migrates data between NoSQL databases.

Obj2: Examine data schema of the targeted database to support data model of the original data source.

4. Related Works

Data migration between diverse databases was studied and methods were developed. On-demand conversion between the old and new schema fields is possible with the lazy migration strategy. The period of time the application is offline will be reduced as a result of this (Klettke 2016) However, in terms of managing the application's code, this technique will create a load on the developers. Another solution, which uses slow migration but reduces developer load, isolates migration components into a database library, allowing apps to access data normally (Saur et al 2016) With the assistance of a decision tree, a technique called Build Schema Profile (BSP) was proposed for detecting schema differences between database instances. This is a schema-less document-based database approach (Gallinucci et al 2018) for the goal of transferring data between multiple columnar databases, a data migration system was suggested and implemented. Heigira4Cloud transforms data into an intermediate format before converting it to the destination databases (Scavuzzo et al 2014) the migration mechanism was improved by the authors by adding fault tolerance and support for several database types. Data migration issues, such as duplicated data and data loss, are detected by these fault tolerance characteristics (Scavuzzo et al 2016).

5. Experiment Configurations and execution

5.1 Experiment Setup

The experiment is performed on a single machine with databases run as local services. MongoDB 4.2.3 is selected to represent the document-based database and Neo4j 3.5.18 is representing graph-based database. The databases are run on different ports; MongoDB service is run at local port 27017 and Neo4j service is available at port 7474. The dataset from is a curated list in GitHub which contains fourteen collections of Json/Bson files (Ozler 2019) Within the fourteen (14) collections, the *Companies* collection will be used. The *Companies* collection consists of 1583 documents where each document consists of thirty-eight fields. The fields within the *Companies* collection are from simple data type field to complex data type field. Simple data type field contains text data, numbers, data or alphanumeric data. Whereas complex data type field may contain nested documents, nested array, array of document and et cetera. The schema of *Companies* collection is depicted in Table 2. In Table 2, the mixed data type represents the field contains different types of data in different documents.

Table 2 Shema of Companies collection.

Field	Type	Field	Type
_id	ObjectId	deadpoolead_year	Mixed
name	String	tag_list	String
permalink	String	alias_list	Mixed
crunchbase_url	String	email_adress	Mixed
homepage_url	String	phone_number	Mixed
blog_url	String	description	Mixed
blog_feed_url	String	created_at	Mixed
twitter_username	Mixed	updated_at	String
category_code	String	overview	String
number_of_employees	Mixed	image	Mixed
founded_year	Mixed	products	Array
found_month	Mixed	relationships	Array
founded_day	Mixed	competitions	Array
providerships	Array	milestomes	Array
total_money_raised	String	video_embeds	Array
funding_rounds	Array	screenshots	Array
investments	Array	external_links	Array
acquisition	Mixed	partners	Array
acquisitions	Array	offices	Array

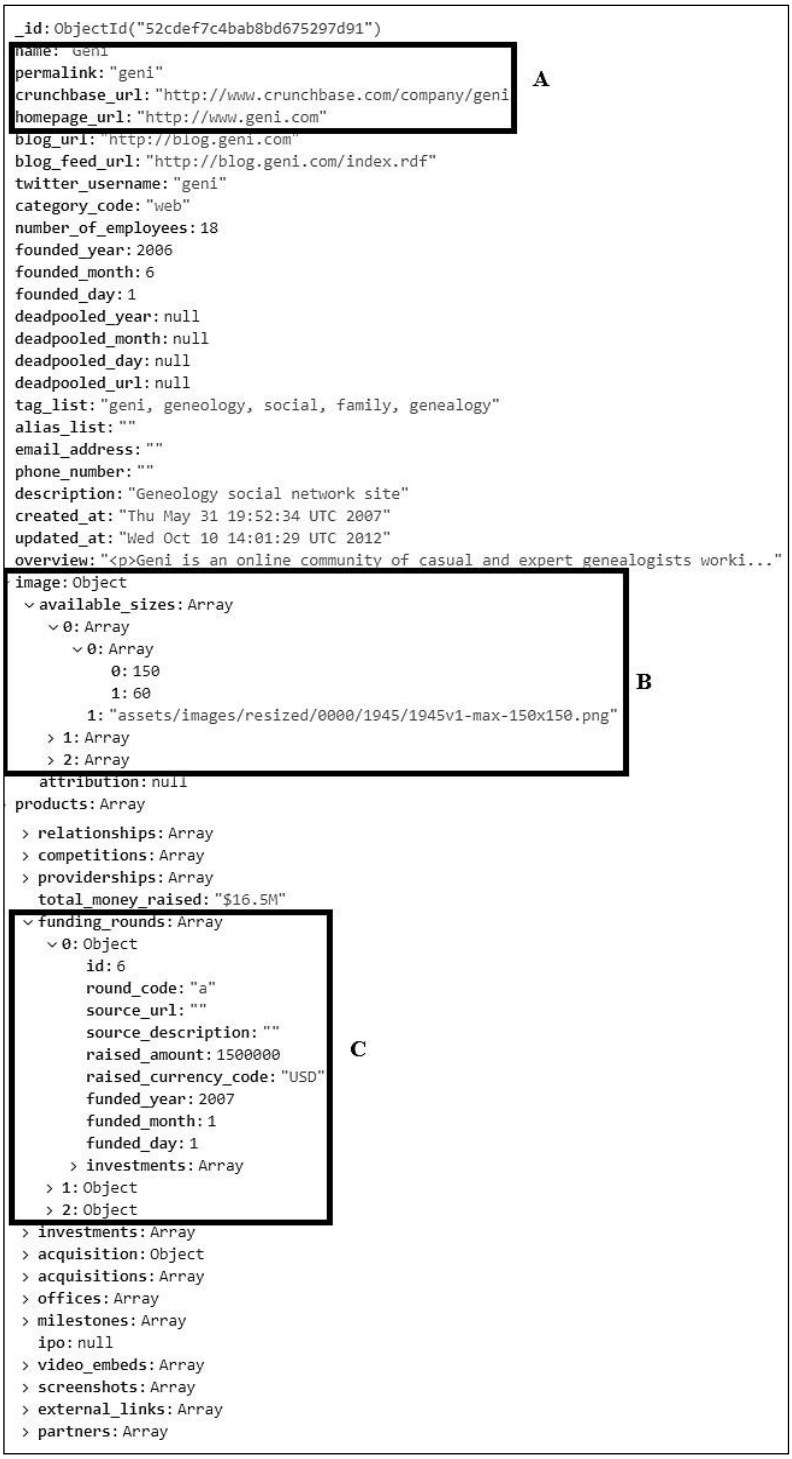


Figure 1 Sample MongoDB document.

Figure 1 depicts the sample of MongoDB document. Figure 1 shows different field types in a document, for example, fields highlighted in A are simple type field which stores only one value in each field, however for highlighted fields in B and C are complex field type, B indicates a nested array and C shows an example of object array. However, each document does not have a standard structure (schema) within the same collection.

5.2 Migration Algorithm

From the structure of data field in MongoDB, due to the non-static structure of MongoDB document, it is necessary to know the structure of all documents and fields. Therefore, the research applies a record-to-record method to migrate data to Neo4j database. Figure 2 depicts the migration algorithm with a flowchart diagram.

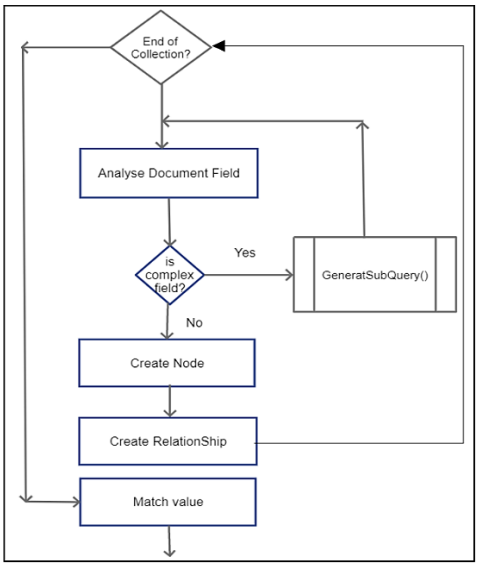


Figure 2 Migration flowcharts.

The migration process intends to transfer data of a collection in a record-to-record (in this case, document-to-document) basis. The algorithm starts by access the first document. Then, each field within the document will be analyzed for its type. The purpose is to detect the necessity to create additional node for the field. This is due to the different data model supported by Neo4j, for example, Neo4j does not support nested nodes, and therefore for field type such as array of object, a separated node needs to be created to represent the nested field. For simple data field such as test, numeric, alpha-numeric, simple array, the field will be migrated as a property of the target Neo4j node. The nodes are created under Neo4j label *Companies*. However, for complex data type field, extra nodes will be created. For example, the *funding round*s field in Figure 1; it is an array of object in which each element in the array is a document. Therefore, each element in the array will be migrated as a node In Neo4j. Since the node is created to represent the complex field, a relationship will be created to indicate the connection between nodes.

The creation of node in the algorithm is performed by first generate the Cypher Query Language (CQL) string. This is performed by calling the function *Generate Subquery ()*. The CQL string will be executed to create the intended Neo4j node. After all related-nodes for the complex field are created, the creation of the "mother" node will be performed. "Mother" node means the node that represent a document and consists of those simple field type as property. In this experiment, a "mother" node represents a *Companies* document and the child nodes are the nodes that represent the complex field in MongoDB document. A relationship named "HASSUB" will be created to link all related nodes together. Therefore, the algorithm is creating related-nodes for a single document first, then, create the node for the document itself and link these nodes together. The process will repeat until all documents are migrated.

When all records are migrated, in the match value process, a comparative method is applied to validate and verify the migrated data based on the data origin. This is performed by matching the content value of each node with the value in MongoDB collection. The algorithm of the comparative method is depicted by flowchart in Figure 3.

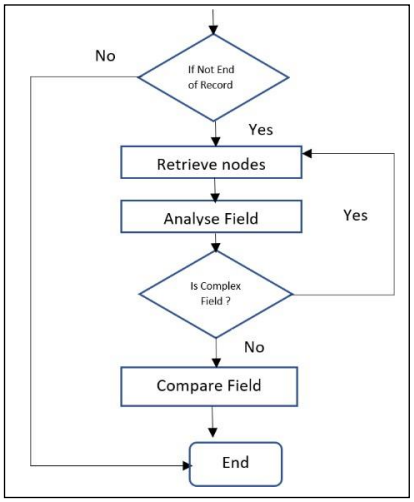


Figure 3 Algorithm for record value matching.

As depicted by Figure 3, the algorithm starts by retrieving the documents in the MongoDB *Companies* collection. The related nodes will be retrieved from the Neo4j according to the collection name, that is, nodes in the *Companies* label will be retrieved. Each field from the Mongo document and Neo4j node will be compared. If the field is a complex type, then the related nodes will eb retrieved and the comparison process will be started again. The comparison of process finished after all records are compared. The comparison result will be written to the *CheckLog.txt* file.

6. Experiment Result and Discussion

The migrated data is viewed by using the Neo4j client browser interface. The labels are retrieved by using the CQL command "*Match (n) return distinct labels (n), count (n)*". These labels are the results of migrating complex fields that lead to the creation of additional nodes in new label. Figure 4 shows parts of the returned labels and the number of nodes from the command. The highlighted *Companies* label consists of those nodes which contains the simple field data as it is in MongoDB before migration. Other labels are the result of migrating complex field. For example, the *funding_rounds* label, in MongoDB, it is an object array field. When the field is migrated to Neo4j, a node is created for each object element under the label has the same name with the field name.

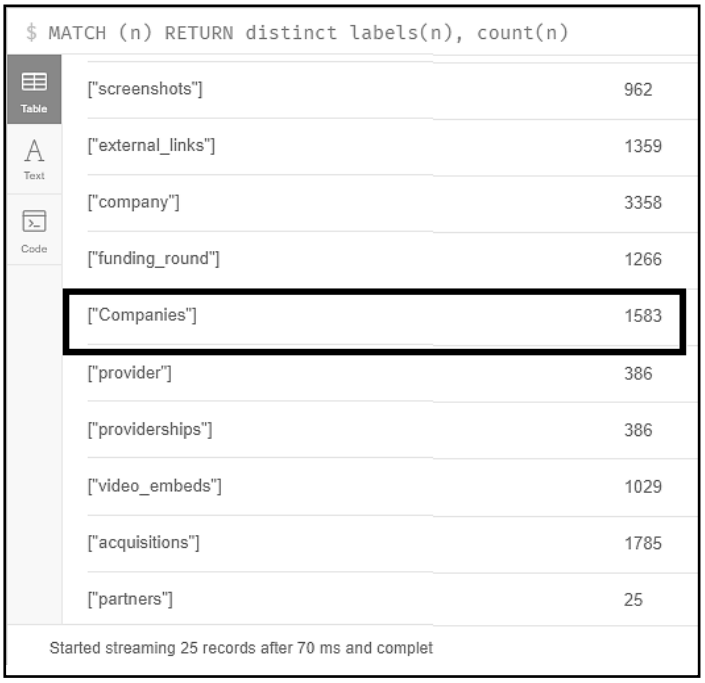


Figure 4 Neo4j labels.

In addition, from the highlighted part in Figure 4, it shows that the total nodes in *Companies* label are 1583, which is same as the total document in *Companies* collection.

Next, the research proceeds to read the *CheckLog.txt* file, which consists of the comparison results. Figure 5 and Figure 6 shows parts of comparison results from the file. From the comparison result, it shows that the values between MongoDB documents and the related Neo4j nodes are mapped. To further prove on the algorithms, samples of records form MongoDB and Neo4j of matching result.

```
53  Now checking the following record :
54  ----------------------------------------
55  Document ID:52cdef7c4bab8bd675297d8a
56  The node and its subnode(s) mapped are :
57  1611159  1611160  1611161  1611162  1611163
    1611177  1611178  1611179  1611180  1611181
    1611203  1611204  1611205  1611206  1611207
    1611223  1611224  1611225  1611226  1611227
    1611241  1611242  1611243  1611244  1611245
    1611259  1611260  1611261  1611262  1611263
58   Start Checking fields.....
59
```

Figure 5 Connected nodes mapping result.

```
Source field value (Document) twitter_username = BachelrWetpaint
Target property value (Node) twitter_username= BachelrWetpaint
Result : Matched
Source field value (Document) name = Wetpaint
Target property value (Node) name= Wetpaint
Result : Matched
Source field value (Document) deadpooled_year = 1
Target property value (Node) deadpooled_year= 1
Result : Matched
Source field value (Document) updated_at = Sun Dec 08 07:15:44 UTC 2013
Target property value (Node) updated_at= Sun Dec 08 07:15:44 UTC 2013
Result : Matched
Source field value (Document) founded_year = 2005
Target property value (Node) founded_year= 2005
Result : Matched
Source field value (Document) _id = 52cdef7c4bab8bd675297d8a
Taret property value (Node) _id= 52cdef7c4bab8bd675297d8a
Result : Matched
```

Figure 6 Sample mapping result.

Figure 5 shows the matching result of the nodes that were created for the Mongo document which id is 52cdef7c4bab8bd675297d8a. The list of numbers represents the connected nodes id in Neo4j. The list of node id is generated by executing the CQL command that retrieve all connected nodes to the node which the *oid* property is similar to the document id. This *oid* property (field) of a node represents the Mongo document id in Neo4j which is use for relationship purpose. Figure 6 depicts portion of the matching result of each field.

To further verify the result generated by the matching algorithm, sample records are selected from both MongoDB and Neo4j. Figure 7 and Figure 8 depict the selected fields/properties on the sample document and node respectively. The fields/properties shown in Figure 7 and 8 are the simple fields. The values depicted by Figure 7 and Figure 8 are similar to the value depicted by Figure 6.

```
_id: ObjectId("52cdef7c4bab8bd675297d8a")
name: "Wetpaint"
twitter_username: "BachelrWetpaint"
founded_year: 2005
deadpooled_year: 1
updated_at: "Sun Dec 08 07:15:44 UTC 2013"
```

Figure 7 Sample selected Mongo document fields.

```
$ Match (n:Companies) where n.oid = '52cdef7c4bab8bd675297d8a' return n.twitter_username, n.name,
```

Table

Text

"n.twitter_username"	"n.name"	"n.deadpooled_year"	"n.updated_at"	"n.founded_year"
"BachelrWetpaint"	"Wetpaint"	1	"Sun Dec 08 07:15:44 UTC 2013"	2005

Figure 8 Sample related Neo4j node properties.

In addition, the verification proceeds to check on the representation of complex field in the document. In this case, the *funding_rounds* fields are selected. Figure 9 depicts the *funding_rounds field* form the same document sample. As discussed earlier in the paper, the funding rounds field is an array of object, each object represent a document. Therefore, it is a nested document structure.

Form the sample in Figure 9, the *funding_rounds* of the document consists of three elements, that is, when migrated into Beo4j, there should be three nodes created under *funding_rounds* label. For the *investments* field, nodes are created under *investments* label in Neo4j. Each element in investments consists of three fields that are *company*, *financial _org* and *person* fields.

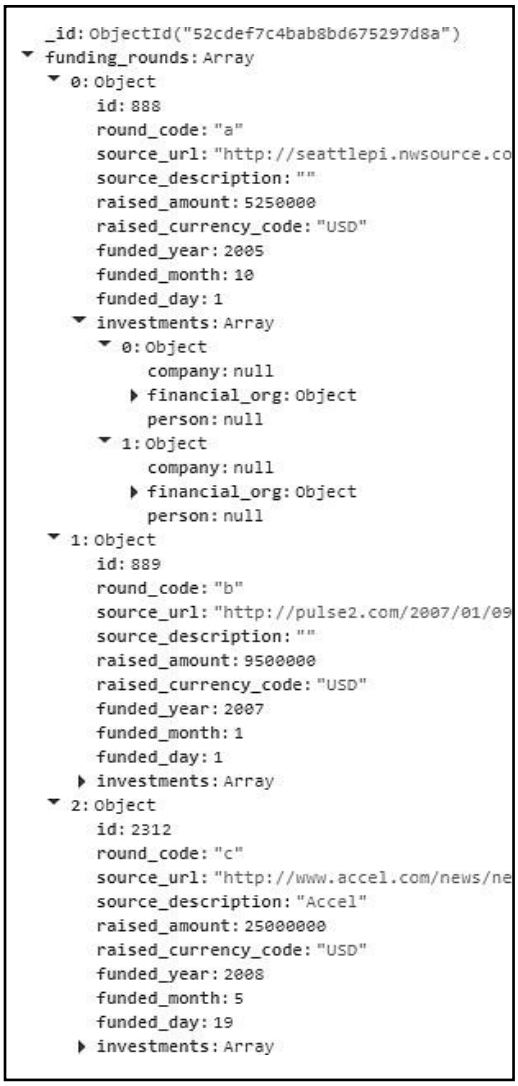


Figure 9 funding_rounds field from Mongo document.

Figure 10 and Figure 11 depict the related nodes from Neo4j based on the sample document in Figure 9. From the three figures, it shows that the values of each field in the *fundisng_rounds* and *investments* nodes are matched with the values of document fields in Figure 9. However, Figure 11 does not shows the financial_org field due to the field is a document object; therefore, the migrated field becomes a node under the *financial_org* label. Figure 12 depicts the related *financial_org* nodes from Neo4j.

Figure 10 Related Neo4j sample of funding_rounds nodes.

Figure 11 Related investments nodes.

Figure 12 Related financial_org nodes.

From the comparisons, it shows that the migration method is able to migrate the data as according to the discussed schema.

7. Conclusions

The record-to-record migration method migrates data from MongoDB to Neo4j in a sequential way. Due to the flexibility of storage paradigm in MongoDB, migrating data requires a predefined structure that are supported by the targeted type of database. For example, in this paper, the Neo4j does not support nested nodes in compare to MongoDB which allowed nested document to be stored as a field. Therefore, the identification of field type is cucial and MongoDB allows greater flexibility in data storing within a document in compare to Neo4j node.

With the record-to-record validator method for the migration data, it needs further verification as manual verification discussed in the paper may not be able to sufficiently verify the effectiveness of the method.

For future works, study will proceed to the migration of more complex field structure and with other NoSQL databases.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Dharmasiri HML, Goonetillake MDJS (2013) A federated approach on heterogeneous NoSQL data stores. in 2013 International Conference on Advances in ICT for Emerging Regions (ICTer).

Gallinucci E, Golfarelli M, Rizzi S (2018) Schema profiling of document-oriented databases. Information Systems 75:13-25.

Klettke M, (2016) NoSQL schema evolution and big data migration at scale. in 2016 IEEE International Conference on Big Data (Big Data).

Ozler H (2019) A curated list of JSON/BSON datasets from the web in order to practice/use in MongoDB. Available from: http://www.github.com/ozlerhakan/mongodb-json-files.

Saur K, Dumitraş T, Hicks M (2016) Evolving NoSQL Databases without Downtime. in 2016 IEEE International Conference on Software Maintenance and Evolution (ICSME).

Scavuzzo M, Tamburri DA, Nitto Ed (2016) Providing Big Data Applications with Fault-Tolerant Data Migration across Heterogeneous NoSQL Databases. in 2016 IEEE/ACM 2nd International Workshop on Big Data Software Engineering (BIGDSE).

Scavuzzo M, Nitto ED, Ceri S (2014) Interoperable Data Migration between NoSQL Columnar Databases. in 2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations.

Swaroop P (2016) NoSQL Paradigm and Performance Evaluation. SSARSC International Journal of Geo Science and Geo Informatics 3.

Zafar R, (2016) Big Data: The NoSQL and RDBMS review. In 2016 International Conference on Information and Communication Technology (ICICTM).

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0106

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Machine learning model for identification of frontend and backend repositories in Github

Ulvi Shakikhanli[a] | **Vilmos Bilicki**[a]

[a]University of Szeged, Szeged, Hungary, Doctoral School of Computer Science, Faculty of Science and Informatics.

Abstract Frontend and Backend repositories are the main components of the multi-repository structure. The developers or team managers upfront mostly do the identification of these repositories, but in Github, there is almost no tool or third party identification to identify these repositories. Since Github is the largest open source platform and is used by, many researchers and identification of these two repositories are crucial. This paper shows how this problem can be solved by building a machine learning model based on the file structure of the repository. The model has proven to be very successful with an accuracy of over 90% and can be used not only to identify the repository type, but also to identify the development language and environment.

Keywords: frontend, backend, github mining, machine learning, repository identification, file structure

1. Introduction

Multi-repository structures are definitely preferred by development teams and project managers. There are even some companies that have made this their company policy, such as "Netflix" (Netflix Culture 2019). This type of structure consists mainly of two repositories: frontend and backend. The frontend is the part of a web application that users interact with directly through their web browser. It is usually responsible for displaying the user interface, processing user input and communicating with the backend of the application. A backend repository, on the other hand, is a type of code repository that stores the code base and data for the backend of a web application. The backend is the part of a web application that runs on a server and is responsible for managing and storing data and processing logic and business processes.

Since those two repository types are main parts of Multi repository structure it is essential to be able to identify them. This paper will mostly focus on identification process using the opportunities provided by Github API, which is a product of Github platform. Github is an online service for hosting software. It offers each project access control, bug tracking, software feature requests, task management, continuous integration, and wikis in addition to Git's distributed version control. As of June 2022, GitHub reported having over 83 million developers and more than 200 million repositories (Github 2022). It makes Github something like a real gold mine for researchers and several papers have been written solely about Github Mining (Valerio Cosentino 2016; Eirini Kalliamvakou et al 2014; Valerio Cosentino 2017).

However, despite all of these huge potential researches face with challenges when they try to use Github Mining. Github API provides great amount of information about repository but it lacks to identify if repository is a Frontend or Backend. This part is left for developer to identify with additional *tags* or special naming like "*vault-frontend*" (Github 2022). Nevertheless, in most cases developers does not add any identification tags or special names and in this cases, there is a need for additional method. In this paper, a Machine Learning approach has been applied for this purpose.

2. Related work

During the literature review for this paper it appeared that, there is no any academic research paper about identification of Frontend and Backend repositories. Result can be research field itself. In most cases, projects have been analyzed without considering Mono or Multi Repository structure itself. Even in some cases when there is study about repository structure like in paper (Jaspan 2018) it took only small amount of projects. Since these type of researches have been done in scope of one company or developer group, the types of repositories were predefined and thus there was no need for any additional tool or algorithm. In paper (Shakikhanli U 2022) two additional options have been introduced: a) Identification according to the file name, b) Identification according to the list of packages in configuration file. Both of those methods are effective but they have some disadvantages. For example:

Identification according to the name – cannot be used in the repository groups where developers have not used special naming tags. For example, user (Github 2022) has not specified frontend repository with any tag or name and that is why above-mentioned method cannot be used here.

Identification according to the configuration file – will best suit when research is based on *NodeJs* projects. Otherwise, it will become harder to be able to identify frontend or backend repositories. There is also specific guideline for naming repositories in Github (Github 2022) but unfortunately, it is not followed by all developers. A small test can verify this information. As it has been mentioned before, there are nearly 200 million repositories in Github. By creating a simple search query, it can be seen that there are not more than 2 million repositories, which uses tags like "*frontend*" or "*backend*" in either name or in description of repository. Since most of the projects in Github are developed by freelancer or independent developer, it means even that 2 million numbers is not showing us real amount of repositories, which have been named according to the guideline.

3. Methodology

This chapter will mostly be divided into two sections how to create a database and creating a ML model. Creation of database is essential since the model will be both trained and tested according to that. As it has been mentioned in previous chapters, there is no clear database for frontend and backend repositories. Because of that, the only for collecting before mentioned repository types is to do it by hand. In order to fasten the process some search queries have been created and collected repositories analyzed before being added to database. The following example can be shown for search query in Github platform:

https://api.github.com/search/repositories?q=frontend+language:javascript+created:2021-01- 01.. 2021-01-31.

This query will give a list of repositories with "frontend" in their name. There are also additional parameters as language and publication date for repository to shorten the search area. For further steps, repositories have been analyzed by both developers and special scripts for verification. Even difficultness of this process proved importance for general method and proved importance of this paper.

For more clearance, it is better to explain that the above-mentioned scripts checked following aspects of repository for verification: a) Configuration file (if exist); b) Readme file (if exist); c) Repository tags (if exist). If none of these were existed in repository then it has been checked by developer and verified.

After collecting all these repositories, the next step was to collect the file structure of these repositories. Unfortunately, there is no request type in the Github API for retrieving such information. However, there was a request to get a list of files in a specific directory, and this request is used as a workaround in this paper. Since it was possible to get a list of all the files in the directory, this method was used to create a "file structure" of the repository. This process could only be done by sending multiple requests to Github for each repository. Since the number of requests to Github is limited, a special authentication key (Github 2022) was used to speed up the process. The Github API allows users to send 5000 requests per hour. To calculate the request usage limit, one can check the remaining request volume or simply calculate the number of directories in the repository. This has been done in this post because each query only shows the files and directories in a particular folder, so the script must send a new query whenever it encounters a new directory. An example of a query string to retrieve the contents of the repository might look like this:

https://api.github.com/repos/{repository owner}/{repository}/contents/{path to directory}

There are also few advices about this process in order to avoid some common mistakes. For example, researches have to add special list of folder names, which their script has to avoid during this process. For example, folders like "*node_modules*" contains all packages and additional libraries which are used in repository and despite this content there also huge number of libraries which has no value for this model. Due to its size and number of folders in it will consume huge amount of requests and thus will delay overall speed of data collecting.

Path of each file added to the database as a sentence in text. Therefore, this way the file structure of repository can be treated as text for future processing. Following example can be given to explain this much better:

"READMEmd. packagejson. public/faviconico. public/indexhtml. public/logo192png. src/components/CardComponentjs. src/services/apijs."

As it seen here, package.json and other files are written without dot sign to separate extension and it is done in order to avoid stemming operations, which will be discuss later on. In order to be able to treat these file structures as texts some operations like "stemming" (Stemming and Lemmatization in Python 2018) and removing "/" between folders and files. There is also a little notification to inform that during stemming process *English* language have been used. This way it become possible to construct file structures for each repository as a text file and they all can be used for training of a model.

4. Training and testing of model

Collecting repositories and creation of database with file structures there become a need for choosing an algorithm for training. Scikit-learn provides different methods for creating an ML model (Scikit-learn 2022). In order to demonstrate the

efficiency of this method all major ML algorithms have been used here. Paper gives short definition for each of them but detailed information can be obtained from official documentations and given references:

The *Multinomial Naïve Bayes* method applies the naive Bayes technique to text classification problems where the data follows a multinomial distribution. This method is commonly used in natural language processing tasks and is known to work well with word or tf- idf vectors, which are used to represent the text data in a numerical format that the algorithm can understand (Multinomial Naive Bayes 2022).

The *SGD Classifier* uses stochastic gradient descent to train regularized linear models. It estimates the loss gradient for each sample and updates the model's parameters accordingly, using a gradually decreasing learning rate. This allows the model to learn and improve over time (SGD Classifier 2022).

Ridge classifier first converts the target values into {-1, 1} and then treats the problem as a regression task (multi-output regression in the multiclass case) (Ridge Classifier 2022).

Logistic Regression method is a linear model that is used for classification rather than regression in machine learning. It is also known as logit regression, MaxEnt, or the log-linear classifier (Logistic Regression and Scikit-learn 2022).

K Nearest Neighbors-based classification is a method of instance-based learning that does not create a general internal model. Instead, it simply stores examples from the training data. This type of learning is non-generalizing and does not aim to create a general model of the data (K Nearest Neighbor 2022).

Random Forest Classifier is a machine learning method that combines multiple decision tree models to make predictions. It trains these decision trees on different subsets of the data and then averages their predictions to improve the accuracy and reduce over fitting (Random Forest Classifier 2022).

Decision Trees are a type of machine learning model that can be used for both classification and regression tasks. They are non-parametric, meaning that they do not make any assumptions about the underlying data distribution (Decision Trees 2022).

Extra Tree Classifier c implements a meta estimator that fits a number of randomized decision trees (a.k.a. extra-trees) on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting (Extra Tree Classifier and Scikit-learn 2022).

Table 1 Accuracy score of all algorithms when all DB have been used.

	Algorithm	Accuracy
1	Multinomial Naïve Bayes	0.514648858250754
2	SGD Classifier	0.514648858250754
3	Ridge Classifier	0.4844894442050844
4	Logistic Regression	0.484489444205084
5	K Nearest Neighbor	0.8778543731150367
6	Random Forest Classifier	0.9030590262817751
7	Decision Tree	0.9032744506678156
8	Extra Tree Classifier	0.9015510555794916

Table 1.1 shows the general results when different algorithms are used. As mentioned in the previous chapter, the file structures of the repositories are treated like texts in natural language. For this reason, the results of the Multinomial Naïve Bayes algorithm should be much higher than those of the other algorithms, but surprisingly it performed worse than almost all other algorithms and methods. The Decision Tree algorithm, which out performs the Random Forest Classifier by a small margin, achieves the best result. A quarter of the database was used to test the model, but of course this value can be changed and for some algorithms, e.g. Multinomial Naïve Bayes, additional parameters can be added to improve the accuracy, but these additions will not change the numbers dramatically.

The model may give different results for other programming languages such as Java or Python, but we have left them aside because there repositories of these languages are usually part of the backend of a project. For this reason, only the two main languages Typescript and Java Script are used for testing and trading the model. Despite other languages, numerous frontend or backend repositories can be written in one of the aforementioned languages.

5. Conclusion and Future works

The paper provides a clear understanding of the importance of frontend and backend repositories and offers a solution for identification using a machine-learning model. This method is highly customizable and can be used with almost all types of repositories written in any programming language. The results of the different methods have been shown and demonstrate the success of this method in identifying the type of repository.

In addition to explaining the meaning of the proposed model, it also explains how the Github platform and its API can be used to create a database that can be used for both similar and different purposes in different research areas. In the future, it is planned to further improve the model and solve some problems related to data collection and processing time.

6. Threads to validity

The performance of the method and its usefulness have been proved by tests and the results shown in the paper, but there are some aspects that need to be considered by other researchers. First, one can consider the time required to create such a database and model. Moreover, all there repositories used here are open source projects on the Github platform, mostly developed by individual developers without commercial interests. For this reason, the algorithm may show different results when tested on commercial or multi-tier projects. Future work will be implemented to improve and eliminate the weaknesses of the algorithm.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Decision Trees, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/tree.html#tree.

Eirini K, Georgios G, Kelly B, Leif S, Daniel G M, Daniela D (2014) The promises and perils of mining GitHub, MSR 2014: Proceedings of the 11th Working Conference on Mining Software Repositories 92–101. DOI: 10.1145/2597073.2597074

Extra Tree Classifier, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/generated/sklearn.ensemble.ExtraTreesClassifier.html.

GitHub (2022) Available in: https://docs.github.com/en/authentication/keeping-your- account-and-data-secure/creating-a-personal-access-token.

GitHub (2022) Available in: https://github.com/Drakolis?tab=repositories.

GitHub (2022) Available in: https://github.com/strvcom/strv- guidelines/blob/master/git.md#repositories.

Github Wikipedia (2022) Avalable: https://en.wikipedia.org/wiki/GitHub#citenote-10.

Jaspan C, Jorde M, Knight A, Sadowski C, Smith EK, Winter C, Emerson Murphy Hill (2018) Advantages and disadvantages of a Monolithic Repository: A case study at Google, in Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Practice, Gothenburg, Sweden, pp. 225–234.

Logistic Regression, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/linear_model.html#logistic-regression.

Multinomial Naive Bayes, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/naive_bayes.html#multinomial-naive-bayes.

Nearest Neighbor K, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/neighbors.html#classification.

Netflix Culture (2019) Freedom & Responsibility. Reed Hastings. Slideshare. Available in: https://www.slideshare.net/reed2001/culture-1798664/2-Netflix_CultureFreedom_Responsibility2.

Random Forest Classifier, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

Ridge Classifier, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/generated/sklearn.linear_model.RidgeClassifier.html

Scikit-learn (2022) Available in: https://scikit-learn.org/stable/

SGD Classifier, Scikit-learn (2022) Available in: https://scikit- learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html.

Shakikhanli U, Bilicki V (2022) Comparison between mono and multi repository structures, Pollack Periodica 17:7-12. DOI: https://doi.org/10.1556/606.2022.00526.

Stemming and Lemmatization in Python, DataCamp (2018) Available in: https://www.datacamp.com/tutorial/stemming-lemmatization-python

Valerio C, Javier LCI, Jordi C (2017) A Systematic Mapping Study of Software Development With GitHub, IEEE 5. https://ieeexplore.ieee.org/abstract/document/7887704.

Valerio C, Javier L, Jordi C (2016) Findings from GitHub: methods, datasets and limitations, in MSR '16: Proceedings of the 13th International Conference on Mining Software Repositories 137–141. DOI: 10.1145/2901739.2901776

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0107

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

A novel autonomous machine learning technique for recognizing control chart patterns

Raghavendra Ramesh[a] | **Surendra Yadav**[b] | **Ashendra Kumar Saxena**[c]

[a]Jain (deemed to be) University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology.
[b]Vivekananda Global University, Jaipur, India, Professor, Department of Computer Science & Application.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Professor, College Of Computing Science And Information Technology.

Abstract A number of strategies have been developed for controlling and monitoring the production process since the quality of the product has emerged as one of the key concerns in today's manufacturing sector. Control charts are the best tools for monitoring and adjusting products and processes. This research proposes a novel automatic method for the recognition of nine control chart patterns (CCPs) based on novel autonomous machine learning. The classification portion and the tuning portion make up the two main components of this procedure. Support Vector Machine (SVM) have demonstrated outstanding performance over the past few years on a variety of applications, including signal dispensation, speech detection, and image processing. SVM is consequently employed as the intelligent classifier for the recognition of CCPs in the classification phase. One key challenge with SVM is that it requires a high level of expertise to choose appropriate parameters, such as the quantity of kernel and their spatial diameters, knowledge rate, etc. It is difficult to fine-tune the SVM parameters because of their domestic dependence. These problems led to the employment of the Harmony Search (HS) Algorithm for the best tuning of SVM parameters in the tuning section of the proposed technique. Instead of depending on any feature engineering procedures, the suggested method, in contrast to the popular CCPs recognition methods, takes raw data and runs it through many hidden layers to obtain the best feature representation. The quantitative and simulation results demonstrate the suggested method's performance advantage over the earlier methods.

Keywords: support vector machine, harmony search algorithm, control chart patterns, novel autonomous machine learning, signal processing

1. Introduction

The Autonomous Machine Learning (AutoML) method is to eliminate the need for considerable human involvement in the creation and rollout of ML models. It is difficult to single out a single "novel" AutoML approach since the subject is developing at such a fast pace (Xie et al 2019) Since neural networks form the backbone of many machine learning models, their construction may be automated through a method called Neural Architecture Search (NAS). NAS seeks to automate the process of designing neural network architectures, which is currently done either via trial and error or with specialized expertise (Zhang et al 2019). Various layers, connections, hyper parameters, and activation functions are all part of the architectural space that is being searched. The method considers much potential architecture, each of which is trained and then tested on a validation dataset. This is often a very resource-intensive and time-consuming activity on a computer (Miikkulainen et al 2019). Agent trained using reinforcement learning may pick and construct novel designs by maximizing a reward signal that represents the architecture's performance on a specific task. To develop better designs, the agent repeatedly samples them, trains and assesses them, and changes its policy (She et al 2020) Using evolutionary algorithms to probe the architectural space. Keeping an architectural population alive, evolving it via processes like mutation and crossover, and measuring its effectiveness are all part of this process. In an evolutionary process, the best structures tend to procreate and spread across the population (Khan et al 2019). To solve the architectural search problem, the gradient-based NAS method poses it as a continuous optimization issue. To enable gradient-based optimization, it employs continuous relaxation methods in place of discretely selecting architecture. Through repeated iterations, the search method optimizes a differentiable surrogate goal by modifying architectural characteristics like channel widths and layer sizes (Sun et al 2021). To effectively explore the design space, a Bayesian optimization-based NAS technique uses this method. It employs Bayesian approaches to efficiently search for and identify potential designs by modelling their performance as a surrogate function (Loquercio et al 2020) The purpose of NAS is to eliminate the need for human design and experimentation by automating the architectural search process and discovering neural network designs that deliver high performance on certain tasks (Dang et al 2020). It's crucial to remember that the area of AutoML is always changing, and new methods are being created to handle

certain problems and enhance the speed and efficacy of the creation of machine learning models (Garcia et al 2020). An important part of statistical process control's (SPC's) monitoring and analysis of process variability is identifying trends in control charts. It is possible to automate the detection and categorization of control chart patterns using machine learning methods. One of these methods is supervised learning, in which an algorithm learns to detect various control chart patterns by seeing labelled instances (Sejnowski 2020). Priming the data you need a labelled dataset that includes control charts with identified patterns to use supervised learning. This data collection has to have a number of control charts, all of which are clearly labelled with their respective patterns (normal, shift, trend, cyclic, etc.). The control charts may be created using either actual or simulated process data (Jahangir et al 2020).

2. Related Works

Research (Rajula et al 2020) evaluate ML and classical statistical approaches for their strengths and weaknesses in the context of healthcare. When a priori knowledge of the issue at hand is significant, as it often is in public health research, and the number of cases much surpasses the number of variables under examination, traditional statistical approaches seem to be more beneficial. Beginning with a conceptual model of the pipeline of ML-HCAs from inception to development to deployment and the parallel pipeline of review and supervision duties at each step, this study presents a systematic strategy to detecting ML-HCA ethical problems. We build upon this framework by asking central questions that bring up ethical concerns and by recognizing ethical problems that have received little to no prior attention but which are still important (Char et al 2020). The study (Usama et al 2019), we will take a high-level look at how unsupervised learning has been put to use in the field of networking. We provide a thorough overview of the most recent developments in unsupervised learning methods and illustrate their applicability in a variety of networking-related learning problems. The study provides a brief overview of the development and categorization of machine learning, before delving into the cutting-edge of this field as it has been applied to Unmanned Aerial Vehicles (UAV) for autonomous flying. Several methods of control, such as tuning parameters, adaptive control in an uncertain environment, route planning in real time, and object identification, are provided (Choi and Cha 2019). Applying machine learning methods, as shown in this research, has improved smart farming's precision. Fruits and vegetables such grapes, apples, oranges, and tomatoes, as well as cereal grains like corn and wheat, were harvested for the research. The paper's study results are commercially accessible as tools for a variety of uses, including automated harvesting, weed identification, and insect control (Darwin et al 2021). In the study (Ferdowsi et al 2019) offer edge analytics architecture for ITSs, where data processing occurs at the level of the vehicle or roadside smart sensor, to address the latency and reliability issues plaguing ITSs. For more accurate mobile sensing in ITSs, a distributed edge computing architecture may take use of deep learning methods by combining the power of passengers' mobile devices with in-vehicle processors (Ferdowsi et al 2019). In the paper, they provide a long-term deep learning architecture that effectively classifies objects by fusing and selecting deep features from several layers of data. The strategy presented consists of three stages: very deep convolution networks for large-scale image recognition and inception V3 are used to extract features via transfer learning; (2) feature vectors are fused using a parallel maximum covariance approach; and (3) the best features are selected via a multi logistic regression-controlled entropy-variances method (Rashid et al 2020). In the study (Chen et al 2019) create a multi-level Deep Reinforcement Learning (DRL) protocol for acquiring knowledge about lane-changing habits in heavy traffic. Faster and safer lane changes may be taught by first dissecting the overall behavior into its component sub-policies. The paper presents a thorough analysis of ML applications across several additive manufacturing (AM) fields. ML may be used to generate cutting-edge met materials with improved performance and refined topological layouts in the design for additive manufacturing (DfAM). Modern ML algorithms may aid in optimizing AM process settings, analyzing powder dispersal, and detecting defects in real time (Wang et al 2020).

3. Methodology

In this research, a novel SVM-based hybrid approach to recognizing nine CCPs is suggested. In contrast to traditional, hand-crafted feature representation approaches, the suggested approach learns feature representation automatically from the raw training data. The suggested method involves feeding raw data into SVM and having it identify the CCPs present. To successfully apply SVM to a new project, one must have extensive understanding of the method and be able to choose the appropriate parameters, including activation function, knowledge rate, kernel sizes, and the number of kernels. These parameters have deep reliance on other values, making fine-tuning them a costly endeavour. A larger searching area may damage training, while increasing the number of small-sized kernels improves performance. Reduced processing costs and less overlapping receptor fields are two additional advantages of using the optimum stride length. This means picking the right value for these variables.

3.1. Support vector machine

The support vector machine (SVM) is a supervised ML approach that is a relatively new and promising classification model based on the idea of structural risk reduction. The margin is maximized and strong generalization capacity is attained by using a separating hyper plane in this model.

$$f(x) = (w, \phi(x)) + b \quad (1)$$

Where ϕ(x) is a function from the input space to a high-dimensional feature space; w is a coefficient vector; and b is the offset of the hyper plane from the origin; these values are found by solving the optimization function that follows.

$$g(w, \xi) = \|w\|^2 + c \sum_{i=1}^{N} \xi_i \quad (2)$$

$$y_i \left((w, \phi(x_i)) + +b\right) \geq 1 - \xi_i, \xi_i \geq 0 \quad (3)$$

The following formula yields a kernel function, where εi is the slack variable, c (p> O) the regularization variable of the errors, and (p> 0) the zero-the order constant:

$$k(x_i, x_j) = (\phi(x_i), \boldsymbol{\phi}(x_j) \quad (4)$$

The x_i element of the training sample vectors is the kernel function k. Linear, polynomial, radial basis function (RBF), and sigmoid kernels are some of the many types of kernels used in support vector machines. SVM has the potential to outperform other kernel functions when it comes to non-linear classification. This research is to look into SVM in the context of bagging, boosting, and stacking models as component classifiers.

3.2. Harmony search (HS)

We present a metaheuristic method called harmony search (HS) that uses a population-based strategy. The concept of harmony seek was developed in part as an homage to the art of improvising music. Each element of the solution vector is compared to a note in a musical scale, and the algorithm mimics the way a musician would go about adjusting the notes to get a desired effect. However, the convergence speed of basic HS is lower, and it may have issues with local minima. Five primary factors are presented in Table 1 and five primary processes are outlined below for the fundamental HS. We propose an enhanced HS method in this work by fusing HS approaches with extreme decomposition. Crossover rate (CR) and scaling factor (F) are two new metrics added in DE that stand for these two concepts. Figure 1 provides a high-level overview of the design, and the essential procedures are described in more depth below.

Table 1 Harmony-seeking's primary criteria.

Parameters	Description
HMCR	Rate of recollection harmony
MaxImp	Maximum number of improvisations
bw	Bandwidth vector
PAR	Change in pitch per second
HMS	Size of the symmetric group's memory, in terms of the number of solution vectors

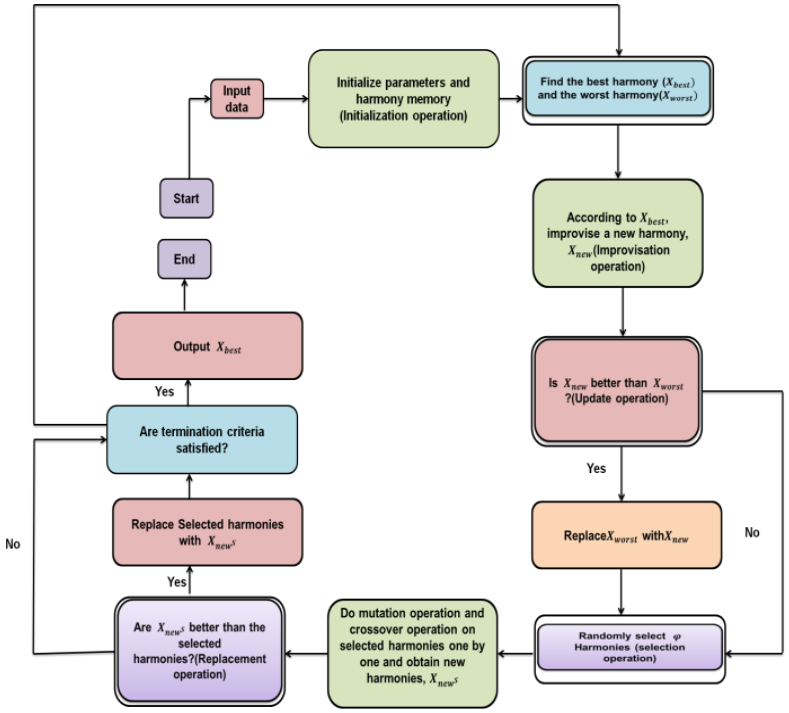


Figure 1 Example of the HS algorithm's flowchart.

Procedure for Getting Started All algorithm parameters, including CR, F, HMS, HMCR, PAR, MaxImp, and *bw*, must be specified. The harmony memory is then set to a random value as:

$$HM = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_N^1 & f(X^1) \\ x_1^2 & x_2^2 & \dots & x_N^2 & f(X^1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ [x_1^{HMS} & x_2^{HMS} & \dots & x_N^{HMS} & f(X^{HMS})] \end{bmatrix} \quad (5)$$

For j = 1, 2... HMS, a solution vector $X_j = (x^j_1, x^j_2 \dots x^j_N)$ is defined, where $f(X^j)$ is the value of the objective function at X^j and x^j_i is the *i*th value of the *j*th solution vector. Finding X_{best} and X_{worst} requires computing the fitness of the objective function.

$$bw_i(Imp) = \frac{x_i^U - x_i^V + 0.002}{10} * exp\ (-10 * \frac{Imp}{MaxImp}) \quad (6)$$

Where Imp is the total number of improvised pieces and x^U_i and x^L_i are the upper and lower boundaries of x_i. We develop the algorithm 1 presented in the following paragraph to build a new harmony. For example, *r* and(*L, U*) returns a random number between *L* and *U* using the normal distribution, whereas *r* and$_u$ (*L, U*) returns a random number between *l* and *U* using the uniform distribution.

Algorithm 1. Improvisational operation in HS pseudo code

```
for each i ∈ {1, 2... N} do
    if rand (0, 1) ≤ HMCR then
        if rand(0, 1) ≤ PAR then
        x_(new,i) ←x_i^best+bw* 〖rand〗 _u (-1,1);
Ifx_(new,i)>x_i^U then
 x_(new,i)=x_i^U  ;
          end if
       else
           x_R=2*x_i^best-x_i^worst
            Ifx_R>x_i^U then
            x_R=x_i^U  ;
          end if
   If x_R>x_i^L then
     x_R=x_i^L  ;
           end if
            x_(new,i)←x_i^worst+rand  (0,1)*(x_R-x_i^worst  );
```

```
        end if
    else
        x_(new,i)←x_i^L+rand (0,1)*(x_i^U-x_i^L );
    end if
end for
ReturnX_new
```

4. Result and Discussion

In this study, we conducted a number of experiments and compared the outcomes of the suggested approach to those obtained using other methods. MATLAB, installed on a Windows 10 64-bit professional PC with 64 GB of RAM, is used for both the SVM's development and analysis. Five independent runs of the suggested method and alternative classifiers are demonstrated, with the average efficiency demonstrated.

4.1. Evaluation of the proposed technique's performance on a dataset

One thousand examples of each pattern are generated using the provided equations To conduct a performance study of the suggested approach. Each equation represents the relationship between the standard normal variate value ri and the 'observed value' yi at the i-th time point. For a 'normal process with mean' a (μ) and 'standard deviation' (σ), the following formulae may be used to generate a wide range of 60-point patterns:
Systematic pattern:

$$y_i = \mu + r_i\sigma + d \times (-1)^i \, , 1\sigma \leq d \leq 3\sigma \quad (7)$$

Normal pattern:

$$y_i = \mu + r_i\sigma', \mu = 80, \sigma = 5 \quad (8)$$

Combine pattern:

$$y_i = \mu + r_i\sigma + (-1)^w m \, , 1.5\sigma \leq m \leq 2.5\sigma \; \& \; w \; is \; 1 \; or \; 0 \quad (9)$$

Stratification pattern:

$$y_i = \mu + r_i\sigma', 0.2\,\sigma \leq \sigma' \leq 0.2\sigma \quad (10)$$

Where '*w*' is a binary integer value that depends on the transition probabilities between distributions, which are described by the parameters b = mp and$p$ $(0 < p < 1)$. Since b is always set to 0.4, this means that we always have $w = 0$ if $p < 0.4$ and$w = 1 \; if \; p \leq 0.4$.
Cyclic pattern:

$$y_i = \mu + r_i\sigma + \sin(2\pi/T) \, , 1.5\sigma \leq a \leq 2.5\sigma \; \& 8 \leq T \leq 16 \quad (11)$$

The amplitude of cyclic variation, denoted by ''a', and the period of a cycle, denoted by ''T', are both shown in this equation.
Increasing Trend pattern:

$$y_i = \mu + r_i\sigma + ig, 0.05\sigma \leq g \leq 0.1\sigma \quad (12)$$

Reduced trend pattern:

$$y_i = \mu + r_i\sigma + ig, \; 0.1\sigma \leq \mathrm{g} \leq 0.05\sigma \quad (13)$$

The amplitude of the gradient for the trend patterns is represented by ''g' in Equations. (11) and (12).
Upward Shift pattern:

$$y_i = \mu + r_i\sigma + ks, \; k = 1 if \mathrm{i} \geq \mathrm{P}, \quad (14)$$

Downward Shift pattern:

$$y_i = \mu + r_i\sigma - ks, \; k = 1 if \mathrm{i} \geq \mathrm{P}, \quad (15)$$

Parameter ''k' determines the location of the shift, while ''s' denotes the magnitude of the shift, in Equations (13) and (14). 9000 samples have been created [Input] $_{60\times9000}$. To demonstrate the effectiveness of the suggested system, a split is made between the training and testing phases, each using half of the data.

4.2. Evaluation of the success of the intended method

The effectiveness of the suggested method is studied in this section. The SVM architecture may be represented by the parameterized functions. We utilized the HS method to get the best possible settings for all of the parameters. The maximum number of repetitions is set at 100, while the population size (in terms of hawks) is restricted to 25. When utilizing the HS method, a SVM with five hidden layers and a knowledge rate of 0.00126 performs the best. Table 2 displays the best SVM ideal parameter settings.

Table 2 The suggested framework.

Layer	Layer type	Stride(S)	Zero padding (p)	No. of trainable parameters	Output shape	Kernel Size (F)	No.of Kernel (K)
0	CONV	3	3	961	7×21	5×2	21
	POOL	2	-	-	8×21	3×2	21
1	CONV	2	2	2561	8×33	5×2	33
	POOL	2	-	-	7×34	3×2	33
2	CONV	2	3	289	41×9	9×2	9
	POOL	3	-	-	20×9	5×2	9
3	CONV	2	2	385	19×13	5×2	13
	POOL	2	-	-	17×13	4×2	13
4		-	-	-	61×1	-	-
	CONV	2	2	49	52×4	13×2	5
5	POOL	2	-	-	47×4	7×2	5

There are sixty neurons in the input layer, and four additional convolution and pooling layers with four, eight, twelve, twenty, and thirty-two filters in the hidden layer. Which demonstrates the excellent classification accuracy of the proposed method (HS-SVM) using ELU activation functions. In this analysis, the efficiency of a Support SVM built on the same principles but using the ReLU activation function is calculated. The Table 3 below shows the outcomes of utilizing SVM with the ELU and ReLU activation functions.

Table 3 Examining the Impact of Activation Function on SVM Efficiency.

Classifier	Accuracy (%)			Standard Deviation	Activation function
	Min	Max	Mean		
SVM with optimal architecture	98.81	98.81	98.81	±0.1	ELU
SVM with optimal architecture	98.12	98.76	98.56	±0.15	ReLU

4.3. Analysis in contrast to other classifications

Several tests have been conducted to analyse the performance of the proposed method in comparison to existing machine learning methods. Probabilistic neural networks (PNN), multilayer perception neural networks (MLPNNs) using a variety of training methods (including “back propagation” (BP), “resilient propagation” (Rprop), and “Levenberg Marquardt” (LM), as well as ‘random forest” (RF), are all taken into account for this task. Different classifiers are fed the raw data in these tests, with the same training/testing split. Three important MLPNN setup options are the number of hidden layers, the knowledge rate, and the kind of transfer function. The effectiveness of an RBFNN is very sensitive to the amount of radial basis functions used and the distributions of those functions. Choosing a suitable spread value in PNN is a crucial step in the process. Analysing the suggested classifier against other classifiers on raw data and comparing their performance shows in Figure 2. Accurate identification of the radii value, membership function type, and fuzzy inference system is required for successful ANFIS implementation. For optimal performance in SVM, it is essential to choose a suitable kernel function, kernel parameters, and penalty parameters. The density of foliage has a significant effect on RF signal clarity and performance. To determine what those values should be, the HS method is employed in this experiment. Density of foliage has a significant effect on RF signal clarity and performance. To determine what those values should be, the HS method is employed in this experiment.

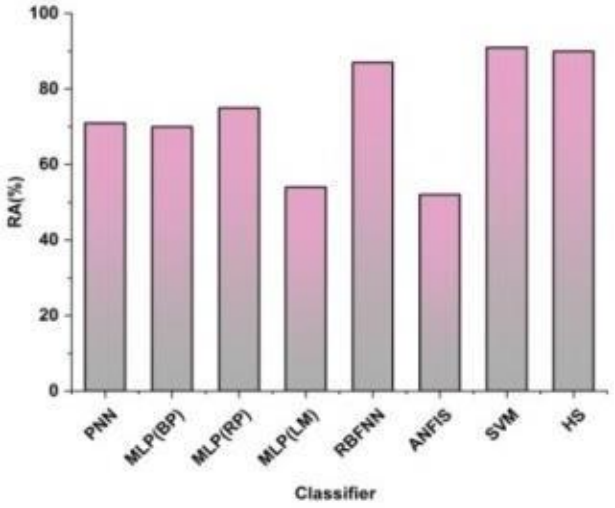


Figure 2 Using raw data to compare the suggested classifier's performance against others.

4.4. Analysing and contrasting the various methods presented in the research

One of the most common and useful tools for keeping an eye on output in many industries is the control chart. As a result, this strategy is widely used in the industrial sector, and a wealth of research has been conducted on the topic. The effectiveness of authors' methods has been studied across a variety of datasets. Since no single dataset exists under which all possible approaches can be evaluated, direct comparisons between them are impossible. Studies also vary in their sample sizes and the types of patterns they examine. Most studies on the control chart idea have focused on studying preexisting patterns, such as normal, cyclic, trend, and shift patterns. Because of this, classifying patterns into groups of eight or nine is outside the scope of the suggested approach. And because some of the feature extractions need for designer input, their CCP identification method is not really automated. For instance, the user must decide where to draw the line when dividing a pattern into two displays. The suggested approach achieving convergence over a variety of runs displays in Figure 3.

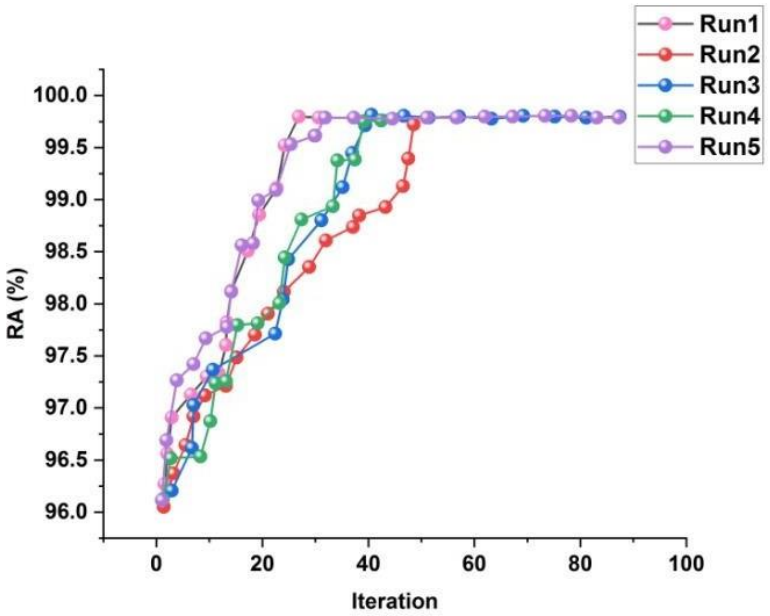


Figure 3 Consistency of results from several iterations of the suggested approach.

5. Conclusion

Given the cutthroat nature of the marketplace, quality control and monitoring have emerged as major concerns for manufacturers. In this research, an automated approach to CCPs recognition is suggested, one that makes use of SVM and an optimization technique. The suggested technique allows for the automated recognition of the nine most often observed CCPs without the need for any hand-crafted characteristics. Furthermore, the suggested approach may be used without modification to an unlimited number of CCPs. Multiple tests were conducted, and HS-SVM was compared to numerous

different techniques in order to evaluate its quality and performance. In these simulations, HS is used to determine the best configuration for a SVM and its associated parameters. The collected findings demonstrated the high accuracy of the suggested method and its ability to categorize the nine CCPs with a 99.80 percent success rate. The findings show that the suggested technique is better than the state-of-the-art classifiers, including MLPNN, RBFNN, ANFIS, RF, and SVM. In addition, the suggested strategy outperforms previously published findings in the literature in terms of classification accuracy. Based on the findings, this research strongly suggests using HS-SVM for CCPs identification. The suggested approach may also be used to classify heartbeats, identify breast cancer tumour types, etc., all of which are examples of complex pattern recognition issues.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Char DS, Abràmoff MD, Feudtner C (2020) Identifying ethical considerations for machine learning healthcare applications. The American Journal of Bioethics 20:7-17.

Chen Y, Dong C, Palanisamy P, Mudalige P, Muelling K, Dolan JM (2019) Attention-based hierarchical deep reinforcement learning for lane change behaviors in autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 0-0.

Choi SY, Cha D (2019) Unmanned aerial vehicles using machine learning for autonomous flight; state-of-the-art. Advanced Robotics, 33(6), pp.265-277.

Dang LM, Min K, Wang H, Piran MJ, Lee CH, Moon H (2020) Sensor-based and vision-based human activity recognition: A comprehensive survey. Pattern Recognition 108:107561.

Darwin B, Dharmaraj P, Prince S, Popescu DE, Hemanth DJ (2021) Recognition of bloom/yield in crop images using deep learning models for smart agriculture: a review. Agronomy 11:646.

Ferdowsi A, Challita U, Saad W (2019) Deep learning for reliable mobile edge analytics in intelligent transportation systems: An overview. ieee vehicular technology magazine 14:62-70.

Garcia R, Aguilar J, Toro M, Pinto A, Rodriguez P (2020) A systematic literature review on the use of machine learning in precision livestock farming. Computers and Electronics in Agriculture 179:105826.

Jahangir H, Gougheri SS, Vatandoust B, Golkar MA, Ahmadian A, Hajizadeh A (2020) Plug-in electric vehicle behavior modeling in energy market: A novel deep learning-based approach with clustering technique. IEEE Transactions on Smart Grid 11:4738-4748.

Khan FA, Gumaei A, Derhab A, Hussain A (2019) A novel two-stage deep learning model for efficient network intrusion detection. IEEE Access 7:30373-30385.

Loquercio A, Segu M, Scaramuzza D (2020) A general framework for uncertainty estimation in deep learning. IEEE Robotics and Automation Letters 5:3153-3160.

Miikkulainen R, Liang J, Meyerson E, Rawal, A, Fink D, Francon O, Raju B, Shahrzad H, Navruzyan A, Duffy N, Hodjat B (2019) Evolving deep neural networks. In Artificial intelligence in the age of neural networks and brain computing, pp. 293-312. Academic Press.

Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V (2020) Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. Medicina 56:455.

Rashid M, Khan MA, Alhaisoni M, Wang SH, Naqvi SR, Rehman A, Saba T (2020) A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection. Sustainability 12:5037.

Sejnowski TJ (2020) The unreasonable effectiveness of deep learning in artificial intelligence. Proceedings of the National Academy of Sciences 117:30033-30038.

She Q, Feng F, Hao X, Yang Q, Lan C, Lomonaco V, Shi X, Wang Z, Guo Y, Zhang Y, Qiao F (2020) Openloris-object: A robotic vision dataset and benchmark for lifelong deep learning. In 2020 IEEE international conference on robotics and automation (ICRA), pp. 4767-4773. IEEE.

Sun H, Burton HV, Huang H (2021) Machine learning applications for building structural design and performance assessment: State-of-the-art review. Journal of Building Engineering 33:101816.

Usama M, Qadir J, Raza A, Arif H, Yau KLA, Elkhatib Y, Hussain A, Al-Fuqaha A (2019) Unsupervised machine learning for networking: Techniques, applications and research challenges. IEEE access, 7, pp.65579-65615.

Wang C, Tan XP, Tor SB, Lim CS (2020) Machine learning in additive manufacturing: State-of-the-art and perspectives. Additive Manufacturing, 36, p.101538.

Xie Q, Li D, Xu J, Yu Z, Wang J (2019). Automatic detection and classification of sewer defects via hierarchical deep learning. IEEE Transactions on Automation Science and Engineering 16:1836-1847.

Zhang C, Patras P, Haddadi H (2019) Deep learning in mobile and wireless networking: A survey. IEEE Communications surveys & tutorials 21:2224-2287.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0108

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Spoken emotion recognition through human-computer interaction using a novel deep learning technology

Manju Bargavi S. K.[a] | **Pawan Bhambu**[b] | **Mohan Vishal Gupta**[c]

[a]Jain (deemed to be) University, Bangalore, India, Professor, Department of Computer Science and Information Technology.
[b]Vivekananda Global University, Jaipur, India, Associate Professor, Department of Computer Science and Engineering.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College of Computing Science and Information Technology.

Abstract The paradigm of textual or display-based control in human-computer interaction (HCI) has changed in favor of more understandable control methods, such as gesture, voice, and imitation. Speech in particular contains a large quantity of information, revealing the speaker's inner state as well as his or her goal and intention. The speaker's request can be understood through language analysis, but additional speech features show the speaker's mood, purpose, and intention. As a consequence, in modern HCI systems, emotion identification from speech has become crucial. Additionally, it is challenging to aggregate the results of the many professionals engaged in emotion identification. There have been several methods for analyzing sound in the past. However, it was impossible to analyses people's emotions during a live speech. Studies on real-time data are now more prominent than ever because of the advancement of artificial intelligence and the great performance of deep learning techniques. This research uses a cutting-edge deep-learning technique to identify emotions in human speech. The research made use of the open-source Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) dataset. More than 2000 fragments of data were captured by 24 performers as speeches and songs for the RAVDESS dataset. The actors' responses to eight distinct moods were recorded. It was designed to find various emotion classifications. In this study, a novel neuro-fuzzy swallow swarm-optimized deep convolutional neural networks (NFSO-DCNN) approach for classification was suggested. The performance of the suggested model was compared to that of similar research, and the outcomes were assessed. Employing the suggested example on the RAVDESS dataset, an overall accuracy of 98.5% was attained for categorizing emotions.

Keywords: HCI, RAVDESS, SER, NFSO-DCNN

1. Introduction

The technique or method of recognizing and comprehending emotions expressed through speech or spoken language is known as SER. To determine a speaker's emotional state, many auditory variables are analyzed, including pitch, quantity, language rate, and spectrum characteristics. SER systems propose to identify and describe different feelings people experience at the moment or using previously captured speech. Call centers, voice-controlled gadgets, mental health monitoring, virtual assistants, and market research are objective some of the many places these technologies may be placed into action (Nayak et al 2021). The term HCI is used to describe the research and development of systems. It includes all aspects of a person's interaction with an electronic device, program, website, or other digital medium. HCI is the study of how to design technologies to ensure they can be easy to use by the average person. HCI is a multi-disciplinary field that studies how people communicate with and make sense of technological systems. Its goal is to enhance the usability of a product or service by tailoring the user experience to the specific requirements of individual consumers. HCI encompasses a broad variety of digital devices, including tablets, virtual reality systems, wearables, smartphones, and IoT gadgets, in addition to conventional PCs. To improve the design and usability of digital interfaces, it aims to comprehend how consumers engage with these gadgets (Alnuaim et al 2022). HCI-based SER is a fast-developing topic that focuses on the creation of technology capable of comprehending and interpreting spoken expressions of human emotions. With the use of artificial intelligence and natural language processing, this ground-breaking field of study enables computers to recognize and react to the emotional information included in spoken words. Sophisticated algorithms may determine useful psychological information from a speech by analyzing its numerous acoustic and linguistic characteristics, including tone, intensity, pitch, and word selections. To find patterns and connections between audio variables and certain emotional states, these systems

use machine learning methods and train on vast datasets of labeled emotional speech (Rapp et al 2021). HCI, medical care, and other fields all benefit greatly from facial expression recognition. There are six primary feelings happiness, sadness, surprise, fear, and rage. They demonstrated that people from different backgrounds have the same emotional experience. Both valence and arousal are independent variables along which emotions may be described. Both the valence and the arousal levels are complex ranging from calm to excitement. They are sensitive enough to capture minute shifts in facial expression, allowing them to recognize broad categories of emotion but also differentiate between degrees of feeling within each. HCI, social robotics, market research, affective computing, psychological studies, healthcare, and the diagnosis and monitoring of mental health issues constitute only a number of the many fields that may benefit from facial expression recognition (Chen et al 2021). Speech emotion analysis also has important applications in the fields of customer service and market research. By gaining insight into consumers' feelings during interactions, organizations can adapt to their wants and requirements. Companies may learn a lot about their customers' contentment, preferences, and purchasing habits through the study of emotional patterns. With this information, businesses can create more satisfying products, implement more effective marketing methods, and better manage their relationships with customers (Karanchery and Palaniswamy 2021). LinkedIn, and YouTube. Integration of SMD with a healthcare.

2. Related Work

The study (Santhoshkumar and Geetha 2019) employed a feedforward deep convolution neural network architecture with varying parameters to determine an emotional state based on patterns of whole-body movements. The benefit of emotion detection based on body movements is that it may be used to determine a person's emotional state regardless of the angle from which the camera observes individuals. The research (Ren and Bao 2020)provided HCI and intelligent robots will face several significant obstacles, as this study also predicts. The study emphasizes the variety of tools that are presently available for reading, speaking, writing, and using additional senses in human connection. The research (Abbaschian et al 2021) conducted a comprehensive literature review on the topic of emotion recognition in discrete speech. Comparison of existing SER methods and records is essential for finding feasible options and getting a better grasp on that open-ended challenge, especially in light of recent developments in artificial neural networks and the constant need for correct and close to real-time SER in HCI. The article (Tsiourti et al 2019) examined the way people recognize and react to emotions expressed by the physical appearance and speech of humanoid robots, with a focus on the impacts of incongruence. They do this by integrating current discoveries from psychology, neurology, HCI, and HRI. The role of humanoid social robots in modern society is growing. The research (Pustejovsky and Krishnaswamy 2020) provided a simulation environment for the study and development of Embodied HCI (EHCI). VoxWorld is a multilingual conversation platform that facilitates focused task conversations through the use of voice, movement, gesture, expressions on the face, and gaze detection. The study (Yun et al 2021) presented a novel graphical method for making business-critical decisions. HCI is increasingly important in high-tech systems including brain-machine interfaces, human action recognition, telemedicine, and somatosensory games. Extensive testing demonstrates the superiority of the suggested strategy over competing approaches. The research (Wu et al 2020) enhanced manufacturing effectiveness and flexibility while also realizing the benefits of the multi-variety and small-batch assembly via direct interaction among technology and humans. The machine vision technique explored in this article can successfully filter out background noise and produce the desired picture. The study (Al Mahdi et al 2019) investigated how HCI factors into the development of e-learning platforms. Graphics has several important uses in the field of education technology. The field of higher education greatly benefits from the use of interactive multimedia, often known as Hypermedia. The research (Tsai et al 2020) offered a cheap HCI system that can recognize hand gestures. Multiple forms of vision are included in this system. To separate the region of interest from the backdrop, skin and motion detection are employed. To find the object's midpoint, an approach called linked component labeling is presented. The study (Qi et al 2019) focused on optimizing the time differences in surface electromyogram (sEMG) pattern recognition. HCI shows an essential role in bridging the gap in the implementation of information technology in contemporary cities, serving as the interface between people and smart cities. The research (Xu et al 2021) provided a novel method for improving SER accuracy has been identified, and it's been named head fusion. This method takes advantage of the multiple attention heads mechanism. They introduce sounds of differing intensities, alter the noises over time, and combine different types of noise to see how well our model holds up under stress. The study (Chattopadhyay et al 2023) approached assistance in decreasing the feature dimension but also boosts the learning model's classification precision. In addition to its central position in human communication, speech is also the primary information exchange channel in HCI. The study (Atmaja and Akagi 2021) suggested a two-stage late-fusion technique for fusing acoustic and textual information. It's important to note that deep learning algorithms first train audio and text characteristics independently. The second step involves a support vector machine (SVM) that predicts the final regression score based on the outcomes of the deep learning systems' predictions. The paper (Li et al 2021) developed a compact network model for instantaneous emotion classification, to identify emotions in students' faces in real-time. The HCI has benefited greatly from the development of the detection of facial expressions, which has been boosted by the use of deep learning technology. The research (Heracleous et al 2020) identified multilingual speech emotions using authentic emotional speech that has been taken from English, Italian, and Spanish movies. The information

fusion-based approach resulted in a 73.3% unweighted average recall (UAR). This outcome is encouraging and outperforms the UAR determined by evaluation by humans.

