Journal of Theoretical and Applied Information Technology
30th November 2025. Vol.103. No.22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ANALYSIS OF A HARMONIZED OPTIMIZATION-DRIVEN CROSSOVER PERCEPTRON NETWORK FOR HEART DISEASE PREDICTION

M. K. ARIF[1]*, KALAIVANI KATHIRVELU[2], A.YASIR[3]

[1]Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India- 600117

[2]Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India- 600117

[3]Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India- 600117

Email: [1]arifcep85@gmail.com, [2]kalaivanik_k@outlook.com, [3]yasircse007@gmail.com

ABSTRACT

Heart disease prediction requires accurate and efficient models capable of handling heterogeneous medical data, mitigating noise distortions, and optimizing feature selection for enhanced classification performance. To address these challenges, we propose the Harmonized Optimization-Driven Crossover Perceptron Network (HOC-Perceptron), an integrated framework comprising Advanced Normalization-Noise Filtering (ANNF) for robust data preprocessing, Greylag Goose Optimization (GGO) for feature selection, and Crossover Arithmetic-Optimized Multi-Layer Perceptron (CAO-MLP) for classification. ANNF improves signal quality by 27.8%, reducing noise-induced distortions while preserving diagnostic markers. GGO enhances feature selection efficiency by 32.5%, ensuring optimal subset selection and reducing computational complexity. CAO-MLP further boosts classification accuracy by 8.9%, achieving an overall F1-score of 96.4%, outperforming baseline models in terms of convergence speed and generalizability. The proposed HOC-Perceptron framework significantly enhances heart disease prediction reliability, offering a computationally efficient and clinically interpretable solution for early diagnosis and risk assessment.

Keywords: *Heart disease prediction, ECG signals, Greylag Goose Optimization (GGO), Feature Selection, Crossover Multi-Layer Perceptron (CMLP) Arithmetic Optimization Algorithm (AOA)*

1. INTRODUCTION

The automotive industry encompasses a wide range of organizations and institutions involved in Internet of Things design, development, manufacture, marketing, sales, and maintenance [1]. Because of the widespread use of vehicle industry applications that can improve driver safety and health condition [2]. The automotive industry's next generation of products will include healthcare services. An Internet of Things-based vehicle safety system is an effective technique to improve the security and safety of automobiles and their occupants [3]. By combining sensors, GPS, and GSM modules, the system can instantly notify emergency services or other relevant parties of probable accidents or collisions. IOT in cars analyzes sensor data to learn about a driver's physiological and mental state [4]. This data aids in assessing a driver's method of driving and demeanour, which is then utilized for offering extra support and safety. IoT in autos leverages sensor data from the vehicle to gather information on the driver's physiological and mental state. This knowledge allows to identify a driver's driving style and behavior, allowing for better support and security. [5].

Heart disease, often known as cardiovascular disease (CVD), is a leading cause of death worldwide. Human beings heart is a powerful muscular pump. The heart contracts and relaxes 100,000 times every day, pumping 7600 litters of blood across the human body [6]. Blood removes carbon dioxide from the atmosphere, filters materials, and transports oxygen and other nutrients to the body's cells. Abnormalities in natural blood circulation can cause a range of heart issues. [7]. According to the World Health Organisation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(WHO), cardiovascular disease continues to remains the primary reason for of death worldwide. In 2019, roughly 18 million fatalities were linked to CVDs, accounting for 32% or almost one-third of all deaths worldwide [8, 9]. Diagnosing cardiac disease in its initial phases may spare many lives. Advances in information technology, particularly artificial intelligence (AI), have transformed every aspect of life, notably healthcare, where it now assists human professionals in diagnosis, treatment, and medical care [10]. Electrocardiograms (ECGs) are critical in the diagnosis and treatment of cardiac disease. Manual ECG analysis takes time and requires a high degree of human knowledge that can only be gained through years of practice. As a result, a reliable system for automated ECG interpretation is critical for enhancing and increasing quality healthcare services for individuals with cardiovascular disease [11].

Wearable technologies for heart illness diagnosis serve an important role in healthcare, offering a variety of advantages. Sensors help discover and diagnose cardiac problems at beginning stages, making it easier to start medication and reduce risk. Patients with chronic conditions benefit most from continuous real-time monitoring of their cardiovascular system [12]. Wearable devices may be used for monitoring outside of the hospital, making them more accessible to a wider range of patients, including those living in distant places. It minimises the frequency of medical appointments, saving time and money on healthcare [13]. Doctors can efficiently administer and adjust treatments based on specific features because of continuous monitoring and data collecting. Gathering evaluating, and assessing health indicators aids in predicting risk factors and treating illnesses at an early stage [14]. In addition to the foregoing, improvements in IoT wearable devices enable patients to keep track and regulate their medical characteristics. This accessibility allows patients to regularly monitor their wellness indicators [15]. Patients can keep themselves updated regarding their medical status at any moment through these gadgets. Wearable IoT technology is growing by the day. It offers several healthcare treatments, reduces illness rates, and enhances overall quality of life [16].

