Hybrid Deep Ensemble Learning with Metaheuristic Optimization for Heart Disease Prediction

B Shamreen Ahamed
Department of CSE, School of computing, sathyabama institute of science and technology, Chennai, India.
Post-doctoral Research fellow, London Metropolitan University, London, UK.
s61537104@gmail.com

Bal Virdee
Senior Professor & Head of Communications Technology Research Centre, London Metropolitan University, London, UK,
b.virdee@londonmet.ac.uk,

Ashish khanna,
Maharaja Agrasen Institute of Technology, Delhi, India,
ashok13raju@gmail.com

R G Vidhya
Research Faculty, Ulm university, Ulm, Germany, Department of ECE, HKBK College of Engineering, Bangalore, India.
vidhya50.ece@gmail.com

S Sivashankar
Department of CSE, KGRCET, Hyderabad, India,
drsivashankars@gmail.com

Abstract -- The major causes of deaths due to heart disease rank among the top for global health concerns, with predicting heart disease for diabetic patients a major challenge. This research proposes a deep ensemble learning with metaheuristic optimization framework for heart disease prediction. The workflow of this model includes performing extensive preprocessing of the provided dataset, followed by an optimization algorithm-based feature selection. Furthermore, the proposed model utilizes a weighted fusion method for combining various health data sources, convolution-pooling, attention, and spatial-temporal learning. Ensemble stacking learning is also used with an SVM learner, with Pufferfish Optimization Algorithm used for fine-tuning. As indicated, the proposed model was successful in achieving 0.980 accuracy, 0.970 precision, 0.990 recall, and 0.950 F1-score for Dataset 1, with 0.990 accuracy, 0.980 precision, 0.970 recall, and 0.980 F1-score for the Cleveland dataset.

Keywords: Optimization, Deep Learning, Disease Prediction, Ensemble Learning.

I. Introduction

The major causes of deaths associated with heart disease are among the major concerns globally. One of the major challenges is associated with the prediction of heart disease among diabetic patients. This, among other challenges, is caused by aspects associated with the complexity of the various health data sources [1]. It is noted that machine learning (ML) and shallow deep learning (DL) algorithms were unable to predict heart disease among various individuals due to an inability to cope with the complexity of the various health data sources [2]. The proposed model involves carrying out extensive preprocessing of given datasets [3, 4], followed by applying the Golden Jackal Optimization Algorithm in selecting features. The proposed model also employs a weighted fusion scheme [5], convolution-pooling, attention, spatial-temporal, ensemble stacking learning with an SVM learner, as well as using a Pufferfish Optimization Algorithm. It is indicated that the proposed model managed to achieve 0.980 accuracy, 0.970 precision, 0.990 recall [6], 0.950 F1-score, 0.990 accuracy, 0.980 precision, 0.970 recollect, as well as 0.980 F1-score [7]. Furthermore, apart from assisting physicians in delivering timely interventions, early diagnosis will enable the system to decrease complications and improve the survival rate of patients [10]. The use of Artificial Intelligence (AI) is also quite promising in the medical field, especially the use of Machine Learning (ML) and Deep Learning (DL) in the recent past [11]. The ability of AI, especially ML and DL, can automatically examine the complex and vast medical records, identify the complex patterns, and create a predictive model while assisting the physicians in decision-making [12]. The use of AI models predicted the risk of serious diseases, including COVID-19, cancer, and heart problems [13], [14]. Additionally, the hybrid ML-DL models can also be used, where the advantages of the models of ML and DL can be combined in a hybrid form and yield positive results [15]. The structure followed for the proposed research work is as follows: A short report on some of the latest studies on the prediction will be presented in section 2. A description of the methodology created will be presented in section 3. The results of the improved model will be presented in section 4. A description of the study's findings will be presented in section 5 [8], [9].

II. Literature Survey

The prediction of heart disease has gained tremendous research interest due to the increased incidence of cardiovascular disease (CVD) among population groups globally [16]. Current research attempts have emphasized the exploitation of ensemble learning, feature selection, and hybrid optimization techniques to improve the prediction performance and avoid redundancy and computation [17].