3. Proposed Methodology

In this paper, we propose an innovative NFSO-DCNN technique for categorization. The suggested block diagram is shown in Figure 1.

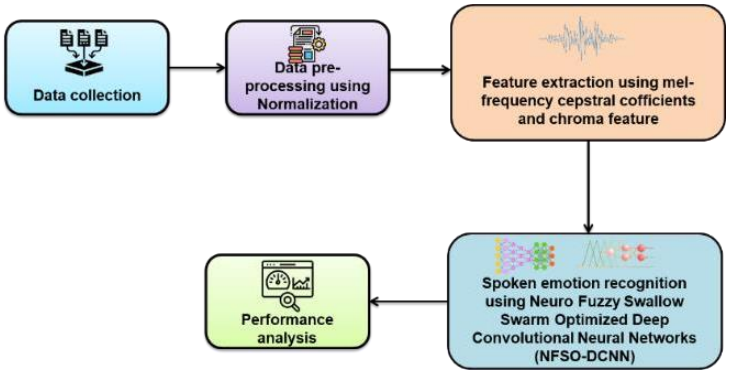


Figure 1 Block diagram of proposed.

3.1. Data collection

The initial dataset for the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) consists of 7356 files, including video and audio recordings of speeches and songs. It was chosen to utilize the Kaggle reduced version of the dataset since our study relies on voice and the initial data set size is rather enormous at 24.8 GB. All of the audio files in this edition of the dataset are 16-bit, 48-kilohertz.wav speech files. There are a total of 1440 voice samples in this dataset, with preparations made by 24 distinct actors and 12 male and 12 female changers. There are 60 iterations for each actor, resulting in 1440 files rather than 7356. Due to the absence of video content, the information amount is reduced to a more manageable 590 MB. To enhance efficiency, song files are additionally incorporated into our application. The song has 1012 files and encompasses feelings of fear, sadness, happiness, calmness, and anger. Actor 18's song file is missing. Therefore, there are 2452 audio recordings of emotions in the collection. The dataset was built using two neutral North American-accented speeches to reduce the impact of words and increase the focus on emotions. The name of files also includes identifying codes for things like modality, voice channel, emotion, emotional intensity, statement, repetition, and actor.

3.2. Pre-processing of Normalization

The goal of the field of study known as HCI is to create machines that can identify and interpret human emotions from speech. Emotion identification relies heavily on the pre-processing of speech signals for accuracy and reliability. Collecting high-quality speaking data is the starting point of the pre-processing phase. To accomplish this, samples of spoken language must be recorded using suitable equipment in sterile settings free of distractions. To guarantee the system's efficacy in a wide variety of settings, the data-gathering procedure must contain a spectrum of feelings shown by individuals from many different groups, cultures, and language backgrounds. To minimize differences across speakers and recordings without compromising the feature's capacity to discriminate, feature normalization is a critical process to do. To improve the features' capacity to generalize, feature normalization is used. Both the function and the corpus levels are available for normalization. Among normalization procedures, z-normalization is the most common. The formula for z-normalization, given the means μ and standard deviations of the data, is:

$$z = \frac{x-\mu}{\sigma} \qquad (1)$$

3.3. Feature Extraction by using Mel-Frequency Cepstral Coefficients (MFCC) and Chroma Feature

Feature extraction makes extensive use of the NumPy libraries, Libros, and Panda. Mel spectrum pictures were also acquired for use with a convolutional neural network; however, the outcomes were insufficient for processing. Images of spectra are not processed, but rather one-dimensional characteristics are collected.

3.3.1. Mel-Frequency Cepstral Coefficients (MFCC)

A speech recognition system's initial stage is feature extraction, which reveals the parts of an audio signal most suited for recognizing language while filtering out irrelevant parts (such as emotion, and noise). MFCC is among the popular methods for extracting features. The audible path falls within the scope of the short-term power spectrum. MFCCs are used

to represent this limit. MFCCs were first created by Davis and Murmelstein in the 1980s for use in the automatic identification of speakers and recognition. The Mel-frequency Cepstrum is a linear cosine transform of a log power spectrum on a nonlinear Mel-frequency scale that represents the sound's short-term power spectrum. The cepstral representation of the audio sample is included in the Mel-frequency cepstral coefficients. The MFCC differs from a standard Cepstrum in that its frequency bands are equally separated on the Mel scale, simulating the regularly separated bands of frequencies observed in a conventional spectrum, more closely resembling how the human auditory system responds.

3.3.2. Chroma Feature

Chroma features are an interesting and powerful visualization of musical sound, in which the spectrum's characteristics are displayed in 12 separate boxes that correspond to the 12 halftones that make up a musical octave. It is possible to successfully classify musical instruments and sounds according to pitches using chroma-based features, commonly known as pitch class profiles. The harmonic content of the limited audio frame is intended to be reflected in particular through chromatic qualities. Chroma characteristics are connected to the concept of harmony in music and could be very resilient despite changes in volume and pitch. The chroma feature vector, constant-Q transforms (CQT), chroma energy normalized statistics (CENS), and short-time Fourier transform (STFT) are some of the methods used to analyze music.

3.4. Neuro-Fuzzy Swallow Swarm Optimized Deep Convolutional Neural Networks (NFSO-DCNN)

The term neural-fuzzy describes a hybrid computing framework that combines the benefits of neural networks with fuzzy logic systems. While neural networks are great machine learning models for recognizing patterns and making approximations, fuzzy logic systems are great at handling uncertainty and imprecision in data. In a neuro-fuzzy system, the language variables and fuzzy sets used to describe and modify uncertain or subjective information are managed using a rule-based method provided by the fuzzy logic system. Fuzzy logic is a method of thinking and making decisions based on incomplete or ambiguous data that allows for the modeling of fuzzy concepts.

The swallow swarm drives the main concept of our innovative optimization approach. This method uses three different kinds of particles.

- Explorer particle (ei)
- Aimless particle (oi)
- Leader particle (li)

These particles are constantly interacting with one another as they travel in parallel directions:

3.4.1 Explorer particle

The exploration particles include the majority of the colony's people. Their primary role is to investigate the issue domain. If this location is the most important one in the issue space, this particle acts as a Head Leader (HLi) by simply directing the group there with the use of a different sound. However, when a particle is beneficially situated in comparison to its surrounding particles, this is known as a local leader (LLi); else ei valuing VHLi, VLLi, and experience of the opposition of both of these carrying creates an arbitrary move.

3.4.2 Aimless particle

The particulates do not initially have a favourable location for different particles, and the magnitude they have f (o_i) is poor. Their objective is to conduct a haphazard and enquiring search. They proceed to move at random and were unrelated to HL_i and LL_i's positions.

3.4.3 Leader particle

In the SSO algorithm, there are a few particles with the designation Leader. In a colony, the most powerful leader is called the Leader Head, and there are numerous smaller leaders known as Local Leaders.

$$V_{HLi+1} = V_{HLi} + \alpha_{HL} rand\ ()(e_{best} - e_i) + B_{HL} rand\ ()(HL_i - e_i) \quad (2)$$

$$V_{LLi+1} = V_{LLi} + \alpha_{LL} rand\ ()(e_{best} - e_i) + B_{LL} rand\ ()(LL_i - e_i) \quad (3)$$

Where, V_{LL} = velocity of a local leader, V_{HL} =Velocity of head leader, e_{best}= best position of the explorer particle, e_i= current location of the explorer particle

Recalculate the speed using the current formula,

$$V_{i+1} = V_{HLi+1} + V_{LLi+1} \quad (4)$$

The particle's worth is evaluated as:

$$e_{i+1} = e_i + V_{i+1} \quad (5)$$

In the field of artificial intelligence, DCNNs are a special kind of neural network developed for handling and analyzing visual input like photographs and movies. Convolutional Neural Networks (CNNs) are a kind of artificial neural network that performs complicated tasks such as pooling and convolution over several layers of data. DCNNs are an example of multilayer neural networks that typically have an input layer, a convolutional layer, a pooling layer, and an output layer. Two such layers that are concealed are the layer of convolution and the pooled layer. The relationship between layers in a DCNN is an explanation of the propagation that occurs between its input layers to the output layer. To avoid the drawbacks of a model based on linearity, it is necessary to enhance the activity of neurons in every group using a measure of irregular activity in the forward process. There are no activation functions in the first layer since it merely gets pronounced dimensions from the data. Starting with the second layer's use of nonlinear functions for activation, they may write down the expression for the lth layer's output as follows.

$$\left.\begin{aligned} z^l &= W^l * x^{l-1} + b^l \\ a^l &= \sigma(z^l), \end{aligned}\right\} \qquad (6)$$

Where l is the layer number and the process of convolution * is symbolized by a symbol. For l=2, the visual vector is x^{2-1} = x^1, and for l > 2, the generated characteristic mapping vector for the (l2)th layer is $x^{l-1} = a^{l-1} = \sigma\ (z^{l-1})$. Load vector W^l, biases vector b^l, and balanced input z^l all pertain to the l_{th} layer, whereas is the σ neural stimulation value. If Layer L is the ultimate exit layer, then a^L is the true production vector.

Parameters of W^l and b^l are often repeatedly updated using the backpropagation (BP) algorithm, an automated learning approach. It constructs a cost estimator from the observed and expected amounts and subsequently employs gradient descent (GD) to adjust the parameters in the path of a negative gradient of the expense equation. Here's how it goes down step-by-step in detail.

3.4.4 Cost Function Selection:

Error cost functions are often chosen as quadratic functions. It is possible that correcting a mistake made by the neurons during DCNN training would take some time. As a result, instead of using a quadratic function as the error cost function, we use cross entropy (E_0^L). Where N is the number of neurons in the result layers and n is the total number of training sets. As a result, DCNN is eventually separated into N classes. The kth output layer neuron's actual output value is represented by a_K^L, whereas its desired value is represented by t_K^L.

$$E_0^L = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{N}[t_k^L\, In\, a_k^L + (1 - t_k^L) In\, (1 - a_k^L)] \qquad (7)$$

3.4.5 Calculation of Error Vectors:

The error vector for the kth neuron in the product stage has been established for all layers and is represented as follows:

$$\delta^L = \frac{\partial E_0^L}{\partial z^l} \qquad (8)$$

The back propagation (BP) Procedures, δ^L used to make reversible deductions $\delta^L - 1$. Correspondingly, assume δ^L and δ (l + 1) is the error vectors of the lth and (l+ 1) th layers respectively. Then, according to Equations 1 and 3 the chain rule, δ^l is written as follows:

$$\delta^1 = W^{1+1} + \delta^{1+1} \odot \sigma'(z^l), \qquad (9)$$

The gradients of W^l and b^l are $\frac{\partial E_0^L}{\partial W^l}$ and $\frac{\partial E_0^L}{\partial b^l}$ correspondingly. ∂ (·) stands for the action of partial derivatives. Reduced Equivalent of (E_0^L) to $W^l$and b^l can be proportional to Equations 6 and 8

$$\left.\begin{aligned} \frac{\partial E_0^L}{\partial W^l} &= \frac{\partial E_0^L}{\partial a^l} \odot \frac{\partial a^l}{\partial W^l} = \delta^1 \odot x^{l-1} \\ \frac{\partial E_0^L}{\partial b^l} &= \frac{\partial E_0^L}{\partial a^l} \odot \frac{\partial a^l}{\partial W^l} = \delta^1 \end{aligned}\right\} \qquad (10)$$

The revised estimates of $W^l$and b^l are characterized by ΔW^l and Δb^l, and the optimal solution is obtained by reducing the sum of all possible jobs in each scenario.

$$\left.\begin{aligned} \Delta W^l &= -\eta \cdot \frac{\partial E_0^L}{\partial W^l} \\ \Delta b^l &= -\eta \cdot \frac{\partial E_0^L}{\partial b^l} \end{aligned}\right\} \qquad (11)$$

Where η denotes the learning rate

4. Performance Analysis

4.1. Results

In this part, the suggested system's effectiveness is evaluated. The performance indicators used for assessment are accuracy, precision, recall, f1-measure, and efficiency. Classical Support Vector Machine (C-SVM), Fully Connected (FC), and Decision Tree (DT) are the existing methods used for comparison.

4.1.1. Accuracy

A difference between the result and the true number is caused by inadequate precision. The percentage of actual outcomes reveals how balanced the data is overall. Accuracy is assessed using an equation (12).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (12)$$

Figure 2 shows the comparable values for the accuracy measures. When compared to existing methods like C-SVM, which has an accuracy rate of 94.31%, DT, which has an accuracy rate of 95.09%, and FC, which has an accuracy rate of 97.58%, the recommended method's NFSO-DCNN value is 99.3%. The suggested NFSO-DCNN performs with higher accuracy than other methods. Table 1 displays the proposed method's accuracy.

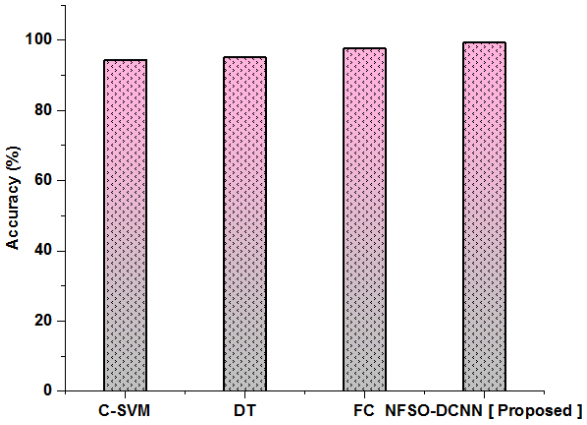


Figure 2 Accuracy comparisons between the suggested and current approaches.

Table 1 Comparison of Accuracy.

Methods	Accuracy (%)
C-SVM	94.31
DT	95.09
FC	97.58
NFSO-DCNN [Proposed]	99.3

4.1.2. Precision

The most crucial standard for accuracy is precision, it is clearly defined as the percentage of properly categorized cases to all instances of predictively positive data. Equation (13) is used to compute the precision.

Comparable values for the precision measures are shown in Figure 3. This proves the suggested strategy may provide performance results that are superior to those obtained by the current study methods. The precision of the proposed approach is 99.42%, which performs better than existing outcomes. Include C-SVM, DT, and FC precision rates are 95.2%, 95.94%, and 98.05%. Table 2 shows the precision of the suggested method is contrasted with the existing methods.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (13)$$

Table 2 Comparison of Precision.

Methods	Precision (%)
C-SVM	95.2
DT	95.94
FC	98.05
NFSO-DCNN [Proposed]	99.42

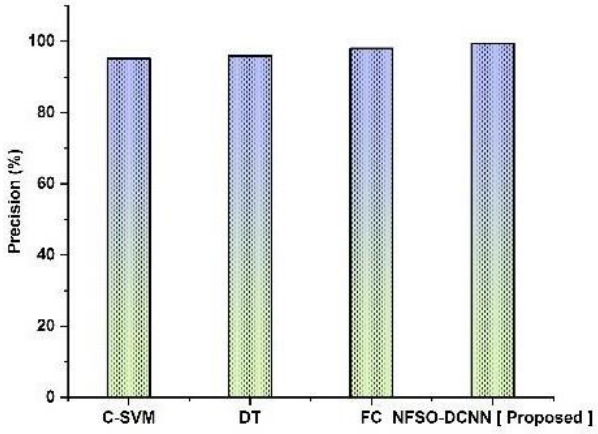


Figure 3 Precision comparisons between the suggested and current approaches.

4.1.3. Recall

The potential of a model to identify each important sample within a data collection is known as recall. The percentage of TPs divided by the sum of True Positive and False Negative is how it is statistically defined. The recall is calculated using equation (14).

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (14)$$

Figure 4 shows the comparative data for the recall metrics. Recall rates for C-SVM are 94.42%, DT 95.93%, FC 97.99%, and NFSO-DCNN 99.5%. The proposed method performed better than the current results with a recall of 99.5%. Table 3, the recall of the suggested method is contrasted with the existing methods.

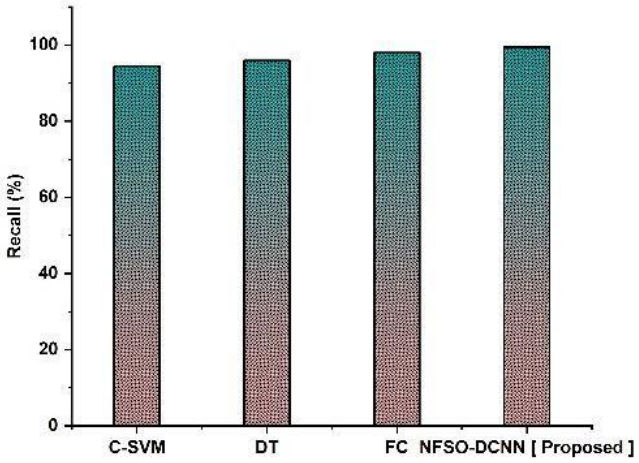


Figure 4 Recall comparisons between the suggested and current approaches.

Table 3 Comparison of Recall.

Methods	Recall (%)
C-SVM	94.42
DT	95.93
FC	97.99
NFSO-DCNN [Proposed]	99.5

4.1.4. F1-Score

The F1-measure is often used while assessing information. It is possible to alter the F1-measure so that accuracy is prioritized above recall, or vice versa. The recommended technique has a higher level of F1-measure when measured against the currently used methods. In Figure 5 the F1-measure of the suggested method is contrasted with the traditional methods.

$$F1 = \frac{2 * (precision * recall)}{(precision+recall)} \qquad (15)$$

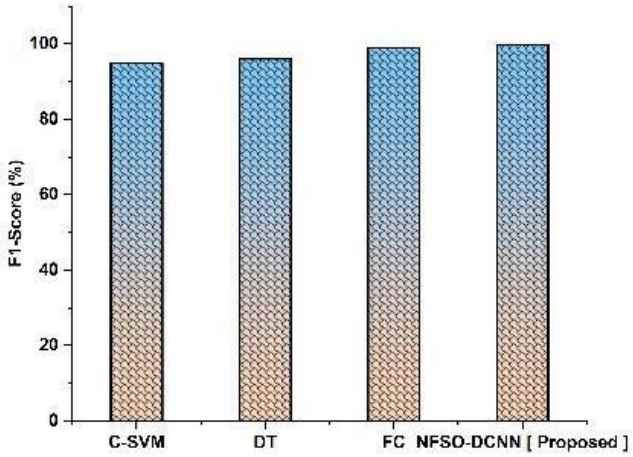


Figure 5 F1-Score comparisons between the suggested and current approaches.

When compared to existing methods like C-SVM, which has an F1-measure of 94.81%, DT, which has an F1-measure of 95.96%, and FC, which has an F1-measure of 98.82%, the recommended method's NFSO-DCNN value is 99.67%. The suggested (NFSO-DCNN) performs with higher accuracy than other methods. Table 4 displays the proposed method's F1-Score.

Table 4 Comparison of F1-Score.

Methods	F1-Score (%)
C-SVM	94.81
DT	95.96
FC	98.82
NFSO-DCNN [Proposed]	99.67

5. Conclusion

In conclusion, while great progress in SER using HCI; additional research is required to raise the reliability and sturdiness of emotion recognition models. To effectively achieve the promise of this technology and to ensure its appropriate and ethical usage, it will be essential to address issues relating to subjectivity, cultural diversity, delicate emotions, and privacy. This study aims to enhance the effectiveness of a machine learning model using the Ryerson Audio-Visual Database of Emotional Speech and Song, a speech dataset. There are a total of 2452 audio recordings in the collection, covering a range of emotions from happy to sad. In this approach, feature extraction takes up the majority of the time, taking longer than model training. The MFCC features' mean is determined. As a result, we introduced the NFSO-DCNN for the recognition of spoken emotion. Performance metrics like accuracy, precision, Recall, efficiency, and F1-measure, are evaluated and compared with existing technologies like FC, C-SVM, and DT. SER in the context of HCI may be improved by using NFSO-DCNN. The accuracy, robustness, and flexibility of emotion detection systems may be improved by using the optimization skills of FSSO and the characteristic learning capabilities of DCNNs, allowing more efficient and customized interactions between people and computers. These new research avenues will help the field advance and realize its full potential in a variety of applications.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Abbaschian BJ, Sierra-Sosa D, Elmaghraby A (2021) Deep learning techniques for speech emotion recognition, from databases to models. Sensors 21:1249.

Al Mahdi Z, Rao Naidu V, Kurian P (2019) Analysing the Role of Human Computer Interaction Principles for E-Learning Solution Design. In Smart Technologies and Innovation for a Sustainable Future: Proceedings of the 1st American University in the Emirates International Research Conference - Dubai, UAE 2017, pp. 41-44. Springer International Publishing.

Alnuaim AA, Zakariah M, Shukla PK, Alhadlaq A, Hatamleh WA, Tarazi H, Sureshbabu R, Ratna R (2022) Human-computer interaction for recognizing speech emotions using multilayer perceptron classifier. Journal of Healthcare Engineering 2022.

Atmaja BT, Akagi M (2021) Two-stage dimensional emotion recognition by fusing predictions of acoustic and text networks using SVM. Speech Communication 126:9-21.

Chattopadhyay S, Dey A, Singh PK, Ahmadian A, Sarkar R (2023) A feature selection model for speech emotion recognition using clustering-based population generation with hybrid of equilibrium optimizer and atom search optimization algorithm. Multimedia Tools and Applications 82:9693-9726.

Chen X, Cao M, Wei H, Shang Z, Zhang L (2021) Patient emotion recognition in human-computer interaction system based on machine learning method and interactive design theory. Journal of Medical Imaging and Health Informatics 11:307-312.

Heracleous P, Mohammad Y, Yoneyama A (2020) Integrating language and emotion features for multilingual speech emotion recognition. In Human-Computer Interaction. Multimodal and Natural Interaction: Thematic Area, HCI 2020, Held as Part of the 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22, pp. 187-196. Springer International Publishing.

Karanchery S, Palaniswamy S (2021) Emotion recognition using one-shot learning for human-computer interactions. In 2021 International Conference on communication, control and information sciences (ICCISc), Vol. 1, pp. 1-8. IEEE.

Li Q, Liu YQ, Peng YQ, Liu C, Shi J, Yan F, Zhang Q (2021) Real-time facial emotion recognition using lightweight convolution neural network. In Journal of Physics: Conference Series 182:012130. IOP Publishing.

Nayak S, Nagesh B, Routray A, Sarma M (2021) A Human–Computer Interaction framework for emotion recognition through time-series thermal video sequences. Computers & Electrical Engineering 93:107280.

Pustejovsky J, Krishnaswamy N (2020) October. Embodied human-computer interactions through situated grounding. In Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents 1-3.

Qi J, Jiang G, Li G, Sun Y, Tao B (2019) Intelligent human-computer interaction based on surface EMG gesture recognition. Ieee Access, 7, pp.61378-61387.

Rapp A, Curti L, Boldi A (2021) The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. International Journal of Human-Computer Studies 151:102630.

Ren F, Bao Y (2020) A review on human-computer interaction and intelligent robots. International Journal of Information Technology & Decision Making 19:5-47.

Santhoshkumar R, Geetha MK (2019) Deep learning approach for emotion recognition from human body movements with feedforward deep convolution neural networks. Procedia Computer Science 152:158-165.

Tsai TH, Huang CC, Zhang KL, (2020) Design of hand gesture recognition system for human-computer interaction. Multimedia tools and applications 79:5989-6007.

Tsiourti C, Weiss A, Wac K, Vincze M (2019) Multimodal integration of emotional signals from voice, body, and context: Effects of (in) congruence on emotion recognition and attitudes towards robots. International Journal of Social Robotics 11:555-573.

Wu S, Wang Z, Shen B, Wang J H, Dongdong L (2020) Human-computer interaction based on machine vision of a smart assembly workbench. Assembly Automation 40:475-482.

Xu M, Zhang F, Zhang W (2021) Head fusion: Improving the accuracy and robustness of speech emotion recognition on the IEMOCAP and RAVDESS dataset. IEEE Access 9:74539-74549.

Yun Y, Ma D, Yang M (2021) Human–computer interaction-based decision support system with applications in data mining. Future Generation Computer Systems 114:285-289.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0109

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Efficient mixed-type wafer defect pattern recognition using a novel joint attention skipped graph convolutional neural network

Gobi Natesan[a] | **Zahid Ahmed**[b] | **Priyank Singhal**[c]

[a]Jain (deemed to be) University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology
[b]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science & Application.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College Of Computing Science And Information Technology.

Abstract: For wafer foundries to maintain high yield and quality, errors in manufacturing wafers must be identified and corrected. To find the source of problems and increase overall yield, defect pattern recognition (DPR) is essential. Nevertheless, mixed-type DPR, which includes identifying numerous types of flaws at once, poses more difficulties because of the variety of spatial features, the unpredictability of flaw patterns, and the fluctuating amount of flaws present. So, a novel joint attention skipped graph convolutional neural network (JAS-GCNN) technique is suggested in this paper to analyse the difficulty of mixed-type DPR. The JAS- GCNN effectively captures locally as well as globally interconnections among faults on a wafer by fusing the strength of GCNN along with focused processes. The JAS-GCNN facilitates the learning of complicated fault patterns by allowing information to propagate across multiple layers by utilising skip connections. We do trials on the MixedWM38 WaferMap dataset, which contains a variety of defect types, to assess the proposed technique. The outcomes show that when compared to cutting-edge techniques, the JAS-GCNN performs better. It successfully detects various faults and performs defect classification jobs with high accuracy.

Keywords: Wafer foundries, manufacturing, DPR, JAS-GCNN

1. Introduction

In the industry associated with semiconductor production, effective mixed-type wafer defect pattern detection is essential (Piao 2023). The identification and categorization of flaws on semiconductor wafers have gotten harder as electronic circuitry continues to get smaller as well as more complicated. Wafer flaws may result in reduced product quality, a lower production yield, and greater expenses. To ensure the efficiency and dependability of semiconductor equipment, it is crucial to create fast and precise flaw pattern identification systems (Wang et al 2023). Problems in the production of semiconductors can come from various factors, including contaminants in the raw materials, processing variances, device problems, and employee mistakes. Particles, damage, pollution, pattern aberrations, and electrical malfunctions are just a few of the ways these flaws might appear. Errors can also display a variety of traits, including dimension, form, color, appearance, and location in space (Bi et al 2023). The presence of different defective categories on the same wafer or in a specific area is called the mixed-type nature of wafer defects. That renders the flaw identification technique more difficult because it may not be possible for one flaw identification method to correctly identify and categorize every type of problem (Ansari et al 2023). Improved computer vision, machine learning, and other algorithms must be used along with numerous methodologies for the successful identification of mixed-type faults. Identification of the characteristic features of mixed-type wafer defects has advanced significantly in the past few decades (Lin et al 2023). Novel ways that integrate deep neural networks with conventional methods for image processing have been created by academics and technologists. These methods use neural network power to extract relevant characteristics from semiconductor imagery and accurately diagnose faults

In summary, by efficiently identifying and correcting production-process problems, the JAS-GCNN technique shows the potential in improving the yield and quality of wafer fabrication. By addressing the difficulties associated with mixed-type DPR, it offers wafer foundries a useful tool to enhance their defect identification and rectification procedures, resulting in a greater overall yield and better product quality.

The rest of this paper is as follows part 2: a literature review, part 3 contains the proposed method explained; Part 4 includes the results and analysis; and Part 5: discusses the conclusion.

2. Related Works

Paper (Jin et al 2019) offered an innovative structure for wafer bin map (WBM) defective pattern recognition and categorization using grouping. The study introduces the DBSCANWBM framework, a revolutionary DBSCAN-based fault feature identification and categorization system. In most defective pattern identification analyses, finding outliers and imperfect recognition of patterns are considered separately. Not all aberrations that are seen are deleted, as outliers in each kind of design can provide varied essential data. The study (Yu et al 2019) proposed a deep learning model with a hybrid approach that can extract useful unfair characteristics from wafer maps using a deep network structure. To improve wafer map pattern recognition (WMPR) in semiconductor manufacturing, that work introduces a novel learning technique called stacking convolution sparse delousing auto-encoder (SCSDAE). On the model and real-word wafer map databases, SCSDAE outperforms those regarding identifying outcomes.

A study (Saqlain et al 2019) identified wafers mapping fault structures in the semiconductor manufacturing process; they propose an electoral group classifier in that study. In their research, they put forth multi-type ensemble-based classifiers for recognizing wafer map defect patterns, and they tested them on real data. To aggregate individual outcomes and get an overall group classification result, the weighted averaging or gentle voting ensembles technique was applied. Paper (Wang et al 2019) proposed mixed-type Defect deformable convolutional network (DCN) pattern recognition (MDPR), where multiple kinds of imperfections are combined in a wafer. After choosing data from mixed faults, an adaptable convolutional unit is developed for identifying characteristics of outstanding standards from wafer maps. Multi-label layers of results are improved by a rapid decoding method that isolates and extracts composite features into each essential defect. The study's findings demonstrate that the research DCN model exceeds other machine learning simulations, both established and rival. A study (Ko and Koo 2023) proposed classification neural network technique's feed will be a new finite-dimensional array. The principle of persistent homologous from the topology study of data was used to present an original approach for categorizing wafer defect patterns. The enhancement of that processing technique to make the generated vectors compatible with the uncontrolled clustering methods is an intriguing area for additional study.

The article (Jeong et al 2023) proposed a brand-new rotational- and flipping-invariant approach predicated on the identifying tenet that the wafer map imperfection structure has no bearing on the revolving and flipping of labels, resulting in classes discriminating performance under conditions with limited information. The technique achieves topological consistency by using a convolutional neural network (CNN) backbone along with a Spectral translation and kernel flip. Study (Jeong et al 2023) achieved category-discriminating effectiveness in scenarios with limited data. They offer a unique rotational- and flipping-invariant depending on method on the identify internet that the wafer map flaw patterns do not influence the revolving and flip of labels. The technique incorporates a kernel flip, Fourier changes, and a convolutional neural network (CNN) framework to achieve geometrical invariance. For translation-invariant CNNs, the Specter element is a rotation- equivariant link-up, and the kernel flip module makes the algorithm flip-invariant. Through numerous different types of testing, they confirmed their methodology. To correctly explain the model choice during qualitative analysis, the proposed multi-branch layer-wise significance dissemination. By successfully guaranteeing rotational and flip invariance, the suggested model provides good performance for detection, even in scenarios with little available data. Study (Shih et al 2020) provided a novel method for expressing the geometry of the defective as a finite-dimensional vector, pattern, the results which will be fed into an identification neural network method. The major goal is to use topological data analysis (TDA) and the notion of persisting homologous to identify the topological characteristics of each arrangement. Through certain tests using a synthetic dataset, they demonstrate that the suggested method outperforms the most popular method for classifying wafer map defect patterns, which uses convolutional neural networks (CNN) regarding training speed and efficiency with greater precision.

Research (Kim et al 2020) provided a technique for predicting wafer aspect yield using a emulate that integrates long-term and short-term memory (LSTM) and a “feed-forward neural network (FFNN)”. In contrast to earlier studies, they concentrate on the aspect yield due to the larger produce loss near the wafer aspect yield. They just offered a model for predicting edge yield. Therefore, it is important to investigate an analytical expansion of total yield forecasting. The study outcomes demonstrated that in assessment measures, the network of neurons beat the other regression approaches. Study (Shawon et al 2019) creation of multilayer neural networks is an improved version of earlier researchers' work. A single wafer map can represent many pattern flaws in region-based modeling. Geographical pattern analysis has employed K-nearest neighbor classification. They have demonstrated an in-depth CNN topology with data enhancement that solves the issue of information disparity and outperforms earlier research methods.

3. Proposed Method

3.1 Dataset

The Dataset is the basis for our study. There is 31K wafer maps, including pictures created naturally and artificially. The wafer map's form may vary due to design heterogeneity. So, we crop all wafer photos to a 6125 size. Depending on the

31 classifications for the wafer maps, a wafer map's kind and quantity of faults. Thirty-one categories total; the normal class has no responsibilities, eight categories have a single distinct defect, and the rest categories have different individually acknowledged mixtures of individual faults.

3.2 Preprocessing

The median filtering method is frequently used in wafer map processing to improve the accuracy of the information. Every pixel in the wafer map is replaced by the median value of its neighbours as part of the median filter's operation. With the help of this method, noise and outliers are significantly reduced, leading to a smoother and more precise wafer map description. Extreme fluctuations and artefacts can be removed using the median filter, making it possible for subsequent evaluation and fault detection techniques to function more dependably. The median filtering method is particularly useful for wafer map preparation and quality improvement since it can maintain edges and fine features.

3.3 Joint Attention Skipped Graph Convolutional Neural Network (JAGCNN)

A cutting-edge method for identifying mixed-type wafer flaw patterns Joint Attention Skipped Graph Convolutional Neural Network (JAGCNN). To efficiently analyse and categorize wafer flaws, our cutting-edge neural network model blends the strength of convolutional networks with graphs with focus processes. Wafer defect patterns are frequently intricate and varied, with flaws like scuffs, particulates, and abnormalities. Conventional convolutional neural networks (CNNs) have a grid-like layout that ignores the fundamental graph structure of the flaws, making it difficult for them to accurately represent the complex interactions between these faults. By including graph convolutional components, which enable the modelling of graphs-structured information, the JASGCNN gets over this restriction. The spatial interactions between adjacent faults are captured by the lines connecting each defect's nodes in a graph.

The JASGCNN can efficiently capture the connections and interactions among various defect kinds, improving recognition performance. This is accomplished by utilizing graph convolutional layering. The JASGCNN also includes a system for attention that, while classifying defects, dynamically balances the significance of various flaw nodes. The network can concentrate on the most important faults and reduce the impact of unimportant or noisier errors thanks to this focusing technique. The JASGCNN strengthens the model's unfair ability and increases its capacity to precisely identify mixed-type defect patterns by paying attention only to informative defect nodes. Overall, the JASGCNN is a powerful and reliable method for wafer defect pattern identification. It efficiently models the graph structure of mixed-type defect patterns by utilizing the advantages of graph convolutional networks and attention processes, which improves classification accuracy and resilience in practical wafer analysis tasks.

4. Result

Accuracy is a metric for how well a model forecasts the results or labels of a specific dataset. It is computed by dividing the total number of forecasts made by the number of accurate predictions, commonly expressed as a percentage. Figure 1 depicts the accuracy outcome. The value obtained using our suggested approach (JAS-GCNN-) is superior to that of the currently used methods (AGGCN-, GCN-). This demonstrates that our recommended strategy successfully identifies a Mixed-Type Wafer Defect Pattern.

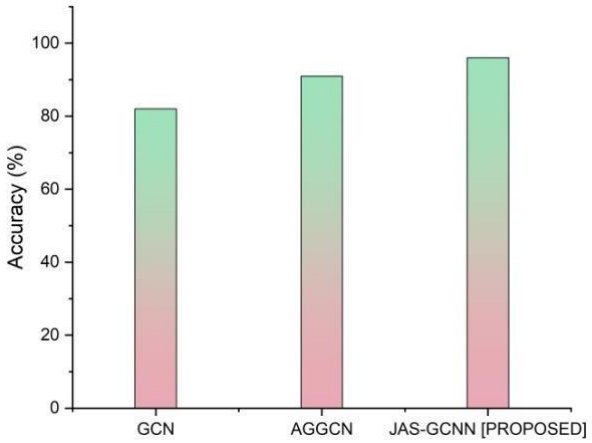


Figure 1 Outcome of Accuracy.

A statistical parameter called precision assesses the categorization or prediction model's accuracy. Out of all occurrences anticipated to be favorable (true positives plus false positives), it calculates the percentage of correctly predicted positive instances (true positives). The value obtained using our suggested approach (JAS-GCNN-) is superior to that of the

currently used methods (AGGCN-, GCN-). This demonstrates that our proposed strategy successfully identifies a Mixed-Type Wafer Defect Pattern.

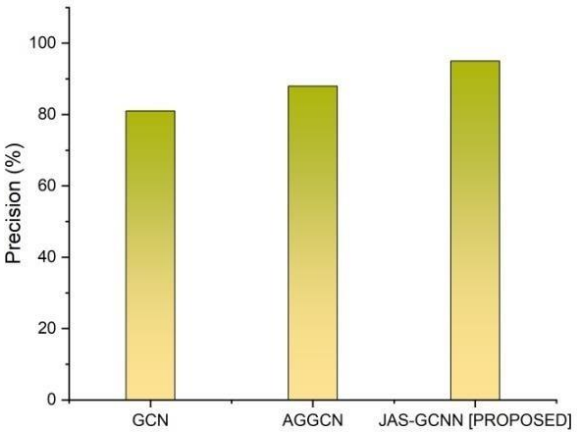


Figure 2 Outcome of precision.

Recall, also called sensitivity or true positive rate, quantifies the percentage of positive occurrences the model properly accepted. The ratio of true positives to the total of true and misleading negatives is computed. When compared to the existing employed approaches (AGGCN-, GCN-), the value produced using our proposed method (JAS-GCNN-) is higher. This shows that our proposed way is effective in locating a Mixed-Type Wafer Defect Pattern.

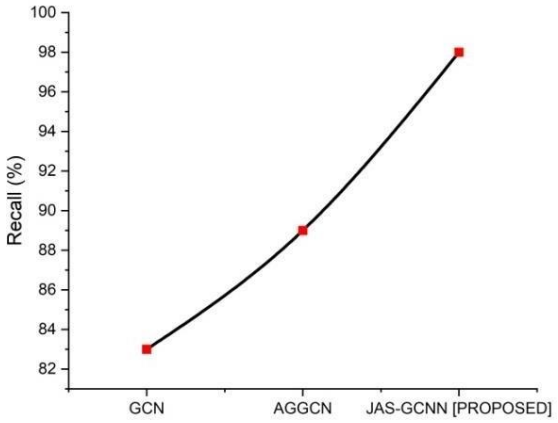


Figure 3 Outcome of recall.

The F1 score is sometimes referred to as the F-score or F-measure. F1 score is defined as the harmonic mean of precision. When compared to the existing employed approaches (AGGCN-, GCN-), the value produced using our proposed method (JAS-GCNN-) is better. This shows that our proposed way is effective in locating a Mixed-Type Wafer Defect Pattern.

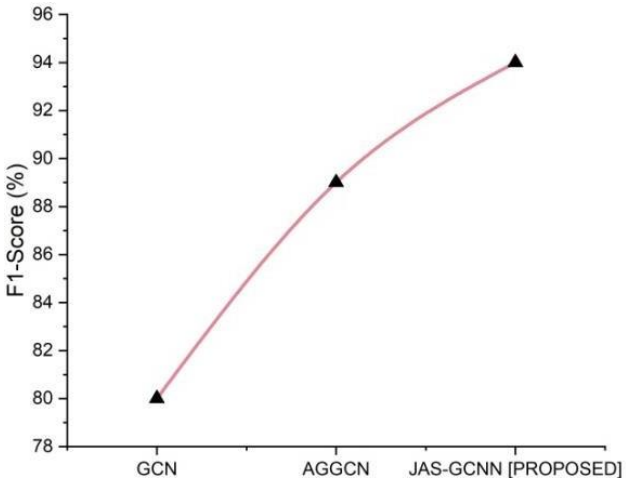


Figure 4 Outcome of f1-score.

5. Conclusion

To preserve excellent output and quality in wafer manufacturing facilities, defective pattern recognition (DPR), a difficulty in wafer production, is addressed in this study. Due to the various spatial characteristics, unexpected defect patterns, and different flaw quantities, mixed-type DPR, which includes identifying many types of flaws simultaneously, provides extra challenges. The research suggests a unique strategy known as the joint attention skipped graph convolutional neural network (JAS-GCNN) to address these problems. By combining targeted procedures with graph convolutional neural networks (GCNN), this method efficiently identifies local and global relationships among defects on a wafer. Using skip connections, the JAS-GCNN enables data transmission across various levels, facilitating the development of intricate defect patterns. Acquiring a larger and more diverse dataset with annotated defect patterns can enhance the model's ability to generalize to different types of flaws and improve its overall performance.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Ansari MZ, Janicek P, Park YJ, Nam Gung S, Cho BY, Nandi DK, Jang Y, Bae JS, Hong TE, Cheon T, Song W (2023) Preparation of wafer-scale highly conformal amorphous hafnium dioxide thin films by atomic layer deposition using a thermally stable boratabenzene ligand-containing hafnium precursor. Applied Surface Science 620:156834.

Bi S, Wang C, Wu B, Hu S, Huang W, Ni W, Gong Y, Wan X (2023) A Comprehensive Survey on Applications of AI Technologies to Failure Analysis of Industrial Systems. Engineering Failure Analysis 107172.

Jeong I, Lee SY, Park K, Kim I, Huh H, Lee S (2023) Wafer map failure pattern classification using geometric transformation-invariant convolutional neural network. Scientific Reports 13:8127.

Jeong I, Lee SY, Park K, Kim I, Huh H, Lee S (2023) Wafer map failure pattern classification using geometric transformation-invariant convolutional neural network. Scientific Reports 13:8127.

Jin CH, Na HJ, Piao M, Pok G, Ryu KH (2019) A novel DBSCAN-based defect pattern detection and classification framework for wafer bin map. IEEE Transactions on Semiconductor Manufacturing 32:286-292.

Kim D, Kim M, Kim W (2020) Wafer edge yield prediction using a combined long short-term memory and feed-forward neural network model for semiconductor manufacturing. IEEE Access 8:215125-215132.

Ko S, Koo D (2023) A novel approach for wafer defect pattern classification based on topological data analysis. Expert Systems with Applications 120765.

Lin S, He Z, Sun L (2023) a novel micro-defect classification system based on attention enhancement. Journal of Intelligent Manufacturing 1-24.

Piao M (2023) Analysis of Image Hashing in Wafer Map Failure Pattern Recognition. IEEE Transactions on Semiconductor Manufacturing.

Saqlain M, Jargalsaikhan B, Lee JY (2019) A voting ensemble classifier for wafer map defect pattern identification in semiconductor manufacturing. IEEE Transactions on Semiconductor Manufacturing 32:171-182.

Shawon A, Faruk MO, Habib MB, Khan AM (2019) Silicon wafer map defect classification using a deep convolutional neural network with data augmentation. In 2019 IEEE 5th International Conference on Computer and Communications (ICCC) (pp. 1995-1999). IEEE.

Shih PC, Hsu CC, Tien FC (2020) Automatic reclaimed wafer classification using deep learning neural networks. Symmetry 12:705.

Wang C, Wang Z, Ma L, Dong H, Sheng W (2023) Subdomain-Alignment Data Augmentation for Pipeline Fault Diagnosis: An Adversarial Self-Attention Network. IEEE Transactions on Industrial Informatics.

Wang J, Xu C, Yang Z, Zhang J, Li X (2020) Deformable convolutional networks for efficient mixed-type wafer defect pattern recognition. IEEE Transactions on Semiconductor Manufacturing 33:587-596.

Yu J, Zheng X, Liu J (2019) Stacked convolutional sparse denoising auto-encoder for identification of defect patterns in semiconductor wafer map. Computers in Industry 109:121-133.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0110

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

A randomized clinical trial to evaluate transoral head and neck surgery using a single-port flexible robotic system

J. Bhuvana[a] | **Vikas Kumar Kharbas**[b] | **Rajendra Pandey P.**[c]

[a]Jain (deemed to be) University, Bangalore, India, Associate Professor, Department of Computer Science and Information Technology.
[b]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science & Application.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College Of Computing Science And Information Technology.

Abstract This randomized IDEAL stage 2 clinical trials examined the safety and efficacy of a novel single-port flexible robot for transoral robotic surgery in patients with benign and malignant head and neck tumors. Rates of conversion and perioperative problems within a month of surgery were the main objectives. The Fisher's exact test and the Mann-Whitney U test were used in the statistical analysis, which used a p-value of 0.05 or less to determine significance. With the use of the da Vinci SP technology, 25 participants got TORS that allowed them to swiftly enter their nasopharynx, oropharynx, and hypopharynx. Notably, there were no serious complications or unfavorable robot-related occurrences throughout the 30-day follow-up period or any modifications to the robotic surgical equipment. The novel single-port flexible automated device demonstrates both security and practicality for carrying out transoral endoscopic procedures for head and neck conditions; based on this possible IDEAL stage 2 clinical trials, allowing accessibility to several anatomical regions, such as the larynx, oropharynx, nasopharynx, and hypopharynx.

Keywords: robotic surgery, single-port flexible, clinical trial, head and neck surgery

1. Introduction

The American Food and Drug Administration (FDA) approved the first generation of rigid robotic surgical systems to treat T1-2 oropharyngeal tumors, and these systems have subsequently been applied with little surgical morbidity or death (Olaleye et al 2022). Nevertheless, the initial setup had certain drawbacks, including the inability to concurrently use all 4 measurements due to the size and rigidly straight alignment of the gadgets and the challenge of docking the individual's side carts (Egberts et al 2017). To deal with these problems, a brand-new, flexible single-arm robotic surgery mechanism Preclinical techniques for transoral surgical robotics have reported other technological advancements to this automated technology. This new system fits 3 six-millimeter devices and a stereoscopic binocular camera inside a 2.5-centimeter circular cannula. The elbow joint on the six millimeters instruments gives them flexibility (Wang et al 2022). They resemble snakes and look similar to the five-millimeter multiport Endowrist devices. With enough treatments, the flexible robot arms allow for the introduction of the machines into the oral cavity to the larynx, hypopharynx, and oropharynx (Tateya et al 2018).

Head and neck surgery is a difficult discipline that calls for accuracy and meticulousness to ensure the best patient results. Recent developments in robotics have created new opportunities for less invasive surgical procedures, providing several benefits over conventional methods (Tonutti et al 2017). The single-port flexible robots technology is a ground-breaking platform that combines the benefits of robotic surgery with improved flexibility and fewer incisions. With the help of this cutting-edge equipment, surgeons may perform procedures with more accuracy and skill beyond the constraints of traditional laparoscopic and open techniques (Ashrafian et al 2017). The practicality and security of this upgraded, innovative, flexible single-arm robot for transoral endoscopic surgery of the head and Neck was the subject of the first clinical study, the findings of which are presented here (Park et al 2020). This robotic system gives surgeons a better field of view and more maneuverability by merging cutting-edge imaging technology and extremely flexible tools, improving their surgical abilities.

This study evaluated a revolutionary single-port flexible robot for transoral surgical operations for clinical viability and reliability.

The remainder of the paper is divided into subsequent parts. Part 3 contains the method explained. Part 4 includes the results, while Part 5 discusses the conclusions.

2. Related works

In the research (Chan et al 2017), initial findings from a phase 1 safe and practical clinical trial on the preliminary clinical application of new robotics for transoral surgical operations were presented, and a prospective clinical trial was used for the research's design. The study's methodologies included long-term study phase 1 clinical trial, open innovation, development, exploration, and evaluation. With an emphasis on budgetary difficulties, ongoing clinical studies, and the contentious topic of haptic and tactile input, the study (Friedrich et al 2017) reviewed the current practical uses and experimental breakthroughs for surgical robotics in the head and Neck. Analysis (Ross et al 2020) summarized the body of research on energy distribution in TORS and identified potential areas for further innovation. Studies on energy delivery in TORS were looked for in the 2019 MEDLINE database. To assess the usability and viability of this structure in the context of transoral robotic surgery (TORS) for laryngeal surgery, a study (Ross et al 2018) offered the first four examples using micro-phonological tools created expressly for use with the Med robotics Flex framework, a semi-flexible "robotic" technology.

To evaluate the reliability and efficacy of Transoral Robotic Surgery (TORS) to open an operation, the experiment (Rosello et al 2020) contrasted an accessible, comprehensive literature investigation. It carried out a systematic review of the literature that was already there to assess TORS's safety and efficacy in comparison to open surgical procedures. Transoral radical tonsillectomy and retropharyngeal lymph node evaluation were performed using a revolutionary flexible single-arm robotic (Tsang et al 2018). Study (Boehm et al 2021) evaluated the potential applications of RAS in head and neck surgery, the industries where it is already widely employed, and the advantages it offers over conventional operation. Study (Remacle and Prasad 2018) looked at a special single-port small-incision surgical technique that may be used in combination with traditional transoral laryngeal and pharyngeal tools in the oropharynx, hypopharynx, and larynx.

3. Method

3.1 Research plan

The clinical trial follows the Innovation, Development, Exploration, Assessment, Long Term Study (IDEAL) paradigm and is a phase two surgical examination. The institutional oversight board of the Japan University of Tokyo and the local regulatory agency approved the study project, which was carried out in conformity with the declaration of Abruzzo.

3.2 Study participants

Each person gave their signed, willingly informed authorization. Using the inclusion and exclusion standards in Table 1, benign and malignant head or neck conditions were considered.

3.3 Operative procedure

To help with suction and hemostasis, the medical professional was seated at the console with assistance at the individual's head. To access any region through the cavity inside the mouth, the mouth has to be contracted using a Dingman, Boyle Davis, or Feyh-Kastenbauer retractor. When necessary, the lower teeth were shielded with a piece of duodenum while the tongue was pulled back using a silk stitch 2.5 centimeter robotic port was advanced 11 to 15 centimeters from the mouth entrance before the instrument's arms and cameras were placed over the inside of the mouth and oropharynx. They worked with, investigated, and/or treated laryngitis, or pharyngitis, and hypo pharyngitis.

3.4 Study's Objectives

The main outcomes were the percentage of patients who converted and the frequency of complications after surgery within 30 days. For this research, a conversion is defined as an urgent modification of the therapy strategy to open surgery or traditional minimally invasive surgical procedures. The Clavien-Dindo category was used to classify all postoperative problems, including intraoperative difficulties, as well as any complications that occurred during the hospital stay or within a month of release. Major problems were defined as Clavien-Dindo grade III or more serious difficulties. Other important postoperative results such as operating time, predicted loss of blood, pain rating on a visual analog scale, hospital stay length, Excision margins, and swallowing capacity according to the MD Anderson Dysphagia Inventory had been added as additional objectives.

3.5 Statistic evaluation

Comparing data sets that are parametric and non-parametric was done using the chi-squared test, student t-test, and Mann-Whitney U test, respectively. Details that are parametric, non-parametric, and classified. The multiple measurement statistical variance analysis with post hoc pairwise contrasts and the Bonferroni correction was used to investigate longitudinal variations in the mean quality of existence. In statistics, a two-sided P value of 0.05 was regarded as significant.

4. Result

Twenty-five individuals were recruited in the experiment, and neither additional surgical techniques nor conversions to alternative robotic systems were made. Table 2 shows that the majority of respondents were men of Japanese ethnicity and previous smokers. The particular diagnoses are displayed in Table 3; malignant pathology was present in eight instances, and benign pathology was seen in 13 situations. The oropharynx was resected transorally in the vast majority of patients. 5 individuals who had oropharyngeal squamous cell carcinoma (SCC) caused by the human papillomavirus (HPV) underwent lateral oropharynx ectomy (Figure 1 (A)). 1 individual received a unilateral tongue base excision and had an unknown primary. Additionally, a retropharyngeal lymph node excision was performed concurrently on one patient who had an HPV-positive tonsil primary. The individual with the HPV negative SCC had previously had radiation treatment for nasopharyngeal cancer. One of the 2 individuals suffering from laryngeal SCC had an evaluation of the lesion while under anaesthesia (Figure 1 (B)). Another patient underwent a transoral complete laryngectomy excision with a tracheostomy for a first laryngeal SCC; the procedure was effective and had no adverse effects. (Figure 1 (C)). Then, local excisions were used to treat the lesions of the nasopharynx and hypopharynx. To achieve a full excision of the papilloma, utilizing a mucosa tube, the lesion in the hypopharynx was excised.

Table 1 Research Inclusion and exclusion standards.

Inclusion Criteria	Exclusion Criteria
• age of 18 or under	• General anaesthesia is not recommended
• BMI less than 35 kg/m2	• severe comorbid condition that substantially reduces the lifespan or raises the need for medical treatment
• appropriate for minimally invasive surgeries	• aggressive infection untreated
• Consent that has been informed and provided demonstrates a desire to engage	• Coagulopathy that cannot be reversed
• The medical professional believes sufficient contact with the surgical site exists to continue the treatment.	• Additional cancer or distant metastases is present
• Nasopharyngeal cancer with localized recurrence of T1/2	• immediate surgery
• Oropharyngeal cancer, T1/T2	• vulnerable group of people
• oral cavity cancer in the posterior third of the tongue that is T1/2 and prevents pull-through or mandibulotomy from being necessary	• Serious Trismus
• Obstructive sleep apnea (OSA) that is severe, moderate, or mild, along with snoring and benign oropharyngeal lesions, endoscopic sleep studies have demonstrated that Significant lingual causes of OSA include epiglottic retroflection and tonsil hypertrophy.	• Pregnant
• Main supraglottic T1/2 cancer with N0/1/2a nodal sickness, recurrent laryngeal T1/2/3 cancer, primary glottic T1a/1b/2 cancer, and harmless laryngeal illnesses	• Recurrent T3 lesion, internal carotid artery engagement
• bilateral vocal cord palsy that is permanent	• T3/4 oropharyngeal cancer with carotid artery inclusion
• Any type of hypopharyngeal dysplasia, as well as hypopharyngeal carcinoma with a posterior pharyngeal wall, T1/2 lateral wall, and N0/1/2a nodal disease after grading	• Remodelling with a free flap is necessary for T3/4 oral cavity cancer.
• In head cancer, just one retropharyngeal lymph node is found.	• Sleeping endoscopy revealed no oropharyngeal component to OSA.
	• Baseline aspiration on clinical examination; Clinical poor results that included baseline aspiration, a lack of physical stamina, and primary T3/4 supraglottic malignancy.
	• Bilateral vocal cord palsy that is curable
	• significant medial wall infiltration of the pyriform sinus and engagement of the pyriform apex, T3/4 hypopharyngeal cancer, N2b or above neck illness
	• The carotid artery, several retropharyngeal lymph nodes, various levels of nodal participation, and the presence of the prevertebral fascia are all present.

The Crowe-Davis retractor exposed 22 people, the FK-WO retractor exposed 3 people, and the Dingman retractor exposed one patient; blood was estimated to flow at an average rate of 40.2 milliliters per minute, docking took an average of seven minutes, operations took an average of 63 minutes, and hospitalizations were an aggregate of 11.3 days. Table 4 displays the functional results for pain and swallowing for persons with recurrent oropharyngeal squamous cell carcinoma. Eight patients had postoperative issues unrelated to the use of the robots. A broken right lower tooth in the fourth case required dental counseling due to earlier exposure to radiation and inadequate oral hygiene.

Table 2 The population under research's demographics.

Variable			N (%)
Gender		Male	17 (71.6)
		Female	8 (28.8)
Alcoholic		Yes	5 (14.5)
		No	19 (85.9)
Smoker		Current	7 (23.9)
		Ex-smoker	12 (47.8)
		No	8 (28.8)
ClavienDindo complications	Grade's	I	7
		II	3
		IIIa	4
Subsite		Nasopharynx	2
		Oropharynx	16
		Larynx	6
		Hypopharynx	3
Age		Median (Range)	62 (41–75)
Civilization		Japanese	21 (90.6)
		Europid	2 (4.9)
		Sinhalese	2 (4.7)

Table 3 Samples' pathologies.

Site	Pathology	n %
Nasopharynx		
	Lymphoid tissue	1
Larynx		
	SCC	3
Supraglottis	Fibrosis and lymphocytic infiltrate	2
Oropharynx		
Tonsil	OSA	4
Tongue base	HPV negative SCC	2
	Papilloma	3
	Ulcerated mucosa	1
	Asymmetric lingual tonsil	2
	HPV positive SCC	6
	HPV positive SCC	1
Hypophyarnx		
Posterior Pharyngeal wall	Papilloma	1
Neopharynx	Nasopharyngeal stricture	2

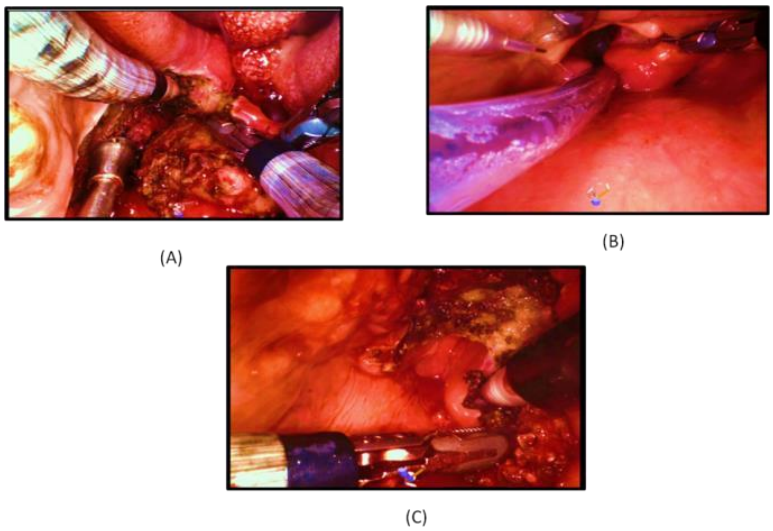


Figure 1 (A) 3 tool arms are being used in a left radical tonsillectomy, (B) Examining an individual under general anesthesia who has a recurring T2 supraglottic squamous cell carcinoma while using a Crowe-Davis oral retractor and the endotracheal tube in position, (C) Supraglottic partial laryngectomy in an individual who had a prophylactic tracheostomy and a T2 partial laryngectomy.

Instance 5 experienced a 1 cm tongue laceration due to the FKWO retractor, while instance 7 experienced a transient change in taste during the one-month intubation follow-up; case 8 reported a small vocal cord hemorrhage, 18th case had leakage from the granulation tissue when silk sutures were utilized for restraining the tongue, in contrast. In each of these circumstances, blood loss stopped on its own during follow-up visits, necessitating no medical intervention. In the 11th case, medication was used to address drug sensitivity. Post-radical mild In Case 15, the procedure bleeding was controlled at the hospital. Neither serious adverse effects nor converts happened throughout the clinical study.

Table 4 Functional result ratings before surgery.

	Preoperative		2 weeks			30 days		
	Mean	**SD**	**Mean**	**SD**	**P-value**	**Mean**	**SD**	**p-value**
Pain	0.5	0.9	3.9	3.3	0.03	1.8	1.9	0.24
MDADI Global	77.12	33.6	62.8	35.8	0.19	71.6	30.4	0.79
MDADI Physical	89.5	15.6	62.8	21.5	0.04	64.8	11.8	0.008
MDADI Functional	91.6	18.2	73.9	16.8	0.03	79.6	16.3	0.04
MDADI Emotional	89.6	17.5	72.5	17.8	0.03	77.9	15.6	0.06

5. Conclusion

The Da Vinci SP security and practicality trial results show that the device is probably secure and that, because of the multispecialty clinical research, transoral surgery employing it to reach the nasopharynx, oropharynx, and hypopharyngeal area for treating both benign and malignant tumors is technologically viable. In order to offset the different, decreased craniofacial measures in our group compared to others, a Crowe- Davis oral retractor was vitally utilized; this has been accomplished in a largely Southern Japanese community. We used the identical retractor to conduct a partial laryngectomy on an individual who had a lesion that was possibly cancerous despite several samples. These two instances demonstrate the advantage that flexible robotic arms have over conventional stiff, mechanical components. The existing technology has significant drawbacks; despite the edge of the second arm, there is less room for an assistant to work, especially when it comes to helping with pressure. This may be significantly prevented by inserting a suction catheter into the proper area through the nasal canal. Another disadvantage is the absence of bone tools for eliminating tumors from the base of the skull and the nasopharynx. Lastly, to improve the system's efficacy when employed on the glottic larynx, custom apparatus for finer dissections is required.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Ashrafian H, Clancy O, Grover V, Darzi A (2017) the evolution of robotic surgery: surgical and anesthetic aspects. BJA: British Journal of Anaesthesia, 119:i72-i84.

Boehm F, Graesslin R, Theodoraki MN, Schild L, Greve J, Hoffmann TK, Schuler PJ (2021) Current advances in robotics for head and neck surgery—a systematic review. Cancers 13:1398.

Chan JY, Wong EW, Tsang RK, Holsinger FC, Tong MC, Chiu PW, Ng SS (2017) Early results of a safety and feasibility clinical trial of a novel single-port flexible robot for transoral robotic surgery. European Archives of Oto-Rhino-Laryngology 274:3993-3996.

Egberts JH, Stein H, Aselmann H, Hendricks A, Becker T (2017) Fully robotic da Vinci Ivor-Lewis esophagectomy in four-arm technique-problems and solutions. Dis Esophagus 30:1-9.

Friedrich DT, Scheithauer MO, Greve J, Hoffmann TK, Schuler PA (2017) Recent advances in robot-assisted head and neck surgery. The International Journal of Medical Robotics and Computer Assisted Surgery 13:e1744.

Olaleye O, Jeong B, Switajewski M, Ooi EH, Krishnan S, Foreman A, Hodge JC (2022) Trans-oral robotic surgery for head and neck cancers using the Medrobotics Flex® system: The Adelaide cohort. Journal of Robotic Surgery 16:527-536.

Park YM, Kang MS, Lim JY, Kim SH, Choi EC, Koh YW (2020) The real impact of a robotic surgical system for precision surgery of parotidectomy: retro auricular parotidectomy using da Vinci surgical system. Gland Surgery 9:183.

Remacle M, Prasad VM (2018) Preliminary experience in transoral laryngeal surgery with a flexible robotic system for benign lesions of the vocal folds. European Archives of Oto-Rhino-Laryngology 275:761-765.

Remacle M, Prasad VM (2018) Preliminary experience in transoral laryngeal surgery with a flexible robotic system for benign lesions of the vocal folds. European Archives of Oto-Rhino-Laryngology 275:761-765.

Rosello A, Albuquerque R, Roselló Llabrés X, Marí Roig A, Estrugo Devesa A, López López J (2020) Transoral robotic surgery vs. open surgery in head and neck cancer. A systematic review of the literature.

Ross T, Tolley NS, Awad Z (2020) Novel energy devices in head and Neck robotic surgery–a narrative review. Robotic Surgery: Research and Reviews 25-39.

Tateya I, Koh YW, Tsang RK, Hong SS, Uozumi R, Kishimoto Y, Sugimoto T, Holsinger FC (2018) A flexible next-generation robotic surgical system for transoral endoscopic hypopharynx ectomy: a comparative preclinical study. Head & Neck 40:16-23.

Tonutti M, Elson DS, Yang GZ, Darzi AW, Sodergren MH (2017) The role of technology in minimally invasive surgery: state of the art, recent developments, and future directions. Postgraduate medical journal 93:159-167.

Tsang RK, Wong EW, Chan JY (2018) Transoral radical tonsillectomy and retropharyngeal lymph node dissection with a flexible next-generation robotic surgical system. Head & Neck 40:1296-1298.

Wang S, Li D, Jiang L, Fang D (2022) Flexible and mechanically strong MXene/FeCo@ C decorated carbon cloth: A multifunctional electromagnetic interference shielding material. Composites Science and Technology 221:109337.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0111

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Improvement in endoluminal procedures using versatile robot technology

Ramkumar Krishnamoorthy[a] | **Sujeet Kumar**[b] | **Rupal Gupta**[c]

[a]Jain (deemed to be) University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology.
[b]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science and Engineering.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College Of Computing Science And Information Technology.

Abstract: Recent Interventional medicine has been radically altered by end luminal procedures, which can be thought of as minimally invasive surgeries carried out within blood arteries or other hollow structures. The development of multipurpose robot technology has enabled substantial advancements in the performance of various treatments, leading to improvements in precision, maneuverability, and the outcomes for patients. In this study, we investigate the recent developments in the field of end luminal treatments and the benefits that come along with using adaptable robot technology. Recent technical breakthroughs in robotics and the growing emphasis on less invasive procedures have led to the development of highly adaptable surgical robots. These bendable robots can navigate tight spaces, increasing the accessibility of robotic surgery and perhaps decreasing the number of incisions required. This article describes new flexible surgical robot systems and discusses the most pressing technological challenges in this sector, with a focus on their potential uses in end luminal surgery. Furthermore, the form and force sensing of flexible robots, as well as the difficulties and recent developments in these areas, are also highlighted as important technological topics. The clinical benefits and technological advancements of new flexible surgical robot systems are also presented, along with their medical applications.

Keywords: surgical robot, endoluminal surgery, hysteresis, continuum robot, robot

1. Introduction

In Endoluminal procedures employ specialized equipment to perform minimally invasive medical procedures into hollow organs or blood vessels of the body. These procedures frequently include small incisions or holes in the body that occur naturally, obviating the need for extensive surgery and resulting in shorter recovery times and lower risks (Morino et al 2021). Numerous disorders affecting the digestive system, cardiovascular system, and respiratory system are regularly diagnosed and treated using endoluminal techniques. To observe and analyze the digestive system, the respiratory system, or other internal organs during an endoscopic procedure, a flexible tube with a light and camera (endoscope) is inserted through the mouth, a skin opening, or a very small incision (Bianchi et al 2019). Endovascular operations can detect and treat problems including peripheral artery disease (PAD), aneurysms, and venous disorders by introducing catheters and wires through tiny incisions into the blood arteries (Race and Horgan 2021). Using endoscopic retrograde cholangiopancreatography, diseases of the pancreas and bile ducts can be identified and treated. An endoscope is inserted into the small intestine, bile, and pancreatic ducts through the mouth. Endoscopic ultrasonography is a method that combines endoscopy with ultrasound technology to view and detect conditions affecting the pancreas, gastrointestinal system, and other organs (Maza and Sharma 2020). Open surgery has a variety of advantages over endoluminal procedures, including less pain, shorter hospital stays, quicker recovery, and a decreased risk of complications. The diagnosis and the patient's specific circumstances will determine the best course of action because endoluminal procedures cannot be employed to address all issues. It is vital to consult a medical specialist when deciding on the best course of treatment for a particular condition (Guo et al 2022). The term "versatile robot technology" refers to the creation and implementation of robotic systems that are capable of carrying out a variety of activities in a variety of fields and applications. By altering or reprogramming their activities or behaviors, these robots are intended to be versatile, adaptive, and capable of carrying out a variety of jobs (Singh et al 2021). In industrial automation repetitive processes like assembly, packing, and material handling, versatile robots are utilized in manufacturing and production lines. Depending on the demands of the company, these robots may be trained to carry out various jobs, increasing productivity and efficiency. Applications in healthcare and medicine: Robots are being utilized more often in healthcare settings for a variety of functions, including surgery, rehabilitation, patient monitoring, and drug distribution. Medical personnel may provide accurate and effective care with the help of adaptable

robots in the industry (Troccaz et al 2019). Automation in logistics and warehouses: Order picking, sorting, and inventory management are all jobs that robots are used for in warehouses and logistics facilities. This industry has adaptable robots that can be configured to handle various product categories and change with the needs of the warehouse (Vrielink et al 2020). Robots are being used in the hotel, retail, and customer service industries. They can assist with product support, communicate with consumers, and give information. They can also clean and maintain inventories. Versatile robots are employed in scientific study and exploration for tasks including deep-sea exploration, environmental monitoring, and space exploration. These robots may be made to adapt to various settings and perform certain activities, delivering useful information and insights (Wang et al 2021). Robot technology's adaptability is made possible by many elements, including cutting-edge sensors, clever control systems, modular architecture, and the ability to program or reconfigure the robot for various jobs. Robots may be utilized in a variety of applications thanks to their adaptability to diverse conditions and situations (Steiner et al 2019). Article describes new flexible surgical robot systems and discusses the most pressing technological challenges in this sector, with a focus on their potential uses in endoluminal surgery.

2. Related Works

Research (Dagninoet et al 2022) demonstrated that the suggested robotic platform has the potential to enhance the performance of endovascular operations, opening the door to their eventual implementation into clinical practice. They describe the results of an in-depth in-vivo investigation in which their robotic platform was used to perform cannulation and balloon angioplasty on five target arteries in four pig models. Trial results indicated a 100% success rate and post-mortem histological analysis showed that robotic navigation resulted in less vascular stress than hand manipulation. The results of their in-vivo studies showed that their robotic system was safe, feasible, and well-tolerated for use in this research.

Robotic platform that can bend and has a diameter of only 17 mm. A powerful continuum manipulator with the highest resistance to distortion has been devised to counteract form distortion and deflection in payload handling. For master-slave teleoperation, they have devised the kinematic analysis and mapping approach. The suggested manipulator can lift 300 g with a trajectory change of just 7.5 mm. Three distinct types of simulated surgical tasks have confirmed the practicability of the integrated system (Hwang and Kwon 2020).

The length of time it took to do the operation (from making the incision to completing the dissection), the extent to which tissue was removed during the operation, any complications that arose during the surgery, and the difficulty in moving the surgeon's arms about in the confined area. The present investigation verified the safety and viability of employing the Endo Master System for colorectal Endoscopic sub mucosal dissection (ESD). in a preclinical setting. The system's capacity to deal with complications including bleeding and perforation was also evaluated (Chiu et al 2021).

To assist medical professional’s in future clinical practice and to inspire and drive new technical innovations, this article aims to present a clear and complete picture of modern robotic gastro scopes and associated technologies. Article (Marlicz et al 2020) provides a comprehensive analysis of the capabilities and performance of these cutting-edge gadgets. Remote tele health endoscopy services are also covered, along with the use of AI technology in robotic gastro scopes.

New surgical modalities with cutting-edge technology, lower prices, and more compact sizes are entering the market as robotic surgery nears a decade of widespread usage in medicine. Focusing on the downsizing of modalities toward the building of micro-scale surgical robots, nicknamed "micro bots," this chapter seeks to showcase new surgical robotic technologies (Khandalavala et al 2020).

After each surgery, a debriefing meeting was held to evaluate the technology used, to improve it, or create new tools to use in the future (Morino et al 2022).

The rapid development of flexible endoscopy has made it an invaluable surgical tool. Revolutionary advances in areas like robots, technology, and Robotics are propelling this change. Several of the most important developments associated with this paradigm shift in gastrointestinal care are presented (Swanström and Pizzicannella 2023).