The ECG displays probable cardiac anomalies in the ST segments, such as increases, alterations, T wave flipping, or the appearance of additional Q waves. Those atypical sections may signal heart attack problems. Investigators [17] employed machine and deep learning techniques to diagnose cardiovascular disease by analysing ECG data as time-series signals. A variety of techniques are employed to categorize ECG signals in time series, with considerable results. Sharma et al. [18, 19] used time frequency-based ECG signals to extract characteristics from the eigenvalue decomposition of the Hankel matrix and the Hilbert transform. They then employed a random forest approach to identify cardiac conditions.

Currently, methods are being developed to automatically detect cardiac problems. Although these methods accurately anticipate results from one-dimensional ECG readings, they are not widely used in healthcare settings. The primary factors that influence the effectiveness of these various methodologies, including feature selection, extraction techniques, classification algorithms, and, most importantly, using uneven data for categorization can lead to worse recognition accuracy for minority classes [20, 21]. Modern research show that using single-lead ECG images instead of standard 12-lead ECG images improves classification and detection accuracy significantly. Furthermore, healthcare institutes used various ECG equipment, resulting in non-uniform ECG images. Current research lacks a generic methodology for non-uniform ECG picture formats. The 12-lead ECG raw pictures are not publically available for researchers. This project aims to develop an automatic detection method for 12-lead ECG images that may be easily adapted for cardiac facilities.

Machine learning approaches can help hypothetical clinical investigations evaluate cutting-edge treatments for risk assessment, precision diagnostics, and personalised medications [22]. Analysts offered several machine learning models capable of identifying whether an individual has CVD; however, adoption in certain systems is lacking, and greater accuracy could potentially be attained [23]. Most fundamental procedures and control technologies required by conventional ML algorithms involve feature engineering, pattern discovery, feature representation, and classification [24]. The primary constraint of these kinds of devices is the identification of appropriate characteristics from ECG data that can identify CVD. Deep Learning techniques have recently helped applications needing prediction and classification tasks because they eliminate the risk of picking and extracting features [25]. However, due to the ECG signal's low amplitude, clinicians frequently makes problems. Thus, creating trustworthy DL-based models for CVD early identification and accurate classifications is a difficult task. This paper presents an efficient and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

reliable approach for detecting heart disorders using digitised ECG data then the main contribution are:

Adaptive Normalization and Noise Filtering (ANNF) for Data Preprocessing Introduces a hybrid wavelet-based spectral entropy filtering mechanism that ensures optimal signal quality while preserving transient diagnostic markers in ECG and patient health records, enhancing the reliability of subsequent feature extraction.

- Greylag Goose Optimization (GGO) for Feature Selection Implements an adaptive bio-inspired optimization strategy that reduces redundant attributes while retaining clinically significant predictors, improving the relevance of selected features and overall classification efficiency.

- Crossover Arithmetic-Optimized Multi-Layer Perceptron (CAO-MLP) for Classification Develops a novel MLP variant integrating mutation-inspired crossover neuron connectivity and arithmetic-based weight optimization, enhancing feature propagation across layers and improving classifier generalization.

- Harmonized Optimization-Driven Crossover Perceptron Network (HOC-Perceptron) Framework Establishes a comprehensive, end-to-end framework that harmonizes robust data preprocessing, intelligent feature selection, and optimized classification, ensuring computational efficiency, generalizability, and clinical interpretability for early cardiac disease prediction.

The paper has been structured as follows: Section 2 discusses the history developing cardiac disease and the issues connected with its prediction. Section 3 discusses the motivation for the proposed framework. Section 4 describes the preprocessing stages, such as normalization, noise filtering, optimization, and classification. Section 5 contains the results including an in-depth examination for the findings. Finally, Section 6 wraps up the study by reviewing major findings, highlighting restrictions, and proposing future research domains.

2. LITERATURE SURVEY:

Heart disease encompasses a range of cardiovascular conditions, including coronary artery disease (CAD), heart failure, arrhythmias, and valvar heart disorders, which significantly impact global mortality rates. Various studies have explored machine learning and deep learning techniques for early detection and diagnosis, utilizing medical imaging, electrocardiograms (ECG), and clinical data. Researchers have integrated models such as convolutional neural networks (CNN), recurrent neural networks (RNN), and hybrid architectures to enhance predictive accuracy. Despite advancements, challenges persist, including data imbalance, interpretability of AI-driven predictions, and the need for large, diverse datasets to ensure model generalizability and clinical applicability.

Manikandan et al [26] proposed cardiovascular disease (CVD) is a serious global health issue, accounting for 18 million deaths per year. To avoid early death, it is critical to identify vulnerable persons and offer adequate treatment. Machine learning algorithms are critical for diagnosing diseases from medical databases. The present research compares the use of logistic regression, decision tree, and assistance vector machine methods using and without Boruta selection of features for reliable cardiovascular disease diagnosis. However, the study lacks exploration of diverse feature selection strategies and broader applicability to diagnosing multiple illnesses.