Further, Gradient Boosting-based Sequential Feature Selection (GBSFS) is also applied to detect the best features [18,19]. By incorporating the features and using stacking, the model resulted in an accuracy of 98.78% using 11 features selected by GBSFS [20,21]. These features can be very useful in improving the efficiency of prediction and eliminating any redundancy in features. Although the model can prove to be accurate and effective, it involves high complexity and good data from healthcare [22,23]. The researchers proposed a coronary heart disease (CHD) risk prediction model named Feature Weighted Hybrid Voting Ensemble (FWHVE). The proposed approach relies on a two-stage system [24].

- From the reviewed literature, the following trends can be identified:
- Techniques such as stacking and hybrid voting always produce better results than standalone classifiers.
- Feature selection and optimization techniques (such as gradient boost selection, forward selection, and genetic algorithms) improve accuracy and reduce redundancy.
- Meta-learning techniques, such as using XGBoost-based stacking models, are being employed for stronger prediction.

Although reported model accuracy is high, challenges remain. Some of these include computational complexity, dataset dependency, non-generalizability due to small dataset size or private data, as well as sensitivity towards various feature engineering methodologies. It is worth noting that only a few articles assess the model's real-time feasibility as well as model explainability in clinical settings. In brief, though recent models for predicting heart disease have shown good performance due to advanced techniques using ensemble and optimization methods, future research areas must be centered on making these models more comprehensible, computationally cheaper, generalized, and useful for clinical settings.

III. Proposed Methodology

In this study, two publicly available heart disease datasets, specifically the large-scale "heart disease" data set on Kaggle and another well-known data set, Cleveland, are used. These two data sets are comprehensive, including extensive clinical characteristics such as age, blood pressure, cholesterol level, and presence of diabetes, as well as lifestyle-related parameters, along with heart diseases or not.

The suggested model for prediction, PO-MIFDELN, has been proposed for the prediction of heart disease for diabetic patients using the concepts of multi-sensor information fusion, deep feature extraction, and ensemble learning with advanced hyperparameters. For both datasets, the initial steps of preprocessing, including missing variable handling, normalization of numerical features, encoding of categorical features, removal of noise, etc., will be performed. For the purpose of improving the efficiency of the model and reducing its dimensionality, the Golden Jackal Optimization Algorithm (GJOA) is utilized for feature selection, wherein the most important features are identified for further processing. The main body of the proposed framework, (Multi-Sensor Information Fusion Deep Ensemble Learning Network (MIFDELN)), is responsible for performing a weighted fusion of multi-source data, where convolutional pooling and attention-based layers are utilized for discriminative feature extraction, and temporal learning is achieved.

Lastly, a number of deep learning base learners, including CNN and LSTM, are integrated using stacked ensemble learning. Then, the outputs from these architectures are combined using an SVM meta-learner for better prediction accuracy and generalization. Lastly, the overall workflow of the provided approach, called POA, is demonstrated in Figure 1. In order to further enhance the predictive capabilities of the model, the Pufferfish Optimization Algorithm (POA) is used to optimizedly tune the parameters of the developed model. This algorithm is based on mimicking the adaptive defense strategies of pufferfish, specifically their adaptability based on changes in their surroundings. In the developed model, the POA is used to intelligently search for optimal parameters that can increase the overall validation accuracy of the model while preventing over-fitting.

There is an exploration of crucial hyperparameters like the rate of learning, batch size, dropout rate, number of hidden layers, number of neurons, and other model-specific parameters. This effectively balances exploration and exploitation for better results. In effect, premature convergence is entirely avoided, and the model converges stably. This adaptive search of parameters ultimately helps the model discover good combinations without wasting much computation.

Henceforth, with the integration of POA, it is ensured that the MIFDELN framework is implemented with hyperparameters that have been optimized, and therefore, robustness and immunity from noise values increase, along with computational efficiency and prediction accuracy. As a result, the optimized model can demonstrate robust and consistent performance even for heart disease prediction on the Kaggle and Cleveland datasets.