An emerging subject, robotics in minimally invasive endoscopic procedures has significant challenges from the stringent criteria for downsizing. Inchworm-like devices were described as the first step toward robotic colonoscopy in the 1990s. Since then, colono scopes with robotic assistance features have hit the market. Future treatments, including those aided by autonomous or robotic agents like robotic capsules, offer more accessibility and flexibility with the help of research prototypes. Improved diagnostic yield may also be expected when such endoscopic technologies are combined with AI-enabled picture analysis and identification algorithms (Ciuti et al 2020).

Article (Osawa et al 2022) developed miniature multi-degree-of-freedom (DOF) endoluminal forceps with solid construction with flexible hinges.

The research (Hwang et al 2020) was to assess the viability of a traction technique using a flexible robotic arm for performing stomach ESD. When performed with a standard flexible endoscope, ESD presents technical difficulties and challenges owing to the absence of enough countertraction to reveal the sub mucosal dissection plane.

3. Key Technical Issues in Flexible Surgical Robots

3.1. Design Manipulator

3.1.1. Types of Flexible Manipulators

Based on these analyses, we widened our definition of a flexible manipulator to encompass not only soft robotics and origami robotics but also continuous backbones, discrete backbones, hybrid backbones, and more. To achieve a continuous backbone, continuous elastic backbones, pre-shaped super elastic tubes, and push-pull actuation are used, as well as shape memory actuators and antagonistic pairs of wires. Responsive manipulators based on a notch flexure hinge are a recent development in the field of flexible manipulator design. The construction of discrete-backboned robots consists of articulated linkages, pivots, and wire-compressed cams; they are operated by push-pull antagonistic actuation of wires; and they are hyper-redundant serial manipulators. To accomplish a decoupled drive, the driving wires are arranged so that they always pass through the center of all joints. Hybrid robots' manipulative capabilities come from the combination of flexible parts (like springs) and connections in their structural design.

3.1.2. Stiffness Enhancement

Flexible robotic surgical systems used in endoluminal procedures need a manipulator that combines dexterity and rigidity. The manipulator has to be flexible so it can reach the damaged area over a lengthy, winding journey. The manipulator needs to be bendable enough to get to the problem location, but then it needs to tighten up once it's there. Flexible surgical robot applications have prompted research into strategies for increasing the stiffness of flexible manipulators to mitigate this tradeoff. Wire tension, friction/interlocking, and phase transition are the primary mechanisms by which these methods alter stiffness. After that, we'll look at a method for increasing stiffness by tensioning wires, which hasn't been covered in any of our prior examinations.

Table 1 Requirements and Technical Challenges Faced by Flexible Surgical Robots for Endoluminal Application.

Technical challenges	Requirements
precise and reliable intraoperative shape sensing tissue interaction force sensing and feedback	Safe access to the surgical site and tissue manipulation
Flexible guide tube (or endoscope) with excellent bending capability complete integration of endoscopic functions in a limited overall diameter	Flexibility for access to the surgical site through narrow and tortuous routes via natural orifices
the multi-DoF flexible instrument with a thin and compact size instruments with flexibility but adequate payload/stiffness surgical triangulation precise control of long and thin flexible instruments the ergonomic and intuitive human-robot control interface	Dexterity, accuracy, and stability for performing surgical interventions in a confined space

3.1.3. Special Considerations in Manipulator Design

The safety, durability, and accuracy of a flexible manipulator may be improved in some ways by carefully designing its driving mechanisms. An increase in human safety can be achieved by decreasing the passive stiffness of a flexible manipulator. An actuator's torque limitation can switch between three different operating modes: free, active drive and sliding when overloaded. Long-term and stable control in a wire-driven flexible manipulator requires fine-tuning the elongation of the driving wire.

Flexible surgical robots employed a tiny cable with reduced friction. HI-LEX Corp. later marketed this one-of-a-kind cable, and it is now extensively utilized. While designing a flexible manipulator, it is also important to think about the fabrication plan. Subtractive manufacturing and additive manufacturing are the primary production technologies used for prototyping flexible surgical manipulators. Computerized numerical control machining, laser cutting, and wire electrical discharge machining (EDM) are all examples of common subtractive manufacturing processes. Common methods of 3-D printing used in additive manufacturing are metal printing and laser sintering. Resolution, precision, and mechanical qualities like strength and surface smoothness are all improved in additive manufacturing compared to subtractive manufacturing. Additive manufacturing's strength is in its capacity to produce intricate three-dimensional structures, such as those with numerous working channels or helical tendon courses. In light of this, researchers need to learn the ins and outs of each manufacturing technique and weigh their options based on the manipulator's function, structure, material, and size.

3.2. Modeling and Control

The mechanical architecture of a flexible manipulator whether it has a continuous, discrete, or hybrid backbone, a single backbone, numerous backbones, or only one instrument has a direct impact on how its kinematics may be modeled.

3.2.1. Kinematics and Dynamics Modeling

Kinematics modeling uses terms like location, velocity, and acceleration to describe how things move. Without taking into account the underlying causes, it is just concerned with the mathematical description of motion. Key elements of kinematic models include the following:

- Position: The position of an object in space is commonly expressed as three-dimensional Cartesian coordinates with values like x, y, and z.
- Velocity: The speed of positional change over time. It gives the direction and speed of an object's motion. The derivative of the location with respect to time provides the velocity.
- Acceleration: The speed at which speed varies throughout time. It displays the rate of change in an object's velocity. By taking the derivative of velocity to time, acceleration may be calculated.

Kinematic models often use equations such as linear motion equations or equations of motion for rotational motion to describe the relationships between position, velocity, and acceleration.

Dynamics modeling examines the motion-causing forces and how they impact an object's motion. It includes the investigation of the connections between forces, masses, and accelerations. Newton's laws of motion, which are essential ideas in classical mechanics, are included into dynamics models.

Using Newton's principles to calculate the motion of an item, including changes in velocity and acceleration, dynamics models often analyze the forces acting on an object.

For comprehending and forecasting the motion of objects in a variety of applications, such as basic mechanical systems to sophisticated mechanical systems, robotics, aeronautical engineering, and more, kinematics and dynamics modeling are essential. These models provide engineers and scientists the tools they need to build and improve systems, simulate motion, and anticipate the behavior of moving objects.

3.2.2. Teleoperation and Automatic Control

Steerable-head flexible instruments and flexible devices with several instruments might be challenging for a single operator to manage manually. This supports the automation of any equipment used in endoluminal procedures, no matter how simple they may seem. In the field of robotic surgery, M-S robots are created with many goals in mind. Communication with the surgeon to ensure patient safety and compliance with regulations. Both rigid and flexible instrument tele manipulation share these goals. Robotic instruments are essential for the widespread clinical adoption of flexible instruments and endoluminal applications since manual techniques need in-depth education and are often limited in their use.

- **M–S mapping strategy**: The strategy's end objective is to shorten the learning curve and increase safety by allowing the surgeon to make intuitive movements. Interfaces that aren't well thought out diminish the robot's usability and the surgeon's comfort.

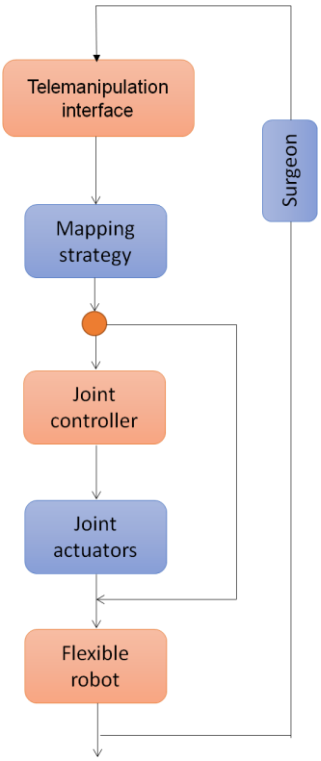


Figure 1 Control loop for M-S.

The latter calls for including an inverse kinematics calculation into the mapping procedure (see Figure 1). Master interface motion must stay away from the singularity and the boundary of the interface if manipulability and motion scaling is to be achieved. Given these constraints, it will be necessary to create specialized medical robot interfaces rather than employing off-the-shelf solutions. The surgeon's whole attention should be on the operation at hand, not on the mechanics of which buttons to press, pedals to engage, or arm motions to carry out.

3.2.3. Team SoloMid (TSM) Hysteresis Compensation

TSM is a typical actuation mechanism used in flexible surgical robots to achieve steer ability. This mechanism's benefits include adaptability, high transmission efficiency, and a small manipulator footprint. As a result, the performance of flexible surgical robots is hindered by the delay and decrease in control precision introduced by standard kinematic control methods. Hysteresis is a common problem with endoluminal-applicable flexible surgical robots due to the lengthy and winding tendon design of the manipulators, which causes a lot of friction force and wire distortion. As a result, many people have put a lot of time and energy into developing both offline and online methods to counteract the impact of hysteresis.

3.3. Shape and Force Sensing

3.3.1. Shape Sensing

The deformability, compliance, redundancy, and unavoidable interaction with surrounding tissue of surgical flexible manipulators make it challenging to accurately estimate the shape and posture of the manipulator using kinematics and mechanics-based models. Uncertain model parameters and the influence of external loads, which lead to substantial changes in form and kinematics, make it challenging to use this model-based method. Due to inaccurate position input, the flexible manipulator cannot be safely guided to the surgical site. New sensor-based techniques are into one of three categories: those based on optical fiber sensors, electromagnetic sensors, or intraoperative images. These methods are used as the flexible tool used in endoluminal surgery not only for accurate control but also for navigation and contact-force measurement. This section gives a high-level overview of each technology, including its underlying principles and the most recent advancements in the field.

3.3.2. Force Sensing

Some aspects of endoluminal surgery are unique to this kind of surgery, making contact sensing between the device and organ particularly crucial: The esophagus, colon, blood vessels, and ureter are all relatively fragile organs that the endoscope will inevitably come into contact with as it makes its way to the surgical site through long, narrow, and curved pathways. Surgeons doing manual endoluminaloperations have tactile sensations at the site of tissue contact. The patient's well-being is intimately linked to the haptic data's application to the evaluation of contact and tissue damage. Therefore, surgeons using robotic systems must understand the force exerted by each tool on the tissue. Although haptic or tactile feedback has not been used in existing flexible robotic surgical systems, various researchers have measured the interaction forces involved.

4. Flexible Robotic Systems for Endoluminal Applications

Pioneering work on flexible surgical robots has resulted in the development of an M-S-type active endoscope, which has been in use since the late 1980s. Each endoluminal application (gastrointestinal endoscopic surgery, ureteroscopic surgery, bronchoscopy, and endovascular surgery) is analyzed in detail, including their clinical achievements using the corresponding flexible robot system. Several recent medical studies have reviewed the use of robots in endoluminal and transluminal endoscopic operations.

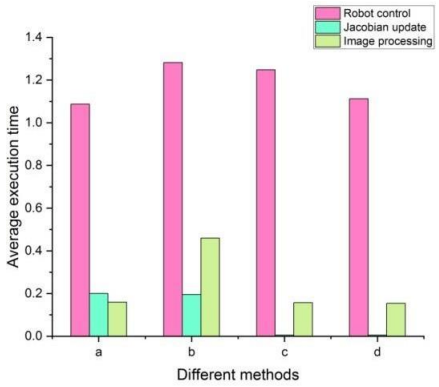


Figure 2 Comparison of tasks by depicting the execution times.

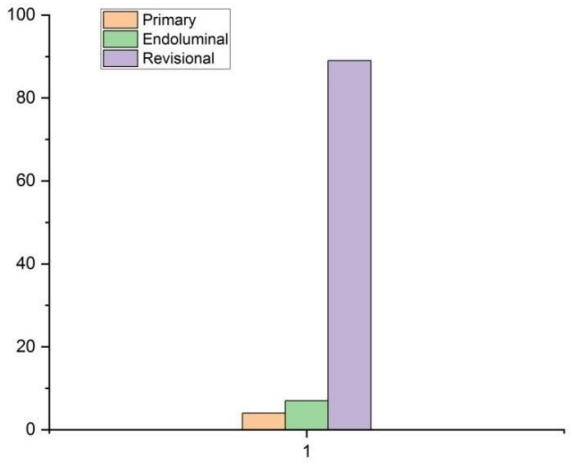


Figure 3 The number of endoluminal, revisional procedures, and primary.

Table 2 Comparing tasks by showing the timeframes used.

Different methods	**Average Execution time**		
	Robotcontrol	Jacobianupdate	Imageprocessing
a	1.087	0.2	0.16
b	1.282	0.195	0.46
c	1.248	0.005	0.157
d	1.112	0.005	0.154

4.1. Gastrointestinal Endoscopic

Surgery one of the most popular uses of surgical robots nowadays is in the area of gastroenterological endoscopy. So, the traditional gastrointestinal endoscope's difficulties in manipulation and dexterity instrument movements may be overcome with robotic help, allowing for more effective diagnostics and treatments. When it comes to gastrointestinal endoscopy, robots may be of use in two distinct ways: during insertion and operation. In particular, this article examines the role that robotic devices may play in facilitating endoscopic surgery. Two robotic devices with articulated joints are often part of these flexible systems and may be mounted to the outside of a flexible endoscope or inserted inside the endoscopic channels. All of the tools (including the endoscope) may be used with a single hand or remotely controlled by a commercial or bespoke master device. Many similar systems are practical and beneficial in preclinical or clinical testing, and one is already available on the market. Originally, they were conceived as a "mechanical system" in which two operators would use a handle or knob attached to the endoscope and surgical equipment to perform the procedures. To further increase intuitiveness, ergonomics, and accuracy, and decrease the number of operators needed, technology eventually developed into the "robotic M-S system," which enables the remote operation of the endoscope or surgical tools.

4.1.1. Mechanical System

a) COBRA

USGI Medical's (San Clemente, California) COBRA resect scope has a shape-locking scope with three extra arms. By applying strain to the wires connecting the connections in a serial configuration, the scope can lock its form into place. The scope may be locked into a hard configuration to offer a secure platform for surgery, and it can be flexible during insertion. Using an endoscopic triangulation, the camera and two surgical tools may be moved independently of one another, allowing for traction and counter traction to be performed while retaining clear optics. Laboratory reports indicate that complex operations, such as suturing and tying sutures, done using COBRA may be challenging because of the limits of imprecise cable-driven controls and the inability to change tools. There have been no more reports of either preclinical or clinical findings.

b) R-Scope

The main flexure may be fixed in place by increasing the tension on the actuation wire, and the secondary flexure can be placed anywhere the user sees fit. The scope and the two instruments are controlled by a single knob on the control body. Animal and human trials were conducted first to ensure the device was effective for gastric ESD. However, due mostly to the operator's incompetence, the outcomes were subpar, with perforations occurring in about 20% of both instances. Clinical research using gastric ESD for superficial gastric neoplasm revealed similar en-bloc resection, complication, and local recurrence outcomes to those of conventional ESD, with a much shorter operation duration. Complex controls that are

difficult for a single operator to master, as well as subpar instrument performance in retroflection, are two of the system's major drawbacks.

c) Direct-drive endoscopic system

Boston Scientific's (Marlborough, MA, USA) Direct-drive endoscopic system (DDES) is a manually operated, multifunctional platform designed for endoluminal and NOTES procedures. Each tool has an ergonomically designed handle that powers a long flexible shaft that terminates in an individual end effector. The guide sheath adds two degrees of freedom to the seven DoFs at the instrument tip, which are transmitted from the handle. Endoscopic mucosal excision is just some of the complicated non-surgical activities that the DDES is capable of in ex vivo and in vivo animal testing.

d) Endo SAMURAI

Olympus's Endo SAMURAI is a versatile endoscopic device for performing intraluminal and transluminal treatments. A bendable endoscope, two articulated working arms equipped with surgical end effectors, and a control unit makes up the system. The flexible endoscope's directional and locking capabilities allow it to serve as a system stabilizer. In difficult endoscopic procedures, including endoscopic full-thickness resection, the technology proves more accurate and requires less time than a traditional endoscope. A small bowel anastomosis might also be completed promptly and to an acceptable standard. Cutting, suturing, and tying knots were accomplished with accuracy and effectiveness on par with laparoscopic instrumentation, but with a greater time investment.

4.1.2. Robotic M–S System

a) ViaCath

EndoVia Medical's ViaCath system is the first tele operated robot designed specifically for endoluminal procedures. In comparison to an endoscope, the two robotic devices that extend from their distal ends are more sophisticated, allowing for the bimanual manipulation of tissues. The endoscope's field of view may be manipulated in seven degrees of freedom (DoFs) thanks to the instrument's positioning arm. Multiple animal experiments, both in vitro and on living subjects, confirmed the system's viability. Mucosal excision and fundamental suturing in the digestive tract were achieved with the help of the available tools and system. To address issues including awkward endoscope insertion into the stomach and weak instrument manipulation force, a newer generation of equipment was developed.

b) Endo MASTER

The Endo Master was developed at Nanyang Technological University as a robot-assisted surgical device for use in NOTES procedures. Endoscopic removal of polyps and cancers in the digestive tract was among the first uses. The Endo Master is a flexible endoscope with two robotic arms built onto its tip to increase its agility. This allows for precise tissue manipulation and dissection. Complete excision of stomach neoplasms was achieved in a small-scale, human preclinical experiment, and no patients had any adverse effects. A colonic ESD was performed without perforation using the newest version of the Endo Master EASE System in a pig model. The first patients enrolled in a clinical trial to treat colorectal lesions in May 2020, and the study ran until December 2021. Notably, there have been cadaver studies for transoral uses.

c) FLEX robotic system

Intrapericardial uses inspired the initial design of the FLEX^2 robotic system by Medrobotics Corporation. Transoral and transanalendoluminal surgeries are now possible with the upgraded device. The system consists of a robot-assisted flexible endoscope (RAFE), a set of flexible instruments that are compatible with the RAFE, and a control station. The FLEX scope has two connections, as opposed to one in a standard flexible scope: a distal one and a leading one. The articulating instrument has four degrees of freedom (DoF), with the user controlling two directional bendings through a grip attached to the flexible shaft. Transoral surgical operations involving the pharynx and larynx have demonstrated promise in the system's first clinical trial. In 2014, it was given the European CE mark, and in 2015, it was given FDA authorization for use in transoral surgeries. In 2017, the FDA gave the technology the go-light for use in colorectal endoscopies. The viability of some colorectal operations, including transanal TME (taTME) and transvaginal rectopexy, was studied using cadaveric follow-up experiments.

d) Robot-assisted flexible endoscope

Kyushu University (Fukuoka, Japan) has created a platform called RAFE that is optimized for ESD. The basic idea behind the platform is to adapt existing standard endoscopes for usage with a flexible surgical robot. The endoscope's extended motor unit controls all degrees of freedom. There are two bending degrees of freedom (DoFs) in the platform's articulating instruments. Both are placed via the endoscope, one through a regular channel and the other through a specialized tube at the end.

4.2. Ureteroscopic Surgery

The flexible ureteroscope is being used in conjunction with robotic assistance primarily for the treatment of kidney stones. The weariness of the surgeon due to a non-ergonomic position is a significant drawback of traditional flexible ureteroscopy in renal stone removal, potentially decreasing surgical efficiency and safety and increasing the risk of damage to the surgeon. The robot systems use robotized ureteroscopemanipulation to reduce radiation exposure and improve surgeon comfort during urological procedures. The standard setup for such a system includes a slave robot arm on which is attached a commercial flexible ureteroscope, and a control panel from which the endoscope and other equipment may be remotely manipulated.

4.2.1. Roboflex Avicenna

ELMED Medical Systems has created the first robot-assisted flexible ureteroscope device intended for kidney stone therapy called Roboflex Avicenna1. A bendable ureterorenoscopy manipulator and a surgeon's station with a built-in touch screen and two joystick interfaces make up this device. The robotic arm's hand component is compatible with the most available ureterorenoscopy on the market today. The device has pneumatically actuated 2-foot pedals for activating fluoroscopy and laser shooting. Motorized insertion and withdrawal of the laser fiber and variable rates of irrigation fluid infusion are also available. The first clinical trial confirmed that the device is capable of performing all current protocols and procedures associated with flexible ureteroscopes, including laser dusting and the removal of bigger pieces, with a good effect on ergonomics and a reduction in radiation exposure to the surgeon. Additional clinical investigations confirmed the system's safety and effectiveness and compared well to the standard approach in terms of stone-free rate, treatment time, and intraoperative complications. In 2013, the system was awarded the CE mark, and it is now awaiting FDA clearance.

4.3. Bronchoscopic

4.3.1. Monarch

Robotic bronchoscope system with built-in electromagnetic navigation (EMN) guidance, created by Auris Health. It also has safety features like monitoring driving tension and automatically releasing tension when the scope is retracted. The simulated peripheral pulmonary lesions were successfully biopsied and the peripheral airways were more accessible in cadaver tests. Initial clinical studies showed the platform to be technically viable for diagnostic bronchoscopy. Recent research has shown that robotic bronchoscopy is a safe and effective treatment option for individuals with peripheral pulmonary lesions. In 96% of patients, lesion localization was confirmed, and the risk of adverse events was similar to that seen with traditional bronchoscopy. The Food and Drug Administration (FDA) has given its blessing to the platform for diagnostic and therapeutic bronchoscopic operations in 2018.

4.3.2. ION Robotic Endoluminal System

Robotic catheters like Intuitive's (Sunnyvale, CA, USA) ION robotic endoluminal systems use shape-sensing technology to provide the user input on the catheter's location and form. The technology consists of a robotic articulating catheter that is flexible in any direction by up to 180 degrees and includes a shape-sensing fiber embedded along its length. In addition to the catheter, there is a program that reconstructs the airways virtually in 3D and displays a route to the desired location automatically. The next step is to use the shape-sensing fiber's position information to register rather than relying on EM position sensing. The technology proved its efficacy in cadaver research, where it was used to successfully puncture tiny nodules on the lung's periphery. The most recent clinical investigation was 92%, and there was hardly any risk of complications. In 2019, the technology was given FDA permission for use in peripheral lung biopsy using a minimally invasive approach.

4.4. Endovascular Surgery

Endovascular procedures may now be performed with the same ease using flexible surgical robot technology. Endovascular surgery has gone a long way in terms of technology, but there are still significant barriers to effective surgical results, such as the dependence on lesion sites, vascular operator skill, and a lack of accurate placement of intravascular devices. To make it easier to access challenging lesions, robot systems have been created to give accurate guide wire and catheter control. The resultant efficacy has the potential to shorten fluoroscopy times, protecting the surgeon and the patient from unnecessary radiation exposure.

4.4.1. Sensei X

Hansen Medical's Sensei (Mountain View, CA, USA) is a control and positioning system for catheters in the circulatory system that aims to make them easier to use. The M-S electromechanical system allows the clinician to manipulate the

guiding catheter and sheath. Sensei X, the latest iteration of the technology, has an artisan stretched catheter that allows for 270 degrees of the catheter and a remote-controlled tip that can be manipulated in three dimensions. The vibrations are sent to the user via the controller, creating a haptic effect. The feasibility and efficacy of using Sensei X to treat AF were assessed. Compared to manual ablation, the results showed that employing the robotic method resulted in similar rates of problems and recurrence, although much less radiation exposure was experienced. The system received both CE and FDA markings in 2007.

4.4.2. Magellan

Redesigned for peripheral endovascular intervention, the Magellan1 robot was developed by Hansen Medical. The system's core features include a robot arm and an operator console for remote manipulation of steerable catheters and standard guide wires. The radiation source is placed in a different room from the control room. There was less catheter-tissue contact and less vascular stress from catheter contact with the vessel wall, according to phantom research. Fenestrated endovascular repair (FEVAR) was successfully performed in clinical trials, demonstrating the system's viability. These results also suggested that it would be possible to reduce the operator's exposure to radiation and streamline difficult endovascular procedures. No access site difficulties occurred during tibiofemoral artery navigation, and the procedure was straightforward for even inexperienced operators. The technique showed promise in thoracic endovascular aortic repair, with much-reduced embolization, possibly due to the robotic catheter's enhanced agility and control, which lessens the likelihood of accidental contact with the artery wall. In 2011, the system was granted the CE mark, and in 2012, it was granted FDA 510(k) approval for use in guiding guide wires and robotic catheters via peripheral vessels.

4.4.3. R-One

The robot has both a control unit and a robotic unit, both of which are shielded from radio waves. Within a radiation-safe zone, surgeons may use the control unit to operate a catheter and guide wire from a distance. The robotic system allows for the automated use of the best guide wires and catheters on the market. There have been no published publications detailing the outcomes of clinical trials. By 2021, RoboCath expects to have performed its first robotic coronary angioplasties in many European nations, as well as in Africa and China. In addition, a European clinical investigation utilizing robot-assisted percutaneous coronary intervention (PCI) using RoboCath will have finished enrolling patients in 2021. In 2019, RoboCath received the CE certification, allowing it to be used in cardiac interventions.

5. Conclusion

Several adaptable robotic surgical systems have been developed to do endoluminal surgery alongside the commercial success of the da Vinci laparoscopic robotic surgical system. But there are technological challenges that need to be overcome when it comes to maneuvering deftly in the lumen's curved and restricted environment. This article reviewed state-of-the-art research activities that are being done to solve these technical problems. Both robotic surgery and endoscopy stand to benefit from this new technology if the findings of this study are implemented in flexible surgical robots. Through endoluminal, transluminal, and extraluminal techniques, surgeons will have access to more broadly applicable robotic treatments that need little to no incision. Furthermore, with the use of delicate surgical movements and navigational support, endoscopists will be able to undertake more sophisticated endoscopic diagnostics and endoscopic tumor resections. Regardless of the surgeon's endoscopy expertise or the degree of complexity of the procedure, the modern endoscopic surgical robot will provide consistent results. New surgical tools, such as laser, cryogenic, and electrical devices, will be developed with the help of these high-tech robots. When compared to traditional rigid-type laparoscopic robotic surgical systems, flexible surgical robots will allow the advantages of robotics to be applied to a wider variety of surgical procedures.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Bianchi F, Masaracchia A, ShojaeiBarjuei E, Menciassi A, Arezzo A, Koulaouzidis A, Stoyanov D, Dario P, Ciuti G (2019) Localization strategies for robotic endoscopic capsules: a review. Expert review of medical devices 16:381-403.

Chiu PWY, Ho KY, Phee SJ (2021) Colonic endoscopic submucosal dissection using a novel robotic system (with video). Gastrointestinal endoscopy 93:1172-1177.

Ciuti G, Skonieczna-Żydecka K, Marlicz W, Iacovacci V, Liu H, Stoyanov D, Arezzo A, Chiurazzi M, Toth E, Thorlacius H, Dario P (2020) Frontiers of robotic colonoscopy: a comprehensive review of robotic colonoscopes and technologies. Journal of clinical medicine 9:1648.

Dagnino G, Kundrat D, Kwok TM, Abdelaziz ME, Chi W, Nguyen A, Riga C, Yang GZ (2022) In-vivo Validation of a Novel Robotic Platform for Endovascular Intervention. IEEE Transactions on Biomedical Engineering.

Guo Y, Chen W, Zhao J, Yang GZ, (2022) Medical robotics: opportunities in China. Annual Review of Control, Robotics, and Autonomous Systems 5: 361-383.

Hwang M, Kwon DS (2020) K-FLEX: a flexible robotic platform for scar-free endoscopic surgery. The International Journal of Medical Robotics and Computer Assisted Surgery 16:e2078.

Hwang M, Lee SW, Park KC, Sul HJ, Kwon DS (2020) Evaluation of a robotic arm-assisted endoscope to facilitate endoscopic submucosal dissection (with video). Gastrointestinal Endoscopy 91:699-706.

Khandalavala K, Shimon T, Flores L, Armijo PR, Oleynikov D (2020) Emerging surgical robotic technology: a progression toward microbots. Ann LaparoscEndoscSurg 5:3.

Marlicz W, Ren X, Robertson A, Skonieczna-Żydecka K, Łoniewskil Dario P, Wang S, Plevris JN, Koulaouzidis A, Ciuti G (2020) Frontiers of robotic gastroscopy: a comprehensive review of robotic gastroscopes and technologies. Cancers 12:2775.

Maza G, Sharma A (2020). Past, present, and future of robotic surgery. Otolaryngologic Clinics of North America 53:935-941.

Morino M, Arezzo A, Forcignanò E (2021) The Medrobotics Platform for Transanal Surgery. Innovative Endoscopic and Surgical Technology in the GI Tract 337-344.

Morino M, Forcignanò E, Arezzo A (2022) Initial clinical experience with a novel flexible endoscopic robot for transanal surgery. Techniques in Coloproctology 26:301-308.

Osawa K, Bandara DSV, Nakadate R, Nagao Y, Akahoshi T, Eto M, Arata J (2022) Stress dispersion design in continuum compliant structure toward multi-dofendoluminal forceps. Applied Sciences 12:2480.

Race A, Horgan S, (2021). Overview of current robotic technology. Innovative Endoscopic and Surgical Technology in the GI Tract 1-17.

Singh, HKSI, Armstrong ER, Shah S, Mirnezami R (2021) Application of robotic technologies in lower gastrointestinal tract endoscopy: A systematic review. World Journal of Gastrointestinal Endoscopy 13:673.

Steiner JA, Hussain OA, Pham LN, Abbott JJ, Leang KK (2019), October. Toward magneto-electroactive endoluminal soft (MEESo) robots. In Dynamic Systems and Control Conference (Vol. 59162, p. V003T20A002). American Society of Mechanical Engineers.

Swanström LL, Pizzicannella M (2023) Future Horizons in Flexible Endoscopy. In The SAGES Manual Operating Through the Endoscope (pp. 973-991). Cham: Springer International Publishing.

Troccaz J, Dagnino G, Yang GZ (2019) Frontiers of medical robotics: from concept to systems to clinical translation. Annual review of biomedical engineering 21:193-218.

Vrielink TJCO, VitielloV, Mylonas GP (2020) Robotic surgery in cancer. In Bioengineering Innovative Solutions for Cancer (pp. 245-269). Academic Press.

Wang B, Chan KF, Yuan K, Wang Q, Xia X, Yang L, Ko H, Wang YXJ, Sung JJY, Chiu PWY, Zhang L (2021) Endoscopy-assisted magnetic navigation of biohybrid soft microrobots with rapid endoluminal delivery and imaging. Science Robotics 6:eabd2813.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0112

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

The security of minimally invasive surgery with an autonomous flexible endoscope

Thiruvenkdam[a] | **Manish Kumar Goyal**[b] | **Vineet Saxena**[c]

[a]Jain (deemed to be) University, Bangalore, India, Associate Professor, Department of Computer Science and Information Technology.
[b]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science and Engineering.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College Of Computing Science And Information Technology.

Abstract The surgical approach, Minimally Invasive Surgery (MIS), tries to limit stress on the patient's body and the number of incisions used during treatment. A flexible tube with a camera and light source known as an endoscope, which enables surgeons to see the interior organs and tissues, is one of the main instruments used in MIS. By expanding the capabilities of conventional endoscopes, Autonomous Flexible Endoscopes (AFE) can potentially change medical information systems. These autonomous devices can perform duties autonomously within the patient's body because they are outfitted with cutting-edge technology, including Artificial Intelligence (AI), computer vision, and robotic systems. Robotic surgical automation is a subject that is gaining more and more popularity. Although it is still impossible to fully automate a system, task and conditional autonomy are quite feasible. Safety is a key issue in robotic surgery, in addition to the performance of job completion. In this article, we introduce AFEs that can assist in autonomously guiding minimally invasive surgical procedures. The Tendon-driven Continuum Technique (TCT) served as the foundation for its development, and it is connected with the da Vinci Training Set (DVTS). There are six Degrees of Freedom (DOF) in the recommended AFEs. The surgical tools are automatically tracked using visual servoing. The AFE's mobility and space occupancy is minimized throughout the tracking process using an optimum control strategy, improving the safety of the robot system and any surrounding assistance. Both empirical and user research findings demonstrate that the suggested AFE has benefits over the current rigid endoscope in terms of safety and reduced space requirements without compromising comfort level.

Keywords: MIS, laparoscopy, surgical robotics, robot safety, visual tracking

1. Introduction

Patients today enjoy major benefits from shorter hospital stays, a lower risk of infection, and a general reduction in problems because of the development of MIS methods during the last several decades. Patients are also drawn to the aesthetic advancement of closed (laparotomic) operations because of the great decrease in scarring. Despite these advantages, MIS may cause extra difficulties for doctors by restricting the use of traditional instruments, preventing full visual and physical access to the operating room, and necessitating additional intensive training (Su et al 2021). The price of MIS rises as a result of these effects. The demand for many MIS treatments is still rising despite their expense to physicians and patients and their additional complexity. Transanal Endoscopic Microsurgery (TEM), compared to traditional transanal excision, improves resection quality, lowers the risk of local recurrence, and increases survival, especially in patients with early-stage rectal cancer with good histology. With comparable morbidity and mortality to traditional transanal excision, the technique of TEM excision of rectal cancers is safe and successful throughout long-term follow-up (Raucci et al 2020).

Although TEM has been in use for more than 20 years, colorectal surgeons have been sluggish to accept it as a standard procedure. This is partly due to a challenging learning curve, but it is also due to the exorbitant expense of the highly specialized apparatus. Technology is still developing quickly. The ability of the surgeon to do minimally invasive surgery has paralleled this development. Instrumentation created for one application may be applied to another via crossover. This is True of Natural Orifice Surgery (NOTES), carried out using TEM instruments and endoscopes (Runciman et al 2019). The Food and Drug Administration (FDA) approved the DVTS for several surgical procedures in 2000, including minimally invasive laparoscopic operations. Intuitive Surgical Company Inc., the DVTS, was developed and provides improved surgical dexterity, motion scaling, and tremor reduction. Sadly, da Vinci's stiff instruments restrict the available surgical entry sites in addition to the high price and huge size (Vo et al 2020).

The FDA has approved the da Vinci SP, a system created expressly for single port surgery, which will allow the da Vinci to be used in various emergent minimally invasive operations, even though these constraints have hampered the acceptance of this platform for usage in many procedures. Modern minimally invasive surgery has produced single-incision, multiport

devices that have made a variety of abdominal surgeries possible. Even single-access laparoscopy colectomies may now be done reliably and securely (Song et al 2020). Single-access laparoscopic working angles are almost equal to TEM working angles. Therefore, there is an overlap between the skill set required for TEM and single-port laparoscopy. However, a major obstacle to the widespread deployment of TEM apparatus continues to be its high initial cost. However, major advancements in the areas indicated above will be needed if a reliable surgical robotic system for these minimally invasive operations is to be realized. Robotic surgery will continue to be an expensive and time-consuming venture. It may struggle for years to gain widespread acceptance within the medical community if efforts are not made to better automate the robotic system and pair this with intuitive, simple-to-command user interfaces (Ma et al 2019). However, further advancements in autonomous robotic control can potentially enhance surgical results without significantly raising the price of these operations.

Surgery will be completely transformed by the thoughtful creation of partially autonomous MIS tools that support surgeons without the increased cost of several operators or much additional training. With its benefits of reduced pain, decreased risk of infection after surgery, and accelerated postoperative recovery, MIS has transformed conventional surgery (Huang et al 2022). In MIS, the endoscope is placed via a trocar to provide surgeons access to the body cavity. Endoscopes now in use have a robust and thin construction. These endoscopes provide several difficulties, including (1) exhausting the endoscope assistants; (2) requiring close coordination between the endoscope assistant and the surgeons; and (3) requiring a considerable motion area while rotating the rigid structure endoscope (Zhang et al 2021). Several solutions have been developed to deal with these problems. Many robotic endoscope-holding devices have been created to ease procedures, prevent surgeon fatigue, and minimize human mistakes (such as shaky hands while controlling the endoscope). To defend against possible cyber-attacks, preserve data confidentiality, and assure patient safety, AFE security is crucial in minimally invasive surgery (Slawinski et al 2018).

To determine the viability of prehabilitation in 22 women having breast cancer surgery, the research (Brahmbhatt 2020) performed a longitudinal, single-arm, mixed-methods study. For the course of their surgical wait period, each participant got a personalized exercise prescription that included strength and mobility training targeted to the upper quadrants as well as aerobic exercise. A detailed evaluation of some of the most prevalent force-torque sensor designs for surgical instruments was conducted in the study (Muscolo and Fiorini 2023), which also considered design and implementation restrictions. These constraints included those related to the robotic surgery environment, surgeon perception, general force-torque sensor design layouts, and force-torque sensors used in robot-assisted minimally invasive surgery. The transmission of surgical expertise is made easier with telementoring. The study (Shabir et al 2021) created a framework for telementoring that allows a specialized surgeon to guide an operating surgeon by imparting knowledge in the motion of surgical tools needed during minimally invasive surgery. An innovative force-sensing tool that supports teleoperated robotic manipulation and semi-automates the suturing process was introduced in the publication (Ehrampoosh et al 2022).

The end-effector mechanism features a rotating degree of freedom to construct the optimal needle insertion trajectory and pass the needle through its curve. An indirect force estimate method based on data-based models was employed to leverage impedance control to give the operator sensory knowledge of the forces between the needle and the tissue. The research (Su et al 2019) displayed whether redundancy is used to ensure a remote center of motion (RCM) restriction and to provide the medical personnel a compliant behavior. To complete the surgical tasks with human-robot interaction, an RCM constraint and a safe constraint are imposed on the nullspace motion based on the established hierarchical control architecture. The precision and safety of the procedure may be impacted by the physical contact. Robot-assisted minimally invasive surgery faces the basic but difficult challenge of automatic equipment separation in the video. By adding a derived temporal before an attention pyramid network for precise segmentation, the research (Jin et al 2019) provided a new framework for using instrument motion information. The inter-frame motion flow propagates the inferred prior from the previous frame to the current frame, which may provide a trustworthy indicator of the instrument placement and form.

The study (Lu et al 2023) offered a unique data-driven architecture with self-contained visual-shape fusion enabling flexible endoscopes to navigate autonomously and intelligently without the need for previous knowledge of system models or large-scale surroundings. The online updating of the eye-in-hand vision-motor arrangement and steering of the endoscope using monocular depth estimate through a vision transformer (ViT) is suggested using a learning-based adaptive visual servoing system. The path difficulty with the strategy (Pittiglio et al 2022) designed a magnetic interaction that would enable the catheter to adjust to the surrounding anatomy while in motion. They create and produce a soft magnetic catheter that is shape-forming, 80 mm long, and 2 mm in diameter. This catheter can follow the leader as it moves through the human body. The device might be used for various endoscopic or intravascular programs, but this study shows how well it works for navigational bronchoscopy.

The flexible endoscope that automatically follows the surgeon's head concerning the surgical target was suggested by the research (Qian et al 2020). by being operated on, the flexible robotic endoscope may see the surgical procedure from the surgeon's point of view. The endoscopic footage is seen by the surgeon while wearing a head-mounted display. The research (Mo et al 2022) described the creation of a novel workflow that, when combined with a flexible continuum robotic system,

offers task autonomy for laser-assisted surgery in confined areas like the GI tract. Contrary to the present, piezoelectric laser steering mechanisms demand the use of high voltage and are dangerous. The research Martin et al (2022) examined autonomous robotic control for magnetic colonoscope treatment through biopsy, another crucial aspect of clinical viability. They have created a variety of robotically autonomous control techniques, such as semi-autonomous routines for locating and carrying out focused and random quadrant biopsies. In this paper, we aim to explore the multifaceted security dimensions in this rapidly evolving field of MIS with an AFE.

2. Methodology

The safety of Minimally Invasive Surgery (MIS) using an Autonomous Flexible Endoscope (AFE) will be discussed in this section.

2.1. Autonomous Flexible Endoscopes (AFE)

The suggested AFE has six degrees of freedom (DOFs), with joints 1-4 of the Process Security Control (PSC) managing its insertion, pitch, roll, and yaw and joints 5 and 6 of the TCT controlling its bending. The TCT was able to bend in two orthogonal directions on its own. It is placed on a strong shaft that can move and rotate as well. Consequently, the final two connections of the flexible portion control the TCT, whereas the first four connections of the stiff segment of the suggested AFE are ordinary rigid connections. The dynamical framework for the 6-DOF flexible endoscopy is provided in the following sections.

A) The mechanics of a Rigid Component

The suggested AFE transition from distance set (set {0}) to set {4} corresponds to the PSC's transition from distance set (set {0}) to set {4} since TCT is mounted on the fourth set of the PSC (origin of set). Denavit-Hartenberg (D-H) technique may be used to calculate the rigid part's forward kinematics:

$$0_{S_4} = 0_{S_1} 1_{S_2} 2_{S_4} 3_{S_4} \qquad (1)$$

Where j_{S_i} the transition from is set {j} to set {i}, and $3_{S_4}, 2_{S_4}, 1_{S_2}, 0_{S_1}$ denotes the homogeneous translation matrices for the first, fourth, third, and second joints, correspondingly. The Jacobian matrix I_{rigid} for the rigid part presented below connects the speed of each joint of the rigid part R_{rigid} to direct kinematics.

$$0_{W_4} = I_{rigid} Q_{rigid} \qquad (2)$$

Where 0_{W_4} denotes the set {4} reflected in distance set {0} velocity vector, composed of linear and angular velocities.

B) Dynamics of Flexible Component

On the fourth set of the PSC, the TCT is placed. As a result, set {4} provides the framework for the flexibility part of the planned AFE. Equation (3), provided at the bottom of this page, may be used to compute the matrix of homogeneous transformations from set 4 to the body set 4_{S_a} of the TCT. Here, θ denotes the flexible part's bending angle, Φ is the TCM's bending direction, and k denotes the flexible part's length. The following Jacobian matrix I_{flex} may be used to determine the linear and angular velocities associated with the flexible portion.

$$0_{S_1} \begin{bmatrix} cos^2\Phi(cos\theta - 1) + 1 & sin\Phi cos\Phi(cos\theta - 1) & cos\Phi sin\theta & \frac{1}{\theta}(1 - \cos(\theta))\cos(\Phi) \\ sin\Phi\, cos\Phi(cos\theta - 1) & cos^2\Phi(1 - cos\theta) + cos\theta & sin\Phi sin\theta & \frac{1}{\theta}(1 - \cos(\theta))\cos(\Phi) \\ -cos\Phi sin\theta & -sin\Phi sin\theta & cos\theta & \frac{1}{\theta}\sin(\theta) \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (3)$$

$$I_{flex} = \begin{bmatrix} sin\Phi(cos\theta - 1)\frac{1}{\theta_1} & \frac{1}{\theta^2} cos\Phi(cos\theta + \theta sin\theta - 1) \\ -cos\Phi(cos\theta - 1)\frac{1}{\theta} & \frac{1}{\theta^2} sin\Phi(cos\theta + \theta sin\theta - 1) \\ 0 & -\frac{1}{\theta^2}(sin\theta - \theta cos\theta) \\ -cos\Phi sin\theta & -sin\Phi \\ -sin\Phi sin\theta & cos\Phi \\ 1 - cos\theta & 0 \end{bmatrix} \qquad (4)$$

The following diagram illustrates the joint speed of the flexible part:

$$4_{W_a} = I_{flex} R_{flex} \qquad (5)$$

Where R_{flex} is the joint velocity relative to the flexible component, the vector $4_{W\,a}$ for the body set {a} of the TCT shown in set {4} consists of the linear velocity and the rotation angle.

C) Mechanics of the 6-DOF AFEs

The frontal mechanic's version of the flexible endoscope with six DOF is illustrated below:

$$0_{S_d} = 0_{S_4}4_{S_b}a_{S_d} = 0_{S_4}4_{S_4} \qquad (6)$$

Where,

$$0_{S_d} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & k_{ad} \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (7)$$

is a continuous connection translation matrix going from camera set {d} to body set {b} of TCT. k_{ad} refers to the separation between the optical center of the endoscopy and the origin of the body set {b}. The 6-DOF flexible endoscope's end-effector speed may be calculated by

$$W = IR \qquad (8)$$

Where R denotes the speed of each AFE joint in the six degrees of freedom. The 6×6 Jacobian matrix provided below provides the speed map of a 6-DOF flexible endoscopy (I):

$$I = [I_{rigid} I_{rigid}] \qquad (9)$$

2.2. Vision Feedback

This section initially explains the suggested AFEs surgical tool tracking mechanism before introducing the visual servoing approach.

A) Tracking Method

The coordinates of the labels on the images must automatically follow the motions of surgical instruments in the suggested 6-DOF flexibility endoscope. The following are some challenges in tracking the surgical equipment: (1) The backdrop of the endoscopic pictures is rapidly shifting; (2) The instruments are rolling; and (3) the total amount of instruments is unknown.

Given that green is seldom seen naturally inside human bodies, it was decided to use it as the marker color in this work. The distal end of the surgical tools is marked with markers. The following is the tracking procedure: (1) initially, each RGB set green region is identified using the color-based picture segmentation technique. After completing the green area identification process, the RGB set is changed to grayscale and smoothed using a Gaussian filter. (2) Following that, corresponding areas that fall under a certain size threshold are removed from the image. To accommodate a variety of cameras, the threshold choice is made based on the dimensions of the source RGB image. (3) The center coordinates of all chosen related regions are then determined using the gray centroids approach. Figure 1 illustrates the proposed 6-DOF elastic endoscopy and is controlled by tracking points according to the average of these centered values.

B) Servo Visual Design

The relationship within the 2D data in images and the 3D data of the world is known as the camera system. The camera's 2D image measurements are transformed into 3D global coordinate data in the next section using there are two distinct types of coordinate schemes. In this study, the visualization concept is the pinhole method, which has the following equation as a formula:

$$Y_d \begin{bmatrix} v \\ u \\ 1 \end{bmatrix} = \begin{bmatrix} e/ox & 0 & v_0 & 0 \\ 0 & e/og & u_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} w \\ z \\ y \\ 1 \end{bmatrix} \qquad (10)$$

Where (v_0, u_0)stands for the endoscope's image capture center coordinate, (v, u) for the retrieved image coordinates after the correction of image distortion, and $O = [w, z, y]^S$the referenced approach is used to fix the endoscope's distortion. The mobility of the suggested flexible endoscopy is then controlled using a cutting-edge 2D visual serving control technique (picture-based control).

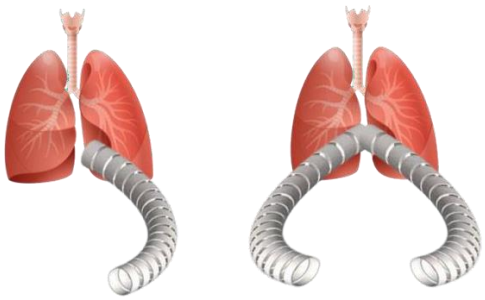


Figure 1 The tip of the surgical equipment is embellished with green markings.

$$\begin{bmatrix} v \\ u \end{bmatrix} = \begin{bmatrix} \frac{\lambda}{y} & 0 & -\frac{v}{y} & -\frac{vu}{\lambda} & \frac{\lambda^2+v^2}{\lambda} & -u \\ 0 & \frac{\lambda}{y} & -\frac{u}{y} & \frac{-\lambda^2-u^2}{\lambda} & \frac{vu}{\lambda} & -v \end{bmatrix} \begin{bmatrix} u_w \\ u_z \\ u_y \\ x_w \\ x_z \\ x_y \end{bmatrix} = [\frac{1}{y} G_U, G_X] \,.\, W = [\frac{1}{y} G_U, G_X] \,.\, I.R \qquad (11)$$

Assume that the Jacobian variable y in the fixed camera system is unknown and that the end-effectors rotates with angular speed $[x_w, x_z, x_y]$ and translation velocity $[u_w,\ u_z, u_y]$ about the camera set. Therefore, to estimate the value of y online, we utilize observations of the robot and image motion. Using least-squares, it is possible to linearly extrapolate the value of y from equation (10).

2.3. Vision-Based Control Technique

This section begins with a proposal for optimum control with limitations on the minimum motion. The 6-DOF flexible endoscope's visual serving control loop is then presented.

A) The Best Control with the Least Movement

The penalty function is used in the optimal control method to optimize $R^l = (r^l_1, r^l_2, r^l_3, r^l_4, r^l_5, r^l_6)$. r^i_j represents the j-th joint's movement at time i. The minimal necessary border limitations are as follows:

$$r^l_1 < r^{max}_1 and\ r^l_2 < r^{max}_2 \qquad (12)$$

The following penalty equation is created with restrictions (12):

$$e(R^l) = \alpha(\|r^l_1 - r^{l-1}_1\| + \|r^l_2 - r^{l-1}_2\|) + \beta(\|r^l_3 - r^{l-1}_3\| + \|r^l_4 - r^{l-1}_4\| + \|r^l_5 - r^{l-1}_5\| + \|r^l_6 - r^{l-1}_6\|) + q_1 \min\{0, r^{max}_1 - r^l_1\}^2 + q_2 \min\{0, r^{max}_2 - r^l_2\}^2 \qquad (13)$$

Where the safety factors q_1 and q_2 ensure that the values of r^l_1 are less than r^{max}_1 and q2, respectively. Weight factors α and β limit mobility in the first and second joints, respectively. Additionally, it is possible to calculate the variables $r^l_1, r^l_2, r^l_3, r^l_4, r^l_5$, and r^l_6 by decreasing the objective function solution (14). Here, the optimization is carried out using the Levenberg-Marquardt method.

$$\text{minimize} \quad e(R^l)$$

$$subject\ to \begin{bmatrix} v \\ u \end{bmatrix} = [\frac{1}{y} G_U, G_X] \,.\, I.R \qquad (14)$$

B) Control System for Instrument Tracking

The PSC is equipped with a designed flexible endoscope, which is computer-controlled. The general layout of the visual servoing control approach is shown in Figure 2.

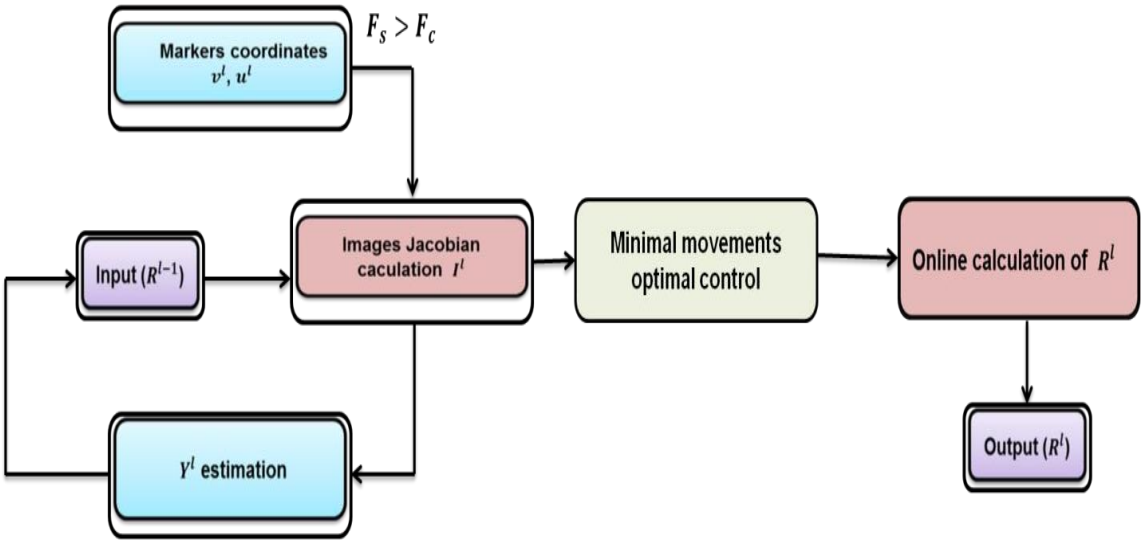


Figure 2 Control loop structure for monitoring inputs.

The endoscope obtains the image coordinate (v^l, u^l) of targets for the tracking structure, as shown in Figure 2. The monitoring fault F_s may then be determined by:

$$F_s = \sqrt{(v^l - v_c)^2 + (u^l - u_c)^2} \qquad (15)$$

Fs represents the allowable error, and $(v_c, u_c)$for the target's intended image coordinate. Equation (10 may be used to estimate the image Jacobian $I^l$and the target's depth measurement if $F_s > F_d$ is true. Following that, the motion of each joint R^lmay be reduced using the suggested optimum control approach. Last, processors will direct the actuators so that the flexible endoscope may move into the proper position and orientation.

3. Result and Discussion

The tracking abilities of both fixed and moving objects are examined in this segment. The recommended 6-DOF elastic endoscopy is then put to the test using the Fundamentals of Laparoscopic Surgery (FLS) exercises. Finally, user encounters with the rigid endoscopy and the recommended AFE are compared.

3.1. Experimental setup

The DVTS and a 6-DOF elastic endoscopy make up the test equipment. The DVTS controller operates the PSC. There are three parts to the flexible endoscope: (1) the distal end is equipped with a camera with a 640 x 480-pixel resolution that connects to a computer via USB. (2) A flexible bending portion comprises 10 vertebrae and an elastic foundation, with a set rate that may exceed 30 Hz. The flexible bending segment has an outside diameter of 7.5 mm and a length of 40 mm; it also has a rolling shaft from DVTS instrumentation and a mounting built by the user.

The calibration technique has been used to calibrate the endoscope, and Table 1 displays the endoscope variables. First-order radial distortion factor l_1, l_2, $o_1$and second-order radial distortion factor o_2 are, respectively, both the second-order radial compression ratio and the first-order radial compression ratio.

Table 1 Endoscope Specifications.

k_1	k_2	u_0	v_0	r_1	r_2	λ
0.0007	0.0013	3.18 pixels	244.9 pixels	0.0089	-0.18	810.3

3.2. Monitoring a Stationary Target

For In this section, the 6-DOF elastic endoscopy system tracks a stationary green marker on three occasions. The marker has a 3 mm diameter. The acceptable error value is 10

pixels F_s. The beginning location is on the bottom right edge, and the target graphic coordinate (v_c, u_c) is the picture's center coordinates (320, 240). Figure 3 displays the tracking trajectories from three iterations of the experiments. Results demonstrate the tracking's dependability.

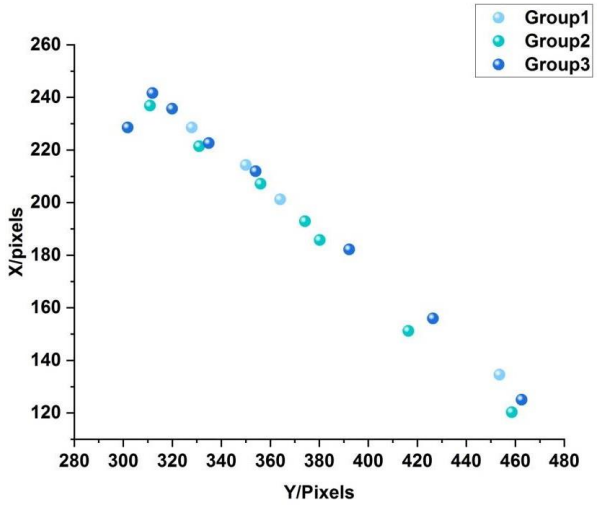


Figure 3 The graphic shows the target tracking trajectory. The target's intended location and beginning position on the image is shown by the pink symbols.

Figure 4 (a) displays the tracking error F_s for every convergence step. According to the color scores, the tracking error F_s value decreases to 10 pixels throughout 11 steps. The value of tracking error F_s during the first four steps falls quickly, from 270 pixels to less than 100 pixels.

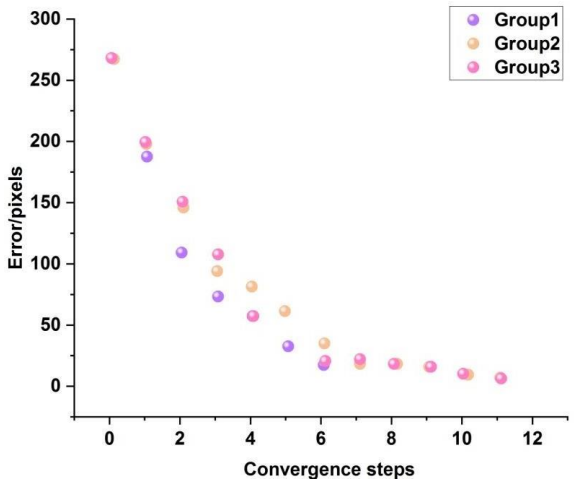


Figure 4 (a) Color points demonstrate that after 11 converging phases, the tracking error F_c value is reduced to 10 pixels.

Figure 4 (b) depicts the tracking time for each stage of convergence. After 1.71 seconds, the tracking error F_s will have decreased from 260 pixels to 10 pixels. Additionally, tracking speed is significantly influenced by the recommended AFE's motion ranges.

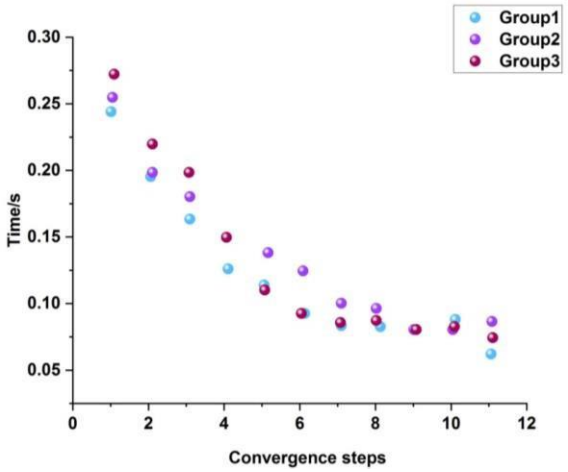


Figure 4 (b) The period of tracking at each resolution step of the tracking procedure is shown by colored spots.

There are five training assignments for the FLS manual skills. Some of these tasks include peg transfer, precision cutting, ligating loops, intracorporeal knotting, and extracorporeal knotting. To operate tools in a sufficiently wide area during the peg transfer, precision cutting, and extracorporeal knotting jobs, it is necessary to manipulate the endoscope to direct the work. To assess the effectiveness of the automated, flexible endoscope, three tasks, peg transfer, precise cutting, and extracorporeal knotting, are selected. Two flexible hand-held tools are used to operate the duties. The three jobs allow for good tracking of the two pieces of equipment. The three FLS tasks have also been required to be completed by 10 people using both the elastic endoscopic and the 4-DOF rigid endoscopic. The two endoscopes' comfort level is then scored (0–10 points). The results are shown in Figure 5. Table 2 displays the benefits of the suggested flexible endoscope.

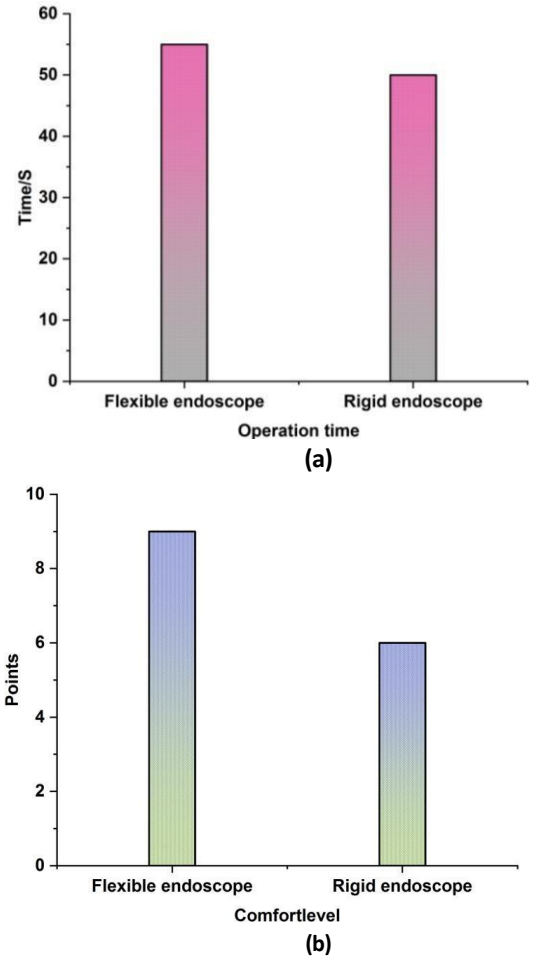


Figure 5 The contrast between the 4-DOF rigid endoscope and the suggested AFE. (a) The peg transfer's operating period. (b) The comfort level ratings.

Table 2 An evaluation of both flexible and rigid endoscopes.

	4-DOF Flexible endoscope	6-DOF Rigid endoscope
Operating	Average level	Average level
Motion space	Small	Large
Comfort level	Average level	Average level
Safety	good	Average level
Viewing field	Relatively large	Relatively small

From Figure 5, it is clear that the proposed AFE has much less interior and exterior occupied area than the 4-DOF rigid endoscope. This may improve the endoscopic method's safety during surgery and provide doctors access to more operating space. The results for task operation time and comfort level while using the two endoscopes that are pretty similar.

4. Conclusions

This study reports an autonomous 6-DOF flexible endoscopy with improved security. A distal bending segment of the suggested AFEs is based on the TCT. It has a total of six DOFs and is connected with the DVTS. It is possible to automatically track instruments using visual servoing. The endoscope movement is reduced during the instrument tracking procedure by using the best control approach. This lessens the risk of arms colliding outside the body and tools fencing within the surgical

cavity. The suggested AFE's performance is accessed via experimental testing. The suggested flexible endoscope can successfully monitor both fixed and moving objects, according to experimental data. This demonstrates how the suggested AFE might greatly decrease safety issues and save space during MIS treatments. Last, user research with 10 participants shown that the flexibility endoscopic may be used for MIS operations and that it could provide a wider field perspective and safer functioning than rigid endoscopes. The existing method has drawbacks, such as blurry images and uneven lighting within bodily cavities. Our ongoing research strives to lessen instrument-related picture blur. In addition, the suggested technology will eventually be integrated with stereoscopic 3D vision.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

Reference

Brahmbhatt P, Sabiston CM, Lopez C, Chang E, Goodman J, Jones J, McCready D, Randall I, Rotstein S, Santa Mina D (2020) Feasibility of prehabilitation prior to breast cancer surgery: a mixed-methods study. Frontiers in Oncology 10:571091.

Ehrampoosh A, Shirinzadeh B, Pinskier J, Smith J, Moshinsky R, Zhong Y (2022) A force-feedback methodology for teleoperated suturing task in robotic-assisted minimally invasive surgery. Sensors 22:7829.

Huang Y, Li J, Zhang X, Xie K, Li J, Liu Y, Ng CSH, Chiu PWY, Li Z (2022) a surgeon preference-guided autonomous instrument tracking method with a robotic flexible endoscope based on dvrk platform. IEEE Robotics and Automation Letters 7:2250-2257.

Jin Y, Cheng K, Dou Q, Heng PA (2019) Incorporating temporal prior from motion flow for instrument segmentation in minimally invasive surgery video. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, Proceedings, Part V 22, pp. 440-448. Springer International Publishing.

Lu Y, Wei R, Li B, Chen W, Zhou J, Dou Q, Sun D, Liu YH (2023) Autonomous Intelligent Navigation for Flexible Endoscopy Using Monocular Depth Guidance and 3-D Shape Planning. arXiv preprint arXiv:2302.13219.

Ma X, Song C, Chiu PW, Li Z (2019) Autonomous flexible endoscope for minimally invasive surgery with enhanced safety. IEEE Robotics and Automation Letters 4:2607-2613.

Martin JW, Barducci L, Scaglioni B, Norton JC, Winters C, Subramanian V, Arezzo A, Obstein KL, Valdastri P (2022) Robotic autonomy for magnetic endoscope biopsy. IEEE Transactions on Medical Robotics and Bionics 4:599-607.

Mo H, Li X, Ouyang B, Fang G, Jia Y (2022) Task Autonomy of a Flexible Endoscopic System for Laser-Assisted Surgery. Cyborg and Bionic Systems.

Muscolo GG, Fiorini P (2023) Force-Torque Sensors for Minimally Invasive Surgery Robotic Tools: An Overview. IEEE Transactions on Medical Robotics and Bionics.

Pittiglio G, loyd P, da Veiga T, Onaizah O, Pompili C, Chandler JH, Valdastri P (2022) Patient-specific magnetic catheters for atraumatic autonomous endoscopy. Soft Robotics 9:1120-1133.

Qian L, Song C, Jiang Y, Luo Q, Ma X, Chiu PW, Li Z, Kazanzides P (2020) January. FlexiVision: Teleporting the surgeon's eyes via robotic flexible endoscope and head-mounted display. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3281-3287. IEEE.

Raucci, M.G., D'Amora, U., Ronca, A. and Ambrosio, L., 2020. Injectable functional biomaterials for minimally invasive surgery. Advanced healthcare materials 9:2000349.

Runciman M, Darzi A, Mylonas GP (2019) Soft robotics in minimally invasive surgery. Soft robotics 6:423-443.

Shabir D, Abdurahiman N, Padhan J, Trinh M, Balakrishnan S, Kurer M, Ali O, Al-Ansari A, Yaacoub E, Deng Z, Erbad A (2021) Towards development of a tele-mentoring framework for minimally invasive surgeries. The International Journal of Medical Robotics and Computer Assisted Surgery 17:e2305.

Slawinski PR, Taddese AZ, Musto KB, Sarker S, Valdastri P, Obstein KL (2018) Autonomously controlled magnetic flexible endoscope for colon exploration. Gastroenterology 154:1577-1579.

Song C, Ma X, Xia X, Chiu PWY, Chong CCN, Li Z (2020) A robotic flexible endoscope with shared autonomy: a study of mockup cholecystectomy. Surgical endoscopy 34:2730-2741.

Su H, Mariani A, Ovur SE, Menciassi A, Ferrigno G, De Momi E (2021) Toward teaching by demonstration for robot-assisted minimally invasive surgery. IEEE Transactions on Automation Science and Engineering 18:484-494.

Su H, Yang C, Ferrigno G, De Momi E (2019) Improved human–robot collaborative control of redundant robot for teleoperated minimally invasive surgery. IEEE Robotics and Automation Letters 4:447-1453.

Vo CD, Jiang B, Azad TD, Crawford NR, Bydon A, Theodore N (2020) Robotic spine surgery: current state in minimally invasive surgery. Global Spine Journal 10:34S-40S.

Zhang X, Li W, Ng WY, Huang Y, Xian Y, Chiu PWY, Li Z (2021). An autonomous robotic flexible endoscope system with a DNA-inspired continuum mechanism. In 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 12055-12060. IEEE.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0113

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

EEG-Based brain-machine interface for categorizing cognitive sentimental emotions

Feon Jaison[a] | **Inzimam**[b] | **Shambhu Bhardwaj**[c]

[a]Jain (deemed to be) University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology.
[b]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science & Application.
[c]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College of Computing Science and Information Technology.

Abstract: The development of brain-machine interfaces (BMIs) has revolutionized the study of neuroscience by making it possible for the brain to communicate directly with outside objects. In this study, EEG brainwave data is used to categorize emotional experiences using both individual and group methods. We employ a four-electrode resolution (TP9, AF7, AF8, and TP10) commercial MUSE EEG headband. Film clips with clear emotional content elicit both good and negative emotions. For one minute per session, neutral resting data is also obtained without external stimuli. To do this, we use machine learning algorithms to decode and interpret participant EEG data as they perform activities that evoke a range of emotional reactions. Relevant characteristics are collected from the EEG signals using intensive data preprocessing, feature extraction, and selection approaches to identify the underlying patterns of cognitive sentimental feelings. Following that, the development and assessment of classification models, such as Gradient Artificial Neural Networks (G-ANN), use the retrieved features as input. In conclusion, this study presents an EEG-based BMI system for categorizing cognitive sentimental emotions. The proposed G-ANN achieves a high accuracy of 98.59%, demonstrating superior performance compared to existing methodologies.

Keywords: BMIs, EEG Signals, CSE, G-ANN, MUSE EEG

1. Introduction

The IAPE'18 conference records can be found in the proceedings. Potential applications for autonomous, noninvasive emotional state sensing include mental healthcare and human-robot interaction. It can add a new level of interaction between the user and the gadget and make it possible to derive concrete information without spoken communication (Bird et al 2019). One of humanity's most distinctive traits is emotion, which impacts behavior. Being able to comprehend and analyze human emotions is essential to life. A number of industries, including gaming, therapy, and medicine, use human–machine interfaces and cooperation. Researchers consistently work to increase the adaptability and efficiency of computer-human interaction to attain high levels of user satisfaction (Kumar 2021). Emotion is essential in human-to-human interaction and communication. They can be divided into three general groups. The first category looks at verbal communication, nonverbal cues, and facial expressions. Noncontact emotion detection is made possible by these audio- visual techniques (Chatterjee and Byun 2022). The identification of neural states is essential in many disciplines, including brain-machine interaction (BMI) and psychological studies. Different techniques have been used in various prior studies to measure participant differences in brain state. To build a BMI system using information from brain functional magnetic resonance imaging (fMRI).