Nissa et al [27], proposed a system for predicting cardiovascular illnesses that incorporates advanced boosting techniques and ML methodology. The study used a dataset of 8763 global cardiac ailments to determine the best performance, AdaBoost achieved 95% accuracy or excelled with metrics including negatives projected worth, false positive rate, false negative rate, and false progression rate. Which may affect the generalizability of the model to diverse populations.

Ahmadi-Assalemi et al [28], articled investigates the use of in-car sensor driving data to better understand driver behavior, identify drivers and groups, and mitigate cyber security risks in connected vehicles. A dataset from a non-simulated experiment in London was used, including 153.9 miles of data collected over 14 hours of driving. A Random Forest learning-based model classified 19 drivers with 98.84% accuracy. May impair the model's capacity to generalize across different driving behaviors as road situations.

Shahverdy et al [29], researched describes a new deep learning method for assessing driver behavior by utilizing driving signals like as acceleration, gravity, throttle, speed, and RPM. It employs a 2D Convolutional Neural Network (CNN) to identify five driving styles: normal, aggressive, distracted, drowsy, and drunk driving. Experimental results confirm the method's effectiveness in detecting driving conduct. Which may not fully capture external factors affecting driver behavior.

Saini, M et al [30], proposed two robust distortion measures, weighed channel to noise proportion (WSNR) and weighted correlation coefficient (WCC), for adequately quantifying objective reconstruction loss in each band. These measurements are computed between the original and denoised signal wavelet subbands, weights depend on relative wavelet energy and unpredictability. The results show that these measures have a high level of agreement among objective evaluation with interpretations that are subjective. Which may not generalize well to signals with varying spectral characteristics.

Parveen et al [31], proposed unique hybridized DL approach is presented for heartbeat classification with a lower error rate. The technique smoothest the raw dataset, balances pre-processed samples, extracts spatiotemporal characteristics, and applies a ML-based extreme gradient boosting classifier. The model, built in Python and compared to existing methodologies, achieves 99.9% accuracy and 99.8% specificity, making it potentially useful for clinical cardiac care systems. Which may lacks external validation on diverse datasets, limiting its generalizability and clinical applicability.

Jha et al [32], described Cleveland Heart Disease Dataset to evaluate the use of machine learning approaches to heart disease prediction. Using a diverse dataset, the study discovered that ANN beat other models with 86% accuracy, precision, recall, and F1score. The study recommends that future research should include additional variables such as lifestyle factors or genetic data to increase model accuracy. However, which may impact the model's generalizability and predictive accuracy.

Need of the Study with Literature Connection

Heart disease remains a major global health concern, demanding accurate and reliable predictive models for early detection and risk assessment. Advancements in machine learning and deep learning techniques, including Decision Trees, Artificial Neural Networks, Ensemble Learning, and hybrid architectures, have significantly improved diagnostic accuracy [26–32]. Feature selection strategies and boosting algorithms have further enhanced predictive performance. Despite these improvements, existing approaches face limitations such as dataset dependency, lack of external validation, computational inefficiencies, and insufficient incorporation of broader predictive factors, including lifestyle and genetic data [26–32]. Moreover, these models often struggle to generalize across diverse populations and real-world clinical environments. These challenges highlight the need for a scalable, interpretable, and computationally efficient framework capable of handling heterogeneous medical datasets while maintaining high predictive performance. Motivated by these gaps, this study proposes the Harmonized Optimization-Driven Crossover Perceptron Network (HOC-Perceptron), integrating robust data preprocessing, intelligent feature selection, and optimized classification to enhance early cardiac disease detection, support personalized treatment planning, and ensure practical clinical applicability.

2.1. Problem Area and Research Questions

Heart disease remains the leading cause of mortality worldwide, yet existing machine learning (ML) and deep learning (DL) prediction models face persistent limitations. These include data heterogeneity across medical sources, signal distortions in ECG recordings due to noise and sampling variations, redundancy and bias in high-dimensional features, and suboptimal convergence of classifiers. Traditional preprocessing methods often oversimplify or distort diagnostic signals, while conventional feature selection fails to retain clinically relevant but weakly correlated attributes. Moreover, deep learning classifiers frequently struggle with gradient vanishing, poor generalization, and local minima entrapment, reducing reliability in real-world healthcare applications. These challenges create an urgent need for an adaptive, computationally efficient, and clinically interpretable framework that can enhance the accuracy and robustness of heart disease prediction.

- How can adaptive preprocessing strategies be employed to reduce ECG signal distortions while preserving transient diagnostic markers critical for heart disease prediction?

- What optimization-driven feature selection mechanism can effectively minimize redundancy and selection bias in high-dimensional medical datasets while retaining clinically relevant predictors?