IV. Results And Discussion

The first dataset being used within this study has a massive heart disease dataset [29], which contains 57,373 records. It contains 18 independent clinical variables, as well as one dependent variable, which is the class label. It is a binary value, where 0 represents a healthy person, and 1 represents a person with heart disease. Splitting this data into training and test data is required for proper model evaluation, keeping in mind the class balance of this data. From this data, 45,898 samples were used for training, and the remaining 11,475 samples were left for testing.

The second dataset used in this work is the Cleveland Heart Disease dataset, which is a commonly used benchmark dataset in the field of cardiovascular science. This dataset features 1,025 records of patients, of which 13 are independent clinical features, while the dependent feature has 2 classes, signifying the absence or presence of heart disease. In keeping with the first dataset, this too has been divided into training and test sets, with 820 examples being used for that purpose, while the test set consists of 205 examples. Table 1 illustrates the distribution of medical records in the two classes in the two datasets.

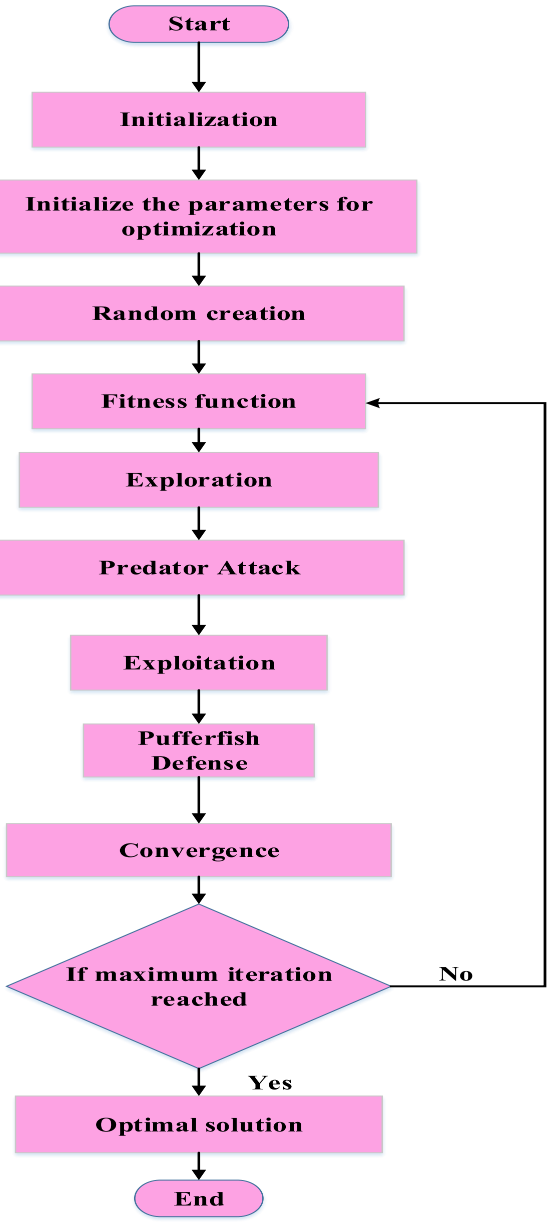


Fig 1: Flowchart of POA

Figure 2 illustrates the hyperparameter sensitivity analysis of the proposed PO-MIFDELN framework on Dataset 1 and the Cleveland dataset. From these results, it can be derived that the model performance is highly dependent on proper hyperparameter tuning. The learning rate ranges to 0.001 as an optimal value, which provides the highest accuracy of 0.940 for Dataset 1 and 0.910 for the Cleveland dataset. Similarly, the batch size study also delivers peak values of 0.940 for Dataset 1 and 0.920 for the Cleveland dataset at an optimal batch size of 32. In the case of dropout rate variation, the best performance was delivered for 0.2, which gave 0.940 and 0.920 for Dataset 1 and Cleveland, respectively. These results reflect that systematic hyperparameter optimization is an important requirement for enhanced generalization and predictive accuracy.