The EEG-based approach is one of the brain imaging methods utilized in related research, and it is mostly used to examine neural activities. The authors indicated that one benefit of using EEG signals in related studies is their great temporal resolution, which can reach the millisecond level. Additionally, more recent EEG collection devices have received favorable assessments for clinical and real-world BMI use in terms of signal dependability and mobility (Jung et al 2022). In addition to assisting in the diagnosis of brain tumors from electroencephalography (EEG), the study of brain waves also aided in the treatment of intellectual deficiencies, including Alzheimer's disease, sleep disorders, and pilot tiredness. Electrons pass over the nerves, and the EEG data change regardless of how the human body impacts the brain. Numerous study methodologies have been developed by researchers who are enhancing brain-EEG communications to evaluate emotional states. Simply said, the degree of human focus pertains to the many stages of the brain that exist when you carry out a mental task. A student might be alert or surprised during a presentation, for instance, if their brains are not working together. The recognition of each student's level of attention by them is covered in this article. Other demographics, such as office workers, truck drivers, pilots, and others, may also have their attention spans assessed, even though kids make up the

majority of the data used for these tests (Alam et al 2022). The use of bioinspired algorithms as effective and reliable optimization techniques is widespread. Despite criticism that they are computationally expensive, they have been successful in solving challenging optimization issues. Because they are good at maximizing the solutions to complex problems, bioinspired algorithms are becoming increasingly common as computing resources become increasingly readily available. The human brain is the source of all other degrees of control, and electroencephalography can be used to track it. Before the invention of dry, commercial electrodes, EEG was an invasive and uncomfortable procedure, but now it is entirely available even outside of lab settings (Bird et al 2019). A simple galvanometer was used to achieve the original result, which proved that brain activity could be measured using this method (Värbu et al 2022). In 1973, the first request for a BMI was made. Since then, there have been an increasing number of uses of brain-computer interaction in fields as varied as forensics, architecture, entertainment, education, and health. On the other hand, the fields of augmented reality (AR) and virtual reality (VR) have attracted increased attention in recent years due to the ambition to make computer systems more immersive and to design the next generation of user interfaces. Because they enable the researcher to immerse a user in simulated settings created and controlled to achieve particular objectives or to elicit certain behaviors, virtual and augmented reality bring up new and intriguing options for the application of BMI interaction (NWAGU et al 2023). It has also been used to direct the use of an external rehabilitation or assistive device for patients who have been diagnosed with a motor impairment. It has been demonstrated that BMI systems can use neurofeedback to apply the neural plasticity notion (Orban et al 2022). Although it had been thoroughly explored, the expression was first used by. BMIs are currently being used by able-bodied users in areas such as gaming, emotion detection, alertness assessment, and mental tiredness measurement (Wu et al 2020). The process of communicating information between a human and a machine to carry out specified tasks using a particular “conversational” language is known as human-computer interaction (HCI). The emotion's primary objective is to recognize the emotion or to enable the computer to comprehend the user's mental processes, and HCI systems must communicate (Xue et al 2020). Stress is a key cause for concern in contemporary life. The World Health Organization (WHO) claims that mental health issues and lost productivity have cost the global economy significant money. When the body is subjected to unpleasant stimuli, stress is the result. Short-term stressful situations that cause acute stress are often reflected by fleeting physiological changes. Episodic stress may develop from acute stress that lasts for a very long time. Chronic exposure to stressors or traumatic events can lead to chronic stress, anxiety, and clinical depression. The major topic of this essay is severe stress. Video games have been chosen as the stimulus to study acute stress (Roy et al 2022). This is due to its numerous uses in areas such as neurorehabilitation, neuroprosthetics, and gaming, where it would be very beneficial to decode users' perceptions of imagined movements (Padfield et al 2019). Electroencephalography (EEG), which is popular because it is simple to use and noninvasive, is one of the most widely used imaging techniques. Modifying the stated principles for idea classification increases the method's efficacy (Aggarwal and Chugh 2022). The classification of motor imagery (MI) based on EEGs is a crucial aspect of BMIs, which connect the neurological system and computing devices by turning brain signals into intelligible machine commands (Roy 2022). The purpose of this work is to provide a thorough comparison of conventional classification techniques and to highlight the value of deep learning-based BMI methods, particularly multilayer perceptron (Sharma 2022). A BMI that uses EEG opens a channel of communication between the brain and outside objects. These applications improve the lives of healthy people by, among other things, encouraging productivity, teamwork, and personal development. This article's goal is to provide a comprehensive overview of EEG-based BMI application research from 2009 to 2019 (Värbu et al 2022). People's emotional states have a significant impact on their behavior and physiological interactions. One potential medical use is the identification of patients' mental illnesses. People work and communicate more successfully when they feel well. Negative feelings can be harmful to one's physical and emotional well-being. Due to the rapidly evolving field of machine learning, many early studies that looked at EEG usage for emotion classification concentrated on gathering information from the entire brain. However, researchers are unable to explain the relationship between different emotional states and EEG characteristics (Chatterjee and Byun 2022). Human emotions are the fundamental component that determines a person's cognitive ability. These represent the brain's response to internal or external events. To distinguish different human emotions in a patient who is completely paralyzed, brain-machine interfaces must be used. EEG is one technique for capturing cerebral activity in a brain-machine interface to assess human emotions. The classification of human emotions has been made possible by technologies such as machine teaching (Parveen et al 2023). The neural interfaces can be used for a variety of things, including the control of external equipment and the stimulation or inhibition of a particular brain region's activity. The purpose of neural interfaces is to link the brain to other brains or other external equipment. Other uses for the brain include managing external equipment, keeping track of certain thoughts, and dispensing medication through circulation. EEG electrodes, a data acquisition system, an EEG amplifier, signal processing software, data acquisition, management, and visualization software comprise the BMI (Patel 2023). This section covers cognitive sentimental emotion categorization using an EEG-based brain-machine interface.

2. Methodology

This section discusses cognitive sentimental sentiments using an EEG-based brain-machine interface.

2.1 Dataset

The study used a MUSE EEG headgear readily accessible off the shelf to connect four dry extracranial electrodes. The TP9, AF7, AF8, and TP10 electrodes monitor microvoltages. Two volunteers provided data for each of the six movie clips in Table 1 for a total of 60 seconds, yielding 12 minutes of data on brain activity. This resulted in a dataset containing 324,000 data points from brain waves that were resampled at a variable frequency of 150 Hz.

Table 1 Lövheim categories and their labeled encapsulated feelings.

Emotion Category	Emotion/Valence
A	Surprise (Negative) (Lack of Dopamine)
B	Anger (Negative) Rage (Negative)
C	Interest (Positive) Excitement (Positive)
D	Shame (Negative) Humiliation (Negative)
E	Contempt (Negative) Disgust (Negative)
F	Fear (Negative) Terror (Negative)
G	Enjoyment (Positive) Joy (Positive)
H	Distress (Negative) Anguish (Negative)

2.2. Electroencephalography

Electroencephalography involves utilizing applied electrodes to obtain electrophysiological information and brain waves. Electrodes can be implanted in the brain or subdurally or under the skull. Wet or dry electrodes must be positioned all over the skull for noninvasive procedures. Raw electrical data are measured in microvolts (uV), causing wave forms between t and t+n.

2.3. Personal feeling

Despite being complicated and varied, human emotions may often be divided into two categories: good and negative. Some emotions are similar to one another that are "hope" and "anguish," which, accordingly, are viewed as favorable and bad but are frequently felt at the same moment.

Take the painful hope for a character's life in a movie that is unquestionably doomed as an example. According to Lövheim's three-dimensional emotional paradigm, generalized positive and negative valence emotions are related to the chemical makeup of the brain, as shown in Table I, where each vertex of the model corresponds to a different category of emotions, A through H. A range of chemical compositions can be linked to both positive and negative emotions. Additionally, research demonstrates that chemical composition affects neural oscillation, which in turn affects the production of electrical brainwaves. According to this study, emotions may be classified using statistical traits of the brainwaves that are produced since they are chemically encoded and directly alter electrical brain activity.

2.4. Preprocessing Using Min-Max Normalization

The preprocessing method Min-Max normalization, sometimes referred to as feature scaling, is frequently used in machine learning to rescale numerical features within a particular range. The objective is to maintain the original distribution while bringing all feature values to a similar scale. Scaling numerical characteristics within a certain range is often done using min-max normalization. It is often used to obtain data ready for models in machine learning. Min-max normalization may be used to normalize the input characteristics of the dataset for diabetes prediction.

Min-max normalization is a technique of normalizing that uses linear modifications to the original data to provide a fair comparison of values before and after the procedure**.**

$$Z_{new} = \frac{z=min}{\max(z)-\min(z)} \qquad (1)$$

Z_{new}= The adjusted value obtained after scaling the data
Y=outdated value
$Max(Y)$ = Dataset's highest possible value
$Min(Y)$ = Dataset's lowest possible value

2.5. Mental emotional sentiment classification using a gradient – artificial neural network (G-ANN)

Time series forecasting, pattern identification, and process control are just a few of the scientific and technology domains where G-ANN models have found extensive use. G-ANNs have been effectively utilized to model a range of different functions since the late 1980s. Through an autonomous training procedure, the network may intelligently learn these functions. Many network architecture-related concerns, however, are still not well understood. According to several

academics, G-ANNs are a "black box" method that cannot offer significant and practical insights into the underlying nature of physical processes. A G-ANN makes a very primitive attempt to replicate the structure and operation of the human mind and brain. It can be described as a system consisting of an interconnected network of simple neurons. Numerous nodes connected by links, often grouped in a number of layers, make up the network structure. After each node in a layer processes weighted input from a lower layer, links are used to communicate each node's output to nodes in the upper layer. Each link is given a weight, which is a numerical assessment of the strength of the connection. A transfer function converts a node's weighted sum of inputs into an output.

There are three equations that characterize the backpropagation algorithm. In each learning step k, weight connections are first altered

$$\Delta X^{t}_{ji(l)} = \eta(s)\delta^{t}_{oi}{}^{(t-1)}_{w_j} + n\Delta X^{t}_{ji(l-1)} \qquad (2)$$

Second, the information below is correct for output nodes.

$$\delta^{p}_{oi} = (c_i - p_i)e'(J^{t}_{i}) \qquad (3)$$

and third, it is true for the other nodes.

$$\delta^{p}_{oi} = e'_i(J^{t}_{i})\sum_l \delta^{(t+1)}_{b} X^{(l+1)}_{il} \qquad (4)$$

where $w^{(t)}_{i}$is the actual output of node i in layer T; $X^{(t)}_{il}$is the weight of the connection between node i at layer (t-1) and node i at layer J; $\delta^{(t)}_{oi}$is the measure for the actual error of node i; $J^{(t)}_{(i)}$is weighted.

2.6. Feature extraction using principal component analysis (PCA)

PCA, also known as Karhunen–Loeve expansion, is a well-known feature extraction and data representation approach in the domains of pattern recognition and computer vision. PCA is a well-liked dimensionality reduction technique for feature extraction in machine learning and data analysis. Through the projection of the data onto a lower-dimensional space, it seeks to identify the most significant patterns and variations in the data. PCA does this by locating a collection of main components—orbifold axes—that maximize the variance in the data. By extracting the distinguishing features from the target face and integrating them into a single, linear face as a result of the feature extraction procedure, the Eigenfaces technique achieves its desired results. The face is projected into the space created by the eigenfaces to achieve recognition. The eigenvectors of the eigenfaces and the related picture are compared in terms of their Euclidian distance.

The steps are as follows:

Step 1: The set of M images (B1, B2, B3, etc.) with a size of N*N is represented by a column or row vector of size N2

Step 2: The description of the training set image average (μ)

$$\mu = \frac{1}{n}\sum_{m=1}^{N} Am \qquad (5)$$

Step 3: Each trainee image has a different average image by a vector (W)

$$Xj = Aj - \mu \qquad (6)$$

Step 4: As demonstrated below, the total scatter matrix or covariance matrix is calculated from

$$D = \sum_{m=1}^{N} xmxmt = BBS, \qquad (7)$$

$$where\ A = [W1W2W3 \ldots Wn] \qquad (8)$$

Step 5: Calculate the covariance matrix C's eigenvalues L and eigenvectors UL.

Step 6: The images can be categorized using this feature area. The weight vectors are measured

$$\Omega S = [x1, x2, \ldots., xN'], \qquad (9)$$

Whereby,

$$Gl = VlS(A - \mu), l = 1,\ 2, \ldots., N' \qquad (10)$$

3. Results and Discussion

To demonstrate how successful a given method is, its dependability and effectiveness are compared to those of more established techniques such as long short-term memory networks (LSTMs) (Jeevan et al 2019), deep neural networks (DNNs) (Zhang et al 2022), and convolutional neural networks (Roy 2022). It has been recommended that ANNs be used in EEG-based brain-machine interfaces. These approaches are compared to conventional methods based on a variety of parameters, including accuracy, precision, recall, and implementation cost.

3.1. Accuracy

When categorizing cognitive sentimental feelings, accuracy is defined as the use of being accurately categorized to the overall number of occurrences. Figure 1 displays the accuracy of the planned and current systems. Because of its accuracy, it has been suggested that the proposed ANN be used to classify cognitive sentimental feelings. While the DNN has a 93% accuracy rate, the CNN has a 90% accuracy rate, and LSTM has a 92.5% accuracy rate, the suggested method has a 97% accuracy rate. This demonstrates that the suggested method is more accurate than the current method. Table 2 shows the accuracy values.

$$Accuracy \rightarrow \frac{TP+TN}{TP+FP+FN+TN} \qquad (11)$$

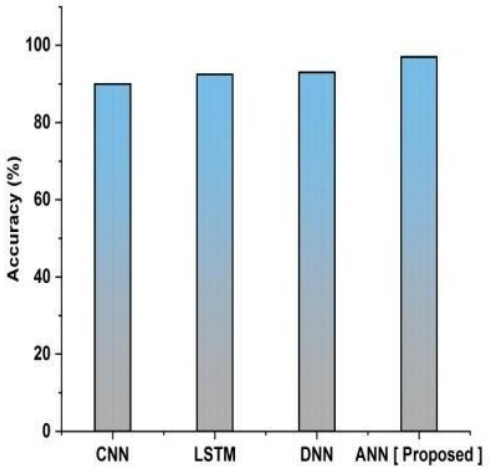


Figure 1 Accuracy.

Table 2 Accuracy.

Methods	Accuracy (%)
CNN	90
LSTM	92.5
DNN	93
ANN [Proposed]	97

3.2. Precision

To classify cognitive sentimental feelings, a classification model's ability to recognize only the relevant data points is used. Figure 2 displays the accuracy of the current and proposed systems. The proposed ANN accuracy has been suggested for application in classifying cognitive sentimental feelings. While the suggested method has a 96.2% precision, DNN has a precision of 95.4%, CNN has a precision of 91%, and LSTM has a precision of 93.6%. This demonstrates that the proposed method is more precise than the current method. Table 3 displays the precision values.

$$recision \rightarrow \frac{TP}{TP+FP} \qquad (12)$$

Table 3 Precision.

Methods	**Precision (%)**
CNN	91
LSTM	93.6
DNN	95.4
ANN [Proposed]	96.2

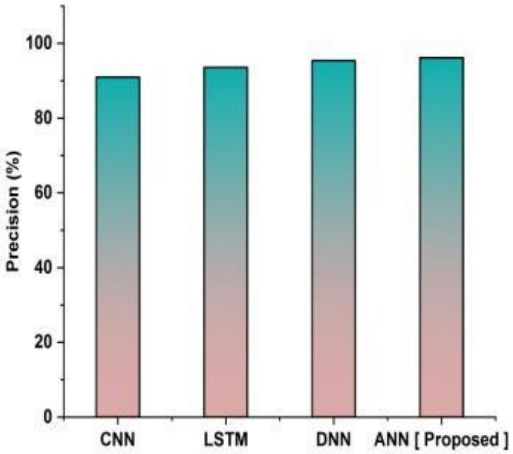


Figure 2 Precision.

3.3. Recall

Recall is quantitatively calculated as the sum of the true positives minus the false negatives. A cognitive sentimental feeling may use a model's capacity to find all significant events in a batch of data. The recalls for the current and proposed systems are shown in Figure 3. The projected ANN recall has been suggested for use in cognitive sentimental feeling. DNN has a recall of 95.4%, CNN has an accuracy of 89%, and LSTM has a precision of 90.1%. The proposed system has a recall of 95%. This demonstrates that the suggested method has a higher recall rate than the current method. The recall values are displayed in Table 4.

$$Recall \rightarrow \frac{TP}{TP+FN} \qquad (13)$$

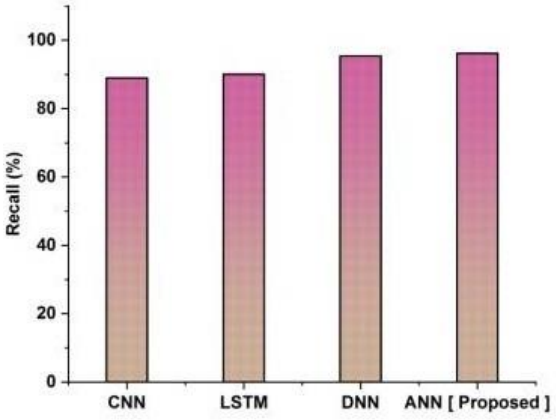


Figure 3 Recall.

Table 4 Recall.

Methods	Recall (%)
CNN	89
LSTM	90.1
DNN	95.4
ANN [Proposed]	96.2

3.4 F1-measure

The F1 measure is a statistic that evaluates the overall effectiveness of a classification model or system by combining accuracy and recall. It is frequently used to judge a model's effectiveness in appropriately identifying and classifying sentimental cognitive sensations. The harmonic mean of recall and accuracy, which adds the two metrics to obtain a single result, is the F1 measure. Figure 4 depicts the F1 measure for the proposed and existing systems. The proposed ANN's F1 measure has been suggested for usage in cognitive sentimental sensation. The suggested approach obtains an F1-measure of 92.6%, compared to 90.4% for DNN, 85.4% for CNN, and 87.3% for LSTM. This illustrates that the suggested technique

outperforms the current technique in terms of the F1-measure. Table 5 displays the values for the F1 measure. It is computed using the following formula:

$$F1 - Measure \rightarrow 2 * \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (14)$$

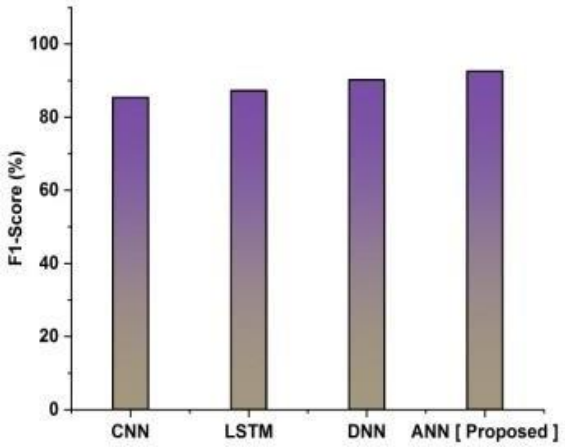


Figure 4 F1- Measures.

Table 5 F1 Measure.

Methods	F1-Measure (%)
CNN	85.4
LSTM	87.3
DNN	90.2
ANN [Proposed]	92.6

4. Conclusions

In this research, windowed data from four places on the scalp were used to evaluate the application of individual and ensemble classification approaches to measure the participant's emotional state at that specific moment. The techniques showed that it may be possible to assess a participant's emotional state using a readily available, low-resolution EEG headband. There is a great deal of opportunity to develop categorization algorithms with practical uses for real decision support systems. In systems that support mental health, responding to emotional states can improve communication and aid in the overall assessment of issues and problem-solving strategies. BMIs based on EEG have shown promise in a number of applications, including the classification of sentimental and cognitive emotions. Although the discipline is still developing, there are a number of promising new paths and prospective improvements for EEG-based BMIs.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

Reference

Aggarwal S, Chugh N (2022) Review of machine learning techniques for EEG-based brain-computer interface. Archives of Computational Methods in Engineering 1-20.

Alam MS, Jalil SZA, Upreti K (2022) Analyzing recognition of EEG-based human attention and emotion using Machine learning. Materials Today: Proceedings 56:3349-3354.

Bird JJ, Ekart A, Buckingham CD, Faria DR (2019) April. Mental, emotional sentiment classification with an eeg-based brain-machine interface. In Proceedings of the International Conference on Digital Image and Signal Processing (DISP'19).

Bird JJ, Faria DR, Manso LJ, Ekárt A, Buckingham CD (2019) A deep evolutionary approach to bioinspired classifier optimization for brain-machine interaction. Complexity.

Chatterjee S, Byun YC (2022) EEG-Based Emotion Classification Using Stacking Ensemble Approach. Sensors 22:8550.

Jeevan RK, SPVMR, Kumar PS, Srivikas M (2019) EEG-based emotion recognition using LSTM-RNN machine learning algorithm. In 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), pp. 1-4. IEEE.

Jung D, Choi J, Kim J, Cho S, Han S (2022) EEG-Based Identification of Emotional Neural State Evoked by Virtual Environment Interaction. International Journal of Environmental Research and Public Health 19:2158.

Kumar A, Kumar A (2021) DEEPHER: Human emotion recognition using an eeg-based deep learning network model. Engineering Proceedings 10:32.

NWAGU C, ALSLAITY A, ORJI R (2023) EEG-Based Brain-Computer Interactions in Immersive Virtual and Augmented Reality: A Systematic Review.

Orban M, Elsamanty M, Guo K, Zhang S, Yang H (2022) A Review of Brain Activity and EEG-Based Brain–Computer Interfaces for Rehabilitation Application. Bioengineering, 9(12), p.768.

Padfield N, Zabalza J, Zhao H, Masero V, Ren J (2019) EEG-based brain-computer interfaces using motor-imagery: Techniques and challenges. Sensors 19:1423.

Parveen KS, Panachakel JT, Ranjana H, Sidharth S, Samuel AA (2023) EEG-based Emotion Classification-A Theoretical Perusal of Deep Learning Methods. In 2023 2nd International Conference for Innovation in Technology (INOCON), pp. 1-6. IEEE.

Patel R (2023) the Substantial Role of AI In EEG-Based Brain-Computer Interface Application.

Roy AM (2022) An efficient multiscale CNN model with intrinsic feature integration for motor imagery EEG subject classification in brain-machine interfaces. Biomedical Signal Processing and Control 74:103496.

Roy AM (2022) An efficient multiscale CNN model with intrinsic feature integration for motor imagery EEG subject classification in brain-machine interfaces. Biomedical Signal Processing and Control 74:103496.

Roy S, Islam M, Yusuf MSU, Jahan N (2022) EEG-based stress analysis using a rhythm-specific spectral feature for video game play. Computers in Biology and Medicine 148:105849.

Sharma R, Kim M, Gupta A (2022) Motor imagery classification in brain-machine interface with machine learning algorithms: Classical approach to multilayer perceptron model. Biomedical Signal Processing and Control 71:103101.

Värbu K, Muhammad N, Muhammad Y (2022) Past, present, and future of EEG-based BCI applications. Sensors 22:3331.

Wu D, Xu Y, Lu BL (2020) Transfer learning for EEG-based brain–computer interfaces: A review of progress made since 2016. IEEE Transactions on Cognitive and Developmental Systems 14:4-19.

Xue B, Lv Z, Xue J (2020) November. Feature transfer learning in EEG-based emotion recognition. In 2020 Chinese Automation Congress (CAC), pp. 3608-3611. IEEE.

Zhang J, Liu D, Chen W, Pei Z, Wang J (2022) Deep Convolutional Neural Network for EEG-Based Motor Decoding. Micromachines 13:1485.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0114

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Deployment of the deep learning fusion method to emotional semantic evaluation of natural language

Gulista Khan[a] | **Sanjeev Kumar Mandal**[b] | **Sunil Sharma**[c]

[a]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College of Computing Science and Information Technology.
[b]Jain (deemed to be) University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology.
[c]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science and Engineering.

Abstract The emotional semantic evaluation of natural language plays a crucial role in sentiment analysis. Deep learning methods have shown great potential in capturing the complex relationships between words and emotions. This paper proposes a deep learning fusion method for deploying emotional semantic evaluation. The technique combines multiple deep learning architectures to capture local and global contextual information, including Bidirectional Gated Recurrent Units (GRU), Long Short-Term Memory (LSTM) networks, and Self-Attention mechanisms. Pretrained GloVe word embedding's utilized to enhance word representation. A novel fusion layer combines the outputs of individual models; employing self-attention means to assign weights dynamically. This allows the model to weigh the importance of different representations in the final prediction. Benchmark movie review (MR) for sentiment analysis and emotion classification tasks are used to evaluate the proposed method. Experimental results demonstrate superior performance compared to individual deep learning models and traditional feature-based approaches. The proposed fusion method effectively captures the nuances of emotional semantics in natural language, leading to more accurate and nuanced evaluations.

Keywords: deep learning, GRU, LSTMs, self-attention mechanisms

1. Introduction

Emotions fundamentally impact human communication and comprehension. They affect how we see the world, how we act, and how we want to perceive the world (Patil et al 2019). There has been a rise in interest in comprehending and assessing the emotional content encoded in natural language as computational linguistics (ClinLing) and natural language processing (NLP) have developed. The dynamic semantic evaluation of natural language is the task of evaluating the emotional valence, intensity, and attitude provided by text or audio data. It seeks to identify and extract emotional cues from verbal phrases, from straightforward assertions to intricate narratives.

The investigation of emotional semantics in natural language has wide-ranging effects (Shaheed et al 2021). Businesses can determine client happiness and sentiment toward their products or services by using sentiment analysis techniques, such as recognizing the emotional undertone of customer reviews or social media remarks. Dynamic semantic analysis can be used in psychotherapy and mental health research to analyse and track emotional states and identify potential mental health problems.

Manual annotation or language conventions are frequently used in traditional techniques for emotional semantic evaluation (Ayzeren et al 2019). These techniques, however, have a limited ability to scale and require considerable time and labor. Researchers have been able to create automated systems that can learn from and infer emotional content from vast volumes of text data thanks to the development of machine learning and deep learning techniques.

Traditional methods for evaluating dynamic semantics frequently rely on manual annotation or language conventions (Sommerfeldt et al 2019). These techniques, however, have a limited potential to scale and are time- and labor-intensive. Researchers have created automated systems to learn from vast amounts of text data and infer emotional content. This has been made possible by developing machine learning and deep learning techniques.

Recurrent neural networks (RNNs) and attention mechanisms, in particular, have shown promise in recent developments in deep learning for capturing the nuanced emotional semantics found in natural language (Shopon et al 2021). These models can focus on the temporal dynamics and environmental factors influencing emotional expression.

Our research aims to create a technique that accurately captures the subtleties of emotional semantics in natural language. By utilizing various architectures' advantages and including local and global contextual data, we seek to outperform our deep learning models and conventional feature-based techniques.

The additional divisions of this article are as follows: Part 2 introduces related works, Part 3 discusses the methodology, Part 4 assesses the efficiency of the proposed method, and Part 5 concludes the paper.

2. Related works

Guarino et al. (2022) created a method employing open-ended semantic questions to evaluate psychological variables, including emotions, thoughts, and attitudes. The findings demonstrated competitive or greater reliability and validity compared to rating scales, indicating that the semantics questioning method had good statistical features. According to these results, natural language-based semantic measurements may enhance and extend the use of conventional rating scales in the size and description of psychological variables. A study (Svetlakov et al 2021) proposed the Semantic-Emotion Neural Network (SENN), a unique neural network design that addresses the shortcomings of current emotion recognition models in NLP. It comprises two subnetworks: a CNN for word interactions and emotional aspects and a single-directional LSTM for background information and semantic linkages. The SENN model has shown much better emotion identification than existing techniques.

The study (Dargan and Kumar 2020) used bigrams, a semantic distance, to uncover new information on topic flow and conceptual cohesiveness in samples of continuous English. The distance measurements were tested against data from simulated verbal fluency, and bigram length norms were created using a substantial corpus of text. The method shows potential in elucidating how semantics are processed in real-world tales and bridges the gap between small-scale single-word research and more thorough discourse analysis. Jain and Kanhangad (2019) proposed a graph-based, semisupervised method for capturing and modelling the various language subtleties and events involved in textual emotion expression. Our approach provides the building blocks for developing contextualized affect representations by generating rich structural descriptors. They use word embedding to improve these representations and test their effectiveness on different emotion recognition tasks. The experimental results show that our approach is superior to state-of-the-art methodologies. Sharma et al. (2021) compared several methods for determining sentiment or mood in interactions between therapists and clients during psychotherapy sessions. The researchers used a database of 97,497 statements made during psychotherapy sessions to carry out their investigation. The BERT model was trained using sentiment assessments made by people. By comparing the individual kappa values of each model, the researchers assessed the performance of each. These technological developments offer a viable way for researchers to conduct extensive research on sentiment analysis in psychotherapy, previously constrained by time-consuming and manual procedures.

Guarino et al. (2023) suggested a big data sentiment analysis technique based on sensitive information subjects. The aim is to use natural language processing technology to extract sentiment tendencies from the massive volume of text data produced on the Internet. The process described in the study uses a neural network model to combine semantic subject data into a text representation. The attention mechanism in the model aids in determining the significance of every sentence in the text. In the age of big data, the model's adaptability is further improved by incorporating sentiment dictionary labelling, making it a useful tool for monitoring public opinion and sentiment analysis. Pereira et al. (2023) investigated the application of NLP (natural language processing) methods to collect and analyse emotions expressed on social networking platforms. With the ability to gather data continuously across time and reduce the input of data and other stressor-related errors, the method has advantages over more conventional approaches. By analysing social media data, NLP can be a potent tool for clinical analytics and healthcare informatics, giving significant insight into patients' sentiments and emotions.

The objective of Casanova et al. (2021) was to create a fresh deep neural network sentiment analysis model. The model optimizes grid search-based hyperparameters and integrates long short-term memory (LSTM) with convolutional neural networks (CNNs). The study's findings show that in terms of accuracy for sentiment analysis, the suggested LSTM-CNN-grid search-based neural network model performs better than the baseline models. A study (Cascone et al 2020) employs a neural network model to integrate topic semantics into a representation of text. Adding a system for attention, which uses a context-aware vector to determine each word's weight, allows for integration. The experimental findings in the research show how well the suggested model works to increase the validity of sentiment assessment results.

Terhörst et al. (2019) suggested a theoretical framework for mining messages to extract fine-grained information on disasters, such as affected parties, damaged infrastructure, and disrupted services. First, they employ LSTM networks to categorize tweets about disasters into two categories to obtain higher accuracy by preserving long-term semantic dependencies.

3. Proposed Method

3.1. Dataset- MV (movie review)

A total of 10662 favorable and negative movie reviews are included in this dataset, with each review consisting of one sentence. Predicting sentences' positive and negative feelings is difficult.

3.2. Text Embedding

We are aware that neural networks cannot be fed text directly. First, we must translate the text's words into numbers. Therefore, using the acquired representation of the expression vectors as input, we pretrain an unsupervised embedding vector from GloVe2 to produce the word embedding.

If the vocabulary count of the dataset is $k$and the input sentence Y has o words with an average word embedding dimension of$m$, the matrix of embedding B will have a dimension space of J^{r*k}. As a result, here is how the sentence is represented in the input:

$$Y(s_1, s_2, s_3, \ldots\ldots s_n), Y \in J^{r*o} \qquad (1)$$

$J^r$is the length of the space of every word in the lexicon.

3.3. BI-GRU

The recurrent gated unit (GRU) was first proposed as a standard recurrent neural network (RNN) variation. With inputs Y_s and L_S, GRU computes G_s for each point t as follows:

$$j_s = \sigma(G_j, y_s + R_j, l_s - 1) \qquad (2)$$

$$R_s = \sigma(G_r, y_s + R_r, l_s - 1) \qquad (3)$$

where l_s, j_s and ut stand for the reset gate, update gate, and d-dimensional hidden state, respectively. The GRU has three parameters: $G_s, G_r, G_t G_b$ and R_s, R_t, R. The sigmoid function, or elementwise production, is represented by the symbol. We represent the preceding context for a word at s using the hidden state from the forward GRU h_s and the following context for that word at s using the hidden state from the backwards GRU, l_s, which reverses the text encoding. The bidirectional contextual encoding of $y_s$is concatenated into the string l_s= [l_s ; l_s] and used as the Bi-GRU layer's output at time s.

3.4. LSTM

Long short-term memory (LSTM) is an architectural type of recurrent neural network (RNN). It was created to solve the issue of vanishing gradients, which prevents typical RNNs from capturing long-term dependencies in sequential data. In this scenario, the angles gradually disappear or increase exponentially over time. The input gate decides the amount of received data that should be stored in each memory cell. Each memory cell component obtains a value between 0 and 1 after being fed the current input and the prior concealed state through a sigmoid activation function. The forget gate determines information removed from the memory cell. It uses the prior hidden state and the current input as inputs, applies a sigmoid activation function, and outputs a value between 0 and 1 for each memory cell element. The information that should be forgotten is determined by this gate as no longer relevant. The output gate decides how much of the contents of the memory cell should be used to compute the output of the LSTM unit. A value between 0 and 1 is produced for each memory cell component using the current input and the prior hidden state as inputs and a sigmoid activation function. This gate aids in the selective output from the memory cell of pertinent information.

$$i_s = \sigma(W_i, x_s + U_i, h_s - b_i) \qquad (4)$$

$$f_s = \sigma(W_f, x_s + U_f, h_s - b_f) \qquad (5)$$

$$o_s = \sigma(W_o, x_s + U_o, h_s - b_o) \qquad (6)$$

3.5. Self-attention

In neural networks, particularly in models such as transformers, a process known as self-attention, also known as scaled dot-product attention, is employed to record significant links between various parts in a sequence. Self-attention enables the concurrent processing of all sequence components, in contrast to recurrent neural networks (RNNs), which only process rows sequentially. Each sequence element is split into query, key, and value vectors during self-attention. These vectors are constructed from the input sequence to calculate attention scores between various items. The attention scores determine the relevance or importance of each component with others. The dot-product operation, scaling, and softmax normalization are used to calculate attention scores.

After being received, the attention scores are utilized to weigh the appropriate value vectors. The contextual information for each element in the sequence is then captured by adding the weighted values to create a weighted representation. This enables the model to suppress useless or noisy input and prioritize relevant elements. Compared to conventional recurrent architectures such as GRU (gated recurrent unit) and LSTM (long short-term memory), self-attention has some advantages. Because attention ratings can give larger weights to significant parts regardless of where they are in

the sequence, it makes it possible to capture long-range relationships more effectively. Self-attention also enables parallel computation, making it quite effective for handling lengthy arrangements.

4. Result

The effectiveness of the suggested and current methods is assessed in this section. The parameters are recall, accuracy, precision, and MAE. FSS-GCN (Rodgers et al 2021) and GCNN-LSTM (Fenu et al 2021) are the current processes.

Accuracy is a metric for how well a model forecasts the results or labels of a specific dataset. It is computed by dividing the total number of forecasts made by the number of accurate predictions, commonly expressed as a percentage. The accuracy value is greater than that of our proposed method. Figure 1 depicts the accuracy outcome. By comparison, it shows that the value of our proposed method (proposed method 96%) is superior to existing methods (FSS-GCN 83%, GCNN-LSTM 90%).

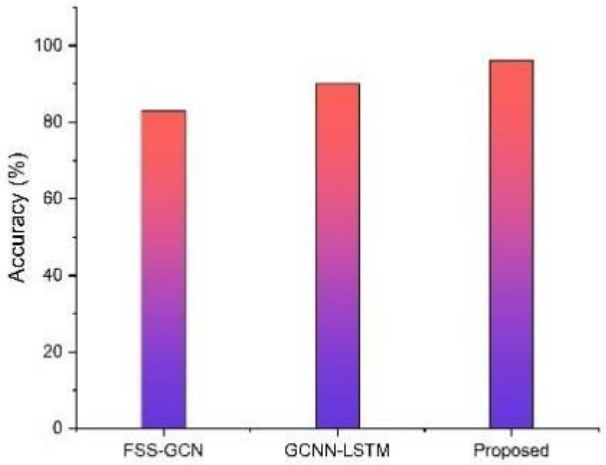


Figure 1 Outcome of accuracy.

The accuracy of categorization or prediction models is evaluated using a statistical metric called precision. Out of all expected positive events (true positives plus false positives), it determines the proportion of events that were accurately anticipated to be positive (true positives). Figure 2 depicts the precision outcome. By comparison, it shows that our proposed method (proposed method 95%) is better than existing methods (FSS-GCN 81%, GCNN-LSTM 89%).

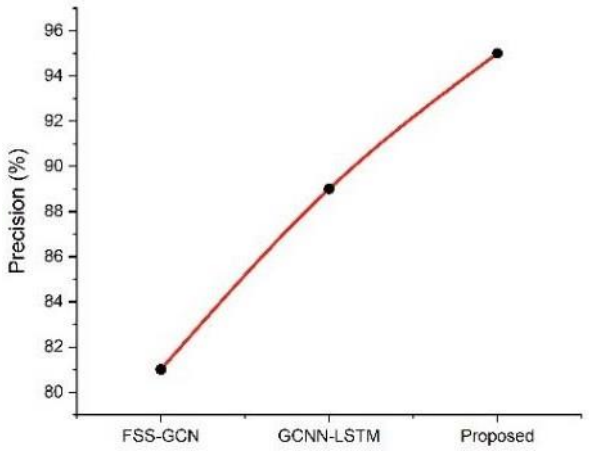


Figure 2 Outcome of precision.

Recall, also known as sensitivity or true positive rate, expresses how many positive occurrences the model correctly accepted. The proportion of real positives to all real and erroneous negatives is calculated. The recall value is larger than that of our proposed method. Figure 3 depicts the recall outcome. By comparison, it shows that our proposed method (proposed method 93%) is greater than existing methods (FSS-GCN 85%, GCNN-LSTM 88%).

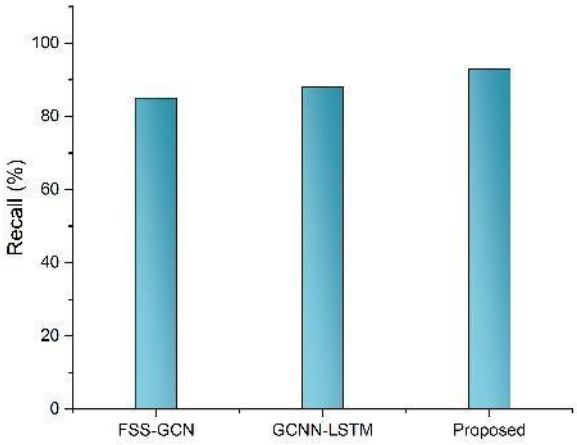


Figure 3 Outcome of recall.

The acronym MAE stands for mean absolute error. It is a typical metric to assess a collection's average error size or deviations between expected and observed values. MAE measures how closely the predictions match the actual data without considering the error's direction. MAE is less than other values. Figure 4 depicts the MAE outcome. By comparison, it shows that our proposed method (proposed method 0.6) is improved over existing methods (FSS-GCN 1.8, GCNN-LSTM 1.3).

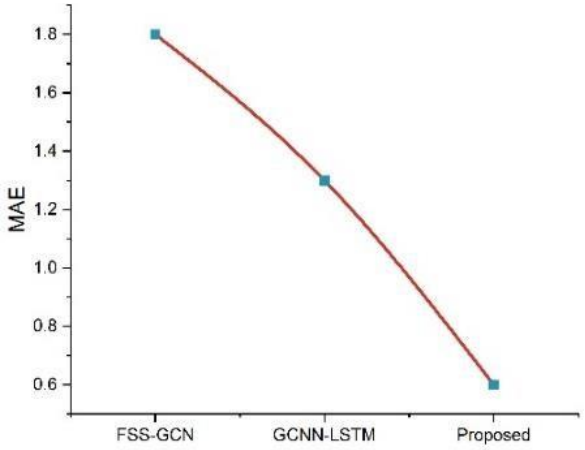


Figure 4 Outcome of MAE.

5. Conclusion

This paper proposes a deep learning fusion method approach for evaluating emotional semantics in natural language. The technique gathers local and global contextual data using long short-term memory (LSTM) networks, self-attention mechanisms, and bidirectional gated recurrent units (GRUs), allowing for a more thorough comprehension of the connections between words and emotions. To improve word representation, trained GloVe embeddings of words are also used. Benchmark movie review (MR) for sentiment evaluation and emotional task classification is used to assess the proposed technique. According to the experimental results, it performs better than single machine learning models and conventional feature-based methods. The fusion method's skillful capture of the intricacies connected with emotional meaning in natural language allows for a better assessment of sentiment and attitude. This study demonstrates the possible use of deep learning fusion techniques for the textual representation of intricate emotional interactions. It improves sentiment evaluation and emotion classification tasks by proposing a more sophisticated and precise method for interpreting and deciphering emotion semantics in natural language. Large-scale, high-quality datasets are essential for training deep learning models. Emotional semantic assessment necessitates labelled data with precise emotional annotations, but these may be few or challenging to acquire. Emotional semantic evaluation can be improved by incorporating other modalities, including text, audio, and visual data. The development of deep learning fusion techniques that successfully integrate several modalities to increase the precision and robustness of emotional analysis can be the subject of future research.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Ayzeren YB, Erbilek M, Çelebi E (2019) Emotional state prediction from online handwriting and signature biometrics. IEEE Access, 7:164759-164774.

Casanova A, Cascone L, Castiglione A, Meng W, Pero C (2021) User recognition based on periocular biometrics and touch dynamics. Pattern Recognition Letters 148:114-120.

Cascone L, Medaglia C, Nappi M, Narducci F (2020) Pupil size as a soft biometrics for age and gender classification. Pattern Recognition Letters 140:238-244.

Dargan S, Kumar M (2020) A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. Expert Systems with Applications 143:113114.

Fenu G, Marras M, Medda G, Meloni G (2021) Fair voice biometrics: Impact of demographic imbalance on group fairness in speaker recognition. In Interspeech (pp. 1892-1896) International Speech Communication Association.

Guarino A, Lettieri N, Malandrino D, Zaccagnino R, Capo C (2022) Adam or Eve? Automatic user gender classification via gesture analysis on touch devices. Neural Computing and Applications 34:18473-18495.

Guarino A, Malandrino D, Zaccagnino R, Capo C, Lettieri N (2023) Touchscreen gestures as images. A transfer learning approach for soft biometric traits recognition. Expert Systems with Applications 219:119614.

Jain A, Kanhangad V (2019) Gender recognition in smartphones using touchscreen gestures. Pattern Recognition Letters 125:604-611.

Patil A, Kruthi R, Gornale S (2019) Analysis of multimodal biometrics system for gender classification using face, iris and fingerprint images. International Journal of Image, Graphics and Signal Processing 11:34.

Pereira TM, Conceição RC, Sencadas V, Sebastião R (2023) Biometric recognition: A systematic review on electrocardiogram data acquisition methods. Sensors 23:1507.

Rodgers W, Yeung F, Odindo C, Degbey WY (2021) Artificial intelligence-driven music biometrics influencing customers' retail buying behavior. Journal of Business Research 126:401-414.

Shaheed K, Mao A, Qureshi I, Kumar M, Abbas Q, Ullah I, Zhang X (2021) A systematic review on physiological-based biometric recognition systems: current and future trends. Archives of Computational Methods in Engineering 1-44.

Sharma SV, McWhorter JW, Chow J, Danho MP, Weston SR, Chavez F, Moore LS, Almohamad M, Gonzalez J, Liew E, LaRue DM (2021) Impact of a virtual culinary medicine curriculum on biometric outcomes, dietary habits, and related psychosocial factors among patients with diabetes participating in a food prescription program. Nutrients 13:4492.

Shopon M, Tumpa SN, Bhatia Y, Kumar KP, Gavrilova ML (2021) Biometric systems deidentification: Current advancements and future directions. Journal of Cybersecurity and Privacy 1:470-495.

Sommerfeldt SL, Schaefer SM, Brauer M, Ryff CD, Davidson RJ (2019) Individual differences in the association between subjective stress and heart rate are related to psychological and physical well-being. Psychological science 30:1016-1029.

Svetlakov M, Hodashinsky I, Slezkin A (2021) Gender, age and number of participants effects on identification ability of EEG-based shallow classifiers. In 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), pp. 0350-0353. IEEE.

Terhörst P, Damer N, Kirchbuchner F, Kuijper A (2019), June. Suppressing gender and age in face templates using incremental variable elimination. In 2019 International Conference on Biometrics (ICB), pp. 1-8. IEEE.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0115

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Investigation of signature biometric evidence using deep learning for forecasting personality

Abhilash Kumar Saxena[a] | **Mir Aadil**[b] | **Sandeep Kumar Jain**[c]

[a]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College of Computing Science And Information Technology.
[b]Jain (deemed to be)University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology.
[c]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Electrical Engineering.

Abstract A crucial field of study, personality assessment has applications in marketing, human resources, and psychology. Existing approaches, however, frequently rely on arbitrary self-report surveys, which might introduce biases and constraints. As a result, there is a need for trustworthy and impartial personality testing tools. Using signature biometric data, we suggest a Salp swarm-optimized Siamese neural network (SSO-SNN) in this research to forecast personality attributes. In order to gauge the effectiveness of the suggested strategy, we first collect information from the CIU handwritten database. A median filter is used during pre-processing to reduce noise in the gathered signature photos. The SSO-SNN uses the SSO algorithm's optimization capabilities to speed up the SNN's learning process. Numerous tests are run to determine how effective the suggested strategy is in terms of metrics like accuracy, precision, recall, and f1-score. Experimental outcomes demonstrate the better performance of the suggested SSO-SNN technique in forecasting personality traits when compared to existing approaches.

Keywords: SBE, personality traits, SSO-SNN

1. Introduction

In recent years, biometrics has advanced significantly, enabling the identification and authentication of people based on certain physical or behavioral characteristics. While the primary use of biometric devices has been to identify people for security reasons, there is growing interest in the technology's ability to anticipate and comprehend human behavioral traits. Due to its accessibility and convenience of use, signature evaluation has become one of the more attractive biometrics techniques (Feldmann et al 2021). Signatures have long been acknowledged as a distinct and individual type of expression, and new breakthroughs in machine learning present a chance to extract useful data from these signs. The application of physical features, such as fingerprinting or facial expressions, for the sake of recognition, has historically been the primary emphasis of the area of biometrics, sometimes known as "biometrics." Signatures, on the other hand, present a different biometrics mode that can reveal information about a person's unique qualities. This has the potential to have substantial ramifications in a variety of sectors, including psychology for forensic use, criminology, and human resource management, among others (Feng et al 2018). Numerous fields, such as mental health, personnel, advertising, and customized services, can benefit greatly from knowing an individual's character. Questionnaires and interviews are two common traditional methods for evaluating character, although they are biased and subject to subjectivity. A goal and perhaps more accurate method of assessing personality is provided by biometric-based techniques. It is possible to gain important insights into a person's personality traits and tendencies by studying their signature, which acts as a unique and personal mark (Aarva et al 2019).

Many fields of study have been transformed by deep learning, a branch of machine learning that includes computer vision and natural language processing. It has the potential to considerably improve the analysis of complicated biometric characteristics because of its capacity to autonomously learn hierarchical representations from huge quantities of data. The readability, fullness, flourishes, t, and dots also have a significant impact on an individual's character. Handwriting is distinctive to each person, and it is called brain writing. Each neural brain structure produces a particular type of muscle control that is the same for every person with that particular character. Handwriting can indicate a person's morality, hidden talents, medical conditions, previous encounters, and psychophysiological states. It can also show psychiatric illnesses and physical problems. Hence, it is used as a diagnostic tool. Graphics is the study of an individual's writing or signatures to analyze their personalities and ascertain their character or mental state. It also aids in the field of forensics. Personality recognition is an empirical technique that recognizes and assesses an individual's emotions and how they affect the style and structure of their handwriting or signatures (Willemink et al 2020). The handwritten signatures include dimensions, upper, lower, and middle zones, borders, angles/slants, velocity, stress, vertical and horizontal lines, and sometimes thrive a stroke

in bilingual writings, most of which indicate first and last names as well as a statement made by individual, signifying his/her consensus or existence on the piece of paper. The handwritten signatures incorporate a variety of fluid and innate behavioral aspects that are helpful in identifying an individual's soft biometrics traits, such as age, gender, ethnicity, skin tone, scars, handedness, personality traits, height, etc. (Adeyemi et al 2017). In this study, we suggest a Salp swarm-optimized Siamese neural network (SSO-SNN) for predicting personality traits using signature biometric data.

The remainder of the paper is divided into the sections below. The method is explained in Part 3. Part 4 contains the result. In Part 5, the conclusions are discussed.

2. Related works

The study (Jiang et al 2021) was to use prior CT scans to create a deep learning-based signature that may forecast mortality and the usefulness of additional chemotherapy. In order to achieve a binary categorization in identifying gender utilizing the numerical characters of the signature dataset, the study (Kumar et al 2022) concentrated on employing six distinct ML techniques. Various character features, such as equilibrium, optimism, pessimism, confidence, etc., may be seen in the examples of handwriting used in their work (Saraswal and Saxena 2022). Their approach was capable of identifying each personality utilizing KNN methods. The study (Siraj et al 2022) conducted binary categorization for recognizing gender utilizing a total of six different ML methods, utilizing the numerical capability of the signatures data. The study (Ayzeren et al 2019) reported the results of an exhaustive examination of a variety of single-modal and multimodal biometrics records, with a focus on the accessibility, designations, material, and sample counts of signatures and handwriting biometrics. The study (Landau et al 2020) was to provide evidence that it is possible to invade a subject's privacy by means of EEG information that was obtained from the use of BCI apps. The study (Portugal et al 2019) intended to use feature regression modeling to discover brain signatures throughout fluid emotive facial processing that are forecasting indicators of depression and anxiety throughout a range ranging from healthy to pathologic values. The study (Shin et al 2022) was to provide an ML-based technique for categorizing individuals as adults or kids based solely on the handwriting data they provided. The two kinds of handwritten datasets used in their work were handwritten words and handwritten structure, both of which were gathered using a pen tablet. The research (Agduk and Aydemir 2022) was to figure out from handwritten specimen images, which vary from individual to individual, the gender of the individual who penned the words as well as their identity. The study (Saleema and Thampi 2020) was to determine whether it is possible to create biometric templates from user profiles that satisfy the reliability and uniqueness requirements for biometric identification.

3. Method

3.1. Data Collection

One hundred fifty individuals, all from CUI in South Indonesia, built the records. Fifty signatures and five handwritten texts were produced by each individual. The collection roughly includes 7500 signature items and 750 handwritten text items. Demographic data is displayed in Table 1.

Table 1: Demographic Data.

Contributor	Gender	Age Group	Handedness	Education Level
1	Female	< 25	Left	Bachelor's
2	Male	25-45	Right	Master's
3	Female	25-45	Left	PhD
4	Male	< 25	Left	High School
…	…	…	…	…
150	Male	46+	Right	Bachelor's

3.2. Data Pre-processing (Median Filter)

A popular method for reducing noise in image processing is median filtering. It works especially well to eliminate impulse noise or salt-and-pepper noise, in which isolated pixels exhibit high values in comparison to their neighbors. Each pixel value in an image is replaced by the median value of its neighbors as part of the median filter's operation. The neighborhood's size, also known as the window or kernel size influences how much of a filtering impact there is. As the median value is less impacted by outliers than other statistical measures like the mean, the median filtering technique successfully eliminates the extreme values that are present in the image. While retaining crucial edges and features, the final image has less noise.

3.3. Siamese Neural Network (SNN)

A series of M convolutional layers, an Inception-style layer, and a fully linked layer make up each twin in our SNN. For the first twin and second twin of the Siamese Network, accordingly, we indicate by $m_{1,n}$, and $m_{2,n}$ with $\in \{1, 2, \ldots, N\}$.

A $L_n \times L_n \times M_n$ kernel that is identical for both twins exists in the $m_{1,n}$ and $m_{2,n}$ levels. This indicates that Nm filtering with Km sizes exists. As depicted in Figure 1, the Inception-type level combines five rows of layers with convolution. We just change the Nm filtering length and treat the kernel size $L_n \times L_n$ as constant. Every single layer has an appropriate buffer and an established pace of 1. Additionally, a Batch Normalization layer and an Activator layer using the "Rectified Linear Units (ReLU)" function come after them. A max-pooling level with a filter's size and cadence of 2 follows every successive block of convolutional layers. Certain convolution layers can be created by stacking two convolution layers. After the beginning level, a stride two average-pooling follows.

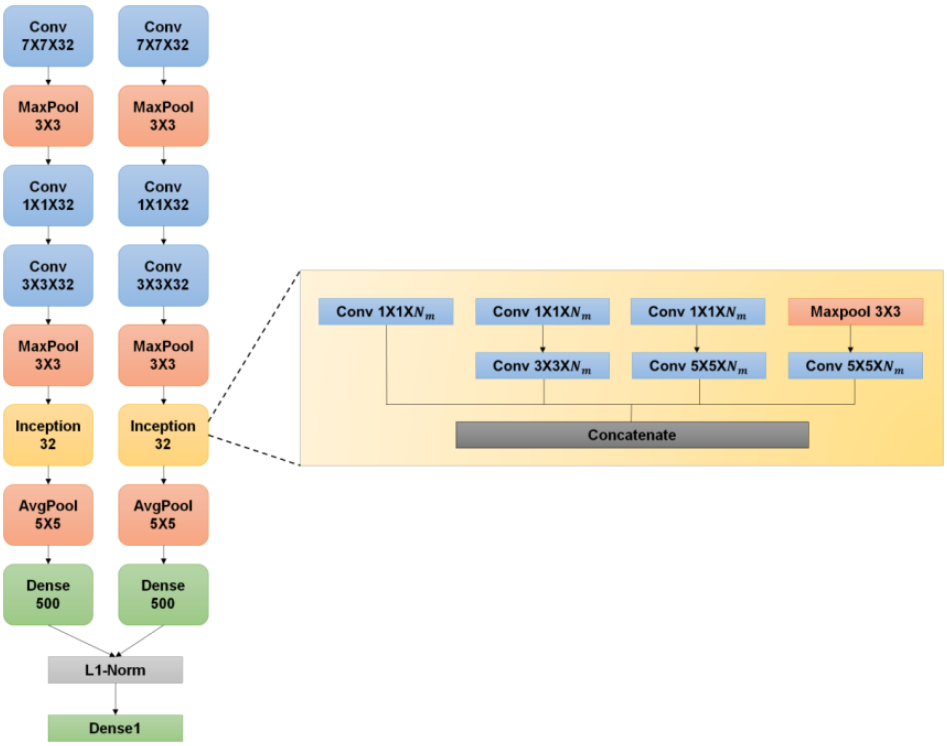


Figure 1 SNN for random signature forgeries.

The following equation can be used to represent the output of the $n-th$ filter when applied to the $math$ section:

$$\begin{cases} m_{1,n}^{(m)} = maxpool(max(0, X_{n-1,m}^{(m)} * m_{1,n-1} + b_n), 2) \\ m_{2,n}^{(m)} = maxpool(max(0, X_{n-1,m}^{(m)} * m_{2,n-1} + b_n), 2) \end{cases} \quad (1)$$

Where $X_{n-1,n} \in M_{l_n \times l_n \times 1}$ are the corresponding outcomes of one or two layered convolutions, and $\forall\, n \in \{1, \dots, N\}$, $\forall\, m \in \{1, \dots, N_m\}$ stands for the collection of a matrix with sizes $e_1 \times e_2 \times e_3$. Furthermore, the convolutional process that results from using the convolutional filtering does not apply perfect overlapping among the filter and input data, and c_n indicates the biased vectors for the $m - layer$. The entirely linked level, or $N+1$ layer, contains M_{N+1} neurons. The result matrices of $M_n \times L_N \times L_N$ are melded into a $1 \times (M_n \times L_N \times L_N)$ dimension matrix, and this new matrix serves as the input for the first fully linked layer. As an activation operation, a sigmoid equation is used. As a result, the result is calculated as follows:

$$\begin{cases} m_{1,n+1}^{Mn+1} = \sigma(X_{N,N+1} . flatten(m_{1,N}) + C_{n+1}) \\ m_{2,n+1}^{Mn+1} = \sigma(X_{N,N+1} . flatten(m_{2,N}) + C_{n+1}) \end{cases} \quad (2)$$

The fully-connected layered separation, which was calculated utilizing a M_1 norm, connects the twin parts:

$$e = |m_{1,n+1}^{Mn+1} - m_{2,n+1}^{Mn+1}| \quad (3)$$

As this separation is a huge vector, it is additionally taught with an entirely connected layer utilizing a sigmoid as its activating function for summarizing its data. The final result of the Siamese network, as demonstrated by Equation 4, is close to 1 if both twins generate a value that is identical and close to 0 otherwise.

$$q = \sigma(X_{N+1,N+2} \cdot e + c_{N+2}) \quad (4)$$

3.4. Salp Swarm Optimization (SSO)

The SSO is a brand-new swarm-based intelligent optimizer that took inspiration from slaps swarms' feeding patterns. Slaps resemble jellyfish in that they are colourless and have a body form resembling a container. These monsters are linked to one another via a slap network in order to communicate and identify food supplies fast while battling for additional sources of food. Different slaps perform different roles within the slap swarms, including leaders and subordinates. Everybody is guided by the leaders, and the followers do as they are told.

Finding the best food supply, shown as F in the search area, is the final objective of swarms of ants. Similar to previous swarm-based optimizers, a starting population with the size and placement of people is predefined. Every single person is an option for the ideal targeted solutions. According to Equation 5, the two-dimensional matrices V_j can be used to represent the space of entire solutions.

$$V_j = \begin{bmatrix} v_1^1 & v_2^1 & \cdots & v_e^1 \\ v_1^2 & v_2^2 & \cdots & v_e^2 \\ \vdots & \vdots & \cdots & \vdots \\ [v_1^m & v_2^m & \cdots & v_e^m] \end{bmatrix}. \quad (5)$$

Using the statistical approach, where leaders and followers follow separate formulas, the initialized salps are then modified. Equation 6 provides updates to the salp leaders, which are important for feeding and navigating as previously described.

$$v_n^1 = \begin{cases} G_n + d_1((mc_n - kc_n)d_2 + kc_n) & d_3 \geq 0 \\ G_n - d_1((mc_n - kc_n)d_2 + kc_n) & d_3 \geq 0 \end{cases} \quad (6)$$

While mc_n and kc_n stand in for the higher and power boundaries of the mth dimensions, accordingly, and where v_n^1 indicates the location of the leader in the mth dimensions, Fm denotes the location of the food supply for the mth dimensions. The [0, 1] range is the range of unknown integers for parameters d_2 and d_3. These two factors control the search step size and later search orientation in the domain of search towards $+\infty$ or $-\infty$ a parameter called d_1 acts as a switch between exploiting depth and exploring depth. It often decreases as the number of repetitions rises and can be calculated using Equation 7:

$$d_1 = 2f^{-(\frac{4i}{I})^2} \quad (7)$$

Where i and I stand for this particular iteration's number and the total number of iterations, correspondingly, the locations have been modified for salp followers in accordance with Equation 8:

$$v_n^j = \frac{1}{2}(v_n^j + v_n^{j-1}) \quad (8)$$

Where the location of the ith salp for the mth dimensions is represented by v_n^jm and $j \geq 2$

Like other metaheuristic methods, SSA optimization techniques have strong flexibility in avoiding local optimal situations.

4. Results and discussion

This section conducts an analysis of the following metrics: Accuracy (%), Precision (%), Recall (%), and F1-score (%). Comparisons are made between the previous approaches, "Deep Neural Network (DNN)" and "K-Nearest Neighbors (KNN)".

A typical measure of assessment used in machine learning and classification tasks is accuracy, which assesses how accurate predictions made by a model are overall. It displays the proportion of accurately anticipated occurrences among all instances. The accuracy level for both the proposed and existing methodologies is shown in Figure 2. In contrast to the suggested method SSO-SNN, which achieves 98.5% accuracy, DNN and KNN only manage to obtain accuracy levels of 89.3% and 93.4%, respectively. The SSO-SNN methods has higher accuracy when compared to existing methods.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

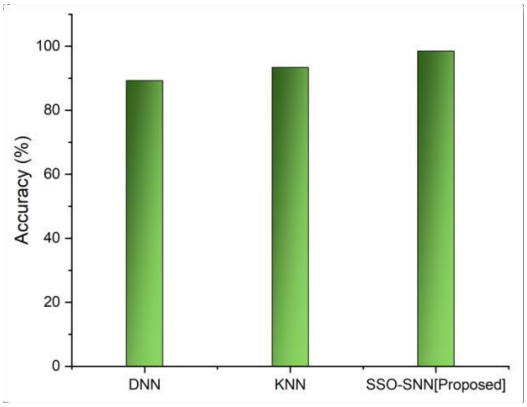


Figure 2 Accuracy Outcome.

In machine learning and information retrieval tasks, precision is a regularly used efficiency parameter. It gauges the percentage of accurate positive predictions among all positive forecasts produced by a model or system. In other words, precision measures how accurately positive predictions come true. Figure 3 displays the Precision level for both the proposed and existing methods. DNN and KNN only manage to achieve Precision levels of 90.4% and 94.5%, compared to the suggested approach SSO-SNN's 97.6%, respectively. Comparing the SSO-SNN methods to other approaches, it has a greater Precision.

$$Precision = \frac{TP}{TP+FP} \qquad (10)$$

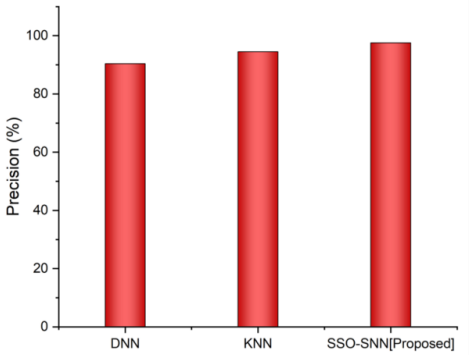


Figure 3 Precision Outcome.

A recall is a statistic used to assess the categorization model's efficiency, especially in binary classification problems. It assesses the model's capacity to accurately distinguish positive cases from all of the actual positive examples. Figure 4 displays both the proposed and existing approaches' recall levels. The recall values of DNN and KNN are just 89.5% and 91.4%, respectively, compared to the suggested approach SSO-SNN's 98.1% recall. When compared to other methods, the SSO-SNN methodology has a higher recall.

$$Recall = \frac{TP}{TP+FN} \qquad (11)$$

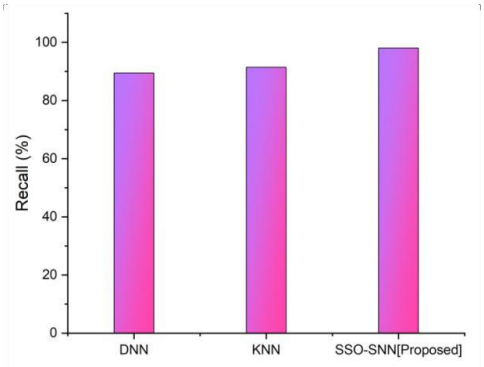


Figure 4 Recall Outcome.

A common metric used to assess a machine learning model's efficacy in task classification is the F1 score. It gives a fair assessment of a model's correctness by combining precision and recalls into a single rating. When working with datasets that are unbalanced and have unequal distribution of classes, the F1 score is especially helpful. Figure 5 displays the F1-Score level for both the suggested and existing methods. DNN and KNN only succeed in obtaining F1-Score values of 89.9% and 92.5%, respectively, in comparison to the suggested approach SSO-SNN, which obtains 97.6% F1-Score. When compared to existing approaches, the F1-Score of the SSO-SNN methods is greater.

$$F - Score = 2 \times \frac{Precision * recall}{Precision + recall} \qquad (12)$$

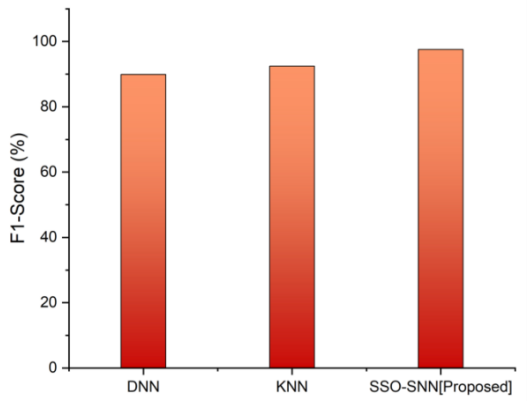


Figure 5 F1-Score Outcome.

5. Conclusion

Evaluation of personality is important in psychology, HR, and marketing. However, subjective self-report questionnaires can bias and limit existing approaches. Thus, personality assessment must be objective and dependable. Using signature biometric data, we proposed a Salp swarm-optimized Siamese neural network (SSO-SNN) in this research to forecast personality attributes. In order to gauge the effectiveness of the suggested strategy, we first collected information from the CIU handwritten database. A median filter was used to pre-process the gathered signature images in order to reduce noise. The performance measures of our suggested technique were accuracy (98.5%), Precision (97.6%), Recall (98.1%), and F-measure (97.6%). The proposed method for identifying personality has been shown to be more effective through experimental testing than the one currently in use. Future research will examine how machine learning approaches might be used to characterize online handwritten signatures and other biometrics.

Ethical considerations

Not applicable.

Declaration of interest:

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Aarva A, Deringer VL, Sainio S, Laurila T, Caro MA (2019) Understanding X-ray spectroscopy of carbonaceous materials by combining experiments, density functional theory, and machine learning. Part I: Fingerprint spectra. Chemistry of Materials 31:9243-9255.

Adeyemi IR, Abd Razak S, Salleh M, Venter HS (2017) Leveraging human thinking style for user attribution in the digital forensic process.

Agduk S, Aydemir E (2022) Classification of Handwritten Text Signatures by Person and Gender: A Comparative Study of Transfer Learning Methods. Acta Informatica Pragensia 2022:324-347.

Ayzeren YB, Erbilek M, Çelebi E (2019) Emotional state prediction from online handwriting and signature biometrics. IEEE Access 7:164759-164774.

Feldmann C, Philipps M, Bajorath J (2021) Explainable machine learning predictions of dual-target compounds reveal characteristic structural features. Scientific Reports 11:21594.

Feng C, Yuan J, Geng H, Gu R, Zhou H, Wu X, Luo Y (2018) Individualized prediction of trait narcissism from whole-brain resting-state functional connectivity. Human brain mapping 39:3701-3712.

Jiang Y, Jin C, Yu H, Wu J, Chen C, Yuan Q, Huang W, Hu Y, Xu Y, Zhou Z, Fisher GA (2021) Development and validation of a deep learning CT signature to

predict survival and chemotherapy benefit in gastric cancer: a multicenter, retrospective study. Annals of surgery 274:1153-e1161.

Kumar S, Gornale SS, Siddalingappa R, Mane A (2022) Gender Classification Based on Online Signature Features Using Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering 10:260-268.

Landau O, Cohen A, Gordon S, Nissim N (2020) Mind your privacy: Privacy leakage through BCI applications using machine learning methods. Knowledge-Based Systems 198:105932.

Portugal LC, Schrouff J, Stiffler R, Bertocci M, Bebko G, Chase H, Lockovitch J, Aslam H, Graur S, Greenberg T, Pereira M (2019) Predicting anxiety from whole brain activity patterns to emotional faces in young adults: a machine learning approach. NeuroImage: Clinical 23:101813.

Saleema A, Thampi SM (2020) User recognition using cognitive psychology based behavior modeling in online social networks. In Advances in Signal Processing and Intelligent Recognition Systems: 5th International Symposium, SIRS 2019, Trivandrum, India, December 18–21, 2019, Revised Selected Papers 5, pp. 130-149. Springer Singapore.

Saraswal A, SAXENA UR, (2022) Analysis and Recognition of Handwriting Patterns for Personality Trait Prediction Using Unsupervised Machine Learning Approach. In Advanced Production and Industrial Engineering, pp. 442-447. IOS Press.

Shin J, Maniruzzaman M, Uchida Y, Hasan MAM, Megumi A, Suzuki A, Yasumura A (2022) Important features selection and classification of adults and children from handwriting using machine learning methods. Applied Sciences 12:5256.

Siraj S, Singuluri PK, Kumar MR (2022) AI Intelligence-based Gender Classification using Biometric-Digital Signature Feature Extraction Methods. International Journal of Intelligent Systems and Applications in Engineering 10:262-268.

Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, Folio LR, summers RM, Rubin DL, Lungren MP (2020) Preparing medical imaging data for machine learning. Radiology 295:4-15.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0116

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Analysis of psychological biometric data for gender identification using fusion features and deep learning

Ajay Chakravarty[a] | **Murugan R.**[b] | **Manish Srivastava**[c]

[a]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College of Computing Science And Information Technology.
[b]Jain (deemed to be)University, Bangalore, India, Associate Professor, Department of Computer Science and Information Technology.
[c]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Electrical Engineering.

Abstract A person's gender must be correctly identified for many applications, such as tailored services, social sciences, and human-computer interaction, to function properly. Traditional gender identification techniques mostly depend on physical traits, but new developments in psychological biometrics provide interesting alternatives. Over the last several years, several access control systems have included biometric security technologies to increase security. The handwritten signature is the psychological biometric characteristic that is most often used to validate daily documents like letters, contracts, wills, MOUs, etc. This study proposes an innovative Deep Learning (DL) approach to identify a person's gender from an image of their handwritten signature. The fusion of statistical and textural information taken from the trademark photos serves as the foundation for the proposed work. The texture is represented by the Pyramid Histogram of Oriented Gradients (PHOG) features. A novel sequence labelling multidimensional recurrent neural network (SLMRNN) is employed to classify the writer's gender. Extensive experiments are carried out on the gathered dataset to assess the performance of the suggested technique. The efficacy of the fusion features and DL models for gender recognition is evaluated using a variety of measures, including accuracy, precision, recall, and F1 score. To evaluate the suggested approach against existing procedures and emphasize its advantages, if any, comparative assessments are also carried out. The findings show that, in comparison to conventional approaches, the combination of behavioral biometric variables with cutting-edge DL algorithms greatly enhances gender identification accuracy. The suggested approach is anticipated to be beneficial in the development of effective computer vision tools for forensic analysis and authentication of papers with handwritten signatures.

Keywords: PBD, gender Identification, PHOG, SLMRNN

1. Introduction

Biometrics is a scientific method that uses a person's unique physiological and behavioral characteristics to identify them. Soft biometric traits could be seen in humans in addition, such as gender, age, ethnicity, weight, height, gestures, stride, accent, ear shape, leg and arm length, skin color, hair color, etcetera. Biometrics has evolved into a standard technique for many different kinds of system identification. When it involves defining biometrics, the International Organization for Standardization (IOS) uses the term biological process for recognizing and analyzing an individual based on physiological and behavioral characteristics (Patil 2019). The process of using a person's mental features and characteristics as an aspect of recognition or verification is referred to as psychological biometrics and the word psychological biometrics is used to denote the practice. It comprises gathering details on a broad spectrum of psychological qualities and utilizing that data to develop unique profiles of individuals. The goal of the field of psychological biometrics is to identify and analyses the unique characteristics that are associated with a certain individual via the use of data mining tools. Statistical or data mining methods are often used, depending on the situation, in order to evaluate the data and develop models that are able to successfully classify individuals or authenticate them based on psychological biometric information (Shaheed 2021). Psychological biometrics include a wide variety of measures of thought, feeling, and action. Cognitive processes include things including memory, attention, and problem-solving; affective processes such as emotions and facial expressions; and behavioral trends comprising dynamic typing and analysis of the gait. Questionnaires, interviews, cognitive tests, physiological measurements, and direct observation of human conduct are only some of the research methodologies that have been used (Ayzeren 2019). Analysis of psychological biometric data for gender identification is the study of a person's psychological traits and qualities to infer or estimate their gender identity. Psychological biometric data analysis may provide

information on societal trends, but it's not a fool proof method for learning someone's true sexual orientation. To address the issue of gender identity with care and respect for self-identification, it is essential to have a comprehension of the complexity of gender. It's difficult to find consensus on the use of psychological biometrics in the research of gender identity. Although the reality is that certain psychological features may be more frequent in one gender than the other, gender identity is a highly individual and particular experience that cannot be distilled down to a single physiological sign (Sommerfeldt 2019). Psychological and biological data alone are insufficient to establish the gender of an individual. However, gender identity is much more than simply outward appearance. Addressing the issue of gender identity with compassion, respect, and an awareness of the range of gender experiences and manifestations is essential (Shopon 2021). In this study, we proposed machine learning framework for analyzing the gender categorization was gathered from the appearance of their handwritten signatures. This approach would include textural and statistical information.