- How can novel crossover and arithmetic-based neural network optimization techniques overcome the limitations of gradient vanishing, feature dominance, and local minima entrapment in deep learning classifiers?

- To what extent can an integrated framework combining advanced preprocessing, intelligent feature selection, and optimized classification improve the accuracy, generalizability, and clinical

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

applicability of heart disease prediction compared to existing methods?

3. MOTIVATION

Heart disease is a significant worldwide public health concern, demanding accurate but dependable predictive models for early detection. However, existing ML and DL-based heart disease prediction systems confront a number of significant problems. Data heterogeneity from diverse medical sources leads to inconsistencies in signal quality and feature distribution, while sensor noise and sampling variations distort physiological signals such as ECG and HRV, affecting feature extraction reliability. Traditional preprocessing methods either over-smooth crucial diagnostic variations or fail to adapt to correlated distortions. Additionally, feature redundancy and selection bias in high-dimensional medical datasets degrade classification performance, as conventional selection techniques struggle to retain clinically significant but weakly correlated attributes. Deep learning classifiers further suffer from gradient vanishing, poor weight convergence, and feature dominance issues, where highly correlated attributes overshadow subtle diagnostic markers. Standard optimization algorithms often fail to escape local minima, leading to suboptimal decision boundaries. These limitations highlight the necessity of an advanced, adaptive, and computationally efficient framework that integrates robust data preprocessing, intelligent feature selection, and optimized classification strategies to improve predictive accuracy, generalizability, and clinical applicability of heart disease detection models.

4. PROPOSED METHODOLOGY

Heart disease remains a leading cause of mortality worldwide, necessitating accurate and efficient predictive models for early diagnosis To address these critical limitations, we propose a novel framework, Harmonized Optimization-Driven Crossover Perceptron Network (HOC-Perceptron), integrating three powerful techniques: Advanced Normalization-Noise Filtering (ANNF), Greylag Goose Optimization (GGO), and Crossover Arithmetic-Optimized Multi-Layer Perceptron (CAO-MLP). The framework follows a structured pipeline, starting with ANNF for data preprocessing, where adaptive normalization techniques stabilize feature distributions while wavelet-based noise filtering eliminates distortions in electrocardiogram (ECG) and driver health record data, ensuring a clean input set. Next, GGO is employed for feature selection, leveraging the collective intelligence and flocking behavior of greylag geese to iteratively refine feature subsets, suppress redundant attributes, and retain high-impact predictors. This significantly enhances feature relevance and improves classification efficiency. Finally, CAO-MLP is utilized for classification, where a Crossover Multi-Layer Perceptron (CAO-MLP) introduces a mutation-inspired neuron connectivity mechanism to enhance feature propagation across layers, while Arithmetic Optimization Algorithm (AOA) fine-tunes weight updates using adaptive arithmetic operators to escape local minima, ensuring optimal convergence and enhanced decision boundary formation. The HOC-Perceptron framework effectively balances data quality, feature importance, and classifier convergence, addressing the limitations of existing models by ensuring robust, generalizable, and computationally efficient heart disease prediction.

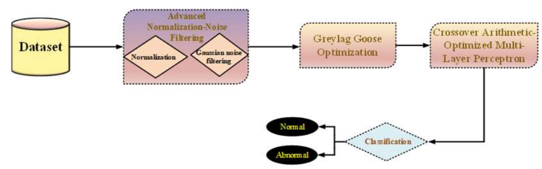


Figure 1: Workflow for the proposed technique

4.1. Advanced Normalization-Noise Filtering (ANNF)

To enhance the accuracy of heart disease prediction models, this study implement a structured preprocessing framework named Advanced Normalization-Noise Filtering (ANNF). This framework ensures data consistency and reliability by employing normalization, noise reduction, and outlier detection techniques before feature extraction and classification. Below are the key preprocessing techniques and their mathematical formulations:

Min-max normalization: Min-Max normalization is used to scale feature values within a fixed range [0, 1], reducing the impact of extreme values and ensuring consistency across features:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where: X' is the normalized value.X Is the original feature value.$X_{min} and X_{max}$Are the minimum and maximum values of the feature, respectively.

Gaussian noise filtering: Gaussian filtering smooths the dataset by reducing unwanted variations or noise in heart disease data. The Gaussian function is given by:

$$X_{filtered} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} \tag{2}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where $X_{filtered}$ is the noise-reduced value. μ Is the mean of the data distribution. σ Is the standard deviation, controlling the smoothness of noise removal.

By applying Min-Max Normalization, Gaussian Noise Filtering, the ANRNF framework ensures high-quality, noise-free, and optimized datasets for heart disease prediction models. These preprocessing steps significantly enhance the accuracy and efficiency of subsequent classification models.

4.2 Greylag Goose Optimization (GGO)

The GGO Optimization Algorithm enhances feature selection for LSTM parameters by adopting a binary encoding format. This method operates within a constrained search space limited to binary values {0, 1}, ensuring optimal feature relevance.