TABLE 1: NUMBER OF MEDICAL RECORDS FOR EACH CLASS IN THE DATASETS

Dataset	Classes	Training Set	Testing Set	Total
Dataset 1	Heart disease	21,898	5,475	27,373
	Healthy	24,000	6,000	30,000
	Total	45,898	11,475	57,373
Cleveland	Heart disease	421	105	526
	Healthy	399	100	499
	Total	820	205	1,025

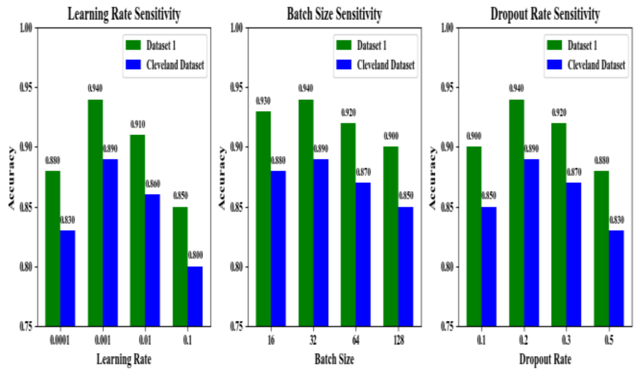


FIGURE 2: SENSITIVITY ANALYSIS OF LEARNING RATE, BATCH SIZE, AND DROPOUT RATE

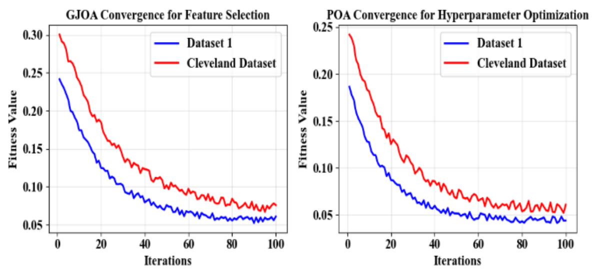
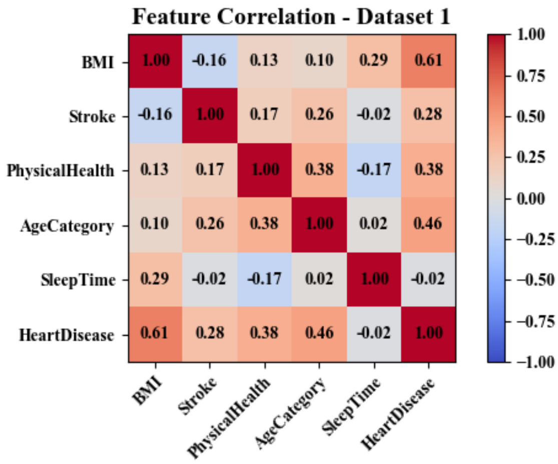


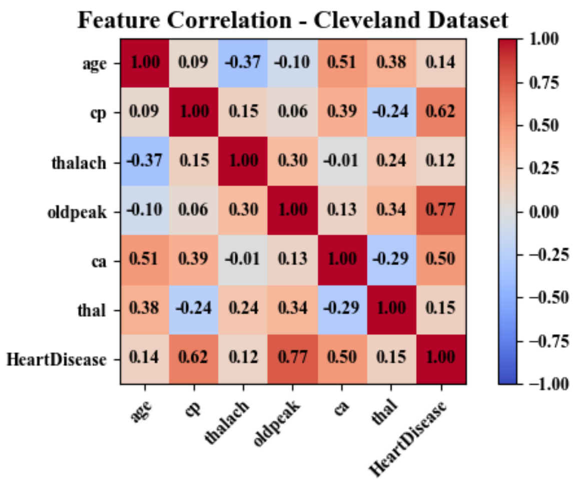
FIGURE 3: CONVERGENCE RATE ANALYSIS

Figure 3 represents the convergence process of these optimization algorithms. The fitness of the Golden Jackal Optimization Algorithm (GJOA) decreases significantly over the initial 30 iterations before slowly converging towards 0.05 fitness values for Dataset 1 and 0.07 fitness values for the Cleveland dataset. Similarly, the convergence of the Pufferfish Optimization Algorithm (POA) has been reported, with steady fitness convergence towards 0.04 fitness values for Dataset 1 and 0.05 fitness values for the Cleveland dataset.

Figure 4 shows the result of the feature importance analysis for both datasets, which identifies the major clinical attributes contributing to heart disease prediction. These results validate the effectiveness, stability, and optimization use of the proposed PO-MIFDELN model.