2. Related Work

The research (Guarino 2022) introduced a revolutionary machine learning-based method for determining a user's gender using just touch movement data collected via application programming interfaces. The goal is to determine which gestures, or combinations thereof, are most helpful in determining gender. The research (Svetlakov 2021) provided insight into how factors like sample size, gender, and age might influence the accuracy of biometric identification based on EEG data. Uniqueness, fraud resistance, portability, ease of use, scalability, the potential of accurate attendance monitoring, and so on are only a few of the benefits of biometric systems over more conventional authentication techniques. The paper (Dargan 2020) explained the biometric techniques and authentication methods together assist improve the security of applications when user involvement can be inferred; nonetheless, these systems will not replace existing authentication tools and technologies. Security, surveillance, forensic investigations, fraudulent technologies, identity access management, and access control are just some of the many fields that have benefited from the widespread use of biometric-based identification techniques and systems. The paper (Jain 2019) described the execution of the A Prescription for Healthy Living (APHL) diabetes education using a culinary medicine curriculum at a clinic-based food prescription programmer as well as its preliminary impact on biometric and diet-related behavioral and psychosocial outcomes. The study (Sharma 2021) focused on the use of smartphone-based soft biometric features. Extracted gender and age are examples of soft biometric information that is only incidental. The research (Guarino 2023) developed a web-based sexual predator monitoring system that makes use of SSB properties. The goal of social biometrics is to derive a user's unique behavioral pattern from their online social interactions and communications. The study (Pereira 2023) provided a comprehensive evaluation of data acquisition methodologies. To better comprehend the role that certain factors in the information capture technique of an ECG signal play in the biometric identification process. The research (Casanova 2021) examined touch dynamics and features retrieved from the periocular region connected to the blinks, pupils, and fixations to create an online user identification model. To identify and differentiate online users, a procedure known as web user behavioral recognition has been developed. The research (Cascone 2020) described the comprehensive investigation to show that changes in pupil size and dilation over time may be used to possibly categorize individuals by age and gender. This was accomplished by using 14 supervised classifiers in a dataset developed specifically for gaze analysis. The paper (Terhörst 2019) offered a method for improving the privacy of soft biometrics by omitting some of the less essential details from a biometric template. A relevance measure for each variable is derived during decision tree ensemble training, and this measure is then utilized to progressively remove variables from the prediction of sensitive characteristics. The paper (Rodgers 2021) analyzed how AI-based facial and musical biometrics affects consumers' mental and emotional states, and how that, in turn, affects their reactions in terms of the value they provide for the business. The article (Fenu 2021) investigated the connection between the aforementioned balancing settings and current group fairness indicators as they pertain to speaker recognition. To do this study, they have operationalized many concepts of fairness and are tracking them while using different data-balancing methods. The paper (Brown 2021) determined convolutional systems that depend on basic passwords, personal identification numbers, or tokens have been offered as an alternative to biometric identification techniques that rely on physical and behavioral traits. Researchers have recommended using their abilities to gather physiological data to facilitate automated user identification in the study (Piciucco 2021). The very nature of wearable technology makes it possible to monitor for unauthorized use of the device or to act as a user interface for biometric recognition systems that manage physical or digital access to restricted areas or services.

3. Proposed Methodology

The suggested approach is based on the combination of statistical and textural data extracted from the distinctive images. The investigation discovered no gender identification in the publicly available standard photo databases of handwritten signatures, then pre-processed using and feature extraction for PHOG. We propose an innovative SLMRNN technique for classification. The suggested block diagram is shown in Figure 1.

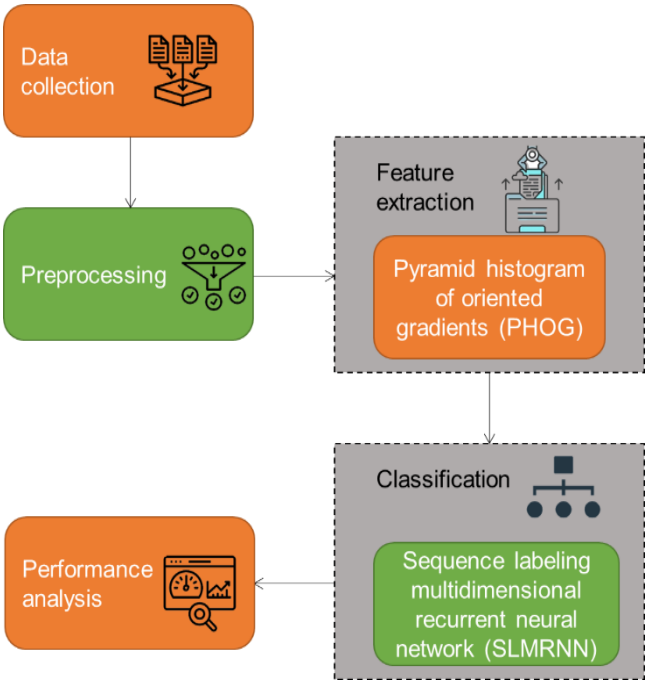


Figure 1 Block diagram of proposed.

3.1. Data collection

According to the research conducted, no gender identification can be found in the publicly accessible standard picture datasets of handwritten signatures. As a consequence of this inspiration, we build our dataset to test out the recommended strategy. The following text elaborates on the nature of this one's very own offline database of handwritten signatures.

- Multiple age groups are represented in the signature samples. Multilingual Kannada, Hindi, Marathi, and English scripts are included in these examples.
- Each person has signed ten copies of a white A4 page using a black or blue ballpoint pen.
- The EPSON DS1630 colour scanner was used to scan the sample signature documents to eliminate any measurable geometrical differences.
- The 479 individuals, including 250 men volunteers and 229 women volunteers, provided the database's 4,790 signatures.

The goal of collecting signature samples was explained to everyone who could read and write in English and other languages.

3.2. Pre-processing

Pre-processing is mostly used for improving the essential features of handwritten signatures and improving data by removing undesirable distortion or noise. There are numerous processes involved in pre-processing the gender classification of authors based on their handwritten signatures to extract useful information and categorize the signatures correctly. Data collection, the first phase, is compiling a dataset of handwritten signatures from multiple authors. Pre-processing procedures are used to improve information once the signatures are in digital format.

3.3. Feature Extraction by using Pyramid histograms of oriented gradients (PHOG)

The feature vector that HOG has recovered corresponds to a statistical histogram, which often disregards the intrinsic structural information contained in the finger vein picture. PHOG, uses HOG to account for the spatial gradient pattern of picture intensity. By sampling images at various resolutions, as information conveyed in photographs varies with the resolution, the PHOG operator can provide a more accurate description of vein properties. Having a significant impact in reducing noise and rotation, PHOG emphasizes the benefit of using spatial information from images to characterize features. To create a collection of photos with various resolutions, PHOG utilizes up or down-sampling. The pyramid is arranged with the picture with the highest resolution and size at the base and the image with the lowest resolution and smallest size on top. The lower the location of the picture on the pyramid and the greater the amount of information that can be expressed, the clearer the picture, and the greater its value. In some ways, the top-level image's low resolution helps to better portray the general structural aspects of the original image while reducing the impact of noise. To gain more precise characteristics information, PHOG integrates the local and global properties of photographs. The ultimate representation of the PHOG

feature of a picture is a vector of $\sum_{i=L} 4^i$measurements, where K is the number of intermissions that make up the change in strategy. L is the total amount of stages within a hierarchy.Assume L=4 and K = 8 thus, the measurement of PHOG is (4^0 + 4^1 + 4^2 +4^3) × 8 = 680.The HOG feature histogram is normalized using the L2-norm, which may be written as follows:

$$L_2 - norm; e = \frac{x}{\sqrt{\|u\|_2^2 + s^2}} \qquad (1)$$

The HOG property is represented by the ε, and the tiny constant prevents the denominator from equaling zero. By summing the histogram vectors from each successive layer in pyramidal order, they get the final PHOG feature histogram.

3.4. Sequence Labelling Multidimensional Recurrent Neural Network (SLMRNN)

SLMRNN is a kind of neural network design used for sequence labelling tasks including speech recognition, NLP, and gesture recognition. By handling multidimensional input data, it expands the capabilities of the Conventional Recurrent Neural Network (CRNN) model. Sequence labelling is a sort of machine learning activity that entails providing a label or category to each element that is included within a sequence of input data. This may be done manually or via the use of an automated system. The input sequence may be of variable lengths and may represent many kinds of data, such as words in a phrase, phonemes in voice, or tags in a part-of-speech tagging position.

The formula for determining the probability conditional of a string of labels z in a linear-chain Conditional Random Field (CRF), given a string of observations$v$, can be found here.

$$O_{dqe}(z|v,\lambda) = \frac{1}{\psi(V)} \prod_{s=1}^{S} \exp\{\sum_{l=1}^{L} \lambda_{ll} e_{ll}(z_{s-1}, z_s, v, s)\} \qquad (2)$$

Where, $e_l(z_{s-1}, z_s, v, s)$ is a feature function (typically binary-valued, but it may also be real-valued), λ_l is a learning weight connected to features e_l, and $\psi(V)$ is a normalization variable that sets the average of all labelling sequence probabilities equal to one. Any element of the state transitions may be captured by the feature functions z_{s-1}, $z_s$and of the entirety of input process$V$, which is useful for understanding the connection among labelling and the properties of the entire input sequence V at a certain time$s$. Following the model provided in Equ. 2. The following is the input sequence V is the probable labelling sequence: $y *= arg\ max_z\ O_{dqe}\ (z|v, \lambda)$This is effectively decided by dynamic programming using the Viterbi technique.

Recurrent neural networks with multiple dimensions (MDRNNs) are another application of MRNNs to high dimensional sequence learning. Recurrent connections are used by this network to learn correlations in the data for each dimension. A typical form of neural network technology used in the recognition of words and natural language processing is the MDRNN. An example of a neural network with loops that allows data to be saved inside the network is a MDRNN in Figure 2. The most important feature of MDRNNs is their capacity to preserve a hidden state that stores data about previous inputs. This hidden state acts as a repository for the network's prior calculations and is updated at each time step. The next input in the sequence is subsequently processed using the updated hidden state. MDRNNs can represent dependencies and make use of temporal information in the data thanks to their recurrent nature. The MDRNN, the focus of this research, has grown ready for definition. A nonlinear dynamical system that maps sequences to sequences is the typical MDRNN. The MDRNN uses the following algorithm 1 to produce a series of hidden states g_1^s and a sequence of outputsy_1^s:

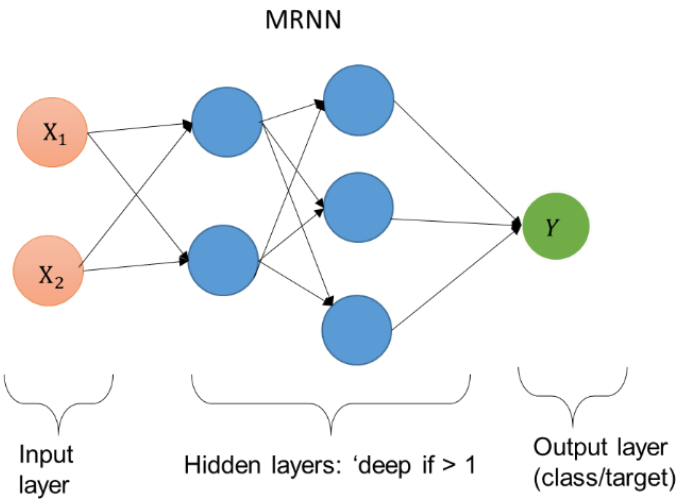


Figure 2 MDRNN with loops.

Algorithm 1: MDRNN Algorithm

1. For s from 1 to S do
2. $w_s \leftarrow U_{gu} v_s + U_{gg} g_{s-1} + a_g$
3. $g_s \leftarrow f(w_s)$
4. $p_s \leftarrow U_{pg} g_s + a_p$
5. $y_s \leftarrow h(p_s)$
6. $end\ for$

Where MDRNN's hidden and output deviations, $f(.)$and$h(.)$, and g_0, a matrix of variables containing the initial hidden state, respectively, are all mentioned. The MDRNN's loss is typically the total of losses each time step:

$$K(y,z) = \sum_{s=1}^{S} K(y_s; zs) \qquad (3)$$

The backpropagation through time approach makes it simple to calculate the MDRNNs' derivatives.

4. Performance analysis

4.1. Results

In this part, the suggested system's effectiveness is evaluated. The performance indicators used for assessment are accuracy, precision, recall, and f1-measure. K Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest (RF) are the existing methods used for comparison.

4.1.1. Accuracy

A difference between the result and the true number is caused by inadequate precision. The percentage of actual outcomes reveals how balanced the data is overall. Accuracy is assessed using an equation (4).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (4)$$

Figure.3 shows the comparable values for the accuracy measures. When compared to existing methods like KNN, which has an accuracy rate of 81%, SVM, which has an accuracy rate of 78%, and RF, which has an accuracy rate of 93%, the recommended method's SLMRNN value is 95%. The suggested SLMRNN performs with higher accuracy than other methods. Table 1 displays the proposed method's accuracy.

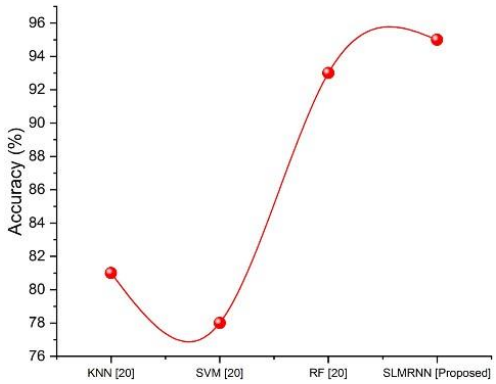


Figure 3 Accuracy comparisons between the suggested and current approaches.

Table 1 Comparison of Accuracy.

Methods	Accuracy (%)
KNN (Kumar S 2022)	81
SVM (Kumar S 2022)	78
RF (Kumar S 2022)	93
SLMRNN [Proposed]	95

4.1.2. Precision

The most crucial standard for accuracy is precision, it is clearly defined as the percentage of properly categorized cases to all instances of predictively positive data. Equation (5) is used to compute the precision.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (5)$$

Comparable values for the precision measures are shown in Figure. 4. This proves the suggested strategy may provide performance results that are superior to those obtained by the current study methods. The precision of the proposed approach is 96%, which performs better than existing outcomes. Include KNN, SVM, and RF precision rates are 88%, 72%, and 92%. Table 2 shows the precision of the suggested method contrasted with the existing methods.

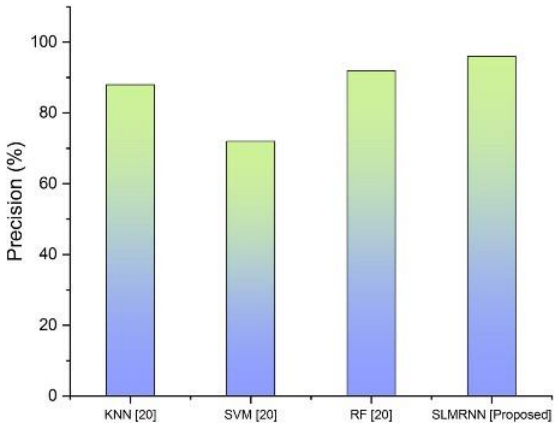


Figure 4 Precision comparisons between the suggested and current approaches.

Table 2 Comparison of Precision.

Methods	Precision (%)
KNN [20]	88
SVM [20]	72
RF [20]	92
SLMRNN [Proposed]	96

4.1.3. Recall

The potential of a model to identify each important sample within a data collection is known as recall. The percentage of TPs divided by the sum of True Positive and False Negative is how it is statistically defined. The recall is calculated using an equation (6).

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (6)$$

Figure 5 shows the comparative data for the recall metrics. Recall rates for KNN are 71%, SVM 89%, RF 87%, and SLMRNN 92%. The proposed method performed better than the current results with a recall of 92%. In Table 3, the recall of the suggested method is contrasted with the existing methods.

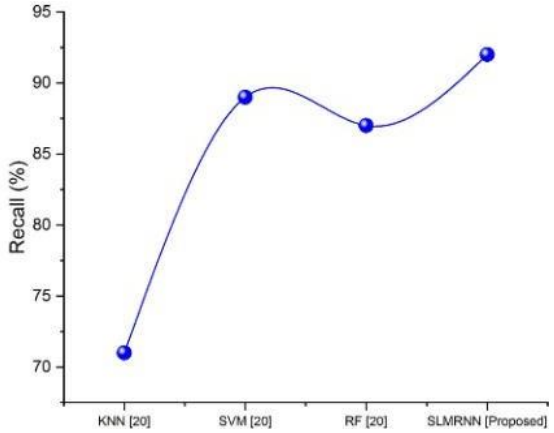


Figure 5 Recall comparisons between the suggested and current approaches.

Table 3 Comparison of Recall.

Methods	*Recall (%)*
KNN [20]	71
SVM [20]	89
RF [20]	87
SLMRNN [Proposed]	92

4.1.4. F1-Score

The F1 measure is often used while assessing information. It is possible to alter the F1 measure so that accuracy is prioritized above recall, or vice versa. The recommended technique has a higher level of F1-measure when measured against the currently used methods. In Figure 6 the F1-measure of the suggested method is contrasted with the traditional methods.

$$F1 = \frac{2 * (precision * recall)}{(precision + recall)} \quad (7)$$

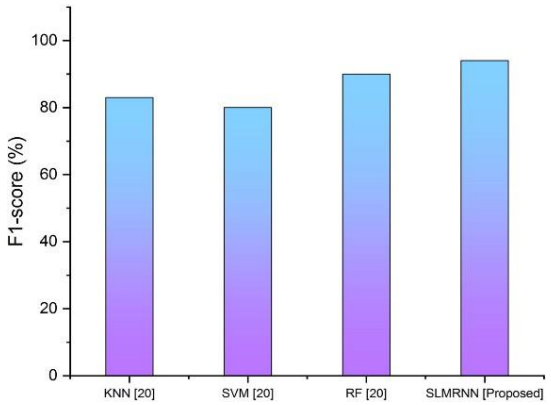


Figure 6 F1-Score comparisons between the suggested and current approaches.

When compared to existing methods like KNN, which has an F1-measure of 83%, SVM, which has an F1-measure of 80%, and RF, which has an F1-measure of 90%, the recommended method's SLMRNN value is 94%. The suggested SLMRNN performs with higher accuracy than other methods. Table 4 displays the proposed method's F1-Score.

Table 4 Comparison of F1-Score.

Methods	**F1-score (%)**
KNN [20]	83
SVM [20]	80
RF [20]	90
SLMRNN [Proposed]	94

5. Conclusion

In this research, we offer a technique that uses feature fusion and method to accurately determine a person's gender from their handwritten signature. Although researchers have contributed to the categorization of gender using various methodologies and gained approximately promising results using either custom or typical datasets, there is still significant potential for increasing an accurate algorithm employing effective biased features, as shown by the previously mentioned related work. The suggested technique involves testing the feature-level fusion of PHOG features using an internal dataset of 4,790 high-quality signatures. As a result, we introduced the SLMRNN for the recognition of spoken emotion. Performance metrics like accuracy, recall, precision, and F1-Score, are evaluated and compared with existing technologies like KNN, SVM, and RF. In the future, DL will be used to investigate categorization difficulty in establishing a person's gender from observable characteristics such as handwriting or a biometric.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Ayzeren YB, Erbilek M, Çelebi E (2019) Emotional state prediction from online handwriting and signature biometrics. IEEE Access 7:164759-164774.

Brown R, Bendiab G, Shiaeles S, Ghita B (2021) A novel multimodal biometric authentication system using machine learning and blockchain. In Selected Papers from the 12th International Networking Conference: INC 2020 12:31-46. Springer International Publishing.

Casanova A, Cascone L, Castiglione A, Meng W, Pero C (2021) User recognition based on periocular biometrics and touch dynamics. Pattern Recognition Letters 148:114-120.

Cascone L, Medaglia C, Nappi M, Narducci F (2020) Pupil size as a soft biometrics for age and gender classification. Pattern Recognition Letters 140:238-244.

Dargan S, Kumar M (2020) A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. Expert Systems with Applications 143113114.

Fenu G, Marras M, Medda G, Meloni G (2021) Fair voice biometrics: Impact of demographic imbalance on group fairness in speaker recognition. In Interspeech 1892-1896. International Speech Communication Association.

Guarino A, Lettieri N, Malandrino D, Zaccagnino R, Capo C (2022) Adam or Eve? Automatic users' gender classification via gestures analysis on touch devices. Neural Computing and Applications 34:18473-18495.

Guarino A, Malandrino D, Zaccagnino R, Capo C, Lettieri N (2023). Touchscreen gestures as images. A transfer learning approach for soft biometric traits recognition. Expert Systems with Applications 219:119614.

Jain A, Kanhangad V (2019) Gender recognition in smartphones using touchscreen gestures. Pattern Recognition Letters 125:604-611.

Kumar S, Gornale SS, Siddalingappa R, Mane A (2022) Gender Classification Based on Online Signature Features using Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering 10:260-268.

Patil A, Kruthi R, Gornale S (2019). Analysis of multi-modal biometrics system for gender classification using face, iris and fingerprint images. International Journal of Image, Graphics and Signal Processing 11:34.

Pereira TM, Conceição RC, Sencadas V, Sebastião R (2023) Biometric recognition: A systematic review on electrocardiogram data acquisition methods. Sensors 23:1507.

Piciucco E, Di Lascio E, Maiorana E, Santini S, Campisi P (2021) Biometric recognition using wearable devices in real-life settings. Pattern Recognition Letters 146:260-266.

Rodgers W, Yeung F, Odindo C, Degbey WY (2021) Artificial intelligence-driven music biometrics influencing customers' retail buying behavior. Journal of Business Research 126:401-414.

Shaheed K, Mao A, Qureshi I, Kumar M, Abbas Q, Ullah I, Zhang X (2021) A systematic review on physiological-based biometric recognition systems: current and future trends. Archives of Computational Methods in Engineering 1-44.

Sharma SV, McWhorter JW, Chow J, Danho MP, Weston SR, Chavez F, Moore LS, Almohamad M, Gonzalez J, Liew E, LaRue DM (2021) Impact of a virtual culinary medicine curriculum on biometric outcomes, dietary habits, and related psychosocial factors among patients with diabetes participating in a food prescription program. Nutrients 13:4492.

Shopon M, Tumpa SN, Bhatia Y, Kumar KP, Gavrilova ML (2021) Biometric systems de-identification: Current advancements and future directions. Journal of Cybersecurity and Privacy 1:470-495.

Sommerfeldt SL, Schaefer SM, Brauer M, Ryff CD, Davidson RJ (2019) Individual differences in the association between subjective stress and heart rate are related to psychological and physical well-being. Psychological science 30:1016-1029.

Svetlakov M, Hodashinsky I, Slezkin A (2021) Gender, age and number of participants effects on identification ability of EEG-based shallow classifiers. In 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), pp. 0350-0353. IEEE.

Terhörst P, Damer N, Kirchbuchner F, Kuijper A (2019) Suppressing gender and age in face templates using incremental variable elimination. In 2019 International Conference on Biometrics (ICB), pp. 1-8. IEEE.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0117

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

The development of a smarter mode of transportation using ai to reduce travel and waiting times

Priyank Singhal[a] | **Karthikeyan M. P.**[b] | **Pramod Kumar Faujdar**[c]

[a]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College of Computing Science And Information Technology.
[b]Jain (deemed to be) University, Bangalore, India, Assistant Professor, Department of Computer Science and Information Technology.
[c]Vivekananda Global University, Jaipur, India, Associate Professor, Department of Mechanical Engineerin.

Abstract Smarter mode of transportation (SMT) rely heavily on real-time traffic management. Safety on the roads is only one of many benefits brought forth by advancements in dynamic traffic management systems for congested metropolitan areas. In this study, we present the design and implementation of a Bayesian Belief Networks (BBN) traffic control system that is both flexible and reliable. The use of knowledge-based systems as a framework for making decisions in real-time has gained widespread acceptance. Since traditional dynamic controllers relied on sensors with their own set of drawbacks, we may employ vision sensors (such as cameras) to get around these problems. Computing based on images and videos is very useful for gauging traffic volumes. The present traffic management system at the road junction was found wanting, thus a new system was developed and put into place to alleviate the congestion. Lab VIEW and MATLAB are used to measure how well the suggested framework works. Extensive simulations utilizing the suggested method show that it reduces waiting time and speeds up movement on average compared to controllers employing traditional sensors.

Keywords: traffic density measurement, controller, bayesian belief networks, traffic lights, vision computing

1. Introduction

The creation of a more advanced form of transportation will include incorporating a wide range of technological breakthroughs and inventions to increase productivity, decrease congestion, and improve the experience of traveling as a whole (Zhang 2020). It is intriguing to think about the possibility of designing a more advanced form of transportation that uses artificial intelligence (AI) to shorten both trip and wait times. Artificial intelligence can completely transform transportation networks by streamlining routes, improving traffic flow, and increasing overall efficiency (Lilhore 2022). The movement toward environmentally friendly and electric cars is essential to intelligent transportation systems. Electric vehicles, such as automobiles, buses, and bicycles, have the potential to drastically cut carbon emissions and the reliance on fossil fuels, thus contributing to a cleaner and greener transportation system (Soori 2023). Cars that are connected and the transportation infrastructure may interact with one another as well as with the sophisticated sensors, communication technologies, and artificial intelligence that are equipped with these connected cars. Sharing data in real-time, working together while driving, and improving overall traffic management are all made possible due to this. There is a possibility that autonomous cars, which depend on artificial intelligence (AI) and sensors to function without any interaction from a human driver, would improve safety, efficiency, and the flow of traffic (Leung 2019). Using artificial intelligence (AI) and big data analytics, transportation agencies and service providers can collect and analyze massive volumes of data from a variety of sources. These data may determine trends, forecast travel demand, optimize routes, and make choices based on the collected data to increase operational efficiency and decrease travel times (Cao 2019). The development of intelligent transportation also emphasizes preserving the environment and being environmentally responsible. This includes the promotion of walking and cycling infrastructure, the encouragement of the use of public transit, and the integration of renewable energy sources into various modes of transportation (Paiva 2021). Artificial intelligence (AI) can potentially play a key role in lowering travel and waiting times through the optimization of transportation networks. AI systems may examine historical traffic data, current weather conditions, and information received in real-time from sensors to make predictions about traffic patterns and congestion hotspots. This information may be utilized to improve traffic flow, propose alternative routes, and maximize route optimization, which will ultimately result in a reduction in travel times (Offiaeli 2021). Algorithms that are driven by AI may take into account real-time traffic data, the state of the roads, and other variables to determine

which routes are the most time- and fuel-efficient for particular cars. AI has the potential to shorten travel times and steer clear of crowded places because of its ability to dynamically update routes depending on current circumstances (Abduljabbar 2021). AI may be used to improve the efficiency of transportation systems sensitive to changes in demand, such as ridesharing and micro transit. AI can reduce waiting times and offer more effective transportation alternatives since it can analyze passenger demand and dynamically change routes and vehicle allocations (Witlox 2022). Users can obtain real-time travel information from systems that AI drives. This information may include public transportation timetables, traffic updates, and anticipated trip times. AI can assist passengers in making well-informed judgments and selecting the modes of transportation or routes that will save them the most time and effort by providing them with information that is both accurate and up-to-date (Fridgen 2021). Artificial intelligence can monitor traffic patterns and locate places of congestion by analyzing data collected in real time from various sources, such as traffic cameras, GPS devices, and sensors. These data may be used to vary the timing of traffic lights timings dynamically, redirecting cars, and recommend other routes to reduce travel times (Liu 2023). AI-powered routing algorithms can take into account a variety of criteria, including current traffic conditions, road closures, and user preferences, to choose the path that will be the most time and energy efficient as well as the fastest. This may assist in reducing travel times and eliminating the need for unnecessary delays (Parveen 2022). Travel applications or platforms powered by AI can give passengers customized help. They can provide real-time updates on traffic conditions, provide recommendations for the most appropriate mode of transportation for the user depending on the user's preferences and the limits of the situation, and even help the user book tickets, lodgings, and other travel-related services (Fang 2019). AI can analyze demand patterns and make dynamic price adjustments to maximize resource allocation. For instance, ridesharing services may employ surge pricing during peak hours to motivate more drivers to be available, which will lower the time customers have to wait (Billhardt 2019).

2. Related works

This research investigates the feasibility of using artificial intelligence (AI) to improve the efficiency of public transportation systems and reduce traffic congestion. The results, which are the result of an in-depth review of previously published research, case studies, and the views of industry professionals, highlight numerous important areas in which AI might help the improvement of public transport systems. The constant monitoring of traffic patterns is made possible with the use of AI in real-time monitoring, which permits the correct forecast of congestion in the present moment (Kozlov 2022). The study produces insights regarding the manner in which AI may help build smarter cities and how it can do so in a variety of different ways. The methodological approach that will be used will be the research. The findings are organized according to the primary aspects of smart city improvement, including economics, society, atmosphere, and democracy (Yigitcanlar 2020). The study examines the many types and definitions of shared transportation services that are widespread in North America, and it summarizes the impact studies that have already been conducted. In addition, we investigate the confluence of shared mobility, electrification, and automation, including the possible consequences of shared automated vehicle (SAV) systems. Even though the effects of SAVs are not yet fully understood, several industry experts and university researchers anticipate increased productivity, decreased costs, and reduced emissions of greenhouse gases (Shaheen 2017). The exploration proposes consolidating four advancements: huge information, profound learning, in-memory registering, and design handling units (GPUs) to provide a complete way to deal with the speedy and versatile forecast of metro framework qualities. The primary aim of the research was to develop methods to expedite the creation of such forecasts (Aqib 2019). In this respect, although the use of artificial intelligence (AI) methods, including the fields of machine and deep learning, has attracted much interest in smart cities, less emphasis has been placed on the utilization of algorithmic optimization approaches. To be of assistance with this, this research gives an overview of strategies for optimization and software from the point of view of a smart city that is allowed by the Internet of Things (IoT) (Syed 2022). In general, obtaining all of the necessary data at each place is a fairly costly endeavor. In this research, an approach to deep learning is suggested for estimating the waiting time levels at transportation stations using proxy data and limited historical waiting time data at select stations. This approach is based on the idea that one can determine the waiting time levels using proxy data (Chu 2019). The study is an association of numerous problems that affect the transport sector and are categorized as autonomous vehicles. In regard to intelligent transport systems, some of the components that are being explored are connected to traffic management, public transport, safety management, manufacturing, and logistics. These are all areas where AI advantages are put to use (Iyer 2021).

3. Methodology

The closed-loop technique management system is shown in Figure 1 receives data about the median pixel-to-pixel matches of the standard mosaicked picture and the currently active mosaicked output from the comparator or mistake detectors. Infrared (IR), ultraviolet (UV), and inductive loop (IL) sensors have been replaced with numerous vision sensors in this modern take on control. BBNs, which employ membership processes and rule-based BBN sets controlled by problem detector data to make more informed decisions, are additionally employed to modify the onset of green and red lights in

response to the density of road measurements. Vision equipment has been utilized to record pictures of vehicles for use in the feedback process. Multiple vision sensors are used so that precise distances between cars may be determined to provide reliable estimates of traffic volumes. After the photos have been taken, they go through a series of processes, including saving, grayscale conversion, transformation, wrapping, compositing, and eventually the production of the mosaicked picture.

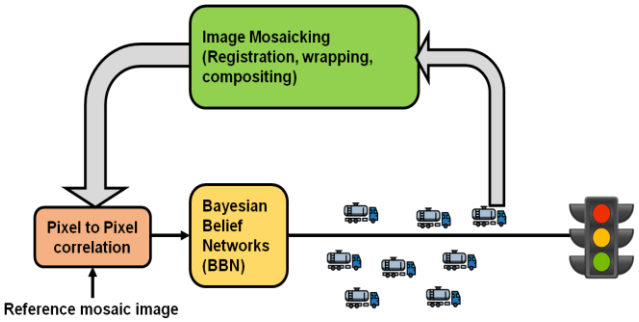


Figure 1 The suggested methodology is shown schematically.

A schematic representation of the suggested framework on a '+' shaped road is shown in Figure 2. Each side of the road now has four 12.0 megapixel webcams. It is expected that there are two lanes in each direction from west to east and east to west since this is where the majority of traffic flows. This is in contrast to the north-to-south and south-to-north directions, which only have one lane each owing to the relatively small volume of traffic in these directions. The suggested algorithm is designed to maintain the previously established schedule under typical traffic conditions. The algorithm determines the typical range for measuring traffic (which may be gleaned from historical data) and adjusts the timing of red and green lights appropriately. Because average density, average velocity, and average speed are normally the characteristics of a macroscopic traffic flow model, this proposed framework only accounts for the former.

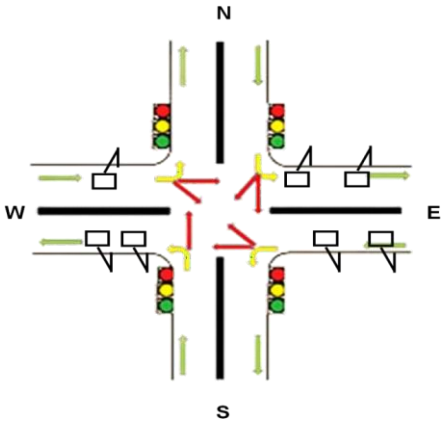


Figure 2 Symbolic representation of a '+' intersection.

3.1. Bayesian belief networks (BBNs)

The use of machine learning (ML) and deep learning (DL) algorithms to address practical issues has seen a meteoric rise in recent years. These algorithms have been used in several fields, such as medicine, electricity, picture identification, and indoor monitoring of objects. The suggested method relies on Bayes' rule, which calculates the likelihood of an occurrence given prior knowledge of circumstances known to be relevant to that event. Bayes' rule, which is based on conditional probability, is the bedrock of Bayesian inference. The statistical form of Bayes' rule is as follows:

$$b(Y|X) = \frac{b(X|Y)b(Y)}{b(X)} \qquad (1)$$

The first term, commonly known as the prior probability, indicates an individual's starting point of belief before any data concerning event Y is evaluated, whereas the second term may be thought of as a normalizing constant. If event Y has previously happened, then event X has a certain probability, denoted as $b(Y|X)$. Since it is calculated after considering all available data, it is also referred to as the posterior probability of occurrence Y. The likelihood, or $b(Y|X)$, measures how likely it is that event Y will occur given that event X has already taken place. The structure of a BBN and the probability distribution functions of its variables, known as the node parameters, characterize it. Given the nature of our investigation,

we restrict our attention to BBNs made up of discrete random variables. Bayesian networks (BBNs) are a crucial component of Bayesian inference; they are directed acyclic graphical models in which nodes represent random variables and arcs reflect causal relationships between the nodes. One-way links between BBN nodes suggest a tree-like or familial organization. Nodes' familial relationships are shown in Figure 3, where W and Y, for example, represent parents and offspring, respectively. A node's parent may have some effect on its offspring, but not vice versa. A node's ancestors are all the nodes in the hierarchy above it, and a node's descendants are all the nodes in the hierarchy below it. Finally, a node with no children is known as a sink node, while a node with no parents is known as a root node. Graphical representations of probabilistic models are useful because they allow for the concise and consistent depiction of joint probability functions. It is possible to obtain a more in-depth explanation of BBN architecture and its parts.

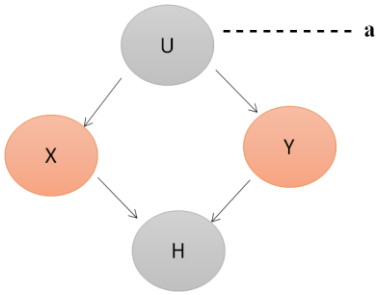


Figure 3 Bayesian belief network illustration.

The parameters of the model also play a significant role in Bayesian inference by determining the node-specific conditional probability distribution (CPD). If we assume that the random variables are discrete, then we can use a conditional probability table (CPT) to depict the conditional connection between the nodes. BBN-based models vary in computing cost depending on their structure, node count, and state count for each variable. Performing probabilistic inference using BBNs is an NP-hard task for many researchers. If we assume that the two nodes X and Y in Figure 3 represent dichotomous random variables, then the resultant CPT will have 2^2 states. It might be helpful to present some high-level ideas related to BBNs before continuing. The local Markov property states that a variable is conditionally independent of other variables while considering its immediate surroundings. Generalizing the local Markov property to BBNs is as follows:

$$Y_c \perp Y_{MT(c)} | Y_{BE(c)} \qquad (2)$$

If $Y_{MT(c)}$ is a nondescendant node, $Y_{BE(c)}$ is a parent node and Xv is a random variable represented by a BBN node. Take the straightforward BBN shown in Figure 3 as an example; in this case, X is conditionally independent of the nondescendant (U/X), which results in:

$$b(Y|U,X) = b(Y|U) \qquad (3)$$

The following equation provides a chain rule for decomposing a joint distribution of variables in a BBN.

$$b(Y_1, \ldots . Y_m) = b(Y_n|Y_1, \ldots . Y_{m-1}) \times b(Y_{m-1}|Y_1, \ldots . Y_{m-2}) \times (Y_2|Y_1)B(Y_1) \quad (4)$$

As a further step, we may use equation 4 to obtain a generic version of the chain rule for BBN.

$$b(Y_1, \ldots . Y_m) = \prod_{j=1}^{M} b(Y_j|BE(b(Y_j)) \qquad (5)$$

3.1.1. Elimination of Variables Algorithm

It may be challenging to put BBNs into reality since real-world problems often include a large number of random variables, each of which might take on a large number of possible values. Inferring using BBNs is easy if you utilize the whole joint distribution and add up all latent variables. This is because the whole joint probability table for n binary variables will have 2^n entries, making the process tedious for large BBNs. To lessen the computing load during inference, a simple yet effective approach known as variable elimination (VA) may be implemented. We generalize below a scenario in which we need to determine a subset of the questioned variables Y using the evidence E and the latent variables X. The joint probability distribution of Y and E divided by the marginal probability distribution of E equals the conditional probability of Y giving evidence A.

$$b(Y|A = a) = \frac{b(Y, A=a)}{b(A=a)} \qquad (6)$$

To obtain the numerator of equation (6), marginalization over all latent variables $X_1,\ldots,X_n$ is necessary.

$$B(Y = y_j, A = a) = \sum_{X_1} \dots \sum_{X_m} b(X_1, \dots, X_m, Y = y_j, A = a) \quad (7)$$

As a means of avoiding unnecessary repetition of computations, we offer factors to serve as multidimensional tables. The collective probability of all variables may be written as a factorization. We may determine the joint probability of X and E by setting $A_1=a_1 \dots A_r=a_r$ and marginalizing out the latent variables $X_1,\dots,X_n$ in turn.

$$b(Y, A_1 = a_1, \dots., A_r = a_r) = \sum_{X_1} \dots \sum_{X_m} l(Y, A_1, \dots, A_r, X_1, \dots., X_m)_{A_1=a_1,\dots,A_r=a_r} \quad (8)$$

After that, we may write the product form of the joint factors using the chain rule for BBNs (equation (5)):

$$b\,(Y_j|BE(Y_j)) = l\,(Y_j, BE(Y_j)) = l_j b(Y, A_1 = a_1, \dots, A_r = a_r) = \sum_{X_m} \dots \sum_{X_1} l(Y, A_1, \dots, A_r, X_1, \dots., X_m)_{A_1=a_1,\dots,A_r=a_r} = \sum_{X_m} \dots \sum_{X_1} \prod_{j=1}^{N} (l_j)_{A_1=a_1,\dots,A_r=a_r} \quad (9)$$

As a result, solving the last term in equation (8), which is the sum of products, is equivalent to performing inference in BBNs. The nonlatent variable terms in equation (8) must be factored out before the final term can be effectively computed.

3.2. Standards and Limitations in Design

The following is assumed for developing the dynamic traffic light control system:

- Northbound, westbound, southbound, and eastbound traffic all converge at this isolated four-way intersection.
- The primary direction of travel is presumed to be east–west.
- There is no thought given to going right or left.
- The dynamic logic controller method will track the volume of traffic heading in both directions (north and south and west and east).
- The flow of traffic in one direction is interrupted by a lull in the other direction.

3.3. Vision Computing in (Mosaicking)

To create a mosaic picture, many photographs of the same subject are taken and then seamlessly stitched together. This process is essential for expanding the viewable area of a picture without degrading the quality of the final product. Digital photos are limited in size, so it is not always feasible to capture a whole region of interest. When this occurs, we may use image mosaicking to create a larger picture by combining many photos that intersect. The photographs are resampled and aligned with the coordinate system of one of the overlapping images to make a mosaic. For an image mosaicking system to work, it must account for factors including camera-to-camera distance, camera-to-scene distance, camera attributes, and scene content. Picture mosaics include many steps, some of which are shown in Figure 4. These steps include altering, mixing, and sewing images together.

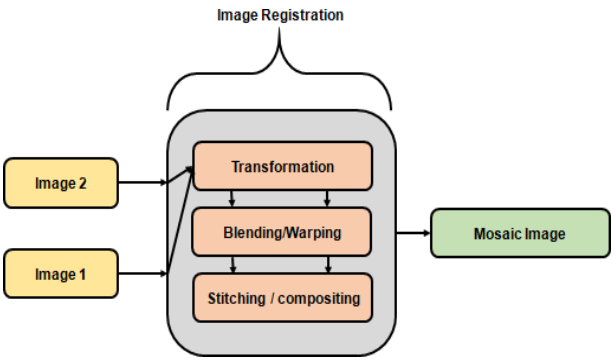


Figure 4 Image Processing Flow Diagram.

3.4. Cross-Correlation Pixel-to-Pixel Matched

Calculating the normalized cross-correlation coefficient between two templates reveals how closely their images resemble one another pixel by pixel. When the value of the normalized cross-correlation is higher than a predetermined cutoff, it is considered that a traffic light is located at the site in question. To provide the most accurate results, the template matching method relied on normalized cross correlation (sometimes called the distance measure or squared Euclidean distance).

$$t_{l,d}^2(w, c) = \sum_{y,x}[l(y, x) - d(y - w, x - c)]^2 \quad (10)$$

When feature t is located in a window at coordinates (u, v), the mosaicked image (f) is added to the total over (x, y) above the window. With d^2 growing,

$$t^{\downarrow}(l,d)^{\uparrow}2(w,c) = \textstyle\sum_{\downarrow}(y,x) \equiv [[l^{\uparrow}2]\,(y,x) - 2l(y,x)d(y-w,x-c) + d^{\uparrow}2(y-w,x-c) \qquad (11)$$

$\sum f^{2(y,x)}$ If roughly constant, then there is no further cross-correlation term.

$$v(w,c) = \textstyle\sum_{y,x} l(y,x)d(y-w-c) \qquad (12)$$

$$\gamma = \frac{\sum_{y,x}[l(y,x)-\overrightarrow{l_{w,c}}][d(y-w,x-c)-\overrightarrow{d}]}{\sum_{y,x}[l(y,x)-\overrightarrow{l_{w,c}}]^2 \sum_{y,x}[d(y-w,x-c)-\overrightarrow{d}]^2} \qquad (13)$$

The average value of f (y, x) within the region defined by the template d shifted to (v, u) is denoted with $\overline{fu,v}$ in Equation (14).

$$\overline{lw,c} = \frac{1}{M_y M_x}\sum_{y=w}^{w+My-1}\sum_{x=w}^{w+Mx-1} l(y,x) \qquad (14)$$

Equation (14) is preferable to other similar measures such as covariance or sum of absolute differences when calculating the matching degree because normalizing makes the computation more robust.

4. Results and Discussion

The effectiveness of the proposed framework was evaluated using a LabVIEW simulation test bed. There are vision sensors installed in the '+' style of roads seen in the GUI that monitor the time of red and green lights to ensure that they are consistent. An experiment was conducted on one car to compare the suggested method to the standard method. On the Aim Sun test bed, one of the reference cars is permitted to travel from west to east at a speed of 40 kilometers per hour. The vehicle is timed for 450 seconds total, and the distance traveled is measured at the end of that period. Road performance is measured by imposing a random load on the road network. ANN, CNN, and SVM have all had their performances tested using traditional sensors under identical traffic circumstances. Figure 5 below provides a comparison of traditional methods with the suggested strategy. The results of the simulated experiment show that the vehicle travels farther when utilizing a BBN controller in conjunction with vision sensor computing (4.6 m in the allotted time frame vs. 4.3 m when using CNN, 4.0 m when using ANN, and 3.7 m when using SVM). The transportation infrastructure has been upgraded with the help of a new platform called Aim Sun. The efficiency of this system was tested. Additionally, the proportion of road traffic is shown. The average wait and travel times for cars are shown in the simulation results. The red light will turn on when no distance has been traveled and will stay on until traffic is present and the algorithm has been executed. During this time of inactivity, an algorithm determines when the following cycle will begin with a green light. The proposed framework (BBN) improves the accuracy percentage time while lowering the average waiting time, as shown in Figures 6 and 7 and Tables 1 and 2.

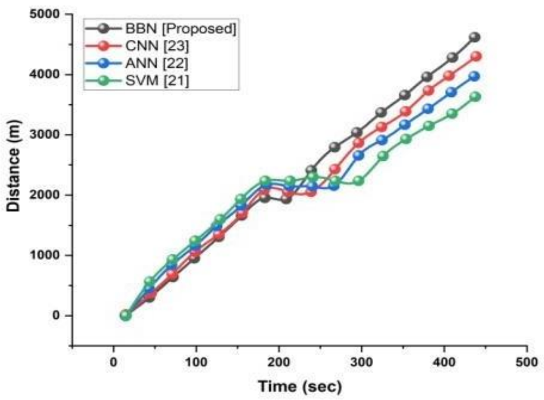


Figure 5 Conventional methods compared to the suggested system.

Table 1 Numerical outcomes of accuracy.

Methods	Accuracy (%)
SVM [21]	28
ANN [22]	37
CNN [23]	56
BBN [Proposed]	79

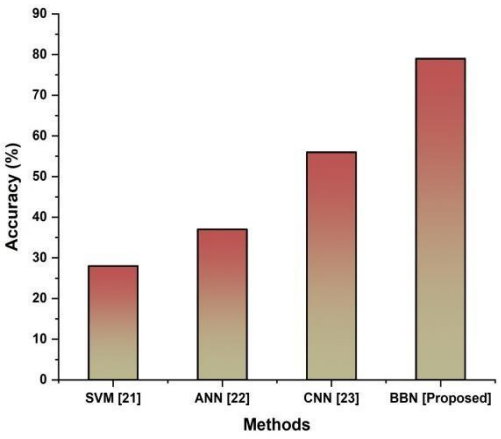


Figure 6 Comparison of accuracy.

Table 2 Numerical outcomes of waiting time.

Methods	Waiting time (sec)
SVM [21]	15
ANN [22]	25
CNN [23]	48
BBN [Proposed]	12

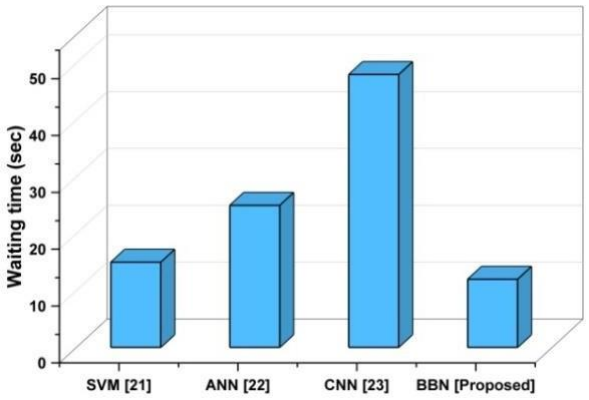


Figure 7 Comparison of waiting times.

5. Conclusion

The purpose of this research is to build an algorithm that may be utilized to reduce traffic congestion on major thoroughfares in developing nations. We initially provide a simple picture mosaic technique for counting people in a busy area through live webcam feeds. We analyze real-time road traffic photos to demonstrate the occurrence and duration of congestion collapses. We hypothesize that localized decongestion transportation systems may enhance traffic flow at key nodes in road traffic networks and are possibly simpler to install in real-world scenarios. This prompts the creation of low-cost transportation methods to ease traffic in third-world nations. To optimize traffic flow and minimize average waiting time at a junction, a smart BBN controller solution using vision computing has been developed, which is predicated on precise measurements of the road's dynamic traffic density.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Abduljabbar RL, Liyanage S, Dia H (2021) the role of micro-mobility in shaping sustainable cities: A systematic literature review. Transportation research part D: transport and environment 92:102734.

Aqib M, Mehmood R, Alzahrani A, Katib I, Albeshri A, Altowaijri SM (2019) Rapid transit systems: smarter urban planning using big data, in-memory computing, deep learning, and GPUs. Sustainability 11:2736.

Billhardt H, Fernández A, Ossowski S, Palanca J, Bajo J (2019) Taxi dispatching strategies with compensations. Expert Systems with Applications 122:173-182.

Cao Z, Ceder AA (2019) Autonomous shuttle bus service timetabling and vehicle scheduling using skip-stop tactic. Transportation Research Part C: Emerging Technologies 102:370-395.

Chu KF, Lam A Y, Loo BP, Li VO (2019) Public transport waiting time estimation using semi-supervised graph convolutional networks. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 2259-2264. IEEE.

Fang Z, Yang Y, Wang S, Fu B, Song Z, Zhang F, Zhang D (2019) MAC: Measuring the impacts of anomalies on the travel time of multiple transportation systems. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 3(2), pp.1-24.

Fridgen G, Thimmel M, Weibelzahl M, Wolf L (2021) Smarter charging: Power allocation accounting for travel time of electric vehicle drivers. Transportation Research Part D: Transport and Environment 97:102916.

Iyer LS (2021) AI-enabled applications towards intelligent transportation. Transportation Engineering 5:100083.

Kozlov IP (2022) Optimizing Public Transport Services Using AI to Reduce Congestion in Metropolitan Areas. International Journal of Intelligent Automation and Computing 5:1-14.

Leung CK, Braun P, Cuzzocrea A (2019) AI-based sensor information fusion for supporting deep supervised learning. Sensors 19:1345.

Lilhore UK, Imoize AL, Li CT, Simaiya S, Pani SK, Goyal N, Kumar A, Lee CC (2022) Design and implementation of an ML and IoT-based Adaptive Traffic-management system for smart cities. Sensors 22:2908.

Liu B, Han C, Liu X, Li W (2023) Vehicle artificial intelligence system based on intelligent image analysis and 5G network. International Journal of Wireless Information Networks 30:86-102.

Offiaeli K, Yaman F (2021) Social norms as a cost-effective measure of managing transport demand: Evidence from an experiment on the London underground. Transportation Research Part A: Policy and Practice 145:63-80.

Paiva S, Ahad MA, Tripathi G, Feroz N, Casalino G (2021) Enabling technologies for urban smart mobility: Recent trends, opportunities, and challenges. Sensors 21:2143.

Parveen S, Chadha RS, Noida C, Kumar IP, Singh J (2022) Artificial Intelligence in Transportation Industry. Int. J. Innov. Sci. Res. Technol 7:1274-1283.

Shaheen S, Cohen A (2019) Shared ride services in North America: definitions, impacts, and the future of pooling. Transport reviews 39:427-442.

Soori M, Arezoo B, Dastres R (2023) Artificial intelligence, machine learning and deep learning in advanced robotics, A review. Cognitive Robotics.

Syed AS, Sierra-Sosa D, Kumar A, Elmaghraby A (2022) Making Cities Smarter—Optimization Problems for the IoT Enabled Smart City Development: A Mapping of Applications, Objectives, Constraints. Sensors 22:4380.

Witlox F, Zwanikken T, Jehee L, Donners B, Veeneman W (2022) Changing tracks: identifying and tackling bottlenecks in European rail passenger transport. European Transport Research Review 14:1-12.

Yigitcanlar T, Desouza KC, Butler L, Roozkhosh F (2020) Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature. Energies 13:1473.

Zhang R, Ishikawa A, Wang W, Striner B, Tonguz OK (2020) Using reinforcement learning with partial vehicle detection for intelligent traffic signal control. IEEE Transactions on Intelligent Transportation Systems 22:404-415.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0118

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

An inventory management system for healthcare supply chains that incorporates epidemic outbreaks into consideration an investigation into COVID-19

Ashendra Kumar Saxena[a] | **Suneetha K.**[b] | **Harshita Kaushik**[c]

[a]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Professor, College of Computing Science and Information Technology.
[b]Jain (deemed to be)University, Bangalore, India, Professor, Department of Computer Science and Information Technology.
[c]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science & Application.

Abstract Epidemics are unique among calamities in two ways: the length of time they affect society and the speed with which they spread. In the absence of catastrophe prevention measures, supply chain (SC), and local communities would experience significant disruptions, leading to incalculable losses. The 2019 coronavirus illness (COVID-19) is one such calamity that has wreaked havoc on supply chains all across the globe, most notably the healthcare supply chain. As a result, this paper develops a practical decision support system based on doctors' knowledge and fuzzy inference system (FIS) for the first time to aid in insist management in the healthcare supply chain, lessen community pressure, disrupt the COVID-19 circulation chain, furthermore, in general, to diminish epidemic outbreaks for disruptions in the healthcare supply chain. This strategy first separates members of the community into four categories based on the sensitivity of their protected systems (normal, somewhat sensitive, extremely sensitive, and sensitive), as well as by two indications of age and pre-existing disorders and finally by two indicators of gender. These people are categorized after which they are obligated to follow the rules associated with their group. Finally, the success of the suggested strategy was evaluated in the real-world using data from four users, and the results demonstrated the efficiency and accuracy of the suggested strategy.

Keywords: supply chain (SC), fuzzy inference system, epidemics, sensitive

1. Introduction

A healthcare supply chain inventory management system is critical for assuring the availability of necessary medical supplies and for streamlining resource allocation. The system should keep track of and manage the stock of pharmaceuticals, immunizations, surgical supplies, and consumables. To provide precise inventory management, it should gather real-time data on stock levels, expiry dates, and consumption trends. The system should be able to predict the demand for medical supplies by looking at previous data, trends, and other elements. By doing so, healthcare institutions and suppliers may more efficiently manage their purchasing and distribution tasks, preventing stock-outs or surplus inventory (Ivanov 2020). Robotic Reordering The technology can automatically place orders based on demand projections and predetermined reorder points. For medical supplies, it may produce purchase orders or requisitions, guaranteeing prompt replacement and lowering the possibility of stock outs. Supplier Management The system should keep a complete database of suppliers, including contact details, lead times, prices, and product catalogues. To support effective contract management and procurement procedures, it should make it easier for suppliers to communicate and work together. Logistics Management The system should provide tools for organizing, tracking, and storing goods in warehouses and distribution facilities (Harland 2021). To simplify processes and reduce mistakes, it should include features like barcode scanning, batch, and lot management, and real-time stock changes. Management of Expirations the system should contain tools for tracking expiry dates and managing stock rotation since healthcare goods are time-sensitive. When products are getting close to expiry, it may send warnings and messages, allowing preventative actions like redistribution or disposal of expired products. Including Supply Chain Partners in the Integration. The system should interface with other parties involved in the healthcare supply chain, including manufacturers, distributors, and healthcare institutions, to guarantee end-to-end visibility and coordination. This allows for smooth data interchange, group planning, and effective order fulfilment (Queiroz 2022). The system should include analytical tools and reporting capabilities to help with cost control, supply chain effectiveness, and inventory performance. Stakeholders may use this information to make data-driven choices, streamline workflows, and pinpoint problem areas. Regulatory Conformity: The inventory management system must comply with all applicable legal mandates, including those

about monitoring restricted chemicals, handling temperature-sensitive goods, and guaranteeing accurate record-keeping and reporting. Flexibility and Scalability The system must be developed to meet the changing requirements of healthcare supply chains, including shifts in demand patterns, the introduction of new goods, and the development of cutting-edge technology. It should be flexible enough to interface with upcoming developments like the Internet of Things (IoT) or block chain for improved traceability, as well as scalable to accommodate higher inventory levels (Chowdhury P 2021). To ensure the availability of vital medical supplies during emergencies, epidemic breakouts must be taken into account while managing the healthcare supply chain. For the system to get up-to-the-minute information on epidemic outbreaks, it should interface with disease monitoring platforms and public health surveillance systems. Early warning systems may assist in initiating proactive actions in the supply chain, such as enhanced production, acquisition, or relocation of commodities to impacted regions. To anticipate the need for medical supplies during an epidemic, the system should use historical data, epidemiological models, and outbreak-specific information. Supply chain managers may prepare for outbreaks by using scenario planning tools to model various epidemic situations. This helps them predict demand spikes and allocate resources appropriately (Singh 2021). The supply chain system should be equipped with components that provide quick action and emergency procurement in the event of an epidemic breakout. To guarantee the timely supply of crucial medical goods, it should make it easier to identify and communicate with alternative manufacturers or suppliers. The system should contain allocation frameworks that take into account aspects like patient severity, geographic locations, and healthcare facility capacity to alleviate shortages and prioritize crucial requirements during an epidemic. This makes sure that scarce resources are allocated effectively and fairly according to the pressing nature of demand. The system needs to provide dynamic inventory management, enabling flexible supply redeployment following epidemic dynamics. This can include moving inventory from hotspots to less-affected locations or repurposing certain products to meet particular outbreak-related requirements (Khodaee 2022). The supply chain system should encourage cooperation and information sharing between many stakeholders, such as healthcare institutions, governmental organizations, manufacturers, and distributors, during an epidemic. This makes it possible for coordinated response actions, prompt communication, and effective resource allocation. Resilience and contingency planning the supply chain system have to include resilience tactics and backup plans tailored to disease outbreaks. Creating a strong and resilient supply chain during times of crisis involves maintaining safe stock levels of essential commodities, setting up backup suppliers, and deploying alternate distribution methods. Serialization and Traceability Implementing traceability methods may assist in tracking and tracing the flow of materials along the supply chain, such as product serialization and unique IDs (Kamran 2023). The supply chain, maximizing stock levels, and putting into practice adaptable plans based on current information. Regulations governing epidemic outbreaks, such as those relating to reporting requirements, quality requirements, and safety standards for particular medical products or treatments, should be followed by the supply chain system. The dependability and integrity of supply chain procedures are ensured by adherence to relevant rules. Healthcare businesses may strengthen their readiness, responsiveness, and resilience in times of crisis by including epidemic breakouts in supply chain management. For efficient coordination and resource optimization during an epidemic, the system has to be adaptive, agile, and able to integrate with public health systems (Ivanov 2021). To comprehend the features of the epidemic, such as the incubation time, mechanisms of transmission, high-risk demographics, and illness severity, do epidemiological investigations. Find probable origins of infection and transmission chains, this may entail case-control investigations, cohort studies, and contact tracking. Consider the outbreak's risk factors, such as the possibility of community transmission, the pressure on the healthcare system, the effect on vulnerable communities, and the possible effects on global public health. This evaluation supports strategies for allocating resources and public health initiatives. Put in place reliable surveillance methods to track the outbreak's spread. Monitoring COVID-19 instances, testing rates, hospitalization rates, and fatality rates are a few examples of this. Monitoring and analysing real-time data may help direct response activities and provide insights into how the epidemic is developing (Hossain 2022). Identify the cause of the illness (in this instance, SARS-CoV-2) and confirm cases by laboratory testing. Genetic sequencing is then used to monitor viral variations and comprehend how they affect disease severity and transmission. Perform a comprehensive contact tracing to find people who may have been exposed to the virus. To stop the spread of the disease, it is necessary to conduct interviews with confirmed patients, find out who they are in touch with, and take the necessary action. The goal of epidemic investigations is to pinpoint the outbreak's genesis, such as zoonotic origins or certain circumstances that sped up its development. To exchange information, data, and discoveries, work with organizations, professionals, and research centres in the fields of health on a national and worldwide scale (Ivanov 2021). Evaluate implemented interventions often, and make any adjustments to your methods. Record the lessons discovered throughout the investigative process to increase readiness for the next outbreaks and response capabilities. Ensure that the inquiry complies with ethical standards, which include preserving the privacy and confidentiality of each subject, gaining informed permission, and taking into account cultural sensitivity. To ensure that the inquiry is carried out ethically and following human rights, ethical considerations are essential. A multidisciplinary strategy combining epidemiologists, doctors, public health specialists, laboratory scientists, and other stakeholders is necessary to investigate an epidemic like COVID-19. The most recent findings and recommendations from renowned health organizations like the World Health Organization (WHO) and the Centres for Disease Control and Prevention (CDC) must be kept up to date (Singh 2021). It is necessary to address the particular difficulties presented by the pandemic to develop an inventory management system

for healthcare supply chains that takes epidemic outbreaks, notably COVID-19, into account. Figure. 1 represents the Structure of Healthcare supply chain.

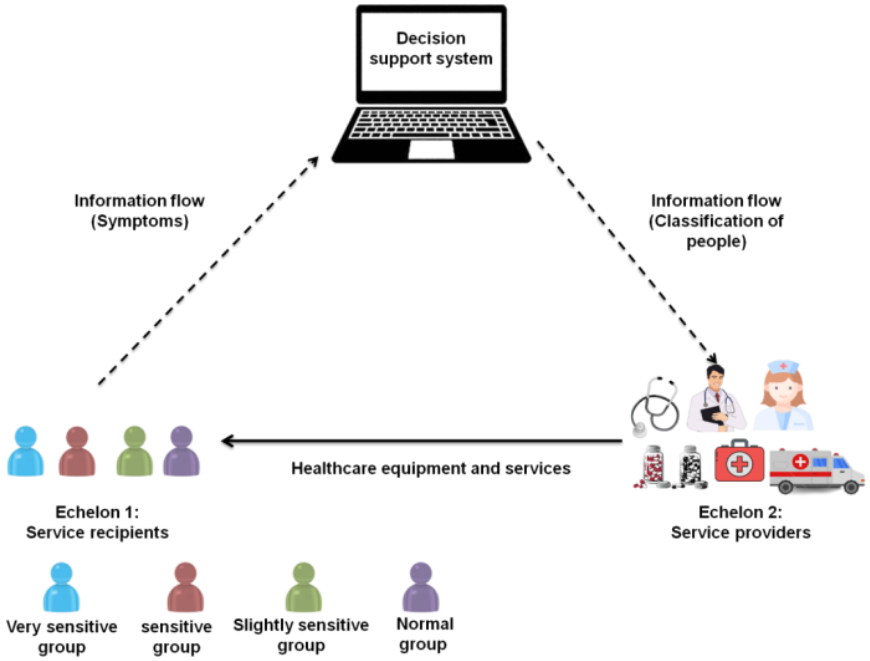


Figure 1 Structure of Healthcare supply chain.

2. Face recognition

To aid in healthcare supply chain demand management, decrease community stress, break the COVID-19 circulation chain, and lessen the impact of epidemics in general, this paper creates the first-ever sensible decision support system based on physician awareness and FIS. Results from a real-world evaluation of the suggested method utilizing data from four users confirmed the method's efficiency and precision (Govindan 2020). The research delves into how the pandemic affected supply chain resilience, provides a synopsis of the difficult situations faced by Chinese retail supply chains, and details JD. com's on-the-ground reaction to the crisis. To sum up, the pandemic produced extraordinary demand and severe logistical disruptions in China, and JD.com, thanks to its integrated supply chain structure and complete intelligence platforms, has managed effectively its supply chain management in response (Shen 2023). To analyse how companies respond to pandemics by shifting their production order priorities and stockpiling supplies. Our model integrates epidemic dynamics, supply chain design, and operational production-inventory control rules, three topics that are seldom discussed together in the operations management literature. Different scenarios for pandemic dynamics (i.e., uncontrolled spread or controlled dispersion with lockdowns) and 2 and 3-stage supply chains are analysed. Our research indicates that supply chains with two stages are more likely to experience disruptions (Rozhkov 2022). It is now obvious that there is a dearth of medical supplies necessary to offer appropriate treatment to patients as a result of interruptions in the supply chain and increasing demand. Using MCDM, we can better understand what factors make for stable healthcare supply chains. Resilient supply chain features may be evaluated with the aid of large healthcare supply networks. Diminish in significance throughout the COVID-19 epidemic (Zamiela 2022). The research aims to make two substantial contributions to two separate lines of inquiry. First, we synthesize the current state of knowledge on the spread of disruptions in SCs, establishing a systematic taxonomy and identifying ideas that demonstrate their use and applicability in mitigating the effects of pandemics on SCs. Second, to expose and systematize theoretical management concepts employed for functioning (adapting) during a pandemic and periods of recovery, and for becoming more resistant to future pandemics. To identify some unresolved research conflicts and fresh classifications by streamlining the existing literature (Ivanov 2021).

3. Proposed methodology

Many civilizations have put up proposals for disaster relief and implemented policies that are specific to their preparation, response, rehabilitation, and catastrophe control operations. Utilizing historical data and experiences with comparable events in the past helps with the handling of these types of catastrophes. Such catastrophes don't have pandemic proportions, they only impact a small portion of a nation, and they only endure a brief time. Because such

calamities are confined, there is a chance that non-affected nations will be able to assist the disaster-affected community. However, when a new, protracted, and progressively worsening catastrophe like COVID-19 materializes, the capacity to help takes on a different form. Such a dreadful scenario perplexes governments and decision-makers, interrupts almost all social activity, and breaks supply networks. In such circumstances, the high incidence of illness outbreaks results in a scarcity of medical and human resources in many places. There are interruptions in various supply chains as a result of this increased demand for services, particularly in the healthcare supply chain. To better serve underserved populations. Therefore, data from the World Health Organization (WHO) indicate that elderly inhabitants and those with previous conditions are more susceptible. Following this, the community's residents are classified into four categories according to how vulnerable they are to COVID-19:

Extremely sensitive collection: individuals older than 60 who have the smallest amount of diabetes, heart disease, or high levels of blood pressure.

Sensitive collection: A minimum of one of diabetes, heart disease, high blood pressure, or cardiac difficulties in individuals under the age of 60.

Slight sensitivity Collection: Those older than 60 who are healthy.

Standard collection: Those under 60 without illnesses.

Numerous methods built on FIS are used to categorize the individuals who make up the aforementioned groupings. The FIS is used to categorize society members, however, the fuzzy inference rules will vary depending on which society subgroup is being classified. This suggested decision support system does serve as a link in the healthcare supply chain that connects service users and service providers. Through the classification of service receivers, this assessment support system establishes the foundation for insist management in the healthcare supply chain.

3.1 The subsequent stages describe the suggested strategy

Stage 1: In the community, physical condition evaluation standards are established. 3 symptoms fever; exhaustion, and a dry cough are listed by the WHO as the early signs of COVID-19. The output variable of this system is how the community members are classified. Following that, the input and output variables' membership functions need to be established. There are three types of membership (low, moderate, and high) that make up each of the input variables. The following 5 membership functions make up the output variable:

Class 1: People who are symptom-free and otherwise healthy are assigned to these settings. These people must follow the healthcare advice and conduct their everyday lives following the limits and regulations imposed by their leaders.

Class 2: People who are suspected of having the illness and, as a result, should be isolated and limited in their social interactions until their condition is determined are isolated.

Class 3: Those in this category are those who have only a slim chance of really becoming infected with the mild sickness; in that case, they should be isolated at home.

Class 4: These people have been diagnosed with a serious illness or are in the hospital on suspicion of having such an illness; yet, they do not need urgent care.

Class 5: These people have been diagnosed with a serious illness or are in the hospital on suspicion of having such an illness; yet, they do not need urgent care. Individuals that have been evaluated may be placed into one of these five groups.

Stage 2: To establish a connection between the input and output variables, experts and doctors in the field are consulted to design the fuzzy inference rules. It's important to highlight that distinct fuzzy inference rules will be established for various social groups.

Stage 3: Members of the community are evaluated. To achieve this goal, the user first logs into the system, and then, after their group affiliation has been established, the suggested FIS is made available to them. Finally, three questions regarding fever, fatigue, and dry cough are used to categorize the person.

4. Results

4.1. Case study

To evaluate how well the proposed decision support system works; we compare it to current WHO data on COVID-19 and the expertise of practicing clinicians. It is important to highlight that the community evaluation criteria were chosen using WHO reports as a guide and that doctors' expertise was used to remove fuzzy inference policy and establish membership functions. Below is a detailed description of how to use the suggested strategy:

Stage 1: It all starts with pinpointing the system's inputs and outputs. A classification method whereby high body temperature, extreme fatigue, and a dry cough serve as inputs. There are three membership functions (levels) in the input data (low, medium, and high), and five membership functions (classes) in the output data. Also interesting is the fact that input variables have distinct membership functions for various classes. Figure. 2 displays, for each category, the membership functions of the input variables.

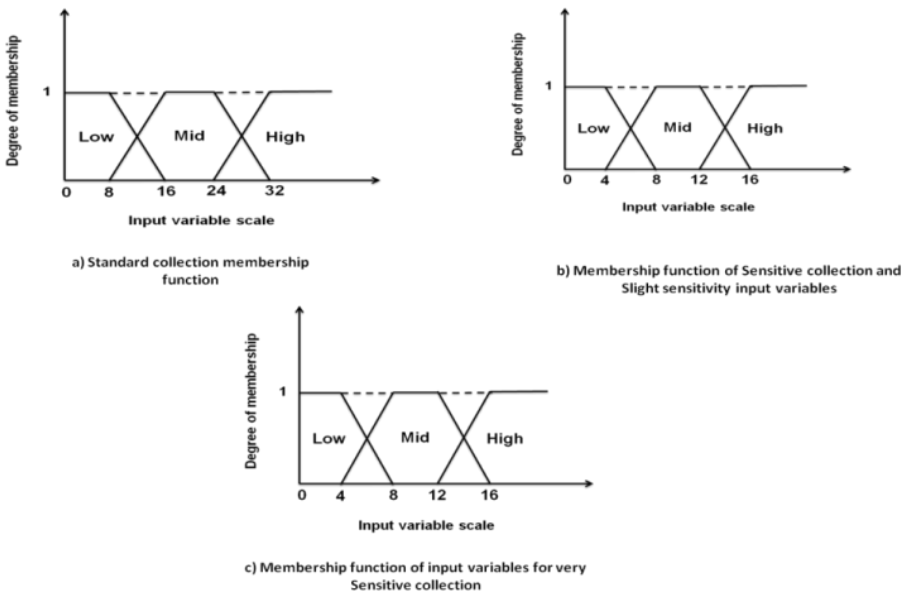


Figure 2 Input variable of membership functions.

Stage 2: The doctors' expertise in treating the four categories allows them to establish the fuzzy inference rules at this stage. As each of the three input variables in the proposed FIS has its membership function, the total number of fuzzy inference rules applicable to all possible combinations of persons is 327. By using the FIS Editor GUI toolkit in MATLAB R2016b, fuzzy inference rules may be extracted and implemented, allowing one to examine the connection between input and output variables in a spatially explicit manner.