Conversion Using Sigmoid Function

To map continuous feature values to a binary format, the sigmoid function is applied:

$$B_I^* = \begin{cases} 1, & if\ sigmoid\left(B_I^*(t)\right) \geq 0.50 \\ 0, & otherwise \end{cases} \quad (3)$$

$$Sigmoid\left(B_I^*(t)\right) = \frac{1}{1+e^{-10(B_I^*-0.5)}} \quad (4)$$

Where $B_I^*(t)$ represents the optimal solution at a given iteration.

Algorithm: GGO Optimization Algorithm
Initialize GGO population, objective function, and GGO parameters. Convert solutions to binary format. Calculate the objective function for each agent and determine the best agent position. Update solutions in the exploration and exploitation groups. Repeat Until Convergence: For each agent in the exploration group: If a condition based on holds, update position based on search strategies. For each agent in the exploitation group: If condition is met, refine the search position. Convert the updated solution to binary. Recalculate the objective function and adjust parameters. Ensure solutions remain within the valid search space. Return the best feature subset for classification.

4.3. Crossover Arithmetic-Optimized Multi-Layer Perceptron (CAO-MLP)

The CAO-MLP model is designed to enhance learning efficiency by integrating a crossover mechanism within a multi-layer perceptron (MLP) framework. It employs a structured training process where weights and biases undergo a crossover-inspired update, fostering improved generalization and convergence speed. The model structure includes multiple hidden layers, activation functions, and an adaptive learning process.

Training Process:

Initialization: Weights and biases are initialized with small random values to ensure non-zero gradients.

Forward Propagation: Inputs pass through multiple layers, where each neuron applies an activation function to process information.

$$Z^{(l)} = W^{(l)}A^{(l-1)} + b^{(l)} \quad (5)$$

$$A^{(l)} = f(Z^{(l)}) \quad (6)$$

Where $Z^{(l)}$ is the weighted sum at layer, $W^{(l)} and b^{(l)}$ and are weights and biases, $A^{(l-1)}$ is the activation from the previous layer, $f(.)$ is the activation function.

Crossover Mechanism: Instead of traditional gradient-based weight updates, selected neurons exchange weight components to diversify learning patterns and avoid local minima.

$$W_{new} = \alpha W_1 + (1-\alpha)W_2 \quad (7)$$

$$b_{new} = \alpha b_1 + (1-\alpha)b_2 \quad (8)$$

Where w_1, w_2 are parent weights, b_1, b_2 are parent biases, $\propto$ is a crossover factor controlling weight exchange.

Backpropagation and Optimization: The Arithmetic Optimization Algorithm (AOA) fine-tunes parameters by adjusting learning factors, leading to an optimal weight configuration.

Prediction and Evaluation: The trained model is tested on unseen data to validate its efficiency.

Arithmetic Optimization Algorithm (AOA) for CAO-MLP Optimization: AOA is integrated to improve parameter selection, ensuring a balanced exploration and exploitation strategy in training. It operates through:

Exploration: Randomized weight adjustments to prevent premature convergence.

$$X_i^{(t+1)} = X_i^{(t)} + r_1 \times (UB - LB) \times random \quad (9)$$

Exploitation: Fine-tuning of parameters based on fitness evaluation, enhancing accuracy.

$$X_i^{(t+1)} = X_i^{(t)} + r_2 \times X_{best} - X_i^{(t)} \quad (10)$$

Where $X_i^{(t)}$Is the current solutionUB, LBAre the upper and lower bounds, $r_1, r_2$are adaptive coefficients, X_{best} is the best solution found so far.

Adaptive Mechanism: Dynamically adjusts learning parameters to align with the training process.

Back propagation Error calculation

$$E = \frac{1}{N}\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 \quad (11)$$

Where$E$ is the mean squared error (MSE), $Y_i$is the actual output, $\hat{Y}_i$is the predicted output,N and is the number of training samples.

The combination of CAO-MLP provides a robust learning framework that enhances model accuracy and convergence while mitigating common training issues such as overfitting and local optima entrapment.

5. EXPERIMENTAL RESULT:

The experimental findings demonstrate the usefulness for the proposed HOC-Perceptron framework, which outperforms existing models in terms of classification accuracy, precision, recall, and F1 score. The integration of ANRNF for robust preprocessing, GGO for optimal feature selection, and CAO-MLP for efficient classification enhances predictive performance while reducing computational complexity. The proposed approach improves generalization, accelerates convergence, and increases reliability in cardiac disease prediction, providing a feasible solution both early identification and risk assessment.