(A)



(B)

FIGURE 4: FEATURE EXTRACTION ANALYSIS

Figure 5 shows the confusion matrices used to assess the classification capability of the proposed PO-MIFDELN framework on these datasets. As can be seen, for Dataset 1, the proposed framework classified 5,500 healthy patients (true negatives) and 3,200 patients suffering from heart disease (true positives) with high accuracy. However, 500 healthy patients were misclassified as suffering from heart disease (false positives), and 800 patients suffering from the disease were incorrectly classified (false negatives). For the Cleveland UHF datasets, the proposed framework classified 450 healthy patients and 400 patients suffering from heart disease with high accuracy and misclassified 50 healthy patients suffering from heart disease and 100 disease patients as healthy patients. The classification capability of the proposed framework can be considered high due to the high classification accuracy compared to the misclassification rate.

Figure 6 presents the ROC curve comparison of Dataset 1 with the Cleveland dataset. Both datasets show an ROC curve above the diagonal reference line, indicating very good discriminative performance. The proposed model achieved an AUC of 99.67 for Dataset 1, and 99.46 for the Cleveland dataset. The near perfect AUC values indicate an exceptional ability of the framework in distinguishing between patients having healthy and heart disease, further validating the effectiveness and generalization capability of the PO-MIFDELN model across different datasets.

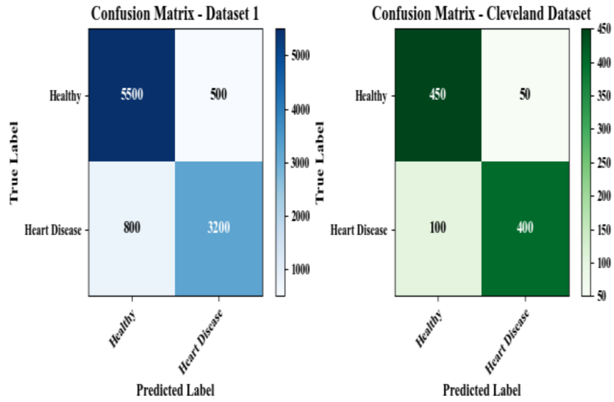


FIGURE 5: CONFUSION MATRICES

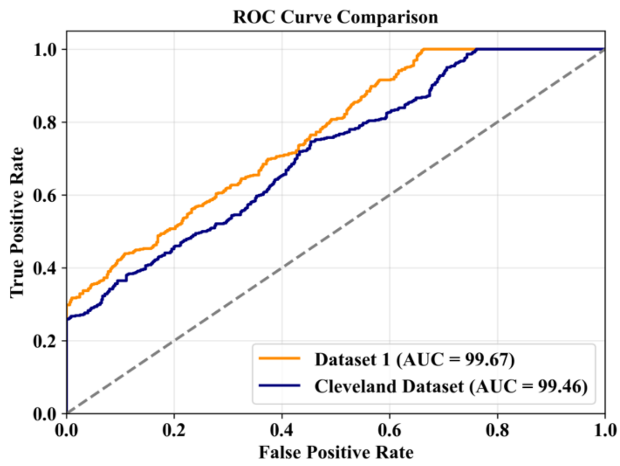


FIGURE 6: ROC COMPARISON

IV. CONCLUSION

The proposed scheme promises an accuracy of 0.990 and an F1-score of 0.980 on the Cleveland dataset and outperformed the majority of the state-of-the-art proposals, including ECSVFM, which, in turn, reported an accuracy of 0.970. These results confirm the diagnostic reliability, stability, and superior discriminative power of the proposed approach. Although the framework yields a promising result, it also suffers from certain limitations. Indeed, the merge of deep ensemble learning and metaheuristic algorithms is computationally complex, and it highly relies on vast quantities of data. Future directions include continued efforts towards designing light and computationally efficient architectures, integration with transfer learning methods, and facilitating real-time deployment in IoT-based healthcare settings. Such methods will help improve the scalability of the model, making it more efficient with reduced resource requirements.