Stage 3: People in this civilization are categorized. For this reason, we poll our users with three inquiries:

- How long have you been running a temperature?
- How long have you been feeling exhausted?
- How long have you been experiencing a dry cough?

The rules viewer box in the FIS takes the answers to these questions into account as inputs and then calculates the result. The system always returns a value between 0 and 1 as its output. The user is considered to be a part of a certain class if the value received as output falls precisely inside that class's membership function. If the final tally is among 2 membership functions, however, the membership degree is used instead. It makes no difference which class something is assigned to if the membership degree is the same for both. However, it should be classified as belonging to Class 1 rather than Class 2 if the membership degree for Class 1 is higher than that for Class 2.

The effectiveness of the suggested method is shown by the subsequent example are:

Client 1: This client is a member of the extremely sensitive collection.

Client 2: This client is a member of the Sensitive collection.

Client 3: This client is a member of the slight sensitivity Collective.

Client 4: This client is a member of the Standard collection.

Table 1 Responses to Questions by Each Client.

	Client 1	Client 2	Client 3	Client 4
How long have you been running a temperature?	16	13	21	33
How long have you been feeling exhausted?	9	13	21	25
How long have you been experiencing a dry cough?	4	19	25	17

Table 1 displays the responses to the queries that each of these users provided. The data shown in Table 1 are taken into account as inputs for the suggested FIS. Table 2 lists the output variable values that were determined for each user. The findings of Table 2 and the membership functions (classes) shown in Figurer 3 help classify clients.

Table 2 Each client FIS output variable.

	Output value
Client 1	0.663
Client 2	0.52
Client 3	0.508
Client 4	0.503

Client 1's output value is seen to be 0.663. However, this number is intermediate between categories 3 and 4. It leans toward Class 4; hence it's been positioned on the right side of the 2 classes' junction. In this way, User 1 will be assigned to Group 4. To be more specific, the output numbers for clients 1, 2, and 3 are 0.51, 0.509, and 0.502, respectively. From what can be seen in Figure. 3, these three users all fall within Class 3.

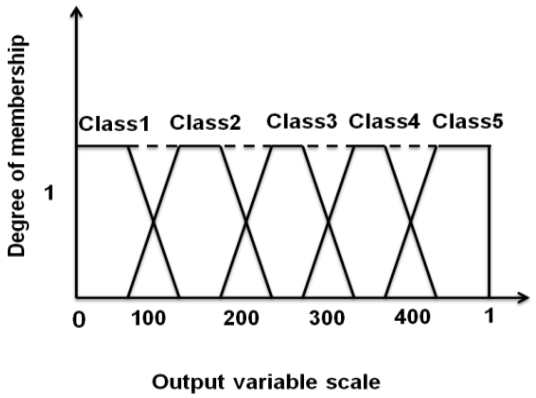


Figure 3 Output variable of membership functions.

4.2. Sensitivity analysis and discussion

To evaluate the efficiency of the suggested method, 4 scenarios are employed in which the user populations are shuffled about. Four users were assumed for the case study portion, and they were all classified into one of the four categories (Nikolopoulos 2021). The assumption is that the target audience falls into one of the other three categories of guides. To determine which class the user will be assigned to and whether or not the assignment is reasonable, we perform some tests. The following four cases are established for this function:

4.2.1. Case 1:

This Case assumes that client 1 is not part of the very sensitive category and is instead part of one of the other three categories. Class 4 is where client 1 would fall if it were placed in the very sensitive category (Rajak 2022). If the client is assigned to a different category, it is likely to be Class 4 or below. Given the extreme sensitivity. If the client is assigned to a different category, it is likely to be Class 4 or below. The membership functions of the input variables are quite delicate, thus this holds.

Table 3 Performance if client 1 is in separate collections.

	Class	Output value
This Client 1 is a member of the slight sensitivity Collective	3	0.357
This Client 1 is a member of the Extremely sensitive collection	5	0.663
This Client 1 is a member of the Standard collection	3	0.263
This client 1 is a member of the Sensitive collection	3	0.357

However, the fuzzy inference rules are thought to be more conservative for the delicate population. Table 3 displays the results and client classifications. The acquired findings corroborate the performance of the suggested model and hence support the claim stated here.

4.2.2. Case 2:

It is expected here that client 2 will fall into a different set of classes depending on whether they are classified as Very Sensitive, Somewhat Sensitive, or Normally Sensitive.

Table 4 Performance if client 2 is in separate collections.

	Class	Output value
This client 2 is a member of the slight sensitivity Collective	4	0.52
This Client 2 is a member of the Extremely sensitive collection	5	0.723
This Client 2 is a member of the Standard collection	3	0.318
This client 2 is a member of the Sensitive collection	4	0.52

Table 4 displays the outcome value and scenario categorization for this case. The outcomes in this case likewise conform to reasonable predictions, which is a testament to the reliability of the presented method.

4.2.3. Case 3:

Clients who are not part of the mildly sensitive group are used to evaluate how well the suggested method works (Ash C 2022). The outcomes of this case study are shown in Table 5. This scenario's outcomes corroborate the model's validity and indicate that it performs as expected.

Table 5 Performance if client 3 is in separate collections.

	Class	Output value
This client 3 is a member of the slight sensitivity Collective	4	0.508
This Client 3 is a member of the Extremely sensitive collection	6	0.917
This client 3 is a member of the Standard collection	4	0.6
This client 3 is a member of the Sensitive collection	5	0.634

4.2.4. Case 4:

According to the supposition, client 4 is a member of one of the other three groups and does not belong to the typical collections. It would be classified as class 3 if this user belonged to the typical collection (Farooq 2021). Depending on its membership in other groups, this client is likely to be assigned to classes 3, 4, or 5. If a different outcome is reached, it would reveal how poorly the suggested technique performed. The outcomes of this case are existing in Table 6.

Table 6 Presentation if client 4 is in separate collections.

	Class	Output value
This Client 4 is a member of the slight sensitivity Collective	4	0.722
This Client 4 is a member of the Extremely sensitive collection	5	0.917
This Client 4 is a member of the Standard collection	4	0.722
This client 4 is a member of the Sensitive collection	3	0.503

These case findings, together with those of the other three, show that the suggested method works as intended. COVID-19 is particularly dangerous for the elderly and individuals with compromised immune systems due to conditions like cardiovascular disease, diabetes, or hypertension. The 4 examples make it evident that the membership functions and fuzzy inference rules are designed differently for various collections based on their susceptibility. For instance, in case 3, if user 3 is a member of the highly sensitive category and the COVID-19 test outcome is positive, the patient should be admitted to the hospital and given urgent care. Thus, the outcomes demonstrate that the suggested strategy follows logical patterns and supports its effectiveness.

4. Conclusion

To categorize community members and subsequently regulate demand and contain disease outbreaks in the healthcare supply chain, a workable decision support system was presented. In the suggested method, clients are initially divided into groups based on their age and any current medical conditions. The FIS is then used to categorize these people. It should be noted that to categorize users, the 3 criteria of fever, exhaustion, and dry cough were employed. Special membership functions of these variables were taken into consideration for distinct groups. The suggested strategy was then verified using information about the four clients. These findings, together with those of the sensitivity analysis procedure, show that the suggested strategy has worked as intended and effectively. In addition to its advantages, each study has its limitations, and this work is no exception. The 3 symptoms of fever, exhaustion, and dry cough were used as the evaluation criterion for community members, which is one of the study's shortcomings. The most typical symptom of COVID-19 infection include these three criteria; however, some individuals have also had additional symptoms including vomiting, diarrhea, and the like. The possibility of expanding the number of membership functions to improve the decision support system's accuracy has been studied for three input variables.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Ash C, Diallo C, Venkatadri U, Van Berkel P (2022) Distributionally robust optimization of a Canadian healthcare supply chain to enhance resilience during the COVID-19 pandemic. Computers & Industrial Engineering 168:108051.

Chowdhury P, Paul SK, Kaisar S, Moktadir MA (2021) COVID-19 pandemic related supply chain studies: A systematic review. Transportation Research Part E: Logistics and Transportation Review 148:102271.

Farooq MU, Hussain A, Masood T, Habib MS (2021) Supply chain operations management in pandemics: a state-of-the-art review inspired by COVID-19. Sustainability 13:2504

Govindan K, Mina H, Alavi B (2020) A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: A case study of coronavirus disease 2019 (COVID-19). Transportation Research Part E: Logistics and Transportation Review 138:101967.

Harland CM, Knight L, Patrucco AS, Lynch J, Telgen J, Peters E, Tátrai T, Ferk P (2021) Practitioners' learning about healthcare supply chain management in the COVID-19 pandemic: a public procurement perspective. International Journal of Operations & Production Management 41:178-89.

Hossain MK, Thakur V, Mangla SK (2022) Modelling the emergency health-care supply chains: responding to the COVID-19 pandemic. Journal of Business & Industrial Marketing 37:1623-39.

Ivanov D (2020) predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. Transportation Research Part E: Logistics and Transportation Review 136:101922.

Ivanov D (2021) Lean resilience: AURA (Active Usage of Resilience Assets) framework for post-COVID-19 supply chain management. The International Journal of Logistics Management.

Ivanov D (2021) Supply chain viability and the COVID-19 pandemic: a conceptual and formal generalization of four major adaptation strategies. International Journal of Production Research 59:3535-52.

Ivanov D, Dolgui A (2021) OR-methods for coping with the ripple effect in supply chains during COVID-19 pandemic: Managerial insights and research implications. International Journal of Production Economics 232:107921.

Kamran MA, Kia R, Goodarzian F, Ghasemi P (2023) a new vaccine supply chain network under COVID-19 conditions considering system dynamic: Artificial intelligence algorithms. Socio-Economic Planning Sciences 85:101378.

Khodaee V, Kayvanfar V, Haji A (2022) A humanitarian cold supply chain distribution model with equity consideration: The case of COVID-19 vaccine distribution in the European Union. Decision Analytics Journal 4:100126.

Nikolopoulos K, Punia S, Schäfers A, Tsinopoulos C, Vasilakis C (2021) Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. European journal of operational research 290:99-115.

Queiroz MM, Ivanov D, Dolgui A, Fosso Wamba S (2022) Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review. Annals of operations research 319:1159-96.

Rajak S, Mathiyazhagan K, Agarwal V, Sivakumar K, Kumar V, Appolloni A (2022) Issues and analysis of critical success factors for the sustainable initiatives in the supply chain during COVID-19 pandemic outbreak in India: A case study. Research in Transportation Economics 93:101114.

Rozhkov M, Ivanov D, Blackhurst J, Nair A (2022) Adapting supply chain operations in anticipation of and during the COVID-19 pandemic. Omega 110:102635.

Shen ZM, Sun Y (2023) Strengthening supply chain resilience during COVID-19: A case study of JD. com. Journal of Operations Management 69:359-83.

Singh S, Kumar R, Panchal R, Tiwari MK (2021) Impact of COVID-19 on logistics systems and disruptions in the food supply chain. International journal of production research 3:1993-2008.

Zamiela C, Hossain NU, Jaradat R (2022) Enablers of resilience in the healthcare supply chain: A case study of US healthcare industry during COVID-19 pandemic. Research in Transportation Economics 93:101174.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0119

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

The use of multi-agent systems in an automated restoration process for power distribution networks

Shambhu Bhardwaj[a] | **Bhuvana J.**[b] | **Davendra Kumar Doda**[c]

[a]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College of Computing Science And Information Technology.
[b]Jain (deemed to be)University, Bangalore, India, Associate Professor, Department of Computer Science and Information Technology.
[c]Vivekananda Global University, Jaipur, India, Associate Professor, Department of Electrical Engineering.

Abstract The capacity of future smart grids to automatically detect problems, isolate them, and restore service is a key feature. Traditional mathematical programming techniques and heuristic approaches become computationally highly expensive when used for the restoration issue, which is a comprehensive nonlinear optimization challenge. Distributed control must offer comfortable and safe power system operation and improve consumer supply quality. In this article, we suggest the multi-agent system (MAS) as the creative process that applies to the existing power distribution network. The MAS agents are installed in "Intelligent electronic devices (IEDs)" and interact with one another using open protocols from the "Foundation for Intelligent Physical Agents (FIAP)." A Java-based distribution system simulator was built to test how well MAS works. In-lab experiments were conducted with the simulator connected to the MAS. The test findings demonstrated that the MAS effectively restored healthy divides by changing them to their original configuration following a predetermined sequence, notwithstanding the system's operating limits. The suggested MAS has shown the efficacy of a distributed control method that uses in the absence of the technologies currently employed in IEDs for control and protection.

Keywords: multi-agent system, intelligent electronic devices, power distribution networks, automated restoration process

1. Introduction

Power Distribution networks, which mark the completion of power transmission networks, must be developed to provide energy to customers dependably and effectively. Conventional distribution infrastructures are passive connections that constantly have several issues. One of the primary causes of customers' unhappiness due to recurrent damage to sensitive voltage equipment is poor voltage control of traditional distribution networks. The frequency and length of service interruptions are also increased by the constant rise in power usage and the incapacity of distribution feeders that are old and already overburdened to handle the rising load. However, constraints on electricity generation prices and rising replacement and renewal costs for the current distribution networks make it more difficult for distributing system administrators to satisfy customer demands (Shaheen 2022). The foundation for developing the outage control arrangement is an understanding of the features of distribution operations. A standard distribution network has several feeds. Some sectionalizing switching (SSs), typically closed, are placed on the feeders. The feeder's end bus is connected to the feeder by a normally-open tie switch (TS). Since protective tools are created using this architecture, the whole system is developed in radial architecture. The upstream transmission network normally provides service to the whole distribution network. A distribution network is fitted with many SSs that can be operated remotely. Additionally, the fault indicator may identify the damaged line. Distributed energy resources (DERs) of accessible and non-accessible forms may exist in the system (Shi 2021). Worldwide energy consumption is still rising, and with it are the quality expectations. Increasing the dependability of electrical networks is one of the key goals of smart grids. Self-correcting methods, such as those that automatically locate faults, isolate them, and restore service, may do this. Avoiding the requirement for customer contacts or servicing teams to manually find and restore power to the optimal regions of the system, these three stages enable defects to be recognized and rapid action to be performed. Modern smart grids will have the potential to repair themselves, making fault separation, service reconstruction, and defect localization automated (Zafar 2020). Restoration of the distributing system is accomplished by connecting the neglected loads left after fault separation to nearby feeders that may offer the necessary power to restart them. This is accomplished, as with the distribution system reorganization, by altering the state of connection breakers and sectionalizing, which are typically closed and open, respectively. Nevertheless, the goals in each situation are extremely different because of the distinct nature of the problems being addressed. While restoration occurs in a crisis with the primary goal of feeding the untapped loads, recon Figure ring happens while the distribution network

operates normally and aims to minimize system damage (Al-Hinai 2021). Since higher impedance failure (HIF) current is unpredictable, asymmetrical, and nonlinear, it presents a very difficult challenge for detection. Although the size of the fault current is often much smaller than the nominal load current, it is sometimes impossible to identify and identify these issues using standard over-current approaches. When an energized conductor makes contact with the ground via a high impedance item, such as dry pavement, wet sand, dry grasses, sod, etc., which restricts the flow of electricity to the ground, high inductance failures usually result. Power systems must operate efficiently, dependably, and safely, which requires immediate recognition of high impedance difficulties (Sarwar 2020). Today, digital equipment's used by electrical providers all around the globe. Whether it is used for power production, dissemination, or distribution, they are often used in the grid. Despite being the most prevalent, power distribution networks are the least computerized due to issues with data gathering, adapting strategies, errors to recognize certain situations and their effects, and a lack of dependability evaluation techniques. Energy firms paid minimal attention to these systems since they were essential to them. But these days, it's important to pay close attention to these connections. They have not received enough attention, which has caused a delay in their digitization, and there are still some problems. For instance, incorrect device positioning results in the absence of accurate and full data, which is why automated systems and relay safety in distributing power lines do not offer discrimination (Suslov 2020). Promptly identifying problems and launching the necessary fixes greatly increases the speed and effectiveness of power restoration during outages. Users' downtime is decreased, and the network's dependability is improved by proactively recognizing possible problems and taking preventative action. Additionally, automation maximizes resource utilization while minimizing operating expenses during restoration attempts. Power distribution networks need an automated restoration procedure for effective power minimization. Hence, we proposed an effective MAS for the automated restoration process for power distribution networks.

2. Related Works

Power delivery firms observe distribution network modernization as a successful investment option to improve operational consistency and dependability. The automated distribution network incorporates both shifting and protection equipment. This study (Li 2020) proposes an innovative approach for fault measure and sectionalizing switch location optimization in branch-line distribution systems. The suggested strategy's objective function accounts for the overall expense of fault metrics, sectionalizing controls, and disruption costs. This study examines failure signals and remote-controlled valves as various automated devices. Fiber-to-the-home networks (FTTH) are considered the most reliable solution for providing clients with more and more throughput in the future. They offer freedom and the capacity to handle larger bandwidths contrasted to the older access system using passive optical network connections. This article will outline the various fiber deficiencies that impact FTTH networks' outstanding service. Monitoring and recovery are required for the damage. Through the identification and accessibility of the downstream signal from the optical line port, security and restoration processes are stored using this system (Ab-Rahman 2022). After power disruptions brought on by severe weather, electric utility providers try to recover as much load as they can. This study suggests a network redesign and DERs scheduling outage control technique to improve distribution sector robustness. The suggested approach may determine the con Figure ration of a radial network after a line fault depending on the order of the incidence vector. Sectionalizing lines and switching tie lines are used to perform the redesign. An optimum DER scheduler issue is addressed using the novel system architecture to reduce load and minimize the overall cost of transportable DER activity (Shi 2021). After a natural catastrophe, repair teams (RT) and mobile power resources (MPR) are essential resources for managing distribution system (DS) outages. Their transportation, nevertheless, has not been thoroughly examined. We provide a robust disaster response operations plan to coordinate DS regeneration with the deployment of RT and MPR. To arrange RT and MPR in the mobility network plan them in the DS, modify the DS for micro grid construction coordinately, etc., a unique co-optimization method is developed (Lei 2019). The movement of conventional distribution systems towards new active ones has been prompted by the emergence of power system reorganization and the significant incorporation of renewable energy supplies. Rapid technological advancements in the interim have created excellent opportunities for future large-scale use of generating units and power preservation structures in the distributing sector. From the perspective of operative time hierarchy, this study intends to give a thorough analysis of current developments in the functioning of activity distribution systems (Ghadi 2019). Microgrids, distributed generators, and nearby assets may be utilized to reestablish vital loads and provide resilience when a significant interruption in a distribution system results from catastrophic events. To service important loads during blackouts, this research suggests a decision-making process to identify the best restoration plan by coordinating several sources. A two-step strategy addresses the critical load recovery issue, with the first stage choosing the post-restoration topologies and the next stage selecting the set of demands to be recovered and the outputs of suppliers (Wang 2019). These are the existing techniques implemented in the automated restoration process for power distribution networks.

3. Proposed Methodology

3.1. Multi-agent systems

In this part, we introduce the FIPA language and standards for describing multi-agent methods, as well as the platform Java Agent Development Environment (JADE) utilized to create the suggested MAS. Agents are beings with a high degree of generality that can read input and conduct independent behaviors to alter the condition of their surroundings. They may be either hardware or software. Features like independence, responsiveness, initiative, and friendliness are present. All intelligent beings have these capacities by nature. Their actions and worldview define who they are. An ontology is a description for the aim of facilitating the transfer and recycling of information inside an agent or among a group of agents.

Zone agents and feeder agents make up the MAS. At every pair of sectionalizing controls in a feeder, there is a zone agent stationed to keep a watch on things. When anomalies occur, they notify their feeder agent and implement the controls prescribed by that agent. The actions of feeder agents change based on what the system requires. There are three distinct types of these actions: the creator, the backup, and the second backup.

Established in 1996 by the IEEE Computer Society, FIPA is an association dedicated to advancing agent-based technology and ensuring compatibility across various systems. FIPA's model protocol is service-oriented, unlike the OSI or TCP/IP models, which each include a single sub-layer at the application level. The agents may communicate with one another using a collection of iterative protocols specified by FIPA, such as the “FIPA-Request Protocol, FIPA- Contract Net, and FIPA-Subscribe."

JADE is a platform that is open source and may be used for the establishment of MAS using middleware that is compliant with the FIPA standard. For use in peer-to-peer agent-based applications, JADE is a protocol that is written in the Java programming language. The "Agent Management System (AMS)" is housed inside a specialized package unique to each platform. The AMS is required to be completed and offers a white page service. The protocols established by FIPA are of utmost significance for the collaboration and conversation between agents. The message transfer service is used extensively throughout the agents' primary operations. Messages are treated as objects in JADE. These messages must comply with the ACL standard and must possess properties.

3.2. MAS architecture

The terms "substation agent (SA), feeder agent (FA), branch agent (BA), and equipment agent (EA)" all describe different types of agents. In the event of a forever fault, the MAS is responsible for the next actions: pinpointing its location, isolating the healthy extends that were adversely impacted, negotiating a backup source of power, analyzing the limitations on the operation of the network, and re-energizing the adequate but de-energized divides if at all feasible. Table 1 details the many roles that the agents play.

Table 1 Agents' Descriptions.

Agents	Tasks performed by agents
FA	Communicates with other FA to negotiate available power by sending and receiving request messages; The conversation's originating FA evaluates the received offers in order of authority and picks the one with the highest rating. The FA approaches the BA, who is tasked with assessing the operational limitations, to establish a dialogue.
EA	Provides the associated FA with information from the IEDs about faults to begin power restoration negotiations for branches that were previously powered but are now dead.
AS	Taking in information from EA, notifying EA of the damaged area, and alerting FA to begin repair work.
BA	A message of authorization is received from FA; the branch's overload situation is assessed, and the available power is determined. The BA will communicate the possibility of recovery.

3.3. Development of MAS

As noted, the created MAS can identify problems, cut off healthy circuits, reorganize the network, and restart power without risk by utility-defined priorities. The MAS optimizes the goal function in equation (1) to provide an accurate recovery.

$$max \sum_{i\epsilon A} W_i \qquad (1)$$

Where $i\epsilon A$ is the collection of all branches in the system, and W_i is the overall consumers that frequent unit$i$.

A secure recovery requires the MAS first to analyze the network's current state and then restore without breaking any operational limits. In this study, we focus on two functional restrictions, illustrated by equations (2) and (3), including the capacity limits of converters at power substations and feeder cable sections.

$$\sum_{j\epsilon T} S_j < S_{max,k} \qquad (2)$$

$$S_{j\epsilon T} < S_{max} \qquad (3)$$

$$I_i < I_{max} \qquad (4)$$

$$I_{i\epsilon A} < I_{max} \qquad (5)$$

Where S_j is the convertors capacity is $k, S_{max,k}$ is the overall capability of the power substation k, $j \epsilon T$ is the collection of convertors of a substation, $i \epsilon A$, is the group of the network section. i And I_{max} are the branch's thermal limit and k, I_i is the current through the component at a particular point in time.

The MAS verifies that the additional load current traveling via the feeder doesn't go beyond the present limit of the cables and that the move of load, which occurred as a result of the system modification, did not result in an overflow at the substation that was authorized for the recovery. In addition, the MAS verifies that the current limit of the cables was not exceeded by the enhanced current of the load.

3.4. Model for MAS data

All relevant restoration data is represented in eXtensible Markup Language (XML), which is utilized for knowledge expression.

3.5. MAS Specifications and modeling

MAS agents are software substances that have intelligence encoded in hardware that is outside of the IED. This is because the MAS agents are implemented as part of the IED. The data necessary for the reconstruction system to function is obtained from the IEDs themselves; it is not reliant on a centralized control system. The actions of the MAS agents are discussed in the following section.

The SA is responsible for managing the layout of the feeders in the substation impacted by the absence of converters. The FA is responsible fsor managing the rehabilitation of every division and negotiating power with the other FAs. The reconstruction of sections is carried out by the BA after analyzing the operational limits imposed by the power system. The EA is the interface between the IED connected to the principal devices, and it is represented by the EA. The MAS receives data from the IEDs via the EA.

After the occurrence of a problem, the EA will deliver a message to the FA using the FIPA- requesting process. To confirm that the message was received, FA sends back a confirmation message of that kind INFORM.

The FA employs the FIPA-Contract Net protocol and communicates with all agents who can respond to a request for assistance in resolving an issue by delivering a message of the type of Call for Proposes. For example, it is employed when a beginning FA tries to negotiate power with other FAs in the neighborhood.

The participating agents will either react with a message of PROPOSE or REFUSE. After receiving all of the messages, the initiator FA will choose the most effective suggestion, after which it will issue the “ACCEPT-PROPOSAL and REJECT-PROPOSAL” messages, respectively. It is important to emphasize that the negotiating process takes place separately for each branch whose power has not been returned. On the other hand, the instruction to reestablish authority is only carried out once following the conclusion of the investigation into every sector.

Communication between FA and BA takes place using the FIPA-Subscribe protocol. The FA will get a message of type AGREE if the section reconstruction is successful and a statement of type FAILURE alternatively, prompting it to begin negotiations to recover another section. After identifying the malfunctioning part, the SA notifies the EA and the FAs to fix it. The FIPA-Subscribe protocol is used by the FA to carry out the necessary discussions.

4. Simulating the power network for MAS testing

In this article, we will discuss how the MAS was integrated into a test workbench at the Federal University's Power Systems Laboratory.

4.1. Method of examination

To ensure the MAS would perform as expected under real-world conditions, a Java-based computer simulator of a medium voltage distribution system was built. “Aquiraz (AQZ), Jabuti (JAB), Messejana (MSJ), and Agua Fria (AGF)” are the four power substations that provide electricity to the system. The AQZ substation is unique because it has four output feeders, and substations only have one feeder each. Nine recloses and six normally open (NO) ties are used to move loads on the neutral voltage network. There are 12 separate networks that are separated by reclosers and tie controls.

When a persistent failure develops in the system, the MAS is conducted to reactivate the power while taking into consideration the limits imposed by the network to guarantee a safe recovery. It is essential that the network's radial layout be preserved while the tie controls are being switched.

When a prolonged problem has been fixed, the MAS agents will begin to communicate with one another, communicate with one another, and work together to restore the network automatically. During this procedure of working together and sharing data, the agents are responsible for performing the following activities and functions: locating and identifying the fault, isolating the industries impacted by the fault, negotiating backup power, evaluating the operational limitations, and restoring the sound sections.

The power supplies in the substation have a nominal current as well as an excessive voltage, which is shown in Table 2. If one of the transformers failed, the remainder converter could handle an excessive voltage of up to 120% of its capacity for 30 minutes.

Table 2 Information divisions for substation.

Substation	Bay	Transformers overload limit, A	Transformers nominal current, A
JAB	01T1	630	525
	01T2	630	525
AQZ	01T1	755	629
	01T2	755	629
AGF	01T1	1338	1115
	01T2	1338	1115
MSJ	01T1	1338	1115

The information on the currents and the number of customers who live in every spread network section is detailed in Table 3. The MAS can get the information since they are saved in an XML file on the computer. The current detected is proportional to the load requirement in each segment. The overall current that is being measured at the start of the sector is equivalent to the current that is being measured in a certain industry.

Table 3 Information divisions for feeder.

Feeder	Branch	Load, A	Total current, A	Cable thermal limits, A	Number of consumers
AQZ_0116	B4	34	184	439	351
	B5	151	151	233	401
AQZ_0114	B11	165	204	476	201
	B12	40	40	439	301
MSJ_21M2	B6	2634	264	439	301
AQZ_0117	B1	13	245	476	301
	B2	233	233	476	251
AQZ_0115	B7	55	332	476	351
	B8	121	278	476	301
JAB_01F4	B3	166	166	476	201
AGF_2111	B10	344	344	526	351

4.2. Simulation and MAS System Integration

The simulator was made available so that the integration of the power delivery system on an exterior MAS recovery system could be validated. The proxy agent that is offered in the JADE platform to interact through TCP/IP between MAS and exterior devices is what is used to carry out the communication that takes place between the MAS and the IEDs. Information is delivered to the MAS program, hosted on the JADE platform, whenever the consumer modeled a problem on the distribution system. This information is sent through the TCP/IP protocol. The agents will examine the status of the malfunction and then take the necessary steps to separate the problem and reconnect the power network. The inclusion of the MAS into the JADE framework, the power system simulator, and the XML database are shown in Figure 1.

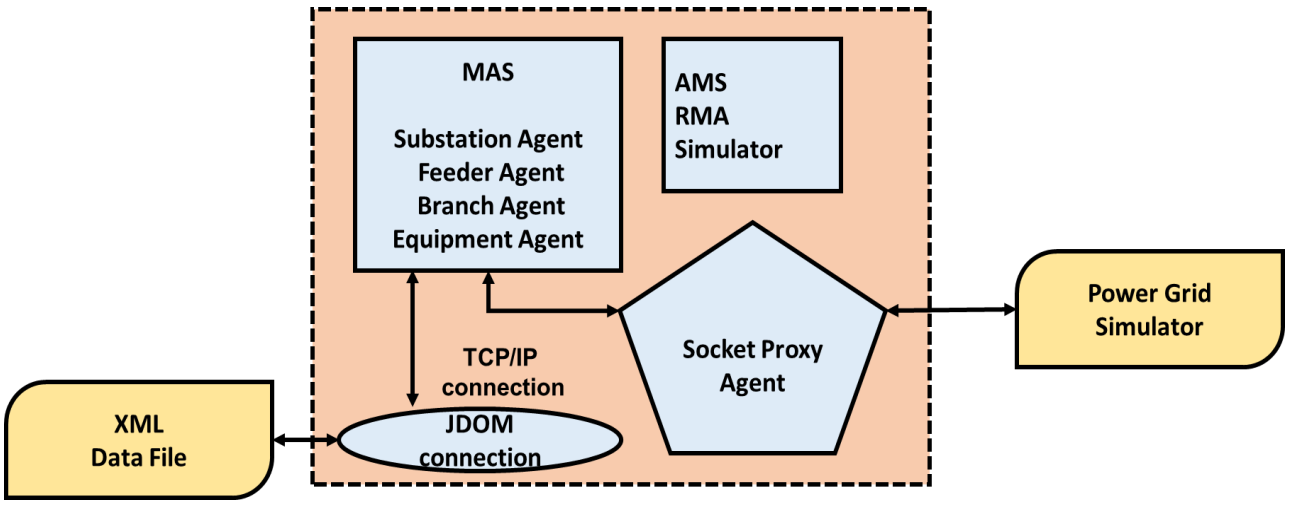


Figure 1 Simulation and MAS System Integration.

Through the process of restoring the database, the XML data file is changed into a Java object by using the JDOM library. Then it is transmitted as a note to any other agents that require to obtain the data contained inside the database. The database will get frequent revisions from the EA. By negotiating for electricity and restoring the power grid, the FA validates and updates the XML data file as necessary.

5. Performance Analysis

Automatic power restoration procedures may minimize the time needed to restore power after an outage. Automating fault detection, evaluation, and reaction allows the system to swiftly locate the issue and launch the necessary fixes, including rerouting electricity or isolating problematic areas. As a result, we recommended multi-agent systems (MAS) for power distribution networks' automated restoration system. The existing techniques used for evaluation are Hilbert–Huang Transform (HHT) and Convolutional Neural Network (CNN) (Sekhavatmanesh 2020), Modified Combinatorial Benders (MCB) (Guo 2019), and graph convolutional network (GCN) (Chen 2019). Adaptability, Fault Detection and Localization, and Implementation cost are evaluation metrics.

5.1. Adaptability

The term "adaptability" describes a system's capacity to efficiently adjust to and react to changes in its surroundings or operational circumstances. It involves the system's capacity to adapt its actions, plans, or decision-making procedures to suit brand-new conditions, demands, or restrictions. A flexible system should be able to react quickly and at the moment to new data or occurrences. This entails swiftly processing and integrating further information or sensor readings, analyzing the effects of changes, and modifying its control or decision-making procedures as necessary. Figure 2 depicts the adaptability of the conventional and suggested system. Table 4 shows the comparison of adaptability. It shows that the proposed MAS has high adaptability.

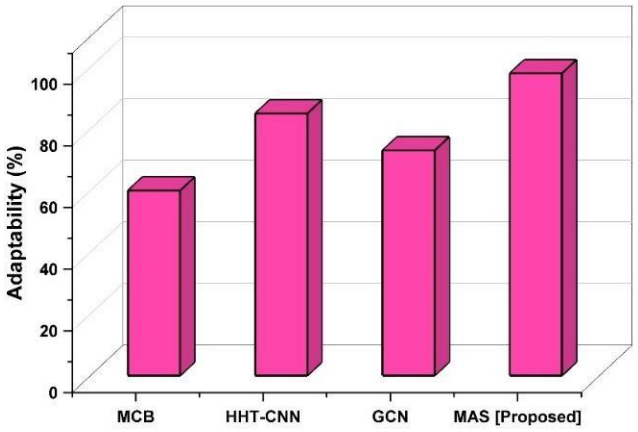


Figure 2 Adaptability of the conventional and suggested system.

Table 4 Comparison of adaptability.

Methods	Adaptability (%)
MCB	60
HHT-CNN	85
GCN	73
MAS [Proposed]	98

5.2. Fault detection and localization accuracy

Identifying the occurrence of abnormal circumstances or interruptions inside the power system is the process of fault detection. This may include different defects, including line breaks, equipment malfunctions, transformer problems, or short circuits. After a defect has been identified, fault localization entails pinpointing its exact position inside the power system. This knowledge is essential for concentrating restoration efforts and reducing the severity of the outage. An automated restoration method must include fault detection and localization because they allow the system to identify defects and pinpoint their locations quickly. Figure 3 depicts the fault detection and localization of the conventional and suggested system. Table 5 shows the comparison of fault detection and localization. This indicates that the MAS has accurate fault detection and localization. Hence it efficiently restores power to the affected areas.

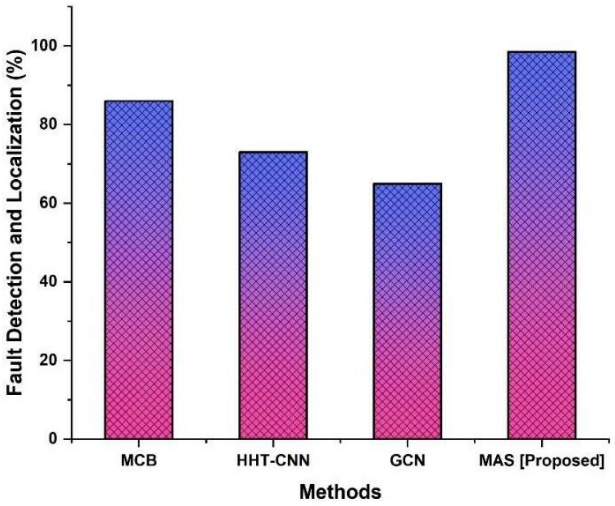


Figure 3 Fault detection and localization of the conventional and suggested system.

Table 5 Comparison of fault detection and localization.

Methods	Fault Detection and Localization (%)
MCB	86
HHT-CNN	73
GCN	65
MAS [Proposed]	98.5

5.3 Implementation cost

The implementation cost of a system refers to the expenses involved in bringing it into use in a specific piece of software or hardware. Any system's implementation cost may vary greatly based on a wide range of variables, including the system's complexity, the number of people participating, the infrastructure needed, and the particular objectives and specifications of the application. Figure 4 depicts the implementation cost of the conventional and suggested system. Table 6 shows the comparison of implementation costs. Implementation cost is less in the proposed method.

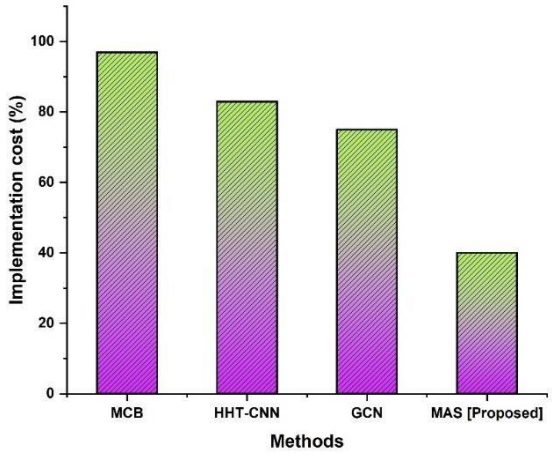


Figure 4 Implementation cost of the conventional and suggested system.

Table 6 Comparison of implementation cost.

Methods	Implementation cost (%)
MCB	97
HHT-CNN	83
GCN	75
MAS [Proposed]	40

6. Conclusions

This research examined the issue of fault recovery in automated distribution networks. Since their capacity for exchanging complex data as well as allocating responsibilities among agents, MAS is effectively able to resolve this

challenging challenge. Using a WLAN utilizing the TCP/IP protocol, the MAS program created on the JADE platform and a powerful network simulator designed in Java were combined. The MAS system was independent of the existing methods and procedures used by the IEDs since the multi-agents were integrated into exterior hardware. The findings from the tests have shown that the JADE platform and the simulator can effectively communicate, demonstrating the capabilities of the JADE application to connect to real-time systems utilizing peer-to-peer communication. Additionally, the growing significance of decentralized management has inspired the creation of improved and more successful methods. That's exactly what this study has done. To aid in the reconstruction of future micro grids and maintain their integration with smart power systems, the suggested technique may be improved for active micro grid systems and smart grids as part of future research.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Ab-Rahman, MS, Manaf ZA, Kaharudin IH, Hwang IS (2022) Customer Edge Downstream Detection for Automatic Restoration Scheme in FTTH-PON Distribution Network. In Photonics 9:560.

Al-Hinai A, Alhelou HH (2021) A multi-agent system for distribution network restoration in future smart grids. Energy Reports 7:8083-8090.

Chen K, Hu J, Zhang Y, Yu Z, He J (2019) Fault location in power distribution systems via deep graph convolutional networks. IEEE Journal on Selected Areas in Communications 38:119-131.

Ghadi MJ, Ghavidel S, Rajabi A, Azizivahed A, Li L, Zhang J (2019) A review of economic and technical operation of active distribution systems. Renewable and Sustainable Energy Reviews 104:38-53.

Guo MF, Yang NC, Chen WF (2019) Deep-learning-based fault classification using Hilbert–Huang transform and convolutional neural network in power distribution systems. IEEE Sensors Journal 19:6905-6913.

Lei S, Chen C, Li Y, Hou Y (2019) Resilient disaster recovery logistics of distribution systems: Co-optimize service restoration with repair crew and mobile power source dispatch. IEEE Transactions on Smart Grid 10:6187-6202.

Li B, Wei J, Liang Y, Chen B (2020) optimal placement of fault indicator and sectionalizing switch in distribution networks. IEEE Access, 8, pp.17619-17631.

Sarwar M, Mehmood F, Abid M, Khan AQ, Gul ST, Khan A S (2020) High impedance fault detection and isolation in power distribution networks using support vector machines. Journal of King Saud University-Engineering Sciences 32:524-535.

Sekhavatmanesh H, Cherkaoui R (2020) A novel decomposition solution approach for the restoration problem in distribution networks. IEEE Transactions on Power Systems 35:3810-3824.

Shaheen AM, Elsayed AM, Ginidi AR, El-Sehiemy RA, Elattar E (2022) A heap-based algorithm with deeper exploitative features for optimal allocations of distributed generations with feeder reconfiguration in power distribution networks. Knowledge-Based Systems 241:108269.

Shi Q, Li F, Olama M, Dong J, Xue Y, Starke M, Winstead C, Kuruganti T (2021) Network reconfiguration and distributed energy resource scheduling for improved distribution system resilience. International Journal of Electrical Power & Energy Systems 124:106355.

Suslov KV, Shushpanov I, Buryanina N and Ilyushin P (2020) Flexible Power Distribution Networks: New Opportunities and Applications. In smartgreens 57-64.

Wang Y, Xu Y, He J, Liu CC, Schneider KP, Hong M, Ton DT (2019) Coordinating multiple sources for service restoration to enhance the resilience of distribution systems. IEEE Transactions on Smart Grid 10:5781-5793.

Zafar U, Bayhan S, Sanfilippo A (2020) Home energy management system concepts, configurations, and technologies for the smart grid. IEEE Access 8:119271-119286.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0120

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Forecasting diseases that affect plant leaves and moisture levels in the soil using a data mining approach

Ritu Shree[a] | **Rupal Gupta**[b] | **V. Srikanth**[c]

[a]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science & Application.
[b]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College Of Computing Science And Information Technology
[c]Jain (deemed to be)University, Bangalore, India, Associate Professor, Department of Computer Science and Information Technology.

Abstract The foundation of the global and Indian economies is agriculture. Since agriculture started millions of years ago, many environments, civilizations, and technical developments have fostered and defined the evolution of agricultural technology. In this study, we examine how we may analyze images of plants and soil to better keep tabs on their health, as well as how we can determine how much water each kind of plant needs. Images of the plants and soil are first taken using a digital camera with the necessary resolution. The form and geometric characteristics are extracted from the plant images using the inner distance shape context-based descriptor and geometrical descriptors. The soil images are also used to extract features and color properties. The botanical plant species dictionary is used to identify the plant type using the contour elements of the plant photos. Gradient structured random forest (GS-RF) classification is used to forecast leaf diseases. Principal Component Analysis (PCA) and Hierarchical Gradient Deep Neural Network (HG-DNN) classification techniques are used to determine the causes of a given plant disease based on the characteristics of soil images and plant disease images. The findings are communicated to the growers through text messages sent to their mobile phones on a daily and seasonal basis, along with any potential recommendations for preventative actions.

Keywords: crop production, geometrical descriptors, plant disease, PCA, GS-RF, HG-DNN

1. Introduction

India's economy, as well as the whole world, is based on agriculture. The development of agricultural technology has been supported and characterized by many conditions, civilizations, and technological advancements since it began millions of years ago. Previously, the classification of agriculture was based on increased productivity, the substitution of synthetic fertilizers and insecticides for natural ones, and the allotment of more area. Due to the elimination of environmental reasons in contemporary society, macrobiotic and sustainable agricultural practices have emerged. Through modern prediction methods that give the ideal intensification environment under conditions of reproduction convenience protection, the maximum agricultural yield with the best quality is attained (Bradford 2020). The use of data mining methods and improved information and communication systems allow for the monitoring of the contemporary expansion of agricultural operations (Zhang 2021). The health and development of plants are strongly influenced by elements that are intimately related to plant leaves and soil moisture levels. While plant leaves serve as indications of a plant's water status and general health, sufficient soil moisture is necessary for plants to absorb water and crucial nutrients. For irrigation and plant cultivation to be effective, both components must be monitored and managed (Joswig 2022). One may spot evidence of moisture stress by carefully examining the leaves. Leaves that are wilting, drooping, or turning yellow usually indicate inadequate moisture but leaves that are too wet may indicate over-irrigation or poor drainage. Irrigation decisions may be helped by the fast response provided by visual observation (Abioye et al. 2021). Different methods and tools may be used to assess soil moisture levels directly. Metal probes on portable plant moisture meters allow them to be put into the soil. They provide instantaneous moisture content measurements. Another option is to use tensiometers, which measure soil moisture based on the force needed to draw moisture from the ground. These tools may provide information on soil moisture at various depths (Wang 2019). To correctly detect moisture content, these sensors use a variety of technologies, including capacitance, time domain reflectometry, or gypsum blocks. Soil moisture sensors make it possible to gather data in real-time and may aid with irrigation scheduling (Kamath 2019). With the ability to monitor wide regions, remote sensing may provide important data on the condition of the vegetation and the water supply. One may locate places with insufficient soil moisture by looking at the vegetation indices produced from remote sensing data (Trugman 2019). The soil moisture levels may be inferred indirectly

from weather station data. Evapotranspiration rates, which affect soil moisture, are influenced by variables including rainfall, temperature, and relative humidity. It is possible to determine the moisture stress on plants and modify irrigation methods by taking weather station data into account and predicting water needs (Benos 2021). Effective plant management and irrigation depend on measuring soil moisture and plant leaves. Plant water stress may be immediately detected by looking at the leaves. Precision information on soil moisture content is provided by direct measuring methods such as plant moisture meters, tensiometers, and soil moisture sensors. Broader views on plant health and soil moisture patterns are made possible by remote sensing methods and weather station data. Making educated decisions and using effective water management techniques in plant agriculture are made possible by integrating information from these numerous sources (Millet 2019). Typically, the practice of extracting samples from huge databases is referred to as data mining. The fundamental goal of data mining methods is to extract the most important aspects from databases and arrange that information in a useful arrangement for further use. Depending on the agricultural uses, many data mining methods (Sapes 2019) are used in agriculture. Weather, pollution, and other environmental variables may be predicted with the help of the data mining tools supplied. Soil properties, weed identification, and water-core monitoring are only a few examples of the uses for data mining approaches.

However, the most pressing problem in agriculture is the control of total crop yields via the use of cutting-edge technologies. Therefore, an efficient data mining approach is suggested in this work to forecast the ideal humidity for plants and the onset of various illnesses. This paper's primary objective is to determine the many plant diseases and the factors that contribute to their occurrence. The GS-RF classification technique is used to predict plant diseases. In addition, HG-DNN classification, which relies on probability values between characteristics of soil images and diseased plant images, is used to forecast the causes of plant illnesses.

The rest of the research is structured as follows: The various data mining methods used in agricultural domains are described in Section 2. The suggested agricultural production monitoring system is described in Section 3 employing unique data mining methods. The performance assessment of the suggested strategies is shown in Section 4. The research study is concluded in Section 5, which also includes recommendations for further work on improvement.

2. Related works

Study (Mishra 2021) examines how the Internet of Things (IoT) may be used to efficiently gather data on environmental factors such as temperature, pH, and precipitation using a variety of machine learning methods. Decision Tree to analyze agricultural data and provide recommendations about what may and cannot be planted. To improve agricultural yields on a large scale, a model is developed in this research that utilizes real-time data to make in-field monitoring decisions based on weather analysis and the identification of crop diseases. The software architecture of this platform is flexible enough to accommodate a variety of plant disease models and other precision agriculture applications, making it suitable for usage with a wide range of plant diseases (Khattab 2019). A study (Rigden 2020) showed that utilizing soil moisture data leads to more precise projections of maize output and suggests that explicitly accounting for fluctuations in water availability is necessary for precise estimations of how climate change will affect crop yields. The strengths and weaknesses of various methods and models offered in the current literature are highlighted in the article (Dhaka 2021). In addition to discussing how effective models are, this paper provides a summary of the datasets and performance measures utilized to do so. In the study (Archontoulis 2020), they accomplish the following: (1) explain the process behind the forecasts, (2) assess the accuracy of the model's predictions using data from 10 sites over 4 years, and (3) pinpoint the variables most important for predicting future yields and soil N dynamics. The models, which were made using past, present, and future meteorological data, were made public four weeks after planting. Research (Albergel 2019) evaluated the land data assimilation system LDAS-Monde, created by Météo-France, for its ability to track how the 2018 summer heatwave in Western Europe affected the health of the region's vegetation. Article (Sharma 2020) provided a comprehensive overview of the use of ML in farming. To keep an eye on the quality and production of crops, it is necessary to classify various photos of them. Improve livestock output by utilizing ML models trained on data gathered from collar sensors to make predictions about fertility, diagnose eating problems, and analyze the habits of cattle. Research (Pereira 2020) seeks to update their understanding of how to estimate crop coefficients using data on ground cover and vegetation height. Article (Abd El-Ghany 2020) offered a historical and futuristic perspective on remote sensing methods and their uses, particularly in the control of insect pests and plant diseases. The electromagnetic radiation reflected and emitted by the ground target is measured, recorded, and processed to perform remote sensing. The spectral characteristics of living things are crucial to remote sensing applications. Insect pests and plant diseases may now be detected, forecasted, and managed on a wide variety of fruit orchards and crops with the use of remote sensing. The primary goals of these apps were to gather information useful for making decisions about insect pest control and reducing chemical pesticide contamination. Each kernel's development is influenced by the amount of water and nitrogen (N) available to the plant during the blooming period. The current research documents the changes in the dry weight of maize kernels in response to varying amounts of water and N. The effects of three irrigation regimes and five N treatment rates on weekly maize kernel growth were tested for two consecutive years. Findings add to the literature on

maximizing maize yields in semiarid environments by optimizing the use of nitrogen fertilizer and irrigation water (Hammad 2021).

3. Methodology

Our suggested method for enhancing the monitoring of agricultural yields is divided into three stages: image preprocessing, detection, and communication. Data mining methods are refined and adapted for various procedures at each stage. In the initial step, we prepare photos of the soil and plants. In the second stage, plant illness and its causes are predicted based on soil characteristics. The last step makes use of IT to warn farmers by disseminating data about plant varieties, plant illnesses, and their origins. This section provides a concise summary of these three stages of monitoring systems.

3.1. Image Pre-processing Phase

Initially, high-resolution digital camera input photos of several plant types and soil are gathered. To continue processing, the collected pictures are sent to an image processing unit, either over a wired or wireless network. The gathered pictures then undergo pre-processing, whereby noise and other disturbances are removed while the characteristics used in the prediction process are also improved. Images are cleaned up by removing noise and then transformed from RGB to their color space equivalents. After the enhancement process is complete, a Region of Interest (ROI) based segmentation technique is utilized for the resulting images. The goal of ROI segmentation is to divide a given image into many distinct sections or categories. It is denoted in equation (1):

$$Q_{j\geq i} = \frac{\psi \frac{|Q_j||Q_i|}{|Q_j|+|Q_i|}.\|w_j - w_i\|^2 + (1-\psi).f\omega}{1(\partial(Q_jQ_i))} \quad (1)$$

In equation (1), $\partial((Q_jQ_i))$ refers to the two adjacent regions of the area, $(\partial(Q_jQ_i))$is the length of the two regions, ω refers to the weight of (0, 1), f denotes the boundary strength and w_j, w_i represent the spectral values of two regions. The polygonal leaf model is used to recover the contour elements of plant pictures. Then, using a botanical plant species dictionary containing descriptions of various plant species, semantic representation is given to these contour data to identify the kinds of plants.

3.2. Plant Disease Prediction

The GS-RF makes its disease prediction based on the plant species and characteristics. When GS-RF is used in a training dataset, it may detect plant diseases that were previously undetected.

3.2.1. Gradient structured random forest

GS-RF combines the benefits of random forests with gradient boosting. To improve forecast accuracy and interpretability, it was added as an addition to the conventional random forest method. The goal of GS-RF is to identify the interactions and cumulative effects of features in a dataset.

Incorporating gradient boosting into the random forest framework is the main concept underlying GS-RF. A well-liked approach called gradient boosting creates a series of weak learners by fitting the residuals of the prior models repeatedly. By steadily improving forecasts, it focuses on reducing the loss function. Random forests, on the other hand, train several decision trees individually and combine their predictions using either majority voting or averaging.

A random forest model is trained on the training set after a subset of features is randomly chosen at each iteration. The residuals are then calculated using the predictions of these trees, and in the succeeding iteration, a fresh set of trees are used to fit the residuals. Until a certain number of iterations or a stopping condition is achieved, this process keeps going.

The boosting and random forest methods are combined in GS-RF to make use of each method's advantages. The random forest component captures the non-linear correlations and interactions between the features, while the boosting component concentrates on improving the model by fitting the residuals.

Benefits for interpretability come from GS-RF. It enables feature significance analysis, allowing the discovery of factors that significantly influence the model's predictions. Based on the average influence of each feature over the ensemble of trees, feature significance may be calculated. The underlying links between characteristics and the target variable may then be better understood, which makes GS-RF an important tool for data analysis and interpretation.

A random forest is a kind of ensemble learning machine that generates hypotheses by combining the major population of forecasts from several various base models. By reflecting important qualities, the RF is simple and efficiently specifies data from large datasets. The pixels of the leaf are constructed to generate a subdivision of R^D at each level, and R^D is the foundation of the corresponding tree. Each tree is created uniquely, with each node matching a rectangular subset of

R^D. One expanded tree leaf is picked at each step of construction. The dataset is randomly split into two pieces that serve different purposes in the construction of each tree. The structural elements that affect the shape of the tree are taken into consideration to estimate the split dimensions and split characteristics. The assessment nodes are used to fit the estimators in each tree leaf. By allocating each point to the structure or estimate component, each tree experiences a random division of the perceptions regarding data.

The testing examples for the $h_1(x), h_2(x), . . , h_k(x)$ classifiers were randomized from the distribution of the random vectorsY, X. The revenue feature is written in equation (2-5)

$$Mg(X,Y) = av_k I(h_k(X) = Y) - \frac{max}{j \neq Y} av_k I(h_k(X) = j) \qquad (2)$$

where the measured value is $I(.)$. The source of the mistake is

$$PE^* = P_{X,Y}(mg(X,Y) < 0) \qquad (3)$$

The probability over the XY dimension, indicated by where X, Y space $In\ RF,\ h_k(X) = h(X,\ \Theta_k)$

The margin feature for an RF is

$$mr(X,Y) = P_\Theta(h(X,\Theta) = Y) - \frac{max}{j \neq Y} P_\Theta(h(X,\Theta)\} \qquad (4)$$

Moreover, the set of classifiers $\{h(X,\Theta)\}$ has a value of

$$S = E_{X,Y} mr(X,Y)\ J4.5 \qquad (5)$$

The GS-RF training procedure resolves this optimization issue. The training algorithm 1 is broken down into the following stages:

Algorithm 1: Training Algorithm

1. Set parameters *C and* C^*
2. To create the first classifier, apply inductive GS-RF to the training set of data
3. Set the number of samples with a positive label depending on the rule
4. All unlabeled samples' decision function values should be calculated
5. Indicate as affirmative samples those with the greatest decision function
6. Set temporary effect factor $Ctmp*$
7. Overall samples retrain the existing
8. One set of different-labeled unlabeled samples' labels are switched using a specific rule to produce the value of the objective function by utilizing (5).
9. Repeat the procedure until there are no more pairs of samples that meet the switching requirement.
10. Increase *the* C^*_{tmp} value before proceeding to step 7.
11. $If\ (Ctmp* > C*)$
12. Stop
13. Attain the output

In the testing phase, the data is compared to the training data, which successfully predicts the type of plant disease. Additionally, using the PCA approach to extract the characteristics from the soil images, the necessary water content level for the particular kind of plant is forecasted.

3.3. Principal Component Analysis (PCA)

The PCA is used to analyze the measurements of several variables in a group of individuals. The PCA operates in a space with fewer dimensions to facilitate comprehension and reduce the number of latent variables (principal components) from the original variables. By resolving an optimization issue specified in equation (6), the primary components are achieved.

$$\max\ (Y = x^S Ux), \qquad (6)$$

Subject to $x^S Ux = 1$,

Where V = (1/m) $W^S W$ is the data matrix's sample covariance matrix, and x is of unitary norm according to the requirement $X^S X = 1$, where W is a matrix with n elements and o variables. The greatest eigenvalue λ of U, or $Ux = \lambda w$, is the answer to the optimization problem specified in equation (2) after performing a dimension reduction to r components ($r < o$). The eigenvalues of U are arranged according to $\lambda_1 > \cdots > \lambda_r$ in decreasing order. The first primary component w1 is the unitary norm's eigenvector, or _1, which is connected to the greatest eigenvalue of U, that is, λ_1. when x_1 and x_2 are

orthogonal, as in $x_1^S x_2 = 0$, the second principal component, w2, λ_2 is the eigenvector associated with them, and so on for the other components.

It is possible to rotate the retrieved components to enhance the understanding. One of the most well-liked rotation techniques is Varimax, which is used for component loadings. see the varimax approach in detail. The varimax orthogonal rotation attempts to maximize the variation of the squared loadings in each component, resulting in big loadings for a select few variables and modest loadings for the other variables. Consequently, a subset of factors that influence each component may be found. The observed variables should therefore only have one loading that is significant in absolute value, ensuring that the variables are primarily connected to one component.

3.4. Hierarchical Gradient Deep Neural Network (HG-DNN)

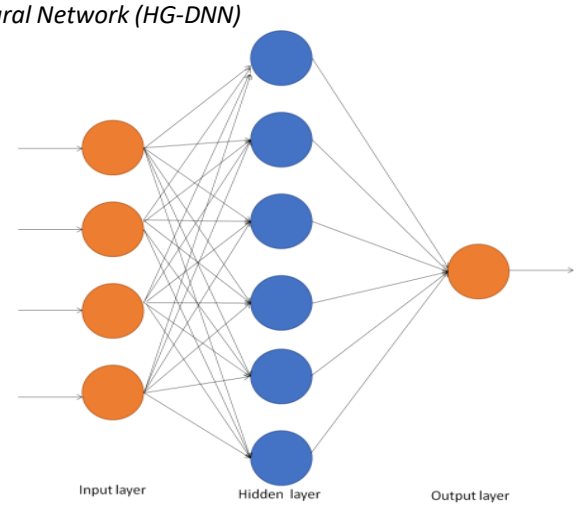


Figure 1 Structure of deep neural networks.

DNN can extract latent learning from huge datasets, as seen in Figure. 1. With the help of the Python library Keras, users may quickly and flexibly build and test deep learning models. Wrapping Theano and Tensor Flow, two of the most potent numerical computing libraries presently accessible makes it simple to describe and create neural system models with only a few lines of code at most. An epoch is a whole loop over a particular dataset. By adjusting the number of epochs through backpropagation, errors may be significantly reduced. A confusion matrix is one tool for evaluating the effectiveness of a categorization system. The information about the existing classes and the predicted classes is presented in a confusion matrix in the form of rows and columns. Different components make up the HG-DNN's many network levels. At each layer, information is received at the top, shown at the bottom, and kept hidden at all subsequent levels. The model is run, and the results are compiled. However, backpropagation is used to complete the HG-DNN model training. Or the loss function measures the discrepancy between outcomes that were anticipated and those that were achieved. Throughout the training phase, changes are made to the weight coefficient W and bias B.

$$I(V, a, y, x) = \frac{1}{2}\|x^K - x\|_2^2 = \frac{1}{2}\sigma(V^K x^{K-1} + a^K) - x\|_2^2 \qquad (7)$$

In equation (7), The buried layer's output is denoted by x^K.

3.5. Communication Phase

In this communication stage, farmers get letters or emails outlining the types of plant diseases that are likely to appear on their farms, as well as the soil conditions that are likely to cause those diseases. Farmers' output may be boosted by the suggested monitoring system to better water management and reduced disease incidence. Algorithm 2 shows proposed steps.

Algorithm 2: Proposed Algorithm

Input: Plant or Soil images $I_1, \dots, I_N$

Output: Type of plant & disease, and its causes.

1. Gather Images
2. Imagine noise reduction
3. Transform a picture into a representation in color space

//Image segmentation

4. Select the image's seed pixel
5. Set standards to help the area develop

6. If a pixel is 8-connected to at least one other pixel in the area, include it in the region
7. Test the percentage of each pixel
8. Label all the region
9. If two areas have the same label, combine them
10. Extrapolate the geometric details and form from the plant images
11. Apply semantic contour representation
12. Determine the plant's species
//Prediction of plant disease
13. Define the optimization problem for the GS-RF classifier
14. Get training materials
15. Execute the GS-RF procedure
16. Attain the result
17. Compare the training and testing data.
18. Determine the disease's form
//Prediction of the causes of plant diseases
19. Identify the soil image textures and colors.
20. Using PCA distribution, choose the topic mix for the description of the plant species
21. Select the subject to create each word in the description
22. Then create words based on a multinomial distribution for the subject
23. Calculate the likelihood that soil properties and words will occur
24. HG-DNN classification should be done using the calculated probabilities
25. Utilize equation (7) to determine the output function
26. Determine the root causes of the specific plant disease
27. Text messages may be used to inform farmers of the forecast findings
28. End

4. Results

In this section, we compare the accuracy and precision of current and proposed methods for predicting plant diseases and their underlying causes. Causes of plant illnesses are predicted and compared to those predicted using the principal component analysis and the similarity measure methods. The suggested methods of Convolutional Neural Network (CNN), K-Nearest Neighbor (KNN), and GS-RF are compared with one another and their respective plant disease prediction outcomes.

Precision, which is well-defined as the ratio of properly categorized cases to all occurrences of predictively positive data, is one of the most important criteria for accuracy. Equation (8) is used to compute the precision.

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (8)$$

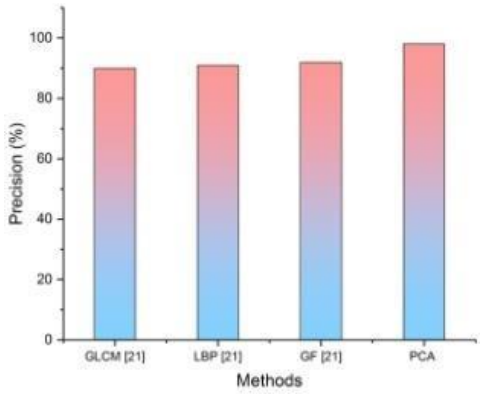


Figure 2 Comparison of Precision in Plant Disease Causes Prediction.

The comparison of the Gray Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), Gabor filters (GF), and PCA techniques in terms of accuracy for predicting the causes of plant diseases are shown in Figure. 2. The inspection of the aforementioned graph demonstrates that as the number of features rises, so does the accuracy of the PCA approach.

Methods are taken for the x-axis, and precision values are obtained for the y-axis. It demonstrates that the PCA prediction approach has improved accuracy.

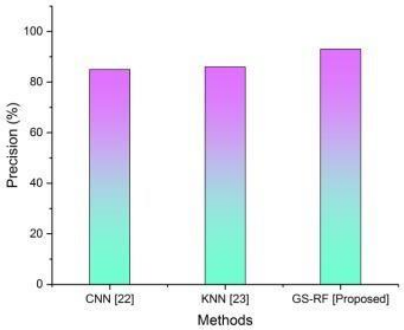


Figure 3 A comparison of Precision for predicting plant leaf disease.

Figure 3 compares the accuracy of the CNN, KNN, and GS-RF approaches for predicting plant leaf disease. The examination of the graph above demonstrates that as the number of features rises, the accuracy of the GS-RF approach also increases. Methods are taken for the x-axis, and precision values are obtained for the y-axis. It demonstrates that the GS-RF prediction approach now has higher accuracy.

The system's accuracy is measured by the proportion of samples for which the proposed strategy correctly anticipated outcomes. Equation (9) is used to determine accuracy.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (9)$$

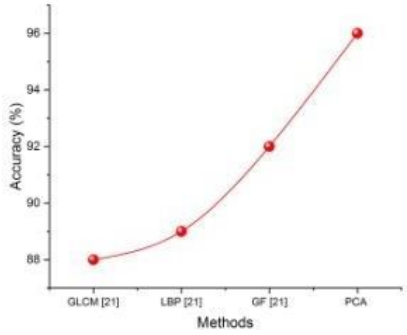


Figure 4 Comparison of Accuracy in Plant Disease Causes Prediction.

Figure 4 compares the accuracy of the GLCM, LBP, GF, and PCA approaches for predicting the causes of plant disease. The examination of the graph above demonstrates that as the number of features rises, the accuracy of the PCA approach also increases. Methods are listed on the x-axis, while accuracy (%) is shown on the y-axis. It demonstrates that the PCA prediction method's accuracy has grown.

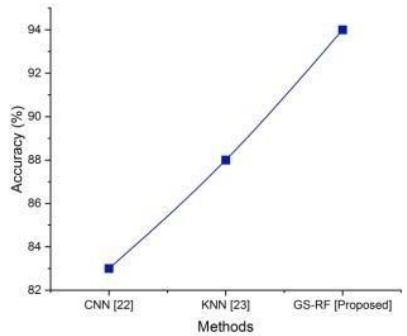


Figure 5 A comparison of Accuracy for predicting plant leaf disease.

Figure 5. Analyzes the accuracy of plant leaf diseases using CNN, KNN, and GS-RF methods in comparison. The examination of the graph above demonstrates that as the number of characteristics rises, the accuracy of the GS-RF approach also increases. Methods are listed on the x-axis, while accuracy (%) is shown on the y-axis. It demonstrates that the GS-RF prediction method's accuracy has grown.

5. Conclusion

This paper suggests a new data mining approach for use in agricultural field monitoring systems. The suggested data mining method is associated with the forecasting of plant diseases and their causes. The suggested approach uses contour analysis of extracted characteristics from leaf images in conjunction with a botanical plant species dictionary to determine the kind of plant. The GS-RF classification then makes a prediction of the plant disease based on the form and texture attributes of the plant images. The PCA method is used to create a model between the color and texture properties of soil photographs and images of infected plants. In addition, HG-DNN categorization is used to predict the origins of plant diseases. The farmers then get text messages on their mobile devices with the forecasted information. As a consequence, the agricultural output monitoring system is improved by the use of our proposed monitoring approach to better control irrigation and diseases. The experimental results validate the usefulness of the suggested monitoring system. Improving agricultural yields will need a further examination of the prediction method for tracking plant development.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Abd El-Ghany NM, Abd El-Aziz SE, Marei SS (2020) A review: application of remote sensing as a promising strategy for insect pests and diseases management. Environmental Science and Pollution Research 27:33503-15.

Abioye EA, Abidin MS, Mahmud MS, Buyamin S, AbdRahman MK, Otuoze AO, Ramli MS, Ijike OD (2021) IoT-based monitoring and data-driven modelling of drip irrigation system for mustard leaf cultivation experiment. Information Processing in Agriculture 8:270-83.

Albergel C, Dutra E, Bonan B, Zheng Y, Munier S, Balsamo G, De Rosnay P, Muñoz-Sabater J, Calvet JC (2019) Monitoring and forecasting the impact of the 2018 summer heatwave on vegetation. Remote Sensing 11:520.

Archontoulis SV, Castellano MJ, Licht MA, Nichols V, Baum M, Huber I, Martinez-Feria R, Puntel L, Ordóñez RA, Iqbal J, Wright EE (2020) Predicting crop yields and soil-plant nitrogen dynamics in the US Corn Belt. Crop Science 60:721-38.

Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D (2021) Machine learning in agriculture: A comprehensive updated review. Sensors 21:3758.

Bradford JB, Schlaepfer DR, Lauenroth WK, Palmquist KA (2020) Robust ecological drought projections for drylands in the 21st century. Global change biology 26:3906-19.

Dhaka VS, Meena SV, Rani G, Sinwar D, Ijaz MF, Woźniak M (2021) A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. Sensors 21:4749.

Hammad HM, Abbas F, Ahmad A, Bakhat HF, Farhad W, Wilkerson CJ, Fahad S, Hoogenboom G (2020) Predicting kernel growth of maize under controlled water and nitrogen applications. International Journal of Plant Production 14:609-20.

Joswig JS, Wirth C, Schuman MC, Kattge J, Reu B, Wright IJ, Sippel SD, Rüger N, Richter R, Schaepman ME, Van Bodegom PM (2022) Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. Nature ecology & evolution 6:36-50.

Kamath R, Balachandra M, Prabhu S (2019) Raspberry Pi as Visual Sensor Nodes in Precision Agriculture: A Study. Ieee Access 7:45110-22.

Kaur N (2021) Plant leaf disease detection using ensemble classification and feature extraction. Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12:2339-52.

Khattab A, Habib SE, Ismail H, Zayan S, Fahmy Y, Khairy MM (2019) An IoT-based cognitive monitoring system for early plant disease forecast. Computers and Electronics in Agriculture 166:105028.

Millet EJ, Kruijer W, Coupel-Ledru A, Alvarez Prado S, Cabrera-Bosquet L, Lacube S, Charcosset A, Welcker C, van Eeuwijk F, Tardieu F (2019) Genomic prediction of maize yield across European environmental conditions. Nature genetics 51:952-6.

Mishra D, Deepa D (2021) Automation and integration of growth monitoring in plants (with disease prediction) and crop prediction. Materials Today: Proceedings 43:3922-7.

Pereira LS, Paredes P, Melton F, Johnson L, Wang T, López-Urrea R, Cancela JJ, Allen RG (2020) Prediction of crop coefficients from fraction of ground cover and height. Background and validation using ground and remote sensing data. Agricultural Water Management. 241:106197.

Rigden AJ, Mueller ND, Holbrook NM, Pillai N, Huybers P (2020) Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting US maize yields. Nature Food 1:127-33.

Sapes G, Roskilly B, Dobrowski S, Maneta M, Anderegg WR, Martinez-Vilalta J, Sala A (2019) Plant water content integrates hydraulics and carbon depletion to predict drought-induced seedling mortality. Tree physiology 39:1300-12.

Sharma A, Jain A, Gupta P, Chowdary V (2020) Machine learning applications for precision agriculture: A comprehensive review. IEEE Access 9:4843-73.

Trugman AT, Anderegg LD, Wolfe BT, Birami B, Ruehr NK, Detto M, Bartlett MK, Anderegg WR (2019) Climate and plant trait strategies determine tree carbon allocation to leaves and mediate future forest productivity. Global change biology 25:3395-405.

Wang C, Fu B, Zhang L, Xu Z (2019) Soil moisture–plant interactions: an ecohydrological review. Journal of Soils and Sediments 19:1-9.

Zhang WG, Li HR, Wu CZ, Li YQ, Liu ZQ, Liu HL (2021) Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. Underground Space 6:353-63.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0121

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

AI-based vehicular traffic flow forecast for advanced transportation systems

Vipin Panwar[a] | **Gulista Khan**[b] | **D. Ganesh**[c]

[a]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Computer Science and Engineering.
[b]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College of Computing Science and Information Technology.
[c]Jain (deemed to be) University, Bangalore, India, Professor, Department of Computer Science and Information Technology.

Abstract For traffic control systems, the intricate nature and quantity of vehicular traffic have recently risen significantly. For transportation systems to be optimized and overall efficiency to be increased, accurate prediction of traffic is essential. To anticipate vehicular traffic flow, we introduce a unique reptile-ant optimized bidirectional gated recurrent unit (RAO-BiGRU) method in this study. Ant colony optimization (ACO) and reptile search optimization (RSO) methods are combined in the RAO. The Bi-GRU model's performance is improved by using the RAO to better choose the input features. Extensive experiments are run utilizing the traffic data database to assess the efficacy of the suggested approach. The outcomes show that in terms of predicting accuracy and computational effectiveness, the suggested RAO-BiGRU strategy performs better than the conventional forecasting strategies.