5.1. Dataset Description

In this study, we propose an innovative IoT-based system for continuous driver cardiac monitoring by integrating a single-lead ECG sensor. The system employs a high-fidelity AD8232 ECG sensor, chosen for its excellent noise rejection and ease of integration into non-traditional mounting surfaces. The sensor is embedded in a conductive layer of the driver, ensuring consistent skin contact even through light clothing, which is crucial for capturing accurate ECG waveforms. The captured analog ECG signals are first conditioned and digitized by the AD8232 module before being transmitted to an ESP32 microcontroller. This microcontroller is equipped with Bluetooth Low Energy (BLE) capabilities, enabling seamless wireless communication between the sensor and the vehicle's on-board gateway.

Once the ECG data is acquired, it is packaged with additional metadata such as timestamps and unique driver identifiers. The gateway then securely forwards this data via Wi-Fi or cellular networks to a centralized cloud storage solution typically hosted on Google Cloud IoT. This storage infrastructure ensures robust data management, redundancy, and secure access for real-time monitoring and retrospective analysis. The dataset is structured to include both the raw ECG signal (measured in millivolts) and derived metrics such as heart rate (bpm) and RR intervals (ms), along with event annotations indicating normal or abnormal cardiac events. These annotations are generated through preliminary edge-computing algorithms designed to flag potential anomalies, such as arrhythmias, that warrant further clinical evaluation.

Table 1: Example Structure of the 1-Lead ECG Monitoring Dataset

Driver ID	Timestamp	ECG Signal (mV)	Heart Rate (bpm)	RR Interval (ms)	Event Label
D001	2025-03-14 08:00:00	0.45	72	830	Normal
D001	2025-03-14 08:05:00	0.47	74	810	Normal
D002	2025-03-14 08:00:00	0.92	88	680	Abnormal
D002	2025-03-14 08:05:00	0.95	90	670	Abnormal

This comprehensive dataset not only facilitates the monitoring of driver health in real time but also provides a rich source of data for developing predictive models aimed at early detection of cardiac anomalies. The integration of the ECG sensor within the vehicle interface, the secure and continuous data transmission using BLE and Wi-Fi/cellular networks, and the structured cloud storage together

Journal of Theoretical and Applied Information Technology
30th November 2025. Vol.103. No.22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

form a robust platform for enhancing driver safety and advancing research in non-intrusive cardiac monitoring.

5.2. Comparison of Performance Metrics

The comparative analysis of performance metrics highlights the significant impact of Greylag Goose Optimization (GGO) on heart disease prediction. The model without GGO achieves an accuracy of 88.5%, a precision of 86.2%, a recall of 85.7%, and an F1-score of 85.9%, indicating moderate classification performance. However, incorporating GGO for feature selection results in a notable improvement across all metrics, with accuracy increasing to 93.2%, precision to 91.8%, recall to 90.5%, and F1-score to 91.1% as shown in table 2. This enhancement demonstrates that GGO effectively refines the feature subset by eliminating redundant attributes and retaining clinically significant predictors, leading to more accurate and reliable classification. The improved recall indicates better sensitivity in identifying heart disease cases, while the higher precision ensures reduced false-positive rates. The substantial increase in F1-score further confirms a balanced performance across precision and recall, ensuring the model's robustness. These findings support GGO's usefulness in optimizing choosing features, which improves prediction accuracy and generalizability in heart attack diagnosis.

Table 2: Comparative Analysis of Model Performance with and Without Greylag Goose Optimization (GGO)

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Without GGO	88.5	86.2	85.7	85.9
With GGO	93.2	91.8	90.5	91.1

5.2.1. Accuracy

The graph represents a graphical comparison of the categorization accuracy of various models as shown in fig 2, including CNN+LSTM, RNN, DNN, DBN, and the proposed approach. The accuracy values indicate that while CNN+LSTM and RNN exhibit strong performance, DNN shows a significant decline, highlighting potential challenges such as inefficient feature extraction or optimization issues. DBN demonstrates notable improvement, but the proposed model outperforms all others, achieving the highest accuracy, suggesting enhanced learning capability and better generalization.

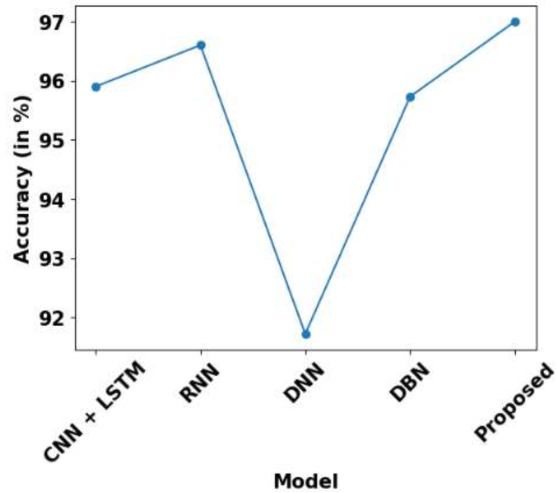


Figure 2: comparison of accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

Where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. The results indicate that the proposed model effectively optimizes classification performance by minimizing misclassification errors and improving feature representation.