REFERENCES

[1] P. Selvam, N. Krishnamoorthy, S. P. Kumar, K. Lokeshwaran, M. Lokesh, et al., "Internet of Things Integrated Deep Learning

Algorithms Monitoring and Predicting Abnormalities in Agriculture Land," Internet Technol. Letters, 2024, doi: /10.1002/itl2.607.

[2] B. Ramesh, V. V. Kulkarni, Ashwini Shinde, Dinesh Kumar J. R, Prasanthi,et al.," Optimizing EV Energy Management Using Monarch Butterfly and Quantum Genetic Algorithms," International Journal of Basic and Applied Sciences" vol.14, no.2, pp. 311-318, 2025, doi: 10.14419/xaqk1294

[3] K. Maithili, A. Kumar, D. Nagaraju, D. Anuradha, S. Kumar, et al., "DKCNN: Improving deep kernel convolutional neural network-based covid-19 identification from CT images of the chest," J. X-ray Sci. Technol., vol. 32, no. 4, pp. 913–930, 2024, doi: 10.3233/XST-230424.

[4] T. A. Mohanaprakash, M. Kulandaivel, S. Rosaline, P. N. Reddy, S. S. N. Bhukya, et al., "Detection of Brain Cancer through Enhanced Particle Swarm Optimization in Artificial Intelligence Approach," J. Adv. Res. Appl. Sci. Eng. Technol., vol. 33, no. 3, pp. 174–186, 2023, doi: 10.37934/araset.33.2.174186.

[5] Wange N. K., Khan I., Pinnamaneni R., Cheekati H., Prasad J., et al., "β-amyloid deposition-based research on neurodegenerative disease and their relationship in elucidate the clear molecular mechanism," Multidisciplinary Science Journal, vol. 6, no. 4, pp. 2024045–2024045, 2024, doi: 10.31893/multiscience.2024045.

[6] Anitha C., Tellur A., Rao K. B. V. B., Kumbhar V., Gopi T., et al., "Enhancing Cyber-Physical Systems Dependability through Integrated CPS-IoT Monitoring," International Research Journal of Multidisciplinary Scope, vol. 5, no. 2, pp. 706–713, 2024. 10.47857/irjms. 2024.v05i02.0620.

[7] Balasubramani R., Dhandapani S., Sri Harsha S., Mohammed Rahim N., Ashwin N., et al., "Recent Advancement in Prediction and Analyzation of Brain Tumour using the Artificial Intelligence Method," Journal of Advanced Research in Applied Sciences and Engineering Technology, vol. 33, no. 2, pp. 138–150, 2023, doi: 10.37934/araset.33.2.138150.

[8] Chaturvedi A., Balasankar V., Shrimali M., Sandeep K. V., et al., "Internet of Things Driven Automated Production Systems using Machine Learning," International Research Journal of Multidisciplinary Scope, vol. 5, no. 3, pp. 642–651, 2024, doi: 10.47857/irjms.2024.v05i03.01033.

[9] Saravanakumar R., Arularasan A. N., Harekal D., Kumar R. P., Kaliyamoorthi P., et al., "Advancing Smart Cyber Physical System with Self-Adaptive Software," International Research Journal of Multidisciplinary Scope, vol. 5, no. 3, pp. 571–582, 2024, doi: 10.47857/irjms.2024.v05i03.01013.

[10] Vidhya R. G., Surendiran J., Saritha G., "Machine Learning Based Approach to Predict the Position of Robot and its Application," Proc. Int. Conf. on Computer Power and Communications, pp. 506–511, 2022, doi: 10.1109/ICCPC55978.2022.10072031.

[11] Sivanagireddy K., Yerram S., Kowsalya S. S. N., Sivasankari S. S., Surendiran J., et al., "Early Lung Cancer Prediction using Correlation and Regression," Proc. Int. Conf. on Computer Power and Communications, pp. 24–28, 2022, doi: 10.1109/ICCPC55978.2022.10072059.

[12] Vidhya R. G., Seetha J., Ramadass S., Dilipkumar S., Sundaram A., Saritha G., "An Efficient Algorithm to Classify the Mitotic Cell using Ant Colony Algorithm," Proc. Int. Conf. on Computer Power and Communications, pp. 512–517, 2022, doi: 10.1109/ICCPC55978.2022.10072277.