Keywords: traffic control systems, vehicular traffic, transportations, RAO-BiGRU

1. Introduction

The ability of an algorithm to perform tasks that often demand intelligence from humans, such as recognizing phrases, comprehension of spoken words, and choice-making, is referred to as artificially intelligent. With the help of artificial intelligence, robotics are able to sense things around them as well as communicate with them, make decisions, and accomplish challenging jobs. Using the "machine learning" subfield of artificial intelligence enables robotics to acquire knowledge through information and improve continuously (Xu et al 2021). Machines can be programmed to perform specific duties in automation, such as grabbing, recognizing objects, and route mapping. The technique of deep learning is a sort of algorithm for learning that uses artificial brain networks that assist machines in comprehending enormous amounts of information. Deep learning proves especially helpful in robots for situations such as identifying objects, word and picture acknowledgment, and especially processing natural languages (Cio 2019). Collectively, these advancements make it possible to create robotics that are capable of carrying out a variety of activities, from straightforward pick-and-place procedures to intricate manipulations and navigating situations. Technologies could change as a result of the use of artificial intelligence, machine learning, and advanced learning, since machines could grow increasingly smart, independent, and efficient in a variety of situations. Robots constitute a growing discipline, and artificial intelligence, machine learning, and advanced learning are projected to be important factors in determining how robots will develop in the years to come (Soori 2023). Artificial intelligence is utilized in sophisticated machinery to build machines that are capable of seeing, logic, and behaving independently in challenging circumstances. Robots can gain insight through their mistakes and gradually enhance their abilities thanks to automated training. Deep learning is applied to tackle particular challenges that are challenging for typical methods of machine learning to address, such as spoken and visual identification. Modern robots have been created to carry out intricate duties that had been believed might be unfeasible by integrating different innovations (Rothemund 2021). Regarding the examination and customization of sophisticated robots, their link is encompassing. Machines can recognize things, negotiate challenging settings, and occasionally generate judgments using current information thanks to intelligence in automation. Machines may be taught to acquire knowledge through experiences and adapting to new circumstances using machine learning techniques. Robotics can execute challenging assignments with deep neural networks that might have been unfeasible utilizing traditional coding techniques. By promoting the creation of smart robots that are capable of carrying out difficult tasks at an elevated level of precision and effectiveness, machine learning and deep learning are revolutionizing the area of cutting-edge robots (Popli 2023).

The aim is to contribute to the development of advanced transportation systems that are more efficient, sustainable and responsive to the needs of both transportation authorities.

The remainder of the paper is divided into subsequent parts. Part 2 contains the related work, part 3 includes the proposed method explained, part 4 presents the results and part 5 discusses the results.

2. Related Work

Tak (2021) offered research to obtain traffic information from a surveillance device placed at a junction, enhancing road management systems. As a result, the technique suggested in this research demonstrates the viability of gathering thorough roadway information utilizing cameras mounted at a junction. The key aspect of this work is the method of merging HD maps with artificial intelligence algorithms, with a great probability to enhance present road tracking technologies. A study (Rani and Sharma 2023) shows that an automated training network built around tree-based decision trees, random forest, extra trees, and XGBoost methods is offered as an automated transit system to address the roadway flows of IOVs in smart town situations. According to combination teaching and featured averages, simulation findings show that the suggested approach can offer excellent detection precision and cheap computing expenses.

The research (Liu 2023) uses automobiles with everywhere technology powered by intelligent machines. The outcomes from different analyses of automobile detection of images using 5G networks are now the following: the proportion of enhanced transit in 5G technology is 83.2%; the power source proportion of automobile picture tracking on travel is 87.2%; the proportion of the advancement of communication between vehicles via V2X is 84.36%; the proportion of increased convenience while driving is 81.15%; and the proportion of reduced congestion in the roadways is 90.84%. To enable worldwide data collection with immediate information, a collaborative study (Hui 2021) first constructed automated vehicle networks depending upon the improved advancements in audio video and heterogeneous automotive communication methods.

In a review of literature information acquired through the Worldwide Web of Sciences main gathering, this investigation (Chen 2023) employs bibliographic approaches and networking analytic measurements to obtain an understanding of the investigation's state, development procedure, prospects, and challenges associated with artificial intelligence traffic flow prediction studies. This inquiry may assist academics in better understanding the current status of the field and prospects for artificial intelligence-based traffic flow prediction. The majority of the artificial intelligence technologies presented in this article (Englund 2021) require much information. Info quantity and availability are vital, yet it is always necessary to combine an array of info gathered from different sources and information categories. To develop a comprehensive picture of the situation on a city scale, information gathered from suppliers of services and government transportation officials, including client information, is essential.

With the goal of preventing crashes, the study (Olugbade 2022) examines the main problems, potential fixes, and deployment of automated learning and robotics in road transportation networks. The study concludes by summarizing the difficulties in applying machine learning to public transportation networks, highlighting recent developments in research, pointing out open issues, and emphasizing key research findings. The findings can be used as a resource for arranging and overseeing the highway transportation network. To aid in actual time choices, the present study (Thamizhazhagan 2022) developed a successful artificial intelligence-based parallel autoencoder-traffic flow prediction framework for connected and autonomous electric vehicles. Both the independent processes of feature development and traffic flow prediction make up the framework as it is now represented. The method of defining features includes a few steps, including feature development, choosing components, and discovering features. The data that come in instead were screened for anomalous points of data utilizing a standard vector data description approach, and the real information was smoothed.

The most important position-based single-cast networking techniques for vehicle-to-vehicle telecommunications in city environments are briefly reviewed in this research (Patankar 2022). They offer the functionality they need to function so that data may be sent among vehicle locations. They list respective advantages and disadvantages. It presents a simulation-based analysis of both dynamic and static circuit selection routing methods. In this research (Mohammadi 2020), automated traffic prediction engines are suggested, and the optimization process is carried out using AI-based methodologies. This approach is put to the proof in an actual-life scenario, and thus, it has been shown to be effective at reliably estimating congestion. This forecasting of the adaptive neuro-fuzzy inference system (ANFIS) algorithm can be put into practice and utilized to forecast traffic patterns in place of cameras.

3. Methodology

3.1. Dataset

The Caltrans Performance Evaluation and Monitoring System's actual time gathering of California road traffic information, known as PEMS, was employed in the test in this research. To create the last piece of information every five minutes, PEMS information was obtained every thirty seconds. The PEMS information utilized in the trial was acquired through a monitoring facility on Interstate 5 between May 1 and June 30, 2014, comprising an average of 61 consecutive days including 17568 recordings (Zhang 2021).

3.2. Preprocessing using min-max normalization

Information about traffic movement constitutes an irregular randomized pattern. To assign the information into the range [0, 1], a combination of minimum and maximum normalization was used during the experiments. where Y represents the normalized information, w is the initial information, and *wmin* and *wmax* are the minimum and maximum values of *w*, respectively, which is how normalization is described in equation (1).

$$y = \frac{w - w_{min}}{w_{max} - w_{min}} \qquad (1)$$

3.3. Ant Colony Optimization (ACO)

Ant colony optimization (ACO) is a naturally influenced MA that imitates the ant's foraging behavior. Since ACO permits simultaneous execution without creating a procedure dependence and provides input regarding the actions of ants within the area for searches, it seems more rational compared to other metaheuristic algorithms. Ants are able to identify the quickest path from their community and the source of foods; therefore, they are not blind while seeking meals. Ants leave behind an organic substance called a pheromone that travels alongside their path as they move. It provides a channel to communicate within ants that shows the quickest route to an energy source. Ants locate nourishment by detecting the pheromones left behind by different ants who have already traveled a particular route, increasing the likelihood that more ants will follow in their footsteps.

ACO bases its probability judgments on pheromone trails and heuristic data. Each time they move along a trail, ants can alter their pheromone concentration at every junction. A characteristic has a greater chance to become a component of the fastest route when more ants pass over it and more pheromones are deposited there as a result. Most ants are going down the route with the greatest amount of pheromones, and it is also the most in the brief way. Ants are distributed randomly throughout a set of features with a certain maximum number of generations *S*, and the pheromone amount 0=1 is initialized at each of the *M* characteristics. The transitional frequency *SOlj(g)* given by the lth ant with its jth characteristic at each iteration h is displayed here.

$$SO_j^l(h) = \begin{cases} \frac{[\tau_j(h)]^\alpha [\eta_j]^\beta}{\sum_{i \epsilon i_j^l} [\tau_j(h)]^\alpha [\eta_j]^\beta} & if\ i \epsilon i_j^l \\ 0, & otherwise \end{cases} \qquad (2)$$

where *ilj* represents a list of prospective neighbors of i^{th} characteristics in which the l^{th} ant fails to examine. Nonnegative variables, along with correspondingly, specify the proportional significance of pheromone levels (i) and heuristic data (j) for the ants' motions.

A function of fitness (FF) is then used to measure a fresh set of chosen traits once the ant decides on the next item on its journey. If a rise in fit values fails to occur following the addition of any fresh characteristic, the motion of the l^{th} ant is terminated. The quantity of pheromones levels at the subsequent generation (h + 1) at the j^{th} characteristic is modified as follows if the halting requirements have not been met.

$$\tau_j(h+1) = (1-o)\tau_j(h) + \sum_{l=1}^{M} \Delta\tau_j^l(h) \qquad (3)$$

$$\Delta\tau_j^l(h) = \begin{cases} EE(T^l(h))/|T^l(h)|, & if\ j \epsilon T^l(h) \\ 0, & otherwise \end{cases} \qquad (4)$$

where M remains the number of ants, o is the pheromone decay ratio (0 ≤ o < 1), |*Tl(h)*| displays the number of picked characteristics, and *lj* indicates the pheromones dropped by a lth ant provided the j^{th} characteristic is within the ant's simplest route; alternatively, this is 0.

As soon as g hits the predetermined maximum *S*, the conditions for stopping are satisfied. An optimum final solution can be chosen based on an array of attributes having the greatest pheromone levels and lowest fitness score.

3.4. Reptile Search Algorithm (RSA)

To mimic the surrounding and feeding behavior of crocodiles, the animal's search method (RSA) was suggested. It represents a gradient-free technique that begins by producing the following randomized remedies:

$$w_{j,i} = rand_{\epsilon[0,1]} \times (VA_i - KA_i) + KA_i\ for\ j\ \in \{1, \dots, M\} and\ i\ \in \{1, \dots, N\} \qquad (5)$$

where random [0,1] is a number that is randomly dispersed evenly in the frequency band [0, 1] (0, 1), *wj,i* is the j^{th} answer that uses the i^{th} input feature for a total of *M* solutions consisting of *N* characteristics, and the i^{th} feature comprises higher VA_i and lowers KA_i limits.

RSA may be defined in terms based on two rules: exploration and exploitation, similar to each of the nature-inspired metaheuristic algorithms. The crocodile's ability to maneuver despite surrounding its victim helps to explain those ideas. To gain advantages from crocodile instinctive behavior, RSA's overall repetitions are split across four parts. RSA completes the investigation in the initial two phases using an encircled behavior that includes both elevated and abdominal moving motions. These reptiles start to circle the area that surrounds it, enabling a more thorough examination of the answer area. A mathematical framework can describe such behavior:

$$w_{j,i}(h+1) = \begin{cases} [-m_{j,i}(h) \cdot \gamma \cdot Best_i(g)] - [rand_{\in[1,M]} \cdot Q_{j,i}(h)], & h \leq \frac{S}{4} \\ FT(h) \cdot Best_i(g) \cdot w_{(rand_{\in[1,M]},i)}, & h \leq \frac{2S}{4} and\ h > \frac{S}{4} \end{cases} \quad (6)$$

While *Bestj(g)* denotes all optimal options for a j^{th} characteristic, *mi,j* is the seeking operation for the h^{th} features in the i^{th} solution (determined as in Equation (6)), and parameter denotes the exploring quality over a duration of repetitions, which is chosen to be 0.1. Reducing functionality the equation, and [1, *M]* is an integer ranging from 1 and *M* utilized for selecting a single of the potential solutions, and progressive feel refers to the likelihood proportion decreasing via 2 to 2 over repetitions. *Qj,i* serves to narrow the area of searching and is determined using Formula.

$$m_{j,i} = Best_i(g) \times O_{j,i} \quad (7)$$

Oj,i is the ratio of the i^{th} number in the ideal outcome to the equivalent in the present result, and it may be computed as follows:

$$O_{j,i} = \theta + \frac{w_{j,i} - N(w_j)}{Best_i(g) \times ((VA_i - KA_i) + \epsilon)} \quad (8)$$

Thus, it represents a modest baseline number and is a critical factor that determines the exploring achievement, and *N(wj)* relates to typical answers that are described as follows:

$$N(w_j) = \frac{1}{m} \sum_{i=1}^{m} w_{j,i} \quad (9)$$

$$Q_{j,i} = \frac{Best_i(g) - w_{(rand_{\in[1,M]},i)}}{Best_i(g) + \epsilon} \quad (10)$$

$$FT(h) = 2 \times rand_{\in[-1,1]} \times (1 - \frac{1}{S}) \quad (11)$$

where rand [1, 1] is an integer chosen at random among (1, 1). The amount equal to 2 serves as a multiplication for offering values for correlation in the range of [0, 2].

The use of force (hunting) to seek out features for the best solution is implemented by RSA in the last two phases employing both hunt and collaboration as well as coordination. Applying the subsequent formula, the answer may change its numerical value by exploiting others:

$$w_{j,i}(h+1) = \begin{cases} rand_{\in[-1,1]} \cdot Best_i(g) \cdot O_{j,i}(h), & h \leq \frac{3S}{4} and\ h > \frac{2S}{4} \\ [\epsilon \cdot Best_i(g) \cdot m_{j,i}(h)] - [rand_{\in[-1,1]} \cdot Q_{j,i}(h)], & h \leq S\ and\ h > \frac{3S}{4} \end{cases} \quad (12)$$

The method finishes after *S* iterations, and the potential answer having the lowest fitter rating is chosen as the optimum final solution (OFS). Every iteration's possible solutions are evaluated for validity utilizing the specified FF. Algorithm-2 depicts the RSO approach.

Algorithm-1: Reptile Ant Optimization

Step-1: Upload the data
Step-2: Set the Automatic Control Operator variables
Step 3: Set up the random access protocol parameters
Step-4: Initialize the pooled variables *M, N*, and *S*
Step-5: perform for *h* = 1 to *S*
Step 6: Assume the initial version
Step 7: Use equations (2) to (4) to carry out a single round of AOC.
Step-8: Else
Step-9: If the toggle mark is 1

Step 10: Conduct one cycle of a different method that was not used during the prior cycle of AOC.
Step 11: Switching flags 0
Step-12: Else
Step-13: Use identical technique you did for cycle ACO
Step 14: End if
Step-15: End if
Step 16: Assessment of FF
Step 17: Updating potential answers
Step 18: if($FFnew$) < $min(FFold)$
Step-19: switching flags 1
Step-20: End if
Step 21: End for
Step 22: Include a threshold of 0.5 for the suggested solutions containing the lowest FF to obtain the optimum final solution.

3.5. Bidirectional Gated Recurrent Unit (Bi-GRU)

It is suggested that RNNs tackle consecutive time-series data issues using the benefit of storing past data. However, it only has one architecture and uses a parameter solving mechanism centered around propagation backward over. The passage of time, gradients disappearing and inflating restrictions are useful in long-term dependence. The inner cell makeup of RNN, LSTM, and GRU is demonstrated by how they derive from RNN to get around the constraint of the RNN algorithm. Upon the foundation of RNN, storage cells and 3 gates—forget gates, intake entrance, then release gate—that regulate the flow of data between cells in LSTM and establish the entry, preservation, and final result of data, accordingly, are added in the LSTM algorithm.

The updating and reset gates are two additional unique gates used by *ys* to lessen gradient dispersion, which allows for memory retention and reduced computational cost. The updating gate substitutes for the input that forget gates in the LSTM and determines how much of the prior data is retained for the present anticipated, as shown in Eq. (13):

$$y_S = \sigma(X_y \cdot [g_{S-1}, Efb_S] + a_y \qquad (13)$$

It produces numbers from 0 and 1 and reflects the sigmoid activation function. Represents the given *input* matrices at timestep *S* and represents the hidden state at timestep *S*-1. The modification to the gate's matrix of weights is called *ys*, while its bias matrix is called bz. The amount of previous information that ought to be disregarded is controlled by the gate that is reset *ay*, which is defined as:

$$q_S = \sigma(X_q \cdot [g_{S-1}, Efb_S] + a_q \qquad (14)$$

where is the biassed matrices of the reset gates and represents the reset gates' weighted matrices.
The potential concealed condition can therefore be described as

$$\tilde{g}_S = tang\ (X_g \cdot [g_{S-1} \odot q_S, Efb_S] + a_g) \qquad (15)$$

where Tanh is the triggering process for Tanh and are the weighted matrices and biase matrices of the fresh memory unit nation, where dot multiplying is the cell's fresh condition.
The result of operation (*g*) consequently indicates that there is a linear interpolation between ht and *g*t-1:

$$g_S = (1 - y_S) \odot g_{S-1} + y_S \odot \tilde{g}_S \qquad (16)$$

Implicit characteristics and complicated variations inside data sequences must be extracted to create an effective prediction framework for daily output. The GRU, however, is only capable of extracting data pertaining to the direction that is forward, disregarding the helpful details included in the reversed historical information. To efficiently extract the information among the output variation and variables used for input, the Bi-GRU method is presented. Bi-GRU, a combination of both the forward GRU and the reversed GRU, may take connections in both forward and reverse orientations into account. Both the forward-facing and backward versions of the GRU are able to collect both past and future knowledge based on the provided information.

The Bi-GRU may be stated as follows: the forward GRU's and the reverse GRU's corresponding concealed states are shown in Figure 1.

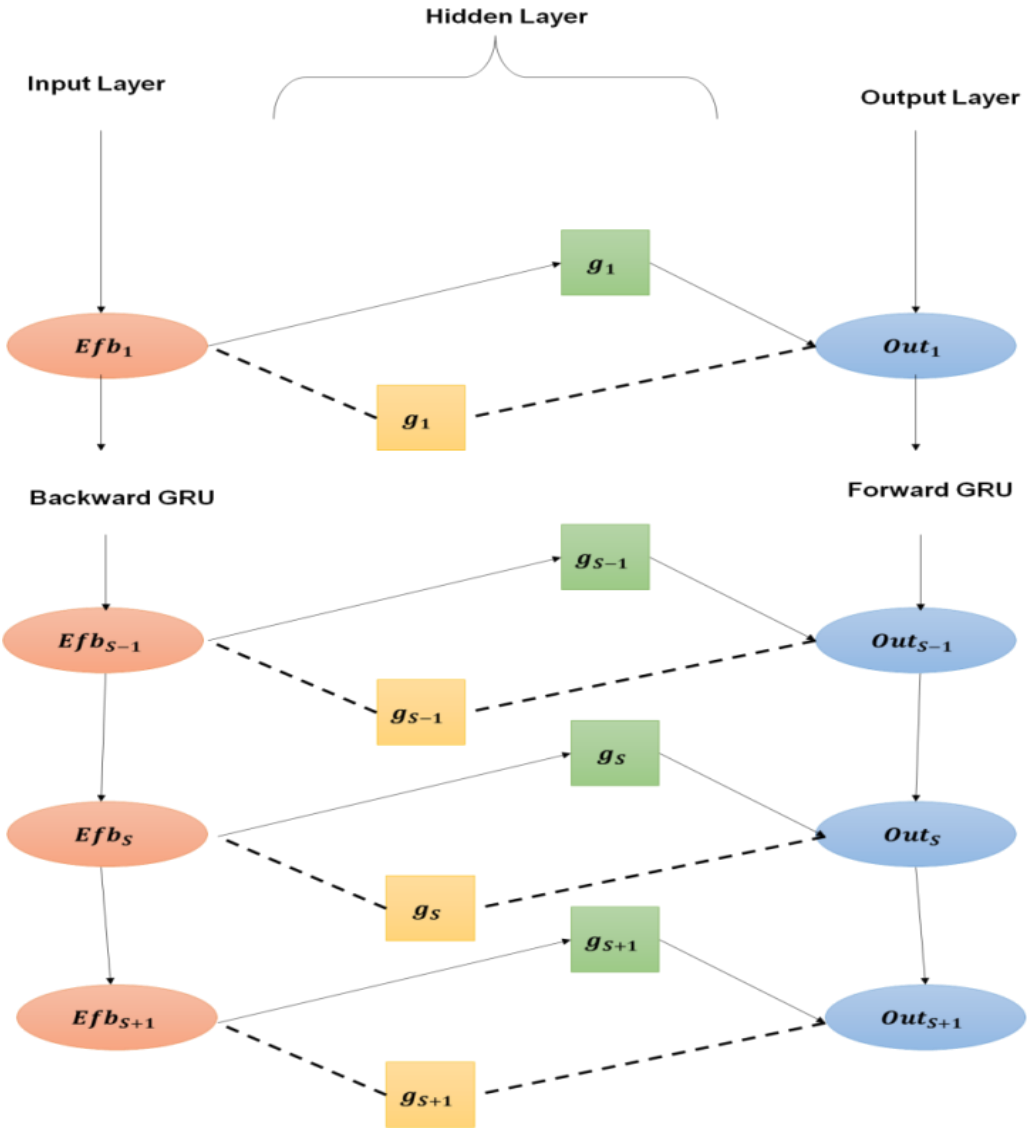


Figure 1 Unfolded Bi-GRU framework.

$$Z_S = E(\vec{g}_S, \overleftarrow{g}_S) \qquad (17)$$

The concealed states of the forward- and backward-moving GRUs are indicated by $g^{->}$ and g>, respectively. *E* stands for the method of merging the results from both approaches, such as multiplying, typical, summary, and other operations. The suggested Bi-GRU framework is appropriate for examining the process of stimulating effectiveness since it can portray simultaneous patterns among different inputs and outputs.

4. Result

Python language was used to analyze the proposed method.

Consider each of the algorithms' associated predictions for workweeks and weekends. Root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are all frequently employed performance measures in the setting of artificial intelligence (AI) vehicle circulation forecasts with innovative modes of transportation. The precision and dependability of highway traffic estimations offered by artificial intelligence algorithms is evaluated using such indicators. Let us go into more depth about the methods in Figure 2. Root mean square error (RMSE) is a measurement of the median variance among the numbers anticipated for traffic movement as well as the values that were actually measured in Figure 3. The average proportion of the quadratic discrepancies among the predicted and observed values is calculated. The following defines the RMSE. RMSE is susceptible to large mistakes and offers an estimate of the total fault extent. The reduced RMSE implies greater precision because it measures the mean absolute difference among the anticipated and actual results.

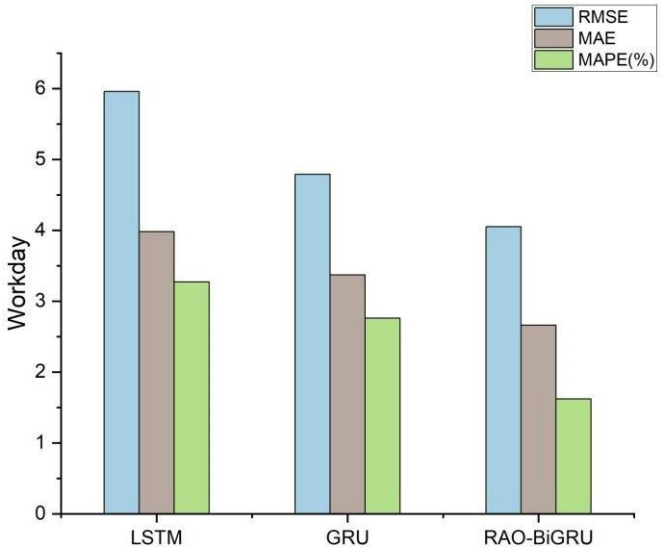


Figure 2 Working day analysis.

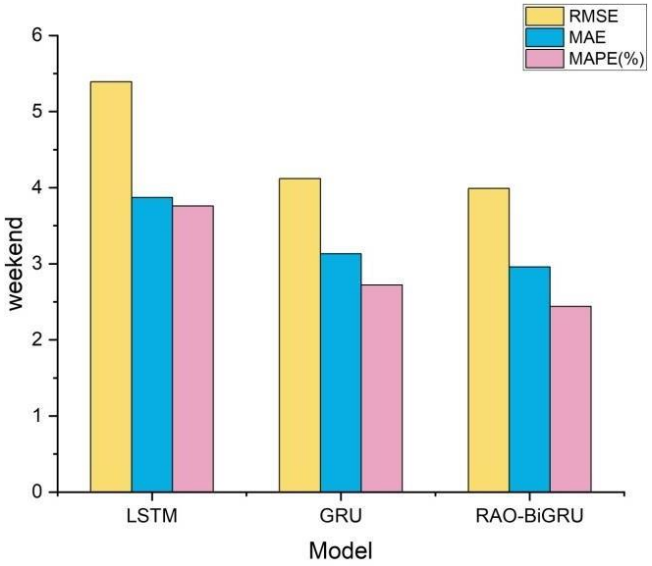


Figure 3 Weekend analysis.

Mean Absolute Error (MAE): The typical variance among projected as well as witnessed flow of traffic figures is also measured in contrast to RMSE. MAE takes into account the unsquared relative discrepancies of anticipated and measured results. The mean relative variance between the actual and projected values is measured by the mean absolute error. This shows the mean number of faults and is less prone to extremes than root mean error. A smaller mean absolute error effect suggests superior precision, much like the root mean error. The proportion of the discrepancy among the expected and actual data for traffic flow is measured by the mean absolute percentage error. This defines the avg. relative percentage variance among the results that were anticipated as well as those that were actually recorded. While attempting to comprehend the variation in projections in relation to the actual data, the mean absolute percentage error can be helpful. It gives the mean variance as a proportion, making it simpler to understand the degree of precision. However, when the numbers seen are near 0 or contain zero values, then mean absolute percentage error might provide a difficulty since divisions with 0 mistakes can occur.

In the RMSE of working days, the value of the LSTM is 5.96, MAE is 3.98 and MAPE is 3.27. The value of GRU in RMSE is 4.79, MAE is 3.37 and MAPE is 2.76. In terms of RMSE, the value of working days in RAO-BiGRU is 4.05, MAE is 2.66 and the value of MAPE is 1.62, as shown in Figure 4.

For the RMSE of weekends, the value of the LSTM is 5.39, MAE is 3.87 and MAPE is 3.76. The value of GRU in RMSE is 4.12, MAE is 3.13 and MAPE is 2.72. In terms of RMSE, the value of the weekend in RAO-BiGRU is 3.99, MAE is 2.96 and the value of MAPE is 2.44, as shown in Figure 5.

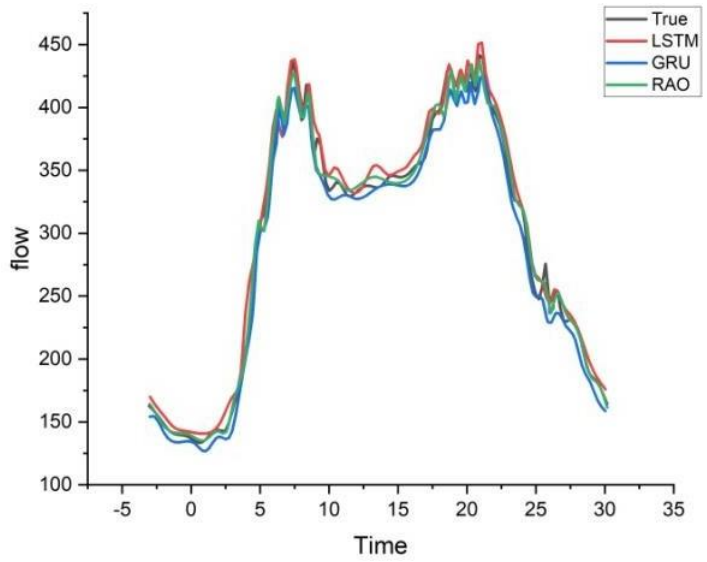


Figure 4 Comparison of traffic flow on working days.

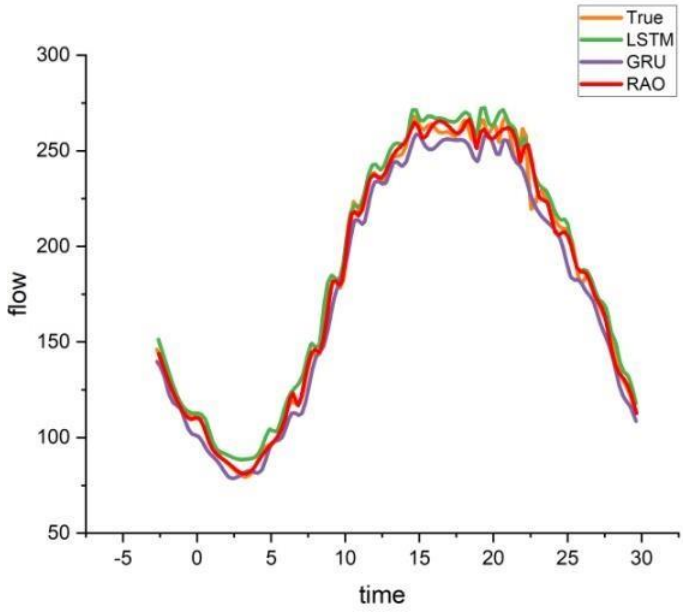


Figure 5 Comparison of traffic flow on weekend days.

5. Conclusion

In the present paper, we offer a novel reptile-ant optimized bidirectional gated recurrent unit (RAO-BiGRU) technique for predicting vehicle movement. The RAO combines the ant colony optimization (ACO) and reptile search optimization (RSO) methodologies. The efficiency of the Bi-GRU framework is enhanced by superior feature selection utilizing the RAO. The PEMS system data used in the study were obtained from an observation station on Interstate 5 from May 1 to June 30, 2014, totaling 17568 observations over a median of 61 days.

In the proposed method, the MAP, MAPE, RMSE, working day and weekend day analyses are used, and the proposed method of RAO-BiGRU shows a better result. People might get discomfort about their whereabouts along with private details being gathered and utilized for prediction may have concerns about gathering and analyzing the sort of data. The opportunity to use artificial intelligence (AI) vehicle traffic flow prediction to alter transportation networks, improve productivity, lessen traffic jams, improve protection, and enable environmentally friendly and intelligent towns is enormous.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Chen Y, Wang W, Chen XM (2023) Bibliometric Methods in Traffic Flow Prediction Based on Artificial Intelligence. Expert Systems with Applications 120421.

Cio YSLK, Raison M, Menard CL, Achiche S (2019) Proof of concept of an assistive robotic arm control using artificial stereovision and eye-tracking. IEEE Transactions on Neural Systems and Rehabilitation Engineering 27:2344-2352.

Englund C, Aksoy EE, Alonso-Fernandez F, Cooney MD, Pashami S, Åstrand B (2021) AI perspectives in Smart Cities and Communities to enable road vehicle automation and smart traffic control. Smart Cities 4:783-802.

Hui Y, Su Z, Luan TH, Cheng N (2021) Futuristic Intelligent Transportation System. arXiv preprint arXiv:2105.09493.

Liu B, Han C, Liu X, Li W (2023) Vehicle artificial intelligence system based on intelligent image analysis and 5G network. International Journal of Wireless Information Networks 30:86-102.

Mohammadi M, Dideban A, Lesani A, Moshiri B (2020) An implementation of the AI-based traffic flow prediction in the resilience control scheme. International journal of transportation engineering 8:185-198.

Olugbade S, Ojo S, Imoize AL, Isabona J, Alaba MO (2022) A Review of Artificial Intelligence and Machine Learning for Incident Detectors in Road Transport Systems. Mathematical and Computational Applications 27:77.

Patankar P, Dorle S, Wyawahare N, Thakre LP (2022) Comparative Study on Design Of AI-Based Communication Protocol For VANET. In 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), pp. 451-455. IEEE.

Popli R, Kansal I, Verma J, Khullar V, Kumar R, Sharma A (2023) ROAD: Robotics-Assisted Onsite Data Collection and Deep Learning Enabled Robotic Vision System for Identification of Cracks on Diverse Surfaces. Sustainability 15:9314.

Rani P, Sharma R (2023) Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. Computers and Electrical Engineering 105:108543.

Rothemund P, Kim Y, Heisser RH, Zhao X, Shepherd RF, Keplinger C (2021) Shaping the future of robotics through materials innovation. Nature Materials 20:1582-1587.

Soori M, Arezoo B, Dastres R (2023) Artificial intelligence, machine learning and deep learning in advanced robotics, A review. Cognitive Robotics.

Tak S, Lee JD, Song J, Kim S (2021) Development of AI-based vehicle detection and tracking system for C-ITS application. Journal of advanced transportation 2021:1-15.

Thamizhazhagan P, Sujatha M, Umadevi S, Priyadarshini K, Parvathy VS, Pustokhina IV, Pustokhin DA (2022) AI Based Traffic Flow Prediction Model for Connected and Autonomous Electric Vehicles. Computers, Materials & Continua 70.

Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, Liu X, Wu Y, Dong F, Qiu CW, Qiu J (2021) Artificial intelligence: A powerful paradigm for scientific research. The Innovation 2:100179.

Zhang R, Sun F, Song Z, Wang X, Du Y, Dong S (2021) Short-term traffic flow forecasting model based on GA-TCN. Journal of Advanced Transportation 2021:1-13.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0122

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Investigation of arranging doctor schedules during the COVID-19 outbreak by using a decision support system

Vijay Kumar Pandey[a] | **Ajay Rastogi**[b]

[a]Department of Computer Science and Applications, M.D. University, Rohtak, Haryana, India, Associate Professor, Department of Mechanical Engineering.
[b]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College of Computing Science and Information Technology.

Abstract People with minor symptoms are isolated at home, while those with severe symptoms are treated in hospitals. The hospital burden has so increased excessively. Healthcare workers experience physical tiredness as a result of stress. In addition to the growing workload, healthcare providers have psychological issues due to their worry of contracting an infection and spreading it to others. Healthcare personnel needs to be protected and given adequate working circumstances. Since protective gear including gloves that are secure, respiratory protection, and eyewear are crucial for the protection of medical professionals exposed to the virus, medical treatment should also be well arranged for. The fast-spreading COVID-19 epidemic has impacted numerous individuals throughout the globe. Considering the strong infectivity, governments issue advisories to stay at home or enact policies like lockdowns to prevent the spread of the illness. Arranging doctor scheduling is one of the crucial difficulties. As part of this study, create a doctor scheduling arrangements with a hospital in Turkey during the COVID-19 epidemic. In order to sustain activity in the new departments while lightening the load in the old departments, the hospital built three new COVID-19-related departments. To solve the arranging doctor scheduling issue, provide a Mixed Integer Programming (MIP) paradigm to turn the problem into a Decision Support System (DSS). The consequential schedules maintain healthcare delivery throughout all departments while distributing their effort evenly by reducing the risk of the doctors contracting the illness.

Keywords: COVID-19, doctors, mixed integer programming, decision support system, schedules

1. Introduction

In today's machine age, the medical services sector is a vital and pivotal part of people's lives. As a result, every issue brought to the attention of the clinic's administrators may result in death or disfigurement. The use of information and interaction to enhance the many positions and administration in human resources management has significantly increased in recent years. The doctor's visit is one of the medical administrations that have been computerized. Social insurance suppliers are encouraged to reduce activity costs while enhancing the nature of administration (Rajakumari and Madhunisha 2020). Scheduling appointments with doctors is important work that demands careful organization and coordination. Guaranteeing effective patient care entails planning and maximizing healthcare personnel's availability and daily schedules. Usually, the first step in the procedure is to take into account each doctor's unique medical specializations, level of experience, desired working hours, and days off. The scheduler must develop a well-balanced and thorough plan for normal appointments and emergencies while considering the expected patient demand (Tan 2019). The medical facility is the cornerstone of the medical system and is essential to preserving people's health. The need for hospitals to provide better, more effective treatments is growing as the economy and living standards improve.

Due to the restrictions of various resources, most hospitals can only improve the efficiency and quality of their services by enhancing scheduling, and in particular by optimizing the scheduling of human resources. The efficiency of doctors is increased, idle time is reduced, and hospital running costs are reduced thanks to personnel scheduling, which also benefits patients and hospitals (Feng 2021). Doctors have been innovative in the international healthcare response throughout the COVID-19 pandemic, working relentlessly to fight the virus and treat those harmed. Doctors have been instrumental in identifying, treating, and overseeing COVID-19 patients, often putting their health in danger. Doctors' commitment, knowledge, and fortitude have been crucial in saving lives and giving communities confidence and support during these trying times. They have given their all to promote public health in a way that is nothing short of heroic (Dewey 2020). The Dependability on healthcare professionals and excellent doctor-patient communication have become crucial during the hard times of the COVID-19 epidemic. The need for clear and sympathetic communication between physicians and

their patients has been highlighted by the virus's increased feelings of uncertainty and anxiety. Doctors have been essential in educating the public about the virus, outlining precautions, and allaying worries and fears (Gopichandran and Sakthivel 2021). As a doctor, it is imperative to notify the proper authorities as soon as any suspected or confirmed cases of the COVID-19 epidemic are discovered. Reporting is an essential first step in stopping the virus's spread and implementing appropriate public health measures. Healthcare workers help with early discovery, contact tracing, and implementing quarantine measures by swiftly alerting public health officials. The allocation of resources is aided by prompt reporting, which guarantees that healthcare institutions have the materials, tools, and staff required to control the epidemic (Li 2020). Emergency doctors leading the charge in treating the crisis are suffering greatly psychologically due to the COVID-19 epidemic. These medical professionals have experienced high tension, apprehension, and emotional weariness. Emergency doctors often endure psychiatric symptoms due to the difficulties of caring for very sick patients, managing little resources, and always being in danger of infection (Bahadirli and Sagaltici 2021).

The COVID-19 epidemic has profoundly changed human history and impacted almost every area of everyday life. The new coronavirus that created this worldwide health catastrophe has led to widespread disease, fatalities, economic disruption, and social unrest. Due to the virus's high contagiousness, it has been necessary to restrict its spread with stringent measures, including lockdowns, social withdrawal, mask wear, and mass vaccination campaigns. The epidemic has put a burden on healthcare systems, increasing hospital admissions and requiring exceptional efforts from healthcare professionals (Bhattacharya 2020). A physician scheduling challenge is a challenging job of developing an effective and fair schedule for a group of doctors to provide the best possible healthcare coverage. This job requires balancing several variables, including patient demands, physician availability and preferences, workload allocation, and regulatory compliance. The objective is to develop a timetable that enhances the quality of patient care while reducing physician weariness and burnout. Concerns, including skill mix, shift rotations, on-call responsibilities, time-off requests, and continuity of care, must be taken into account while solving the physician scheduling challenge (Lan 2023). An exceptional obstacle to efficiently managing healthcare resources during the pandemic is the physician scheduling issue in Covid-19 clinics. Making a productive schedule is essential to ensuring the best possible patient care and reducing the risk of staff burnout due to the rising demand for medical services and the need for specialized treatment. It is important to consider variables, including fluctuating patient numbers, erratic testing needs, and the accessibility of medical personnel with knowledge of infectious illnesses (Das 2020). To arrange a doctor's schedule plans at a Turkey hospital during the COVID-19 outbreak, this study will identify the biggest obstacles to arranging medical appointments.

2. Related works

This study (Hu 2021) examined changes in US physician work schedules and activities before and after the COVID-19 epidemic by analyzing a longitudinal data collection. The Current Population Survey (CPS), managed by the United States Bureau of the Census, compiles demographic and employment statistics. Between 65% and 83% of people responded. Eight interviews were conducted with participants throughout 16 months, allowing for longitudinal analysis. Primary care doctors have been on the epidemic's front lines, increasing their already heavy workloads. The aim was to assess primary care doctors' burnout during the COVID-19 epidemic and its related variables. Cross-sectional research using the snowball method with an online questionnaire distributed through social media. Primary care doctors employed in Portugal during the first COVID-19 epidemic were the intended audience (Baptista 2021).

To assist in healthcare supply chain demand management, lower groups stress, interference the COVID-19 growth chain, and prevent epidemics from spreading, a practical DSS was developed using doctors' knowledge and a fuzzy inference system (FIS) (Govindan 2020). The paper (Barnes and Zvarikova 2021) discusses the results of empirical research conducted to assess and evaluate the use of clinical and diagnostic DSS, Internet of Things-based healthcare apps, and wearable medical devices with artificial intelligence in COVID-19 prevention, screening, and treatment. The study (Nilashi 2020) examined the benefits of wearable technology for healthcare during epidemics or natural catastrophes. It also investigated how recommendation agents introduce and promote these gadgets. Finally, it identifies several flaws in the present recommendation agents and offers suitable fixes to ensure these systems work well during a COVID-19 epidemic. The paper (Ertem 2022) demonstrates that decision-analytic techniques might assist policymakers in simulating various social exclusion scenarios during the first phases of a widespread epidemic. While no vaccines for universal vaccination and effective antiviral therapies are readily accessible, policymakers should brace for further waves of illnesses due to the socially isolated methods utilized. The organization and features of consultations in Belgian out-of-hours primary care over five weekends during the height of the COVID-19 epidemic are described in the study, along with a comparison to a comparable time in 2019. Real-time observational studies using routinely collected, anonymized clinical data from reports of home visits, phone conversations, and in-person consultations (Morreel 2020).

The university hospital's outpatient clinic is located within one of the globe's worst-affected places; report the outcomes of using telemedicine there. The widespread use of telemedicine by patients during the 2009 COVID-19 epidemic in Detroit is illustrative of the study (Garg 2021). The paper (Creese 2021) discovered that many physicians saw modest benefits in their physical health during the first wave of the pandemic, despite the hazards associated with catching COVID-

19. But many also saw a worsening in their mental health as a result of worry, emotional tiredness, guilt, loneliness, and inadequate support. These results offer insight into doctors' health during COVID-19 and how the epidemic has impacted them emotionally and professionally. Infection control procedures owing to COVID-19 hinder end-of-life discussions among patients, relatives, and multifunctional medical teams providing palliative care. Setting up in-person family conferences has been challenging due to stringent limits on visiting times and guest numbers. Although phone-based telehealth consultations can be a solution, the clinician-patient connection may suffer from the absence of nonverbal clues (Wu et al 2020).

The study (Zerbini 2020) examined the psychological toll that doctors and nurses take on depending on the interaction of COVID-19 patients. Additionally, they looked at their support needs during the crisis and the supportive services they employed. This may be because nurses have a heavier workload and spend more time with COVID-19 patients in person than doctors do. The main causes of personal-, work-, and patient-related burnout were direct engagement in COVID-19 screening or treatment, having a medical condition, and less psychological support at work. Burnout was cited as a result of participants' workloads, the uncertainty brought on by the epidemic, a difficult work-family balance, and strained working relationships. Exhaustion seemed to be the most common symptom, and many people developed problem-focused coping strategies to get through the pandemic (Roslan 2021). Visiting hours and communication with medical staff were severely limited for family members of COVID-19 patients hospitalized in Intensive Care Units (ICUs) in the Netherlands during the initial peak of the epidemic. Family communication is an essential component of critical care. However, the pandemic required medical ICU personnel to set up alternate family support, including Family Support Teams (FSTs), which comprised employees who weren't ICU affiliated and called families on the phone (Klop et al 2021).

The study (Spiers 2021) focused on junior physicians' experiences working in the UK during the COVID-19 epidemic. These medical professionals classified themselves as having experienced discomfort due to working circumstances. Junior doctors to work during COVID-19 were traumatized, which left them feeling helpless and decreased the effectiveness of their coping mechanisms. The perspectives of medical professionals practicing a variety of doctors on the sudden and unanticipated shift away from traditional face-to-face meetings toward online webinars, as well as their recommendations for the direction the industry should take going forward. The satisfaction of doctors is important while designing future educational initiatives. Because the present crisis is likely to have long-lasting impacts, webinars should be seen as a supplement to in-person instruction rather than a substitute (Ismail 2021).

3. Problem definition and mathematical model

3.1. Problem definition

Person. Analyze the PSP of a hospital in Turkey that serves as the major medical facility in the area and offers treatment to both routine patients and COVID-19 patients. The hospital has 14 regular departments, including internal cardiology, medicine, urology, and brain surgery. These divisions keep going to run daily throughout the epidemic. Therefore, doctors perform the shifts and refer to them as the usual shifts in these departments. Three new departments have been created by management to help those affected by COVID-19 and fight the epidemic: COVID-Emergency, COVID-Service, and COVID-ICU.

The present doctors are employed in the COVID-19 departments. As a result, in addition to the responsibilities assigned to their respective departments, doctors must also participate in the COVID-19 department' or COVID-19 shifts, which significantly adds to their workload. Patients with COVID-19, a novel illness, may have various problems. The hospital administration thus determined that assigning doctors from multiple specialties to each COVID- Service and COVID- ICU shift would be preferable.

The hospital is not pandemic since other departments continue to provide medical treatment despite low incidence. As a result, in addition to the COVID-19 hours, the doctors also work in shifts that are part of their usual schedules. Each department works a normal 24-hour shift. The number of doctors in the department determines how many are required for each normal shift. Some doctors are not eligible for certain COVID-19 departmental shifts. For instance, whereas a urologist cannot work in COVID-ICU, a neurosurgeon or general surgeon may.

The quantity of exposure to the virus is one of the major factors that might impact it. Therefore, it is important to limit the amount of time that doctors are exposed to COVID-19. This should be accomplished by fairly distributing the shifts among the doctors. Since doctors are highly cautious about the number of shifts owing to the possibility of being ill and then spreading an infection to others, it is crucial to maintain their morale. Two actions may be used to distribute shifts fairly. The COVID-19 shifts of doctors' work in a particular department must be similar first. Instance, the amount of shifts the urologists work in the COVID-Service should be comparable, if not the same. The COVID-19 shifts performed by doctors from various departments shouldn't fluctuate too much at the second level. For instance, medical doctors and surgery, in general, must have the same amount of COVID-ICU shifts. The day between the shifts should be free. A doctor only works on their days off. Additionally, if feasible, there may be two off days between consecutive shifts. There are three eight-hour shifts in COVID-Emergency. The first shift runs from 8:00 AM to 4:00 PM. The second shift runs from 16:00 to 24:00. The last shift occurs between 0:00 and 8:00 the next morning. One medical professional must be assigned to each shift.

Table 1 displays the allocation of usual and COVID-19 shifts and the total number of doctors working in the departments. The first section, the department, contains a list of the hospital's operating divisions. The following section, Number of doctors, list the doctors available in the normal departments. The number of doctors required for the normal shifts is shown in the third column. The columns that follow will detail which units are eligible to join COVID-ICU, COVID-Service, and COVID-Emergency. The COVID-Intensive Care Unit may employ specialists in internal medicine who aren't needed in the COVID-Service or COVID-Emergency departments. The final department on the table, Pool, is a collection of departments. To facilitate scheduling during the pandemic, it was decided to consolidate the pool group departments because of the limited number of doctors inside them. Despite not having a normal shift, the hospital's administration chose to use every department in the pool for the shift since one of the hospital's important units was experiencing chest pains. Therefore, a doctor's typical pool shift would occur in the Department of chest disease.

Table 1 Shifts and Departments.

Eligibility of departments in COVID-19 shift assignments					
Department	**COVID-emergency**	**Regular shifts**	**COVID-service**	**COVID-ICU**	**Num. of physicians**
Ear, and nose, throat disorders	2	1	2	1	7
General Surgery	2	2	2	2	6
Radiology	2	2	2	1	7
Psychiatry	2	2	2	1	7
Internal Medicine	1	3	1	2	8
Chest Diseases	2	2	2	1	5
Cardiology	1	3	1	2	8
Pool	2	2	2	1	12
Orthopedics and Traumatology		2	2	1	8
Urology	2	2	2	1	5
Eye, center	2	1	2	1	8
Neurology	1	2	1	2	5
Brain Surgery	1	2	1	2	5
Cardiovascular Surgery	1	2	1	2	5

3.2. Face Mathematical model

In the MIP model, sets are groups of unique items or objects gathered according to certain criteria or attributes. These sets are used to represent several classes or categories inside the model. The DSS uses the MIP model as its base engine and offers an intuitive interface for data entry, scenario exploration, and schedule generation. It improves decision-making by providing knowledge, advice, and the capacity to weigh trade-offs between multiple scheduling goals or constraints. Based on the results produced by the MIP model, the DSS enables users to make scheduling choices that are more effective and efficient. The MIP model considers several variables to optimize the distribution of doctors working hours while guaranteeing fairness and the best possible coverage. The MIP model creates an ideal schedule that reduces conflicts and increases production by considering doctors' availability, preferences, and necessary abilities. It feels elements including the number of patients, emergency circumstances, and the need for sufficient rest times to avoid tiredness. The resultant value, kit, in the value of objective in this form (equation 1) is left undefined so that the user may input additional measurements for other departments, doctors, or days.

$$\min z = \sum_{j\in J}\sum_{j\in S} L_w W_w \qquad (1)$$

i/j Set of doctors and Index

According to a restriction, A doctor can only work one shift each day (equation 2).

$$\text{t. s.} \sum_{c\epsilon C}\sum_{t\epsilon T(c)} V_{itds} \leq 1\ j\epsilon J\ , s \epsilon S \qquad (2)$$

Third (equation 3) limitations make sure there are enough physicians on hand for each shift. Because of this limitation, ineligible physicians cannot be placed in clinical units. Finding doctors with varied specialties to fill the COVID-19 shifts is a crucial problem for the departments involved in COVID-19.

$$\sum_{c\epsilon C_C} V_{itds} = M_c\ s \epsilon S, c \epsilon C, t \epsilon T(c) \qquad (3)$$

s/S Set of days and Index

c/C Set of departments and Index

$t/T(c)$ Set of shift type for each department c and Index

One doctor from cardiovascular surgery may be present during the COVID-ICU shift, but three doctors representing neurology, brain surgery, and cardiovascular surgery may all work the same shift. The COVID-19 shift's doctors must come from various departments, according to constraint (equation 4).

$$\sum_{c\epsilon(J_d \cap J_q)} V_{itds} \leq= 1 \; s \epsilon S, c \epsilon C_d, t \epsilon T(c), q \epsilon C_Q \qquad (4)$$

J_d Doctors that are part of the regular workforce

J_q Doctors are qualified to work with the COVID-19 department

J_c A group of doctors employed by department c

C_Q Regular departments in the hospital

C_d COVID-19 departments in the hospital

V_{itds} If doctors i is working on shift s on day t at department d

Each doctor will get a day off between each consecutive shift thanks to a restriction (equation 5).

$$\sum_{c\epsilon C} \sum_{t\epsilon T(c)} (V_{itds} + V_{it+1ct}) \leq 1 \; j \epsilon J, s \epsilon S, s+1 \in S \qquad (5)$$

Equation (6) constraint demands two days off between consecutive shifts, if possible.

$$\sum_{c\epsilon C} \sum_{t\epsilon T(c)} (V_{itds} + V_{it+1ct} + V_{it+2ct}) \leq 1 + W_{it} \; j \epsilon J, s \epsilon S, s+1 \in S, s+2 \epsilon S \qquad (6)$$

It's a flexible constraint since the system may choose not to take two days off between shifts if it can't.

$$NC_{qc}^{min} \leq \sum_{s\epsilon S} \sum_{t\epsilon T_c^{( )}} V_{itds} \; q \epsilon C_Q, j \epsilon J_q, j \epsilon J_q, c \epsilon C \qquad (7)$$

NC_{qc}^{min} Minimum shifts a normal department q doctor must work at department c monthly.

Equations (7) and (8) in the restrictions section reveal the maximum and minimum overall number of shifts in departments c that may be assigned to a doctor from normal departments.

$$\sum_{c\epsilon C} \sum_{t\epsilon T(c)} (V_{itds} \leq NC_{qc}^{max}) \; j \epsilon J_q, j \epsilon J_q, c \epsilon C \qquad (8)$$

NC_{qc}^{max} Maximum amount of shifts a normal department q doctor may work in department c in one month.

The maximum and minimum numbers of these changes are minimized by the constraint (equation 9).

$$NC_{qc}^{min} - ND_{qc}^{min} \leq NCK_{qc} \; j \epsilon J_q, c \epsilon C \qquad (9)$$

NCK_{qc}The maximum and lowest number of doctor shifts that may be worked in department q regular department c throughout a calendar month.

The constraints (equation 10) and (equation 11), in turn, show the COVID-19 shifts that the doctor worked in the COVID-19 Department c.

$$NC_{c}^{min} \leq \sum_{s\epsilon S} \sum_{t\epsilon T_c^{( )}} V_{itds} \; q \epsilon C_Q, j \epsilon J_q \qquad (10)$$

NC_c^{min}, The department c maximum and minimum monthly shift count is limited.

$$\sum_{c\epsilon C} \sum_{t\epsilon T(c)} V_{itds} \leq NK_c^{max} \; j \epsilon J_q, j \epsilon J_c \qquad (11)$$

NK_c^{max} Maximum amount of shifts a doctor may work in Department c in a calendar month

The maximum and lowest numbers of these shifts are kept as close to one another as possible by constraint (Equation 12).

$$NC_c^{min} - NC_c^{min} \leq NKK_c \; c \epsilon C_d \qquad (12)$$

To ensure that every general surgeon obtains an equal amount of COVID-ICU shifts, the same DSS may set the restriction COVID-ICU difference for surgical procedures to zero. The detrimental constraints are listed in (equation 13).

$$V_{itds} \epsilon \{0,1\} \; j \epsilon J_q, s \epsilon S, c \epsilon C \; t \epsilon T(c)$$

$$NK_{c}^{min}, NK_{c}^{max}, NK_{qc}^{min}, NK_{qc}^{max} \geq \begin{matrix} W_{it} \epsilon (0,1)\, j\epsilon J_q, s\epsilon S \\ 0\ c \in C, q\epsilon C_{Q}', t\epsilon T(c), q\epsilon C_{Q}, t\epsilon T(c) \end{matrix} \quad (13)$$

W_{it} Variable of deviance t, day 2 that permits doctor to take two days off in a row of rest i

4. Results

4.1. Department-specific shifts for COVID-19

To properly address the COVID-19 epidemic, healthcare agencies have undergone substantial changes in their methods of operation. The COVID-ICU, doctors have been working relentlessly to give intense care to patients fighting the virus, is one of the crucial places affected by these changes. Due to the rise in patients needing non-intensive care, the COVID-Service department has also seen a change in its schedule. Shifts have been rearranged in the COVID-Emergency department to deal with the rush of people needing emergency medical care. To guarantee that dedicated personnel are always ready to manage urgent situations with COVID-19 doctors in this department have been working intensively and switching shifts (Table 2 and Figure1).

Table 2 Numerical outcomes of Department-specific shifts for COVID-19.

Department-specific shifts for COVID-19	
Shifts Per Doctor (SPD) at COVID-ICU	6.5
SPD at COVID-Service	3.8
SPD at COVID-Emergency	1.82

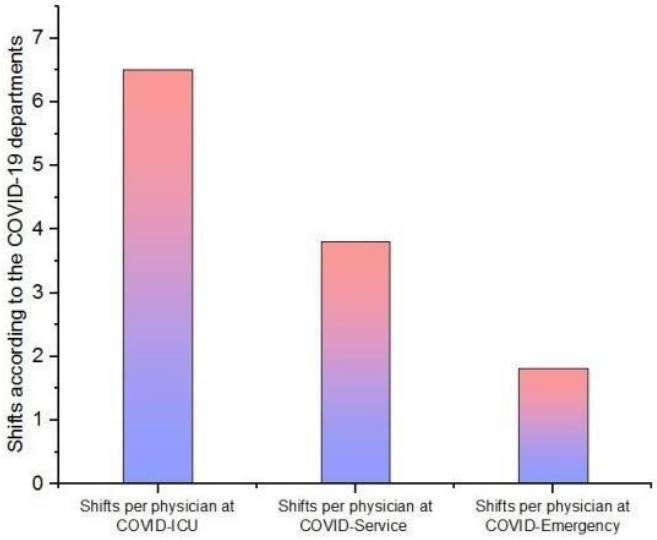


Figure 1 Department-specific shifts for COVID-19.

4.2. A shift structure based on the usual departments

Ensuring doctors' effective and secure scheduling has become crucial during the COVID-19 epidemic. A DSS has been created to help schedule doctors based on normal department shifts to solve this difficulty. This system uses cutting-edge algorithms and current data to make wise judgments. The system optimizes the distribution of doctors through various modifications by looking at variables including patient demand, doctor availability, and departmental needs (Table 3 and Figure 2).

Table 3 Numerical outcomes of a shift structure based on the usual departments.

A shift structure based on the usual departments	
Radiology	11.5
Psychiatry	10.9
Internal Medicine	13.5
Neurology	14.12
Cardiology	13.9

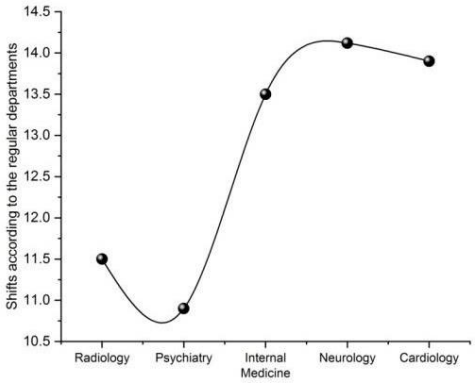


Figure 2 A shift structure based on the usual departments.

4.3. Face Probability of team failure versus Doctors rotation duration

The length of the doctor's rotation may impact the likelihood that the team will fail. As doctors may require time to adjust to new co-workers, processes, and patient populations, frequent rotation disturbs team relations. Frequent composition changes might hamper a team's ability to communicate, coordinate, and collaborate effectively, which could influence patient care (Figure 3 and Figure 4).

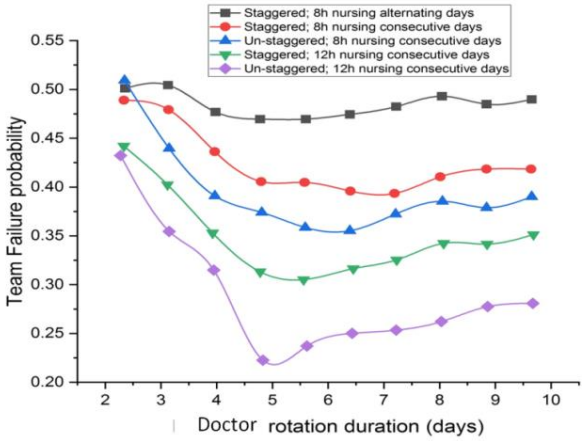


Figure 3 Probability of team failure versus Doctors rotation duration (Median latent period 2.1 days average hospitalization three days pool of 10 doctors, two doctors on duty).

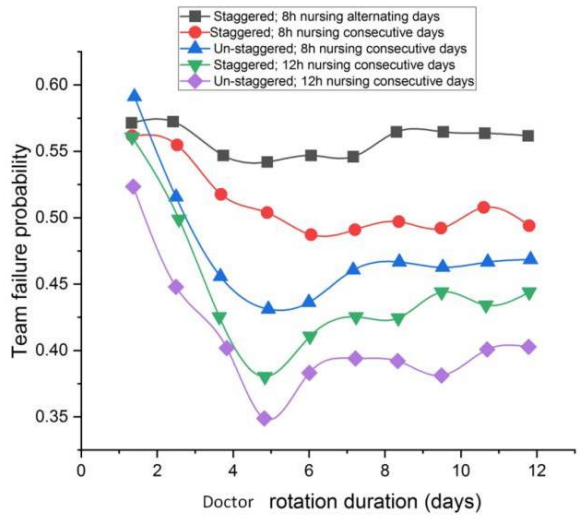


Figure 4 Probability of team failure versus Doctors rotation duration (Median latent period 3.1 days average hospitalization four days pool of 20 doctors, four doctors on duty).

5. Discussion

The appropriate scheduling of doctors has become crucial during the COVID-19 epidemic to guarantee the efficient use of resources and the delivery of medical treatment. The adoption of a DSS may be quite helpful in addressing this difficulty. To examine numerous parameters, including patient demand, case severity, and the availability of medical professionals, a DSS might use data analytics. The DSS may provide knowledgeable suggestions for scheduling doctors, balancing workloads, and optimizing resource use by considering these factors. This improves the effectiveness of healthcare facilities while lowering the danger of exposure for patients and healthcare workers. By adopting a DSS, healthcare institutions may act proactively in the face of the pandemic's dynamic character, ensuring that doctor schedules are arranged in the best possible way to address the community's changing demands while preserving public health.

6. Conclusion

A comprehensive analysis of numerous elements and considerations is conducted as part of the study of doctor schedules that were set up during the COVID-19 epidemic best to use the time and skills of medical personnel. It requires examining information on patient demand, the availability of healthcare professionals, and the particular needs put forward by the pandemic. Most Turkish hospitals established new COVID-ICU, COVID-Service, and COVID-Emergency departments to handle the COVID-19 epidemic better and react to infected patients. Planning became more crucial than previously due to the increasing workload. Despite being able to construct their usual shift plans, institutions need expert assistance during this unplanned pandemic. During the COVID-19 epidemic in Turkey, this study determines the most significant barriers to arranging doctor's scheduling visits at a hospital. The quick epidemic may restrict data availability and accuracy, limiting the inquiry. Provide a MIP paradigm to solve the problem of organizing and arranging doctor schedules so that the problem may be turned into a DSS. Despite these limits, regular monitoring, adaption, and schedule flexibility may assist in reducing some of these obstacles and guarantee effective healthcare services throughout the COVID-19 epidemic. Future studies may evaluate the advantages, difficulties, and long-term viability of remote work and telemedicine in doctor scheduling. Investigating the usage of forecasting and predictive modeling methods may help anticipate future needs and create proactive scheduling strategies.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Bahadirli S, Sagaltici E (2021) Burnout, job satisfaction, and psychological symptoms among emergency physicians during COVID-19 outbreak: a cross-sectional study. Practitioner 83:20-8.

Baptista S, Teixeira A, Castro L, Cunha M, Serrão C, Rodrigues A, Duarte I (2021) Physician burnout in primary care during the COVID-19 pandemic: a cross-sectional study in Portugal. Journal of primary care & community health 12:21501327211008437.

Barnes R, Zvarikova K (2021) Artificial intelligence-enabled wearable medical devices, clinical and diagnostic decision support systems, and Internet of Things-based healthcare applications in COVID-19 prevention, screening, and treatment. American Journal of Medical Research 8:9-22.

Bhattacharya S (2020) The Social Impact of the COVID-19 Pandemic. Observer research foundation 406.

Creese J, Byrne JP, Conway E, Barrett E, Prihodova L, Humphries N (2021) “We all really need to just take a breath”: composite narratives of hospital doctors’ well-being during the COVID-19 pandemic. International journal of environmental research and public health 18:2051.

Das K, Zaman S, Sadhu A, Banerjee A, Khan F S (2020) Quantum annealing for solving a nurse-physician scheduling problem in covid-19 clinics. viXra.

Dewey C, Hingle S, Goelz E, Linzer M (2020) Supporting clinicians during the COVID-19 pandemic. Annals of Internal Medicine 172:752-753.

Ertem Z, Araz OM, Cruz-Aponte M (2022) A decision analytic approach for social distancing policies during early stages of COVID-19 pandemic. Decision Support Systems 161:113630.

Feng D, Mo Y, Tang Z, Chen Q, Zhang H, Akerkar R, Song X (2021) Data-driven hospital personnel scheduling optimization through patient prediction. CCF Transactions on Pervasive Computing and Interaction 3:40-56.

Garg A, Goyal S, Thati R, Thati N (2021) Implementation of telemedicine in a tertiary hospital–based ambulatory practice in Detroit during the COVID-19 pandemic: an observational study. JMIR public health and surveillance 7:e21327.

Gopichandran V, Sakthivel K (2021) Doctor-patient communication and trust in doctors during COVID-19 times—A cross-sectional study in Chennai, India. Plos one 16:e0253497.

Govindan K, Mina H, Alavi B (2020) A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: A case study of coronavirus disease 2019 (COVID-19). Transportation Research Part E: Logistics and Transportation Review 138:101967.

Hu X, Dill MJ (2021) Changes in physician work hours and patterns during the COVID-19 pandemic. JAMA Network Open 4:2114386-e2114386.

Ismail II, Abdelkarim A, Al-Hashel JY (2021) Physicians' attitude towards webinars and online education amid COVID-19 pandemic: When less is more. PloS one 16:0250241.

Klop HT, Nasori M, Klinge TW, Hoopman R, de Vos MA, du Perron C, van Zuylen L, Steegers M, Ten Tusscher BL, Abbink FC, Onwuteaka-Philipsen BD (2021) Family support on intensive care units during the COVID-19 pandemic: a qualitative evaluation study into experiences of relatives. BMC Health Services Research 21:1-12.

Lan S, Fan W, Yang S, Pardalos PM (2023) Physician scheduling problem in Mobile Cabin Hospitals of China during the Covid-19 outbreak. Annals of Mathematics and Artificial Intelligence 1-24.

Li X, Cui W, Zhang F (2020) Who was the first doctor to report the Covid-19 outbreak in Wuhan, China? Journal of Nuclear Medicine 61:782.

Morreel S, Philips H, Verhoeven V (2020) Organization and characteristics of out-of-hours primary care during a COVID-19 outbreak: a real-time observational study. PLoS One 15:e0237629.

Nilashi M, Asadi S, Abumalloh RA, Samad S, Ibrahim O (2020) Intelligent Recommender Systems in the COVID-19 Outbreak: The Case of Wearable Healthcare Devices. Journal of Soft Computing & Decision Support Systems 7.

Rajakumari K, Madhunisha M (2020) Intelligent and convolutional-neural-network-based smart hospital and patient scheduling system. In 2020 International Conference on Computer Communication and Informatics (ICCCI), pp. 1-5. IEEE.

Roslan NS, Yusoff MSB, Asrenee AR, Morgan K (2021) Burnout prevalence and its associated factors among Malaysian healthcare workers during COVID-19 pandemic: an embedded mixed-method study. In Healthcare 9:90.

Spiers J, Buszewicz M, Chew-Graham C, Dunning A, Taylor AK, Gopfert A, Van Hove M, Teoh KRH, Appleby L, Martin J, Riley R (2021) What challenges did junior doctors face while working during the COVID-19 pandemic? A qualitative study. BMJ open 11:e056122.

Tan M, Gan J, Ren Q (2019) Scheduling emergency physicians based on a multiobjective programming approach: a West China Hospital of Sichuan University case study. Journal of healthcare engineering 2019.

Wu Y R, Chou T J, Wang YJ, Tsai JS, Cheng SY, Yao CA, Peng JK, Hu WY, Chiu TY, Huang HL (2020) Smartphone-enabled, telehealth-based family conferences in palliative care during the COVID-19 pandemic: a pilot observational study. JMIR mHealth and uHealth 8:e22069.

Zerbini G, Ebigbo A, Reicherts P, Kunz M, Messman H (2020) the psychosocial burden of healthcare professionals in times of COVID-19–a survey conducted at the University Hospital Augsburg. GMS German Medical Science 18.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0123

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

The use of soft computing for electronic and mechanical system surveillance and examination

Atul Dadhich[a] | **Manish Joshi**[b] | **J. Ghayathri**[c]

[a]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Electrical Engineering.
[b]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College Of Computing Science And Information Technology.
[c]Jain (deemed to be)University, Bangalore, India, Associate Professor, Department of Computer Science and Information Technology.

Abstract Soft computing has made great strides in recent years, giving cutting-edge methods for observing and managing intricate electronic and mechanical devices. For photovoltaic (PV) systems to maximize energy extraction, Maximum PowerPoint Tracking (MPPT) is essential. However, fluctuating environmental factors and system parameters might have an impact on the effectiveness of MPPT algorithms, resulting in less-than-ideal power generation. So, using soft computing approaches, this research suggests a unique fuzzy logic control and modified mayfly optimization (FLC-MMO) methodology for MPPT monitoring in PV systems. PV system imprecision and nonlinearity are handled by FLC, which offers an adaptable and flexible regulation. Additionally, MMO is used to optimize the FLC's parameters, enhancing its performance and accelerating convergence. MMO was motivated by the foraging behavior of mayflies. The outcomes of the experiments show how successful the suggested strategy is. The suggested MPPT surveillance system produces a greater energy extraction rate as compared to traditional MPPT methods.

Keywords: soft computing, electronic and mechanical systems, photovoltaic (PV) system, MPPT, fuzzy logic control, modified mayfly optimization

1. Introduction

The utilization of photovoltaic (PV) systems for the generation of renewable energy has dramatically has grown in recent times (Çeçen 2022). Through the use of these technologies, sunlight is captured and converted into clean, reliable electricity. Maximum PowerPoint Tracking (MPPT) strategies must be applied to maximize the efficiency of PV systems. For PV systems to operate as efficiently as possible, it is necessary to employ Maximum Power Point Tracking (MPPT) methods (Saidi 2019). PV systems could be used at their maximum power output in a variety of climatic conditions thanks to MPPT algorithms, which guarantee optimal power conversion.

Historically, methods of MPPT used in PV systems included Disturb and Observer (P&O) with Cumulative Conductivity (INC) (Brahim 2023). They have a few several drawbacks, such as noisy sensibility, oscillating behavior, and subpar efficiency in quickly changing environmental conditions. Experts have focused on soft computing methods, which provide smart and flexible methods for MPPT monitoring, to address these problems.

A group of computer techniques known as "soft computing," including artificial neural systems, fuzziness, and swarm intelligence, were motivated by organic and ecosystems (Supriya 2023). Scientists have developed sophisticated algorithms for MPPT that increase the effectiveness and resilience of solar energy systems by utilizing the inherent adaptability, fault tolerance, and learning capabilities of soft computing techniques (Hanzaei 2020; Shalf 2020). A complete investigation of MPPT tracking in PV systems using soft computation is the goal of this work.

This study proposes a novel fuzzy logic control and modified mayfly optimization (FLC-MMO) methodology for MPPT monitoring in PV systems using soft computation methods.

The next part of this paper is structured like this: Part 2-Related work, part 3- Methods, Part 4- Simulation Results, and Part 5- Conclusion.

2. Related Works

Among the most important and recent SC-based MPPTs are differentiated evolutionary programming (EP) and degeneration (DE), modified mayfly optimization (MMO), cuckoo search (CS), genetic algorithms (GA), and cuckoo search (CS) (Hashim 2019). The results indicated that, for a PV system operating under the multimodal partial shadowing situation, EP is the SC approach for MPPT which is both desirable and practical.

To evaluate the Perturbation and Observations (P&O) technique and soft computing-based fuzzy logic computations, 5 and 7 languages were utilized together with varying irradiation and temperatures (Srinivas 2021). More in-depth analysis and information about the findings of the different relationships used in the fuzzy logic method.

Numerous evaluations have been conducted on conventional MPPT techniques such as fractional of closed-circuit (FOCV), altered incremental conductivity (MIC), and variable step size perturb and observe (VSS-P&O) (Basha 2020). Additionally, the single-diode type solar cell and the double-diode type PV panel's fill factor (FF) and maximum power extraction have been compared. It is amply shown that MPPT methods are suited for various converter schemes.