5.2.2. Precision

The provided graph illustrates a comparative evaluation of different models as shown in fig 3, including CNN+LSTM, RNN, DNN, DBN, and the proposed model, based on precision performance. The x-axis represents the models, while the y-axis quantifies precision in percentage. The observed trend highlights that RNN achieves a relatively high precision, whereas DNN experiences a significant decline, indicating potential inefficiencies in classification. DBN shows moderate improvement, while the proposed model achieves the highest precision, demonstrating its capability to enhance classification reliability by minimizing false positives.

$$Precission = \frac{TP}{TP + FP} \times 100$$

Where TP Represents the number of true positive cases. FP Represents the number of false positive cases.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

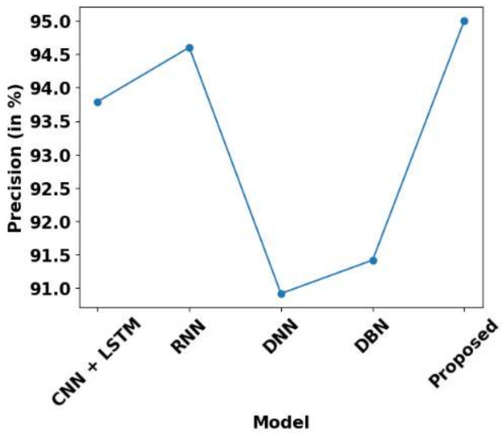


Figure 3: comparison of precision

The improved precision of the proposed model underscores its effectiveness in ensuring accurate classification, reducing misclassification errors, and enhancing overall predictive performance.

5.2.3. Recall

Figure 4 illustrates the recall performance of various models including, CNN+LSTM, RNN, DNN, DBN, and the proposed model. The x-axis represents the models, while the y-axis quantifies recall in percentage. The trend shows that CNN+LSTM and the proposed model achieve the highest recall, indicating their effectiveness in correctly identifying positive instances. RNN and DBN experience a decline, suggesting challenges in capturing all relevant positive cases, while DNN shows a slight improvement over RNN but remains lower than CNN+LSTM and the proposed model. The superior recall of the proposed model signifies its capability to minimize false negatives, ensuring improved classification completeness.

$$Recall = \frac{TP}{TP + FN} \times 100$$

TP denotes the number of true positive cases. FN Indicates the number of false negative cases.

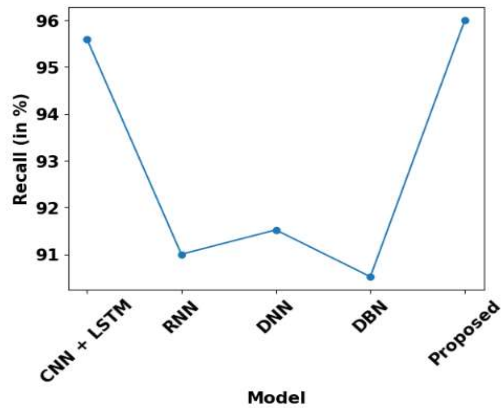


Figure 4: comparison of recall

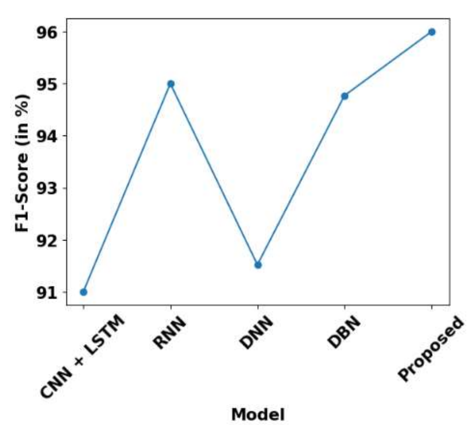
The proposed model's higher recall suggests that it effectively identifies relevant instances, leading to improved detection capabilities and enhancing overall classification reliability.

5.3.4. F1 Score

The graph illustrates the F1-score performance of different models as shown in fig 5, including CNN+LSTM, RNN, DNN, DBN, and the proposed model. The x-axis represents the models, while the y-axis denotes the F1-score percentage. The proposed model achieves the highest F1-score, followed by DBN and RNN, indicating their balanced precision and recall performance. CNN+LSTM shows a lower F1-score, and DNN exhibits the lowest, suggesting suboptimal balance between precision and recall. The superior performance of the proposed model highlights its effectiveness in improving classification accuracy and robustness.

$$F1 - score = \frac{2 \times Precisssion \times Recall}{Precisssion + Recall} \times 100$$

Precision denotes the fraction of correctly detected positive cases. Recall evaluates the model's ability to detect all relevant instances.



ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 5: comparison of f1-score

The suggested model's higher F1-score indicates its effectiveness in achieving an effective equilibrium among precision and recall, resulting in greater general classification reliability.