[13] Sengeni D., Muthuraman A., Vurukonda N., Priyanka G., et al., "A Switching Event-Triggered Approach to Proportional Integral Synchronization Control for Complex Dynamical Networks," Proc. Int. Conf. on Edge Computing and Applications, pp. 891–894, 2022, doi: 10.1109/ICECAA55415.2022.9936124.

[14] Vidhya R. G., Rani B. K., Singh K., Kalpanadevi D., Patra J. P., Srinivas T. A. S., "An Effective Evaluation of SONARS using Arduino and Display on Processing IDE," Proc. Int. Conf. on Computer Power and Communications, pp. 500–505, 2022, doi: 10.1109/ICCPC55978.2022.10072229.

[15] Kushwaha S., Boga J., Rao B. S. S., Taqui S. N., et al., "Machine Learning Method for the Diagnosis of Retinal Diseases using Convolutional Neural Network," Proc. Int. Conf. on Data Science, Agents & Artificial Intelligence, 2023, doi: 10.1109/ICDSAAI59313.2023.10452440.

[16] Maheswari B. U., Kirubakaran S., Saravanan P., Jeyalaxmi M., Ramesh A., et al., "Implementation and Prediction of Accurate Data Forecasting Detection with Different Approaches," Proc. 4th Int. Conf. on Smart Electronics and Communication, pp. 891–897, 2023, doi: 10.1109/ICOSEC58147.2023.10276331.

[17] Mayuranathan M., Akilandasowmya G., Jayaram B., Velrani K. S., Kumar M., et al., "Artificial Intelligent based Models for Event Extraction using Customer Support Applications," Proc. 2nd Int. Conf. on Augmented Intelligence and Sustainable Systems, pp. 167–172, 2023, doi: 10.1109/ICAISS58487.2023.10250679.

[18] Gold J., Maheswari K., Reddy P. N., Rajan T. S., Kumar S. S., et al., "An Optimized Centric Method to Analyze the Seeds with Five Stages Technique to Enhance the Quality," Proc. Int. Conf. on Augmented Intelligence and Sustainable Systems, pp. 837–842, 2023, doi: 10.1109/ICAISS58487.2023.10250681.

[19] Anand L., Maurya J. M., Seetha D., Nagaraju D., et al., "An Intelligent Approach to Segment the Liver Cancer using Machine Learning Method," Proc. 4th Int. Conf. on Electronics and Sustainable Communication Systems, pp. 1488–1493, 2023, doi: 10.1109/ICESC57686.2023.10193190.

[20] Harish Babu B., Indradeep Kumar, et al., "Advanced Electric Propulsion Systems for Unmanned Aerial Vehicles," Proc. 2nd Int. Conf. on Sustainable Computing and Smart Systems (ICSCSS), pp. 5–9, 2024, doi: 10.1109/ICSCSS60660.2024.10625489.

[21] Jagan Raja V., Dhanamalar M., Solaimalai G., et al., "Machine Learning Revolutionizing Performance Evaluation: Recent Developments and Breakthroughs," Proc. 2nd Int. Conf. on Sustainable Computing and Smart Systems (ICSCSS), pp. 780–785, 2024, doi: 10.1109/ICSCSS60660.2024.10625103.

[22] Sivasankari S. S., Surendiran J., Yuvaraj N., et al., "Classification of Diabetes using Multilayer Perceptron," Proc. IEEE Int. Conf. on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), pp. 1–5, IEEE, 2022.

[23] Anushkannan N. K., Kumbhar V. R., Maddila S. K., et al., "YOLO Algorithm for Helmet Detection in Industries for Safety Purpose," Proc. 3rd Int. Conf. on Smart Electronics and Communication (ICOSEC), pp. 225–230, 2022, doi: 10.1109/ICOSEC54921.2022.9952154.

[24] Reddy K. S., Vijayan V. P., Das Gupta A., et al., "Implementation of Super Resolution in Images Based on Generative Adversarial Network," Proc. 8th Int. Conf. on Smart Structures and Systems (ICSSS), pp. 1–7, 2022, doi: 10.1109/ICSSS54381.2022.9782170.