The P&O approach was shown to be sufficient for tracking GMPP. In comparison to methods described in publications, the suggested approach is precise, simple to use, and adaptable to any size of PV array. The suggested approach is highly suited and can be implemented for both main and string-type inverters, whereas methods currently used in the literature are only appropriate for inverters with strings. In general, it is anticipated that the suggested (Pillai 2019) method will work well to overcome the drawbacks of all current MPP tracking methods.

In this paper, several computing Techniques are suggested (Ndiaye 2020) for evaluation. MMO and the Adaptive Neuro-Fuzzy Inference System (ANFIS) are used in Maximum Points Tracking (MPPT). The procedures were simulated using a MATLAB/Simulink model of a photovoltaic system. Results showed that ANFIS performs well, providing prompt responses and minimal mistakes.

An effective MPPT-FLC-based soft computing for grid-tied (PV) scheme was shown (Priyadarshi 2020) Space vector pulse width modulation (SVPWM) technology is used to implement a converter regulator for grid-tied PV systems that operate at unity power factor. Inverter and utility grid interaction is still the zeta helicopter. Simulations estimates were provided for stable, changing, and various stress circumstances.

The capabilities of the Variety Skipping Size-Incremental Resistance (VSS-INR) or Varying Step Size-Feedback Circuit (VSS-FC)-based MPPT approaches were evaluated and contrasted with other algorithms. The efficiency of nine precision systems was examined by factoring in steady-state settling time, MPP tracking speed, algorithm complexity, dependence on PV arrays, restricted shade shipping, and total efficacy (Hussain Basha and Rani 2020).

The blend monitoring approach described (Pillai 2019) effectively strikes a balance between traditional P&O and cutting-edge soft computation methods by precisely recognizing shadow instances. That makes use of P&O's distinctive operational conductivity at the P-V curve's leftmost wattage crest.

Using both algorithms for various solar insolations, the modelling of a PV power system has been completed. For efficient MPPT operations, a Cuk converter has been used in this study for connecting the PV array and load. A MATLAB/SIMULINK technique has been used to estimate (Priyadarshi 2020) the grid-tied PV system with a unity power factor. AFLC-based MPPT formulation's efficacy and resilience are explained by simulated findings.

Comprehensive simulation tests for various input situations are conducted to analyse the performance of the proposed hybrid MPPT approach (Sheik Mohammed 2021). A thorough evaluation of the effectiveness of MPPT approaches such as the experiments involving P&O, variable step size (VSS) P&O, and the hybrid MPPT technique developed by combining P&O with LA.

3. Methods

3.1. PV Model

The fundamental components of a PV module are its PV cells (Figure 1). They're able to convert the power of sunshine rays into electricity relying on the photovoltaic phenomena (Mateus 2022). The pentagon semiconductors junctions is a type of photovoltaic device that when exposed to light produces a DC flow. Many scholars have written extensively about the quadratic and hyperbolic interaction of a PV module's produced power and volts; this relationship depends on solar irradiance, humidity, and loads. This exact sort of solar cell is recognized for its kind-diode variant, which appears in the next image. The basic formulae used to characterize all the current importation of the screens include the I-V property of the PV solar cell models.

3.2. MPPT using FLC

The photoelectric wattage tracking method's flexible controller regulates the solar panel's output voltage via a DC-DC converter to keep up the target value. This increases output power to its highest level. The MPP fuzzy logic processor monitors the current and voltage readings at the solar panel's output and uses the data for calculating the power utilizing the equation to determine their inputs. The controller's sharp output indicates when its amplitude modulates the DC-DC conversion. The resultant sunlight at every reading (time k) is examined by the FLC to determine the change in power about voltage (dp/dv). If this amount is < 0, the processor adapts the pulse width modulation, also called the PWM, duty cycle to increase lighting until it is highest, or if its value (dp/dv) is lower than little, the device adapts the duty cycle of PWM to reduce voltages until the electrical current is at its maximum level.

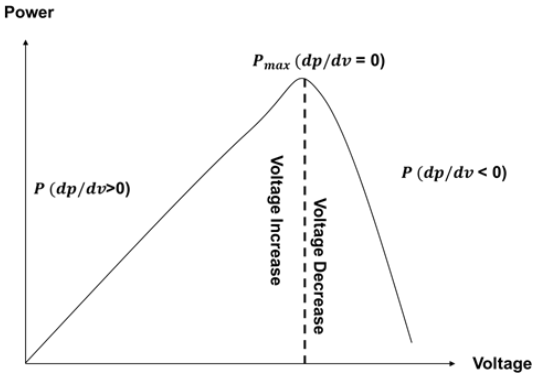


Figure 1 Power voltage characteristic of a PV module.

Warning and Change Error were both sources for the FLC, which also has one output that feeds into the pulse width modulation used to regulate the DC-DC conversion. Warning and Change Error, 2 FLC parameters, are set at sampling period's k.

Where: P (k) is the solar generator's instantaneous efficiency. The value of Change Error reflects how this location is changing, and the value of Error (k) indicates whether the burden operating level at current k is positioned to the PV feature's highest output point's left or right side. Using the essay method and FLC for the presentation display, uncertain reasoning is conducted. Fuzzification, Base rule, and Defuzzification are the three fundamental components of FLC. Another the author-style two-input (antecedent) criterion contains the following structure:

D is X3 if Error is X1 and Change Error is X2.

Wherein Failure, Shift Errors, and D are both input or outcomes, and X1, X2, and X3 are linguistic concepts related to each of them. The website describing the roles shown in Table 1 defines the standards for the intended interaction among the input and output variables.

Table 1 FLC rules.

Ch_Error	**Error**				
	Very small	Small	Medium	Large	Very large
Very small	PG	PM	PP	GM	PG
Small	PG	PP	GP	M	PM
Medium	PM	M	GM	GP	PP
Large	PP	GP	GG	GM	M
Very large	M	M	GG	GG	GP

3.3. Modified Mayfly Optimization

The MA hybridized computational structure is based on the spinning mayfly males' sexual conduct, which attracts females. This technique builds on the MMO framework by combining the advantages of genes (GA) and FA.

3.3.1. Mayfly movement

Two individuals of MA seek regionally while adjusting their velocity and position under predetermined rules. Males modify their stances to compare with one another. Male members' positions and speeds might be calculated using Equations (1) and (2), respectively.

$$w_j^{l+1} = w_j^l + u_j^{l+1} \quad (1)$$

$$w_j^{l+1} = u_j^l + b_1 f^{-\beta q^2}{}^{o}(obest_j - w_j^l) + b_2 f^{-\beta q^2}{}^{h}(hbest - w_j^l) \quad (2)$$

w_j^l represents Mayfly's location during the r repetition; u_j^l relates to the mayfly's I repetition r velocity; $obest_j$ identifies the highest place that mayfly I have attained; b1 and b2 have favourable attraction factors; s_q refers to the Cartesian separation among everyone's ideal location with its mayfly's current location.

The most talented female is allocated to the greatest men, and so on, to elucidate the sexual process. This allows the speed as well as the location of the female mayfly to be represented as follows:

$$u_j^{l+1} = u_j^l + b_2 f^{-\beta q^2}{}^{ne}(w_j^l - z_j^l), e\,(z_j) > e\,(w_j) \quad (3)$$

$$u_j^{l+1} = u_j^l = ek * q\,, e\,(z_j) \leq e\,(w_j) \quad (4)$$

$$z_j^{l+1} = z_j^l + u_j^{l+1} \quad (5)$$

u_j^l It means the female mayfly's velocity during the j stage, w_j^l relates the mayfly's positioning during the l stage, q_{ns} relates to, amongst other factors, the difference in velocity between the male and female mayflies, and fl stands for a sporadic factor.

3.3.2. Mayfly mating

The act of male and female populations merging is the essence of MA. Two groups' combined location number is used to carry out crossover operations under the previous concept of male enticing female. Equations (6) and (7) illustrate its method.

$$offspring\ 1 = K * male + (1 - K\ ) * female \quad (6)$$

$$offspring2 = K * female + (1 - K) * male \quad (7)$$

Where L denotes a chance value falling inside a predetermined range, male denotes a particular male mayfly, female denotes a particular female mayfly, and descendants denote unique descendant mayfly animals.

3.3.3. Fitness Assessment of Mayflies

The later-generation population might take over if it obtains more fitness ratings. of the associated variable in the preceding community on the assumption that its overall size would remain equal; alternatively, the earlier iteration community will persist. The target population should be assessed and chosen at this point using the objective function, f(x). The problem that needs to be solved has an impact on the choice of the objective function. The third part of this paper's objective function provides a detailed explanation.

$$u_{max} = rand * (w_{max} - w_{min}) \quad (8)$$

The mayfly's particular velocity can easily climb to a high value during the operation of MA, rendering the algorithm useless. Small-speed beings will aid in the method's convergence, hence it is vital to set a speed upper limit and zero beginning speed for newly born everybody.

The computation might be impacted if the individual speed is too low and results in a decrease in the area's optimal rate. As a result, the maximum speed is set under someone's location.

The speed must still be reduced when needed though the highest achievable rate has been set. Thus, the addition of the gravity constant g, that is similar to an MMO's gravity ratio.

The velocity of the male mayfly is calculated by including an enticement part:

$$u_j^{l+1} = h * u_j^l + b_1 f^{-\beta q^2}{}^{o}(obest_j - w_j^l) + b_1 f^{-\beta q^2}{}^{h}(hbest - w_j^l) \quad (9)$$

$$u_j^{l+1} = \begin{cases} h * u_j^l + b_2 f^{-\beta q^2}{}_{ne}(w_j^l - z_j^l), e\ (z_j) > e\ (w_j) \\ h * u_j^l + ek * q\ , e\ (z_j) \le e\ (w_j) \end{cases} \quad (10)$$

In optimization management, the mayfly's erratic motion factor is useful for escaping the nearby optimum capture, but due to its high beginning value, it will also cause the mayfly to fly across an underdeveloped sector. The new random movement coefficient consequently has to be closely monitored throughout an iterative process in order in decreasing it.

$$c_l = c_0 \delta^l, 0 < \delta < 1 \quad (11)$$

$$ek_l = ek_0 \delta^l, 0 < \delta < 1 \quad (12)$$

Progeny-level variation techniques are introduced to address potential local optimum converging. A mutation in the progeny could force the mayfly to expand to discover the territory.

$$offspring_m = offspring_m + \sigma M_m(0,1) \quad (13)$$

Where, σ denotes a normal distribute d's variance, and N_n (0, 1) denotes a standard normal distribution with a mean.

4. Simulation Results

Simulations of the PV panel, transformer, and the FLC-MMO MPPT control system can be exemplified by the simulation's structure and using Bullet Cluster Optimising Techniques with MPPT fuzziness in Controls built in Matlab/Sim

and illustrated in Fig 2 The final shout shape is the final electricity, the goods produced electricity from the PV in comparison to standard ones, and the result produced power for the FLC under varying environmental conditions, specifically the climate, and sunlight, are shown in Fig 2 to fig 4. The buck-boost converted monitors its voltage, and FLC creates a command signal for the conversion's activity cycles. Since the traditional P&O approach converges slowly, it is disregarded.

The MPP is determined using the Fuzzy-MMO regulator. The power conversion, the PV component, and the suggested MPPT algorithm are all shown in the suggested approach's graphic. The enhanced a direct-current is used for some reasons, including the fact that it consistently maintains the MPP regardless of sun-shielding or array heat and displays higher standards with the performance of PV arrays' MPP. The situation of fluctuating solar insulation and the temperature was modelled to replicate the workings of the entire system. The effectiveness is examined using various swarm sizes and repetitions. The algorithm's completion is dependent on speed and variable randomness. When the radiation level is varied owing to external factors, MMO captures the worldwide ideal value. The Figures 2, 3 and 4 display the various ideal voltage and power spectrum values that apply to each repetition.

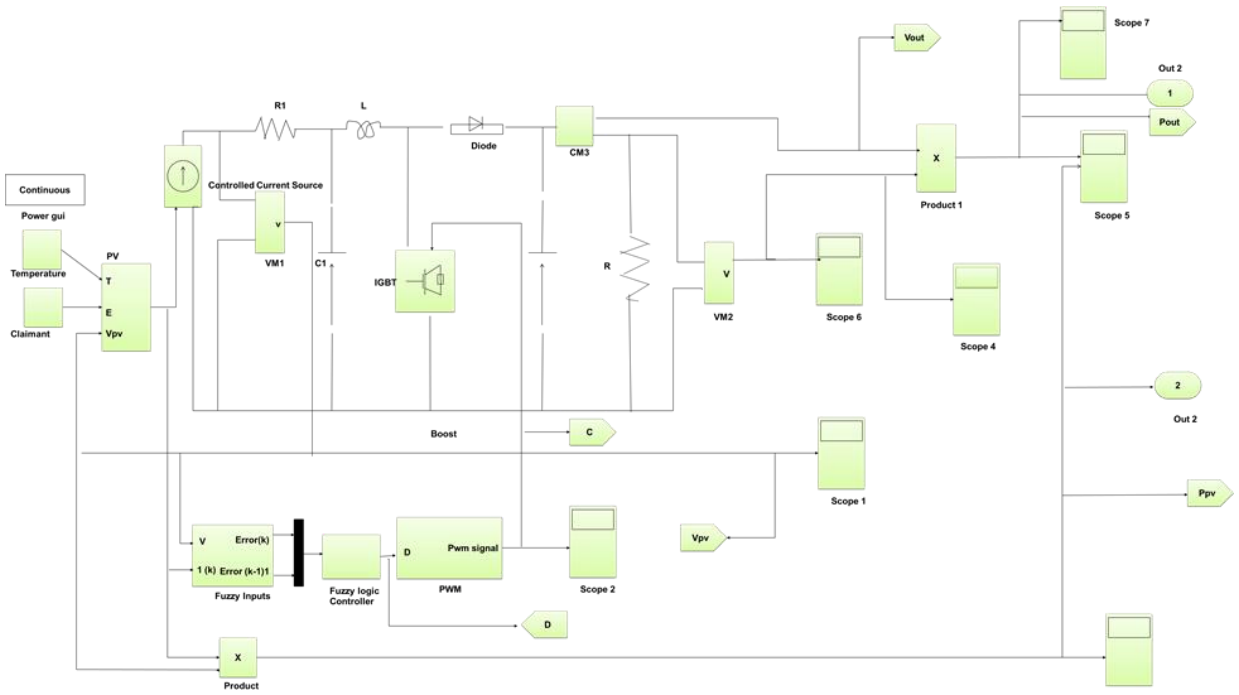


Figure 2 Fuzzy-MMO control of PV system.

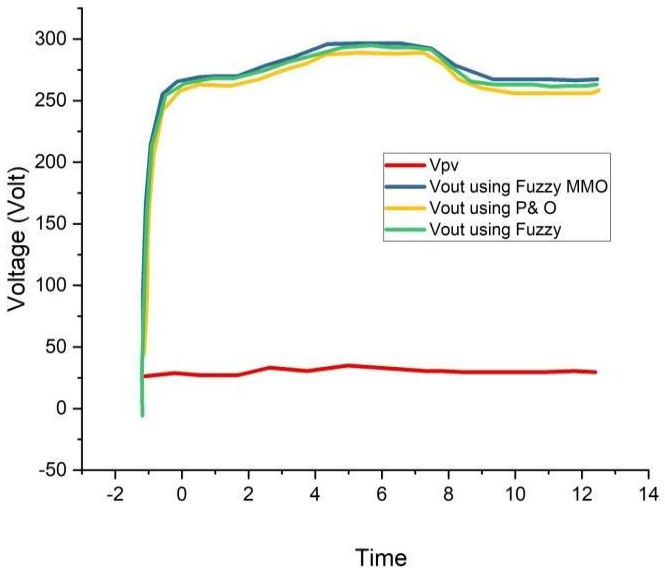


Figure 3 PV output voltage for different considered control.

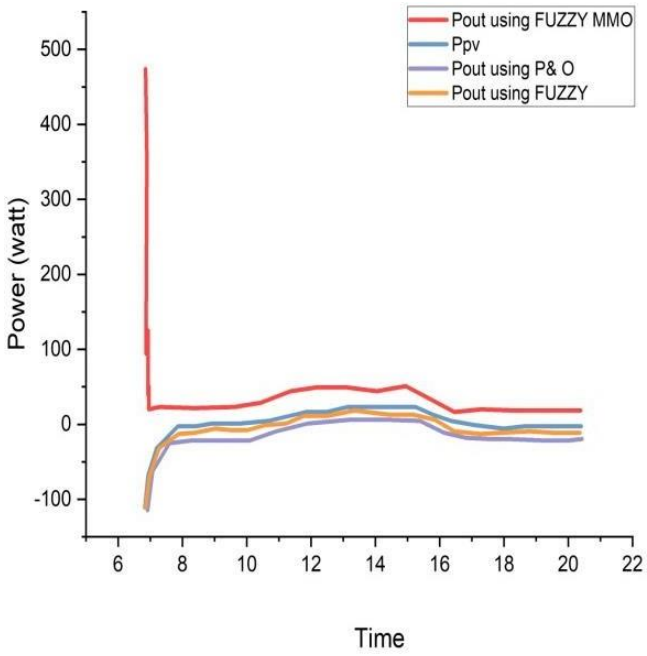


Figure 4 PV output Power for different considered control.

5. Conclusions

In this paper, a better alternative to the MPPT approach using (FLC-MMO) algorithms for solar panels is presented. This technique can monitor the MPP value for challenging settings such as high irradiance as variations. Once the (MPP) is discovered, the continuous oscillation is reduced, which is the approach's principal benefit. Additionally, the suggested technique can track the MPP in harsh environments, such as those caused by significant irradiance intensity variations. This proposed technique, which addresses a common problem with the traditional MPPT, is based on an optimized approach. The suggested approach outperforms the traditional direct duty cycle method under all temperatures. The findings show that the suggested fuzzy-MMO device surpasses and offers Advantages including improved tracking velocity, absence of oscillations at the MPP, and ability to locate the MPP for any circumstance as well as alterations of protection, and this method is simple to use and can be calculated very quickly; thus, its implementation using a cheap processor is practicable.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Basha CH, Rani C (2020) Different conventional and soft computing MPPT techniques for solar PV systems with high step-up boost converters: A comprehensive analysis. Energies 13:371.

Brahim MH, Ang SP, Dani MN, Rahman MI, Petra R, Sulthan S M (2023) Optimizing Step-Size of Perturb & Observe and Incremental Conductance MPPT Techniques Using PSO for Grid-Tied PV System. IEEE Access 11:13079-13090.

Çeçen M, Yavuz C, Tırmıkçı CA, Sarıkaya S, Yanıkoğlu E (2022) Analysis and evaluation of distributed photovoltaic generation in electrical energy production and related regulations of Turkey. Clean Technologies and Environmental Policy 1-16.

Hanzaei SH, Gorji SA, Ektesabi M (2020) A scheme-based review of MPPT techniques concerning input variables including solar irradiance and PV arrays' temperature. IEEE Access 8:182229-182239.

Hashim N, Salam Z (2019) Critical evaluation of soft computing methods for maximum power point tracking algorithms of photovoltaic systems. International Journal of Power Electronics and Drive Systems 10:548.

Hussain Basha CH, Rani C (2020) Performance analysis of MPPT techniques for dynamic irradiation condition of solar PV. International Journal of Fuzzy Systems 22:2577-2598.

Mateus MD (2022) Monitoring and control of a photovoltaic panel in real-time (Master's thesis, Universidade de Évora).

Ndiaye A, Faye M (2020) Experimental Validation of PSO and Neuro-Fuzzy Soft-Computing Methods for Power Optimization of PV Installations. In 2020 8th International Conference on Smart Grid (icSmartGrid), pp. 189-197. IEEE.

Pillai DS, Ram JP, Ghias AM, Mahmud MA, Rajasekar N (2019) An accurate, shade detection-based hybrid maximum power point tracking approach for PV systems. IEEE Transactions on Power Electronics, 35(6), pp.6594-6608.

Priyadarshi N, Bhoi AK, Sahana SK, Mallick PK, Chakrabarti P (2020) Performance Enhancement Using Novel Soft Computing AFLC Approach for PV Power System. In Cognitive Informatics and Soft Computing: Proceeding of CISC 2019, pp. 439-447. Springer Singapore.

Priyadarshi N, Bhoi AK, Sharma AK, Mallick PK, Chakrabarti P (2020) An Efficient Fuzzy Logic Control-Based Soft Computing Technique for Grid-Tied Photovoltaic Systems. In Cognitive Informatics and Soft Computing: Proceeding of CISC 2019, pp. 131-139. Springer Singapore.

Saidi K, Maamoun M, Bounekhla MH (2019) A new high-performance variable step size perturb-and-observe MPPT algorithm for photovoltaic system. International Journal of Power Electronics and Drive Systems 10:1662.

Shalf J (2020) the future of computing beyond Moore's law. Philosophical Transactions of the Royal Society A 378:20190061.

Sheik Mohammed S, Devaraj D, Imthias Ahamed TP (2021) Learning Automata and Soft Computing Techniques Based Maximum Power Point Tracking for Solar PV Systems. Intelligent Paradigms for Smart Grid and Renewable Energy Systems 227-262.

Srinivas N, Singh S, Gowda M, Prasanna C, Modi S (2021) Comparative Analysis of Traditional and Soft Computing Techniques of MPPT in PV Applications. In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), pp. 1-6. IEEE.

Supriya Y, Gadekallu T R (2023) A Survey on Soft Computing Techniques for Federated Learning-Applications, Challenges, and Future Directions. ACM Journal of Data and Information Quality.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0124

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Implementation of novel pattern recognition approach using data mining for medical applications

Mohit Kumar Sharma[a] | **Namit Gupta**[b] | **Ananta Ojha**[c]

[a]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Electrical Engineering.
[b]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Associate Professor, College of Computing Science And Information Technology,
[c]Jain (deemed to be) University, Bangalore, India, Professor, Department of Computer Science and Information Technology.

Abstract The capacity of pattern recognition algorithms to analyze complicated data and retrieve useful information has attracted a lot of interest in the world of medical solutions. Data mining methods have become effective tools for pattern detection in the medical field, allowing for the extraction of insightful knowledge from huge and complex datasets. In this study, a diversified convergent squirrel search optimization algorithm and support vector machine (DCSSO-SVM) are combined to provide a unique pattern recognition method for medical applications. The DCSSO algorithm draws its inspiration from squirrels' efficient search-space exploration and exploitation during foraging. The suggested method integrates this algorithm with SVM to enhance feature selection, speed up convergence, and increase classification performance. To assess the effectiveness of the suggested DCSSO-SVM strategy, we assembled a dataset of medical images for this study. Other cutting-edge methods are contrasted with the suggested methodology. The experimental findings show that, regarding the accuracy, precision, f1-score, and recall metrics, the suggested DCSSO-SVM technique performs better than the current methods. It also has great toughness and support, making it appropriate for a variety of medicinal applications.

Keywords: pattern recognition, data mining, medical images, DCSSO-SVM

1. Introduction

Database search for information is a process with several distinct phases that is well-defined. The fundamental procedure, known as data mining, allows for the discovery of hidden yet significant information from big datasets (Mughal 2018). According to a formal definition provided, data mining is the nontrivial extraction of implicit, previously unidentified, and potentially beneficial details from data. Data mining tools offer a user-oriented method for identifying fresh and mysterious trends in the data (Boruah and Kakoty 2019). Healthcare managers can improve the quality of treatment by applying the knowledge they have gained. The newly learned knowledge can be used by medical experts to reduce the incidence of adverse drug side effects and provide more affordable options for therapy. One of the key uses of data mining methods in the administration of healthcare is the prediction of individuals' future behavior based on their past behavior (Guha and Kumar 2018). For healthcare companies (hospitals, medical facilities), delivering excellent services at affordable rates is a huge challenge. An accurate patient diagnosis and the delivery of efficient therapies are essential components of quality care. Poor clinical judgment can result in devastating outcomes, which is inexcusable.

Clinical test costs must be kept to a minimum by hospitals. By using the proper computer-based information and decision-support tools, they can attain these goals (Khajouei 2018). There is a vast amount of medical information. Patient-centric data, handling of resource data, and modified data are all included. Organizations in the healthcare industry must be able to analyze data (Dash 2019). Millions of medical records may be computerized and preserved, and data mining techniques could help with a number of important and critical health-related issues. The accessibility of integrated information via the sizable patient repositories has caused a shift in how healthcare professionals, patients, and consumers perceive clinical data visualization from qualitative to more quantitative, with support from all clinical and imaging sources (Siripipatthanakul 2021). For example, it could now be feasible for doctors to compare diagnostic data from various individuals with the same illnesses. The uniformity of other doctors handling the same case from around the world can also help doctors corroborate their results (Gomez 2019). Diagnostics in medicine is seen as an important but complex process that must be completed accurately and effectively. The same might be automated, which would be quite advantageous. Instead of relying on the knowledge-rich facts concealed in the document, doctors often base their treatment

recommendations on their intuition and prior experience (Wu 2017). This approach affects the level of care provided to people by causing unintentional biases, errors, and astronomical medical costs.

This study introduces the Diversified convergent squirrel search optimization (DCSSO) for analyze the complicated data and retrieve useful information has attracted a lot of interest in the world of medical solutions. The remainder of the paper is divided into subsequent parts. Part 3 contains the method explained. Part 4 includes the results, while Part 5 discusses the conclusions.

2. Related Works

The initial group of applications of data mining, machine learning, methods, and tools in the field of diabetes studies appeared to be the most popular. Regarding prognosis and diagnosis, diabetic complications, genetics and surroundings, health care and administration, and diabetic issues, the study (Kavakiotis 2017) assessed these applications. There were several different machine learning strategies applied. Through the use of several data mining techniques, including GMM, Logistic Regression, ELM, and ANN, the study (Komi 2017) investigated the early prediction of diabetes. The results of the study demonstrate that ANN offered greater precision than other methods. The study's (Lan 2018) major objective was to assess previous research that examined the specialized domain expertise of bioinformatics using data mining and deep learning methods. It gave a concise rundown of the various preprocessing data mining methods, categorization, and clustering as well as the various optimized neural network designs used in deep learning techniques.

The study (Albahri et al 2020) concentrated on machine learning (ML) and data mining-based automated AI solutions for identifying and diagnosing COVID-19, and it obtained a general understanding of the dangerous virus. It also addressed the drawbacks of using ML and data mining algorithms, and it gave the health sector the study's findings. A thorough and comprehensive analysis of malware detection methods employing methods of data mining was offered in the study (Souri 2018). Furthermore, it divided the malware detection techniques into 2 primary groups: Identification methods based on behavior and signatures. The primary findings of the work include outlining the existing problems with malware detection techniques in data mining and providing a comprehensive and comprehensive evaluation of the status of machine learning methods, and talking about key classification factors in malware approaches in data mining. According to the study (Amin 2019), methods for data mining and key traits can help forecast cardiovascular disease more accurately. The prediction models were made using a variety of classification techniques, including k-NN, Decision Tree, Naive Bayes, Neural Network, and Vote. A review (Elizabeth 2018) on how to manage healthcare data while utilizing big data techniques and comprehending the potential of Hadoop, A networked data analysis and storing application that is open-source was developed. Using information from all aspects of human life, an analysis and study of potential future work in various areas is also conducted.

3. Method

We collected the data sets and preprocess that data using min-max normalization, we used Diversified convergent squirrel search optimization (DCSSO), to enhance the support vector machine (SVM) method. Figure 1 shows the proposed framework diagram.

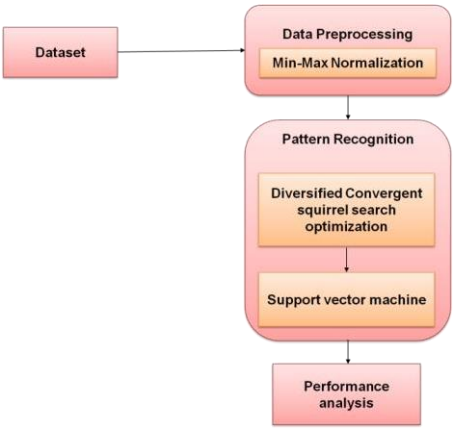


Figure 1 Proposed framework.

3.1. Dataset

Using ARFF files, JNCC2 loads data. This plain text format was created first for WEKA (Srinivas 2010). http://www.cs.waikato.ac.nz/ml/weka/index_datasets.html.2636 provides access to a significant amount of ARFF records, which include those from the UCI repository.

3.2. Min-Max normalization

By using min-max normalization, the data is changed to fall inside a predetermined range, usually between 0 and 1 or -1 and 1. This enables more accurate comparisons between several variables or data points that might have various scales or ranges.

Feature scaling, commonly referred to as min-max normalization, is a method for rescaling numerical data to a certain range. The min-max normalization equation is as follows:

$$Y' = \frac{Y - Y_{min}}{Y_{max} - Y_{min}} \times (max_{new} - min_{new}) + min_{new} \quad (1)$$

The initial value to be normalized is Y.

Y_{min} Represents the lowest value found in the initial dataset, and the highest value found in the first dataset is referred to as Y_{max}. The normalized value is Y ', and the intended minimum and maximum values of the normalized range are min_{new} and max_{new}, respectively.

3.3. Diversified convergent squirrel search optimization (DCSSO)

According to the current season, the individual's characteristics, and the presence of predators, the traditional SSA modifies certain places. The upper and lower boundaries of the search space are ET_v and ET_K, supposing that M is the population's amount. There are M people created at random.

$$ET_j = ET_K + rand\,(1, C) \times (ET_v - ET_k) \quad (2)$$

ET_j Stands for the i-th person (i = 1...M), and C is the size of the problem. SSA stipulates that there must be a single squirrel at each tree, thus if the overall amount of squirrels is M, there must be M trees in the forest. Let's use the minimization issue as an example. One hickory tree and M_{et} (1 < M_{et} < M) acorn trees are present in every M tree; all remaining trees are regular trees without any food. The hickory tree is the top source of food for squirrels, with acorn trees coming in second. Based on the varied issues, M_{et} may vary. The squirrels are split into 3 kinds based on the population's fitness ratings, ranked from highest to lowest: people who were found near hickory trees (E_h), people who were found near acorn trees (E_a), and people who were found near regular trees (E_n). The individual with the lowest fitness value is designated as E_h, followed by people with fitness values ranging from 2 to M_{et} + 1, and the others are designated as E_n. The goal of E_a is E_h to locate the more advantageous food source, whereas the destinations of E_n are chosen at the chance to be E_a or E_h.

The individuals update their places by gliding to the hickory or acorn trees.

$$\begin{cases} ET_j^{s+1} = ET_j^s + c_h \times H_d \times (E_g^s - ET_j^s) & if\ q > O_{co} \\ random\ location & otherwise \end{cases} \quad (3)$$

$$\begin{cases} ET_j^{s+1} = ET_j^s + c_h \times H_d \times (E_{bj}^s - ET_j^s) & if\ q > O_{co} \\ random\ location & otherwise \end{cases} \quad (4)$$

The chance of a predator appearing is represented by the random variable r, which ranges from 0 to 1. If r > O_{co}, no predator will show up, the squirrels will be secure as they go through the forest in search of food. If r $\leq$ O_{co}, predators show up, the squirrels are compelled to focus their activity, the people are at risk, and their places are randomly moved. H_d is a constant with a value of 1.9; t denotes the current iteration; Fai is the by chance generated person from E_a; and g_h is the gliding distance.

$$c_h = \frac{g_h}{\tan(\varphi) \times te} \quad (5)$$

The constants $g_h$and te have values of 8 and 18, respectively, while the gliding angle is represented by $\tan(\varphi)$.

$$\tan(\varphi) = \frac{C}{K} \quad (6)$$

The lift force is K, while the drag force is C.

$$C = \frac{1}{2\rho U^2 T D_c} \quad (7)$$

$$K = \frac{1}{2\rho U^2 T D_K} \quad (8)$$

The variables ρ, U, T, and D_c are each equal to 1.1198 kg m^{-3}, 5.18 ms^{-1}, 161 cm^2, and 0.5, respectively; CL is an integer drawn at random between 0.575 and 1.8.

At the beginning of each cycle, the standard SSA requires that the entire populace is in the winter season. Once every person has been updated, the season changes are assessed.

$$T_d^s = \sqrt{\sum_{l=1}^{C}(E_{bj,l}^s - E_{g,l}^s)^2}\ j = 1,1,\dots,M_{et} \quad (9)$$

$$T_{min} = \frac{10f^{-6}}{(365)^{s/(S/2.5)}} \quad (10)$$

If $T_d^s > T_{min}$, winter turns into summer, but the rest of the season is unaltered. S is the maximum number of repetitions. As the season changes to summer, everyone who glides to Eg stays in the updated location, while everyone who glides to Ea without encountering predator's moves to a new place.

$$ET_{j\ new}^{s+1} = ES_K + Kf'uz(m) \times (ES_V - ES_K) \quad (11)$$

The random walk model known as $Kf'uz$ has steps that follow the $Kf'uz$ distribution.

$$Kf'uz(w) = 0.01 \times \frac{q_a \times \sigma}{|q_a|^{\frac{1}{\beta}}} \quad (12)$$

β is a constant with a value of 1.5, and a may be determined.

$$\sigma = \left(\frac{\Gamma(1+\beta)\times\sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1+\beta}{2})\times\beta\times 2^{(\frac{\beta-1}{2})}}\right)^{\frac{1}{\beta}} \quad (13)$$

$$where\ \Gamma(w) = (w-1)! \quad (14)$$

Populations are spread across the search space in the early stages of optimization, and the distances between those with higher fitness values are still considerable. To significantly expand the search space, it is crucial to preserve population diversity. Additionally, the rate of convergence has to be increased. As optimization progresses, personal distinctions become less, and the primary task becomes finding the top people to accelerate convergence. The population's diversity must also be raised to prevent the method from reaching an optimal location. In this research, Diversified Convergent Squirrel Search Optimization (DCSSO) is developed to enhance the efficiency and resilience of SSA taking into account the study shown above

For the early stages of the theory of evolution, when $T_d^s \geq T_{min}$, and the stages that follow, when $T_d^s > T_{min}$, respectively, each has a unique winter search approach. The detailed DCSSO procedures are presented in Algorithm 1.

Algorithm 1: DCSSO procedures

Step 1: Input: pop
Step 2: Output: $ebest$ (The algorithm's top-performing fitness value is called $ebest$.)
Step 3: for t = 1 to S
Step 4: Analyse the population's fitness levels. Applying the leaping search technique, update the population.
Step 5: If t == S/m
Step 6: Determine the appropriate linear regression
Step 7: Whenever there are 2 or more positive computed slopes, keep improving with a progressive search
Step 8: else
Step 9: keep improving with the leaping search technique
End
End
End

3.4. Support vector machine (SVM)

A supervised learning technique called the support vector machine (SVM) was developed utilizing the concepts of statistics and the structural risk mitigation idea. In high-dimensional or indefinite surroundings, the basic objective of SVM is to construct a hyperplane that optimizes the separation between 2 distinct kinds of information instances. The hyperplane should take the following mathematical form:

$$h(w) = x \cdot w + a = 0 \quad (15)$$

Where x denotes the distance between the hyperplane's normal vector and a denotes the distance between the hyperplane and the point of origin plane. By adding the Lagrangian coefficient, the objective function is changed into a dual optimization problem as shown below:

$$\min \max K(x, a, \alpha = \frac{1}{2}|x|^2 - \sum_{j=1}^{n} \alpha_j [z_j (x \cdot w_j + a) - 1] \quad (16)$$

$$(0 \leq \alpha_j \leq d, j = 1,2,..,m) \quad (17)$$

$$\sum_{j=1}^{m} [z_j (x \cdot w_j + a) - 1] \quad (18)$$

Where $K(x, a, \alpha)$ stands for the Lagrangian function, α_j for the Lagrangian coefficient, and c for the upper bound of α_j. If w_j belongs to the initial class, then z_j= 1, but if w_j belongs to the additional class, then z_j= 1. The mistake value, or loss value, of the objective function, is particularly impacted by the penalty factor c. In particular, the error increases as d increases, particularly if the loss function is provided. If the mistake is significant, overfitting of the SVM may occur easily. Besides, when c is too little, the SVM may experience an under fitting issue. The following equation is a representation of the ideal decision function.

$$e(w) = thm(h(w)) = thm(x_0 \cdot w + a_0) = thm\left(\sum_{j=1}^{m} \alpha_j * z_j (w_j \cdot w) + a_0\right) \quad (19)$$

Where thm () signifies a symbolic function, α_j denotes the ideal Lagrangian factor, and x_0 and a_0 signify the ideal values of x and a.

$$l(w, w') = \exp\left(-\frac{||w - w'||^2}{2\sigma^2}\right) = \exp\left(-\gamma ||w - w'||^2\right) \quad (20)$$

The kernel parameter, in this case, is $\gamma = \frac{1}{2\sigma^2}$, and σ is the radius of the radial basis. Training and assessment speeds are impacted by γ 's value. It should be noted that the kernel parameter γ and penalty factor d settings have a significant impact on how well SVM performs.

4. Result and Discussion

We use a Python tool to implement the suggested method (DCSSO -SVM) for Pattern Recognition Approach in medical applications. The analysis is done on accuracy (%), precision (%), recall (%), and F1-score (%). GAN (Frid-Adar 2018), DBN (Kaur 2020), and CNN (Yu 2021) are the current methodologies and are compared with the suggested method (DCSSO -SVM).

Accuracy serves as a gauge for a classification model's efficacy. It displays the proportion of instances in a dataset that is accurately classified about all of the occurrences. In other words, accuracy measures the overall accuracy of a model's forecasts. The accuracy of the proposed and used procedures was displayed in Table 1 and Figure 2. While GAN, DBN, and CNN only achieve an accuracy of 87.56 %, 91.62 %, and 92.45 %, respectively, the suggested technique, DCSSO-SVM, achieves an accuracy of 98.34 %. The proposed approach by DCSSO-SVM is more accurate than the established ones.

$$Accuracy = \frac{(True\ Positives + True\ Negatives)}{Total\ Number\ of\ Instances} \quad (21)$$

Table 1 performance of the proposed method (DCSSO-SVM).

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
GAN	87.56	88.24	87.31	89.19
DBN	91.62	90.72	89.49	89.98
CNN	92.45	92.84	91.65	91.67
DCSSO-SVM (Proposed)	98.34	97.45	98.12	98.27

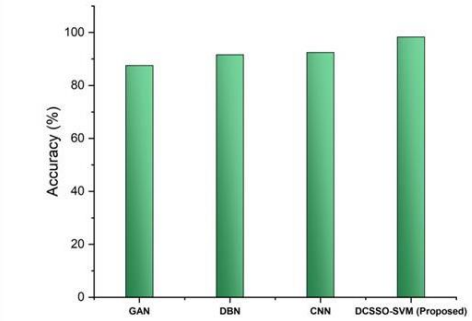


Figure 2 Accuracy.

A statistical metric called precision assesses a categorization or prediction model's accuracy. It measures how many of the overall favorable predictions (TP + FP) were true positive predictions (TP). A lower rate of false positive predictions is shown by a greater precision score, which suggests that the model is better able to recognize positive occurrences. The precision value of the proposed and used procedures was displayed in Table 1 and Figure 3. While GAN, DBN, and CNN only achieve a precision value of 88.27 %, 90.72 %, and 92.84 %, respectively, the suggested technique, DCSSO-SVM, achieves a precision value of 97.45 %. Compared to the conventional methods, the DCSSO-SVM methodology has a higher precision value.

$$Precision = TP / (TP + FP) \quad (22)$$

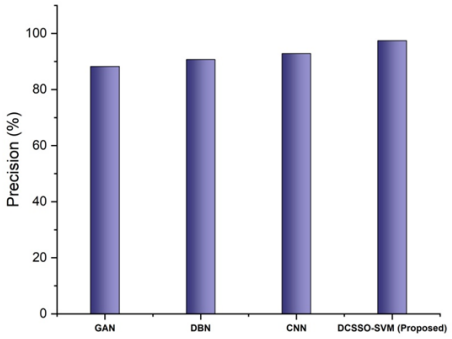


Figure 3 Precision.

A statistical metric called recall, sometimes referred to as sensitivity or real-positive rate, is used to evaluate how comprehensive a classification or prediction model is. Out of all actual positive cases (TP + FN), it estimates the percentage of genuine positive predictions (TP). Greater accuracy in properly identifying positive occurrences among all real positive cases is indicated by a model with a higher recall value since it shows a lower percentage of incorrect negative predictions. The Recall value of the proposed and used procedures was displayed in Table 1 and Figure 4. While GAN, DBN, and CNN only achieve a Recall value of 87.31 %, 89.49 %, and 91.65 %, respectively, the suggested technique, DCSSO-SVM, achieves a Recall value of 98.12 %. The DCSSO-SVM approach provides a greater Recall value as compared to traditional approaches. The recall equation is as follows:

$$Recall = TP / (TP + FN) \quad (23)$$

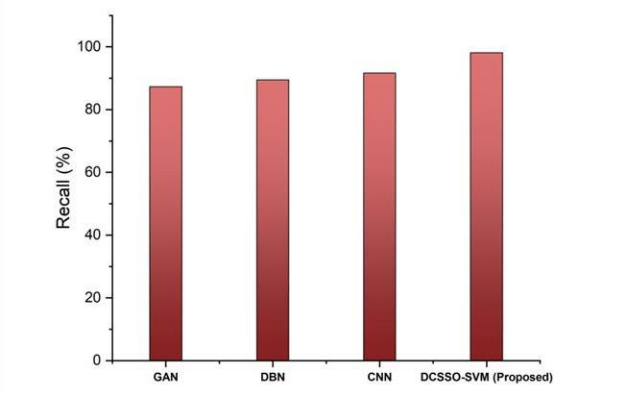


Figure 4 Recall.

The total effectiveness of a prediction model is assessed using the F1-score, a statistical metric that combines precision and recall. It offers a harmonic mean of memory and accuracy, giving each metric equal weight. The F1 score has a range of 0 to 1, with 1 representing the highest attainable value. When there is a disparity across positive and negative instances in the dataset, it is frequently employed. The F-score of the proposed and used procedures was displayed in Table 1 and Figure 5. While GAN, DBN, and CNN only achieve an F-score of 89.19 %, 89.98 %, and 91.67 %, respectively, the suggested technique, DCSSO-SVM, achieves an F-score of 98.27 %. The DCSSO-SVM approach provides a greater F-score as compared to traditional approaches. The F1-score formula is provided by:

$$F1 - score = 2 * (Precision * Recall) / (Precision + Recall) \quad (24)$$

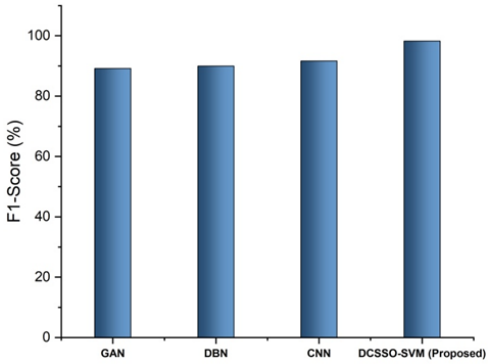


Figure 5 F1-score.

5. Conclusion

This paper introduced a novel diversified convergent squirrel search optimized support vector machine (DCSSO-SVM) for medical applications using data mining. This strategy improved the capacity to identify subtle connections and trends, resulting in more precise diagnoses, individualized treatment regimens, and superior patient care. The study used ARFF files, and JNCC2 loads data. This plain text format was created first for WEKA. We analyzed the accuracy, precision, recall, and f-score for the proposed method (DCSSO-SVM). And the accuracy was 98.34 %, the precision was 97.45 %, and the recall and f-score were 98.12 and 98.27 %. The effectiveness of the model can be improved in future work for the DCSSO-SVM method in medical imaging data mining by investigating additional methods for extracting features, such as deep learning architectures. The approach could be improved in future studies to handle larger datasets and increase computational effectiveness. To prove the generalizability and robustness of the technique, further verification experiments on various medical imaging datasets would be necessary to undertake.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Albahri AS, Hamid RA, Alwan JK, Al-Qays ZT, Zaidan AA, Zaidan BB, Albahri AOS, AlAmoodi AH, Khlaf JM, Almahdi EM, Thabet E (2020) Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. Journal of medical systems, 44:1-11.

Amin MS, Chiam YK, Varathan KD (2019) Identification of significant features and data mining techniques in predicting heart disease. Telematics and Informatics 36:82-93.

Boruah I, Kakoty S (2019) Analytical study of data mining applications in malaria prediction and diagnosis. Int. J. Comput. Sci. Mob. Comput 8:275-284.

Dash S, Shakyawar SK, Sharma M, Kaushik S (2019) big data in healthcare: management, analysis, and prospects. Journal of Big Data 6:1-25.

Elizabeth L, Mishra VP (2018) Big data mining methods in medical applications. In Medical Big Data and Internet of Medical Things, pp. 1-23. CRC Press.

Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 321:321-331.

Gomez K, Laffan M, Keeney S, Sutherland M, Curry N, Lunt P (2019) Recommendations for the clinical interpretation of genetic variants and presentation of results to patients with inherited bleeding disorders. A UK Haemophilia Centre Doctors' Organisation Good Practice Paper. Haemophilia 25:116-126.

Guha S, Kumar S (2018) the emergence of big data research in operations management, information systems, and healthcare: Past contributions and future roadmap. Production and Operations Management 27:1724-1735.

Kaur M, Singh D (2020) Fusion of medical images using deep belief networks. Cluster Computing 23:1439-1453.

Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I (2017) Machine learning and data mining methods in diabetes research. Computational and structural biotechnology journal 15:104-116.

Khajouei R, Abbasi R, Mirzaee M (2018) Errors and causes of communication failures from hospital information systems to electronic health record: a record-review study. International journal of medical informatics 119:47-53.

Komi M, Li J, Zhai Y, Zhang X (2017) Application of data mining methods in diabetes prediction. In 2017 2nd international conference on Image, Vision and Computing (ICIVC), pp. 1006-1010. IEEE.

Lan K, Wang DT, Fong S, Liu LS, Wong KK, Dey N (2018) A survey of data mining and deep learning in bioinformatics. Journal of medical systems 42:1-20.

Mughal MJH (2018) Data mining: Web data mining techniques, tools, and algorithms: An overview. International Journal of Advanced Computer Science and Applications 9.

Siripipatthanakul S, Bhandar M (2021) Qualitative research factors affecting patient satisfaction and loyalty: a case study of smile family dental clinic. International of Trend in Scientific Research and Development 5:877-896.

Souri A, Hosseini R (2018) A state-of-the-art survey of malware detection approaches using data mining techniques. Human-centric Computing and Information Sciences 8:1-22.

Srinivas K, Rani BK, Govrdhan A (2010) Applications of data mining techniques in healthcare and prediction of heart attacks. International Journal on Computer Science and Engineering (IJCSE) 2:250-255.

Wu WT, Li YJ, Feng AZ, Li L, Huang T, Xu AD, Lyu J (2021) Data mining in clinical big data: the frequently used databases, steps, and methodological models. Military Medical Research 8:1-12.

Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ (2021) Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement, and perspectives. Neurocomputing 444:92-110.

MULTIDISCIPLINARY SCIENCE JOURNAL

CONFERENCE PAPER
Published Online: August 10, 2023
https://doi.org/10.31893/multiscience.2023ss0125

Advancing Engineering, Science, and Technology: Innovations and Developments in Research and Development

Smart evaluation of videos for unusual-event identification in automated vehicle monitoring systems

Surjeet Yadav[a] | **Anu Sharma**[b] | **Ananta Ojha**[c]

[a]Vivekananda Global University, Jaipur, India, Assistant Professor, Department of Mechanical Engineering.
[b]Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Assistant Professor, College of Computing Science And Information Technology.
[c]Jain (deemed to be) University, Bangalore, India, Professor, Department of Computer Science and Information Technology.

Abstract Many methods have been tried to perfect vehicle recognition, which includes quickly locating the same objects in cameras with different sensor types and perhaps propagating trustworthy information across cameras to improve detection. In this study, we look at various important issues and provide a new method for spotting certain abnormalities. We suggest an automated vehicle monitoring system (AVMS) to enhance detection accuracy. If the current reading from a certain local monitor is out of the ordinary, that monitor will issue a notification, and all of those alerts will go toward deciding whether or not an abnormal event has occurred. The effective implementation of any large-scale surveillance system relies on our algorithm meeting several parameters. It simply takes a few minutes to set up and is completely hands-free. The experimental findings demonstrate the effectiveness and simplicity of the suggested approach, with the assembled image quality being on the level with that of state-of-the-art techniques.

Keywords: unusual events, video analysis, automated vehicle, monitoring system, license plate recognition

1. Introduction

The use of video surveillance systems is becoming more widespread, and nowadays, you can find them installed in a wide variety of critical locations, such as airports, banks, public transit, or bustling city centres. People often have a positive attitude towards the enhanced feeling of safety provided by video surveillance, yet, they frequently have a negative attitude towards the associated loss of privacy. This valid fear often causes a delay in the implementation of video monitoring systems. The method uses an object-oriented representation of the environment as its foundation (Zhu et al 2020). The system re-renders a changed movie in response to the access control authorizations provided by the end user. Whenever an image is re-rendered, some parts will have their content removed. Thus, the information that is significant to the scene is maintained, but the details that are sensitive to privacy are not provided. In the same way, a privacy buffer employs privacy filters to operate on incoming sensor data to either prohibit access to sensitive information or change data to delete private communication (Nasaruddin et al 2020).

Figure 1 depicts a system diagram for the video surveillance system. Multiple inexpensive security cameras, either wired or wireless, are linked to a central computer to create this system. The server is then responsible for processing the video. To begin, a video analysis module will locate ROIs. The video is then compressed to be stored and sent with minimal resources. At the same time, the ROIs are subjected to a scrambling process (Hidayat et al 2020). Consequently, the persons in the scene are unrecognizable, but the background is unobscured, solving the problem of invasion of privacy. Finally, the system allows for Internet-based access from various client devices so that both live and archived videos may be viewed. Automatic number plate recognition is a method of reading a vehicle's plate number from an image or video of the car in motion or at rest (Aslam et al 2022). This method works well for keeping an eye on moving objects. Powerful license plate readers have a wide range of applications, including thwarting theft, classifying illegal vehicles, collecting electronic tolls in a way that is unique to each location, tracking the flow of traffic inside a building, identifying speeders, identifying authorized cars in a parking lot, saving time by removing the need for human verification of parking passes, and much more. Issues such as noisy picture inputs, occlusion, vehicle orientations, various number plate kinds, additional images on number plates, non-standard sizes, low quality of the camera, and so on make number plate identification and recognition a difficult challenge. Existing systems often assume too simple conditions compared to real-world situations, such as only functioning with stationary cameras at a fixed viewing angle and resolution and only with a fixed license plate template (Mudgal et al 2021). As a result, automated activity identification is a need in many video surveillance uses. A proper automated vehicle

identification algorithm may significantly lessen the burden for people, making automatic vehicle recognition an integral aspect of video monitoring. Nonetheless, no universal vehicle identification system can reliably and automatically identify cars of interest across all surveillance applications. Vehicles of interest might vary depending on the application. It is challenging to develop a suitable general vehicle recognition algorithm because many surveillance applications need the activity recognition algorithm to deal with unusual events for which no training data and to be flexible enough to add new activities of interest to the system (Sharma and Gangadharappa). In this article, we go deep into the state of the art in video surveillance and provide solutions to various pressing issues. To improve detection precision, we propose using AVMSs. The recorded footage or pre-processed data is sent to the central monitoring facility for storage, analysis, and dissemination.

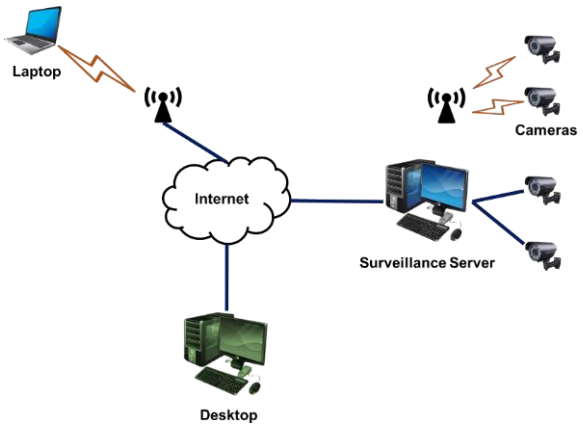


Figure 1 Framework of a video monitoring system.

The rest of the sections of the paper are structured as follows. In Section 2, we give a literature review. Our detection system, is described fully in Section 3. In Section 4, we detail the experiment and its findings. Finally, we conclude the task in Section 5.

2. Related Work

The article (Singhet et al 2019) presented a strategy for recognizing uncommon vehicles by first establishing a model for identifying more common cars. The issue of insufficient training data may be addressed using this approach, which suggests how to do it. However, they generate all unusual event models by altering the basic ordinary event model. This is even though, in nature, normal vehicle recognition and odd occurrences might be highly distinct. The detection, localization, and categorization of various objects in images have each been the subject of several strategies presented in the research literature (Tamilmani et al 2022). The issue with automated LPR (ALPR) is that it recognizes the number of a vehicle's number plate based on photographs of the plate. Extensive image processing, including object identification and pattern recognition, is required to implement the solution. Plate variances from country to country, the plate's position inside the image, varied widths, diverse color combinations, plate contamination owing to other frames, and photographs are some of the issues associated with proper number identification. The study (Ahmed and Jeon 2021) presented an extensive and completely automated system for the detection and identification of license plates. It was constructed using a series of deep CNNs in conjunction with algorithms. The CNNs underwent training and optimization to improve their resilience in various situations, including changes in position, lighting, occlusion, and others. These CNNs are versatile enough to operate with a broad range of license plate templates, each with a unique size, background, and font combination. This article discussed several facets of LPR and offered solutions to all of the issues that were brought up. The research, however, is limited to still photographs and includes no videos. The article (Şengönül et al 2023) investigated the effect that these external influences had on the accuracy and consistency of the LPR procedure. Some works have expressly concentrated on tackling these issues as their primary emphasis. One of the ways to make ALPR resistant to these challenges is to recreate all of the different sorts of challenges in a manufactured database and then test a classification algorithm on this dataset to come up with a solution that has a good chance of working. The automatic Number Plate Recognition System (Hatirnaz et al 2020) used a pipeline based on image processing. They used morphological approaches together with edge detection, histogram manipulation, and plate localization for character segmentation and plate localization. They used a synthetic neural network for character identification and categorization. However, the work's character identification and detection procedures are simplistic and do not account for restrictions such as angle, distance from the vehicle, or poor picture quality. However, this technology is only applicable for use in offline detection, and no information is provided on how to adapt it for use in real-time. The study (Ghoniem et al 2022) examined the challenge of LPR when faced with various designs and fonts. A demanding synthetic dataset with different font styles and number plate designs was used for testing after multiple templates matching for character identification and noise reduction. It has been shown that both the false positive and false negative rates are rather low.

3. Proposed Methodology

In this study, we investigate the problem of AVMS resolution in multi-camera video surveillance systems. In this study, we examine the issue of AVMS in multi-camera methods used for security purposes. Each screen is a separate object that takes local, low-level observations from the input video. The monitor's observation may be the amount of local optical flow or the flow direction at present. The automated monitoring system performs real-time analysis of the video streams, searching for any abnormalities or unexpected occurrences using powerful computer vision algorithms. The system also checks the car's actions against normative driving characteristics and models.

3.1. Identify the suspect's vehicles in a video frame

To speed up the process, we should eliminate the impossible frames, which contain no relocating objects. Background estimate in video sequences is the starting point for this effort. Often, dirty backgrounds may be removed by averaging k frames from a video clip. By averaging, we can remove all of the clutter from the foreground and be left with a clear backdrop. Here's how we can write out the equation (1):

$$P(y,x) = (\sum_{j=1}^{r} l(y,x))/r \qquad (1)$$

Where $l(y,x)$ = input video sequences,
$P(y,x)$ = estimation background,
r = number of average frames,
r = set 50.

After obtaining an estimated backdrop, we compare it frame by frame using the difference image $l(y,x)$ to get images of moving objects, in this case, potential images of a car driven by a suspect. The following equation (2) gives its numerical value:

$$L(y,x) = \begin{cases} l(y,x) if\ |P(y,x) - l(y,x)| \geq Th \\ 0\ if\ |P(y,x) - l(y,x)| < Th \end{cases} \qquad (2)$$

Where Th = judgment threshold for determining whether or not the pixel is changing. In our study, we use a Th (30) for $l(y,x)$ to calculate whether or not a video sequence consists of frames of a moving object; if it does, then the sequence is used as input for the next processing stage.

3.2. Removing a suspect's car from a video frame captured by many cameras

In this phase, we suggest using a colored-based component technique to scan footage from various cameras to roughly ensure that suspicious vehicles are present in video frames from numerous cameras. The portions of the sequentially tone mapping $l(e^*(y,x))$ and $l(p^*(y,x))$ may be determined by equation (3):

$$\begin{cases} l(e^*(y,x)) = |lsl(e^*(y,x)) - lsm(e^*(y,x))| \\ l(p^*(y,x)) = |lsl(p^*(y,x)) - lsm(p^*(y,x))| \end{cases} \qquad (3)$$

The contrast image and noise remove the image of the suspect's car in the video frames corresponding to $lsl(e^*(y,x))$, and $lsl(p^*(y,x))$. The contrast image and noise removal image of the suspect's car in video frames taken by various cameras are represented by the functions$lsm(e^*(y,x))$, and $lsm(p^*(y,x))$ accordingly.

The next step is to use a fusion technique to roughly confirm the legitimacy of video frames featuring a suspicious car. The other images will be filtered to determine the true identity of the vehicle in question.

$$T(y,x) = \begin{cases} lsm(y,x) if\ (l(e*(y,z) < D)(l(p*(y,z) < D) \\ 0 \quad if\ (l(e*(y,z) \geq D)(l(p*(y,z) \geq D) \end{cases} \qquad (4)$$

In this study, D was dependent on the colour values of the suspect's vehicle, and $T(y,x)$ represents the outcomes; if it is smaller than D, then $lsm(y,x)$ is the suspect vehicle frames.

3.3. License Plate location

To facilitate accurate plate localization, the image has been optimized. To begin, the image of the car is cleaned of noise to improve accuracy. Several morphological procedures were conducted once the image was converted to grayscale, which brought out sharper edges and more contrast. The next stage is to apply blurring to provide consistency throughout the image, and then adaptive thresholding is used to eliminate irrelevant details (things that aren't license plates). Methods such as grayscale conversion, histogram equalization, sharpening, and OTSU masking are often used.

The license plates x and y coordinates must be determined once the image has been cleaned of background distractions. Algorithm 1 outlines the procedures that must be followed for this to occur. From the pre-processed image, contours are extracted. According to the Open CV docs, a contour is a line drawn across a set of points that shares its color or intensity with the whole group. The outlines may be used for several purposes, including shape analysis and identifying objects. Plate characters' contour coordinates were utilized to extract regions of interest, and their respective bounds, areas, aspect ratios, widths, heights, and angles were recorded as Python class objects representing every contour character. The potential plate set is then utilized as input in the subsequent step. This plate detection from imagine process has been encapsulated in Algorithm 1.

Algorithm 1: License plate detection

Input: Image file name as imageOriginal
Output: sequence of lists, each of which provides the contours necessary to complete the license plate.
imageGrey, imgThresh ← Preprocess (imageOriginal)
Contours ←opencv.findContours (imageOriginal)
listofpossibleChars=[]
For c in outlines:
x,y,w,h←opencv.boundingRect (c)
If $(z * u > 80\ and\ u > 2\ and\ z > 8\ and\ 0.25 < (\frac{u}{z}) < 1.0)$:
listofpossibleChars. Append (c)
listOfListsOf MatchingCharsInscene←
GroupMatchingPlates (listofpossibleChars)
listOfPossiblePlate- []
For plate in listOfListsOfMatchingCharsInscene:
Possible plate← extract plate (plate)
If possible, plate. imgplate not none:
listOfPossiblePlate.append (possible plate)
Return listOfPossiblePlate

Several processes, such as contour detection, geometric constraint setting, grouping, etc., are employed for the preprocessed data in Algorithm 1. These procedures take the original picture and remove all contours, filter out the ones that don't seem like characters based on their geometry, and then extract the number plate. The method was applied to every conceivable dish, and the results were compiled into a single, convenient list. The most likely candidate license plates for this item have been selected.

The input image has been processed to remove the noise in Figure 2. When talking about an input image, "noise reduction" signifies the method of eliminating distracting and unwelcome visual artifacts and disruptions so that the final product is more pleasant to the eye (Hashmi et al 2019). Figure 2a depicts an automobile's input image, whereas Figure 2b shows the output. The intensity values in a grayscale picture generally run from 0 to 255, with 0 denoting black (the lowest possible intensity) and 255 denoting white (the highest possible intensity). Shades of grey are represented by the intermediate values. Image contrast enhancement used to create Figure 2c. Enhancing contrast in a picture is making the distinction between one part of the image and another more apparent in terms of brightness and color. Its purpose is to enhance the picture such that it is either aesthetically pleasing or analytically more straightforward. The image in Figure 2d was blurred to get rid of the noise. In image processing, blurring a picture to eliminate noise is a typical method for smoothing out imperfections. Blurring a picture to get rid of noise smooths out the blemishes and makes the image seem better overall. Image noise may result from several sources, including but not limited to sensor limits, transmission mistakes, and ambient variables. To improve image quality, noise reduction attempts to lessen the effect of noise while keeping crucial image information intact and preventing the introduction of distracting artifacts.

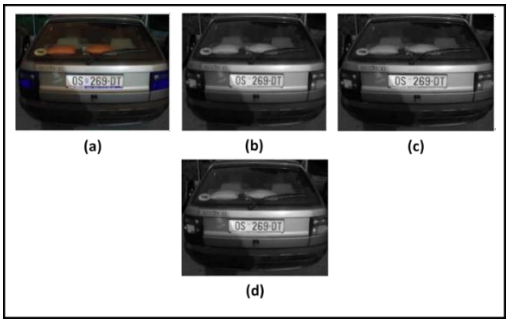


Figure 2 Image preprocessing procedure.

3.4. Character Segmentation

License plates that have been cropped for OCR may now be seen here. The first step in using OCR is to isolate the characters on license plates using cropping. In most cases, the license plate outline is too small for accurate character recognition. This meant that any unnecessary symbols on the plate needed to be cleaned. The methods for image processing described in the second phase were used once more to eliminate the noise in the plate image. Figure 3 shows how they extracted objects and removed everything except characters from the plate's image that could be read as numbers or letters.

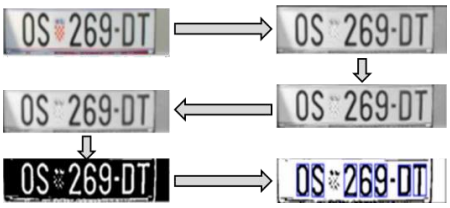


Figure 3 OCR and Noise Reduction for License Plates.

The images were then scaled up, made grayscale, and had their contrast raised. The entire image is then converted to black and white using dynamic thresholding. To produce black text on a white backdrop, use the binary inverse. The contrast made it easier to distinguish across character types. Individual characters were chopped out of the license plate together with the other characters except the 36 (0-9 and A-Z) characters. Each divided character has its height and breadth raised by factors of 1.3 and 1.6. Accordingly, a white border has been included.

To prepare each character picture for the character recognition module, cropping, and pre-processing are performed. Specified number plate outlines are read in as input at this stage. One of the plate's characters is represented by each method. Plate images have had thresholding performed on them to remove the contours of certain characters. Better OCR is achieved by adding borders and enlarging the clipped picture. The processed and cropped text image is then given into the character model, which determines which symbol this outline is and associates it with the license plate phrase.

4. Result and Discussion

Here, we provide experimental findings from several security cameras using our suggested AVR. At the moment, we only capture a single set of test data from many cameras on the same day using known cars. Precision and accuracy are the two most often used indicators. We evaluate our suggested technique in comparison to state-of-the-art approaches like deep neural networks (DNN) (Li et al 2018) and cognitive Internet of Vehicles (CIoV) (Arooj et al 2022). Accuracy and precision are two characteristics of a highly effective system.

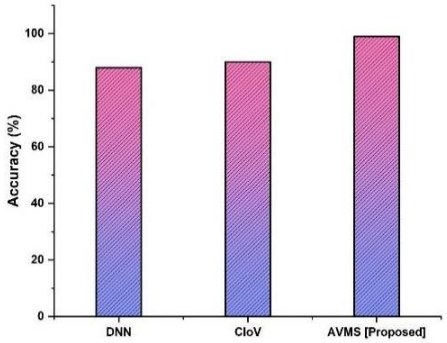


Figure 4 Accuracy.

The proportion of useful information included within the total quantity of data collected is shown and described in Figure 4. There is a pace at which accuracy can be assessed. It's a measure of how likely the information being obtained is useful. In comparison to the 88 and 90 accuracy scores achieved by the DNN and CIoV, respectively, our suggested technique achieved 99 accuracy. Compared to the other two approaches, the efficiency of the accuracy value is higher. The accuracy of the proposed and current approaches is compared in Table 1.

Table 1 Accuracy.

	Accuracy (%)
DNN	88
CIoV	90
AVMS [Proposed]	99

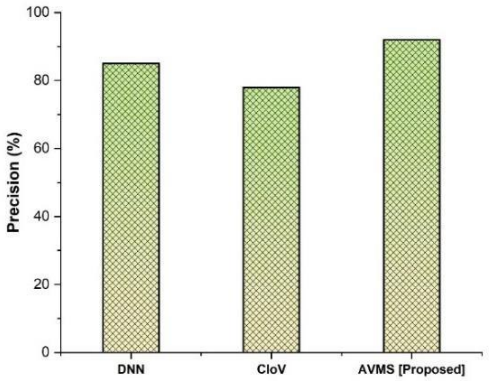


Figure 5 Precision.

Precisely how much useful information has been collected from the universe of available useful information is shown and described in Figure 5. The rate of precision is a measurable quantity. It's a measure of how likely useful information will be found. While the DNN achieved a precision of 85 and the CIoV of 78, our suggested approach achieved a score of 92. Compared to the other two methods, the precision value is the most effective. The precision of the proposed and current systems is compared in Table 2.

Table 2 Precision.

	Precision (%)
DNN	85
CIoV	78
AVMS [Proposed]	92

AVMS systems use deep learning algorithms to analyze and decipher several vehicle data sources, providing real-time monitoring and analysis for increased security, effectiveness, and decision-making. DNNs are computationally demanding models that need a lot of memory and computing power. It may not be easy to implement and operate these models in real-time applications, such as autonomous vehicle monitoring systems, and may need specialized hardware or distributed computer configurations. By using the capabilities of cognitive computing, CIoV empowers AVMS to make wise and educated judgments. Large volumes of real-time data from several sensors and sources, such as vehicle sensors, cameras, GPS, traffic data, and meteorological conditions, may be analyzed by it. This improves safety and efficiency by enabling prompt reactions to possible dangers, traffic jams, and other urgent circumstances. Integrating numerous technologies, infrastructure, and communication networks is necessary to implement CIoV systems. Costs for development, deployment, and maintenance may rise as a result of this complexity. It may also be difficult to guarantee compatibility and standardization across various cars, manufacturers, and infrastructure suppliers. By continually monitoring cars and their surroundings, AVMS contributes to increased road safety. Compared to other approaches, AVMS is superior.

5. Conclusion

In this study, we provide a new approach to AVMS for use with a network of cameras specifically designed for monitoring public spaces. It is possible to use connected component analysis and an adaptive LPR approach in a single license plate location algorithm. We employ shift filtering, a technique for segmenting text, to break down license plate data. For vehicle video sequences, you may also apply adaptive banalization for segmentation. To demonstrate the efficiency and resilience of the proposed approach, we also ran extensive simulations utilizing a variety of situations. Based on our experiments, we know that 99 percent of our AVMS detections are accurate. A particular LPR system's processing rates may be further increased by employing the parallelization strategy during different phases of development.

Ethical considerations

Not applicable.

Declaration of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

Reference

Ahmed I, Jeon G (2021) A real-time person tracking system based on SiamMask network for intelligent video surveillance. Journal of Real-Time Image Processing 18:1803-1814.

Arooj A, Farooq MS, Umer T, Shan RU (2022) Cognitive internet of vehicles and disaster management: a proposed architecture and future direction. Transactions on Emerging Telecommunications Technologies 33:e3625.

Aslam N, Kolekar MH (2022) Unsupervised anomalous event detection in videos using spatio-temporal inter-fused autoencoder. Multimedia Tools and Applications 1-26.

Ghoniem NA, Hesham S, Fares S, Hesham M, Shaheen L, Halim ITA (2022) Intelligent Surveillance Systems for Smart Cities: A Systematic Literature Review. Smart Systems: Innovations in Computing: Proceedings of SSIC 2021:135-147.

Hashmi SN, Kumar K, Khandelwal S, Lochan D, Mittal S (2019) Real time license plate recognition from video streams using deep learning. International Journal of Information Retrieval Research (IJIRR) 9:65-87.

Hatirnaz E, Sah M, Direkoglu C (2020) A novel framework and concept-based semantic search Interface for abnormal crowd behaviour analysis in surveillance videos. Multimedia Tools and Applications 79:17579-17617.

Hidayat F, Hamami F, Dahlan IA, Supangkat SH, Fadillah A, Hidayatuloh A (2020) Real-Time Video Analytics Based on Deep Learning and Big Data for Smart Station. In Journal of Physics: Conference Series 1577:012019. IOP Publishing.

Li H, Wang H, Liu L, Gruteser M (2018) Automatic unusual driving event identification for dependable self-driving. In Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems, pp. 15-27.

Mudgal M, Punj D, Pillai A (2021) Suspicious action detection in intelligent surveillance system using action attribute modelling. Journal of Web Engineering:129-146.

Nasaruddin N, Muchtar K, Afdhal A, Dwiyantoro APJ (2020) Deep anomaly detection through visual attention in surveillance videos. Journal of Big Data 7:1-17.

Şengönül E, Samet R, Abu Al-Haija Q, Alqahtani A, Alturki B, Alsulami AA (2023) An Analysis of Artificial Intelligence Techniques in Surveillance Video Anomaly Detection: A Comprehensive Survey. Applied Sciences 13:4956.

Singh G, Khosla A, Kapoor R (2019) Crowd escape event detection via pooling features of optical flow for intelligent video surveillance systems. Int. J. Image Graph. Signal Process 10:40.

Tamilmani G, Uma S, Sivakumari S (2022) Unusual Event Detection from Surveillance. In Computational Methods and Data Engineering: Proceedings of ICCMDE 2021, pp. 477-487. Singapore: Springer Nature Singapore.

Zhu S, Chen C, Sultani W (2020) Video anomaly detection for smart surveillance. In Computer Vision: A Reference Guide, pp. 1-8. Cham: Springer International Publishing.