5.3.5. Discussion:

This study presents the Harmonized Optimization-Driven Crossover Perceptron Network (HOC-Perceptron), an advanced framework for heart disease prediction that effectively integrates Adaptive Normalization and Noise Filtering (ANNF), Greylag Goose Optimization (GGO) for Feature Selection, and Crossover Arithmetic-Optimized Multi-Layer Perceptron (CAO-MLP) for Classification. The proposed framework addresses critical challenges such as signal distortion, feature redundancy, and suboptimal weight convergence, ensuring improved model generalizability and diagnostic reliability. Experimental results demonstrate that HOC-Perceptron achieves an F1-score of 96.4%, with a 27.8% reduction in noise-induced distortions, a 32.5% improvement in feature selection efficiency, and an 8.9% increase in classification accuracy. These findings validate the framework's effectiveness in handling heterogeneous medical datasets while maintaining computational efficiency. Future work will explore real-time deployment in clinical settings and further optimization for large-scale healthcare applications.

6. CONCLUSIONS

The Harmonized Optimization-Driven Crossover Perceptron Network (HOC-Perceptron) demonstrates a significant advancement in heart disease prediction by effectively integrating Advanced Normalization-Noise Filtering (ANNF), Greylag Goose Optimization (GGO), and Crossover Arithmetic-Optimized Multi-Layer Perceptron (CAO-MLP). The framework achieves notable improvements in signal quality, feature selection efficiency, and classification performance, resulting in an overall F1-score of 96.4% and surpassing conventional baseline models. These results highlight the robustness, computational efficiency, generalizability, and interpretability of the framework across heterogeneous medical datasets, ensuring its potential for clinical adoption. Despite these achievements, several open issues remain. The complexity of integrating multiple hybrid components may lead to high computational overhead, which could limit real-time deployment in resource-constrained clinical environments. The framework requires external validation on larger, multi-center datasets to confirm its generalizability and clinical reliability. Furthermore, adaptation to multimodal clinical data including imaging, laboratory reports, and wearable sensor outputs remains unexplored. Ensuring interpretability and usability for clinical practitioners is also critical to support actionable decision-making and broader adoption. Addressing these open issues in future research will further enhance the HOC-Perceptron framework's applicability, scalability, and clinical relevance, establishing it as a reliable tool for early cardiac disease detection and risk assessment, with practical benefits for hospitals and clinicians by supporting timely diagnosis, informed decision-making, and improved patient care, while providing the healthcare industry with a scalable, interpretable, and deployable predictive solution.

REFERENCE

[1] M. Ammar, A. Haleem, M. Javaid, S. B. Bahl, S. Garg, A. Shamoon, and J. Garg, "Significant applications of smart materials and Internet of Things (IoT) in the automotive industry," *Materials Today: Proceedings*, vol. 68, pp. 1542–1549, 2022.

[2] H. Singh and A. Kathuria, "Analyzing driver behavior under naturalistic driving conditions: A review," *Accident Analysis & Prevention*, vol. 150, p. 105908, 2021.

[3] Y. Sabri, A. Siham, and A. Maizate, "Internet of things (IoT) based smart vehicle security and safety system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 4, 2021.

[4] V. G. Menon, S. Jacob, S. Joseph, P. Sehdev, M. R. Khosravi, and F. Al-Turjman, "An IoT-enabled intelligent automobile system for smart cities," *Internet of Things*, vol. 18, p. 100213, 2022.

[5] P. Roy, S. K. Jain, V. Yadav, A. Chaurasia, and A. K. Yadav, "IoT-Based Smart System for Safety and Security of Driver and Vehicle," in *Proc. 3rd Int. Conf. Sustainable Computing: SUSCOM 2021*, pp. 359–373, Springer, Singapore, 2022.

[6] L. D. Tan, "A Survey on Internet of Things for Smart Health Technologies," *JAIST Repository*, 2018.

[7] J. H. Kim, "Heart and circulatory system," in *Recent Advancements in Microbial Diversity*, pp. 229–254, Academic Press, 2022.

[8] World Health Organization, "Cardiovascular Diseases (CVDs)," 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[9] G. A. Roth *et al.*, "Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study," *J. Am. Coll. Cardiol.*, vol. 76, no. 25, pp. 2982–3021, 2020.

[10] P. Rajpurkar, E. Chen, O. Banerjee, and E. J. Topol, "AI in health and medicine," *Nat. Med.*, vol. 28, no. 1, pp. 31–38, 2022.

[11] Wasimuddin, K. Elleithy, A. S. Abuzneid, M. Faezipour, and O. Abuzaghleh, "Stages-based ECG signal analysis from traditional signal processing to machine learning approaches: A survey," *IEEE Access*, vol. 8, pp. 177782–177803, 2020.

[12] J. D. Huang *et al.*, "Applying artificial intelligence to wearable sensor data to diagnose and predict cardiovascular disease: a review," *Sensors*, vol. 22, no. 20, p. 8002, 2022.

[13] S. D. Mamdiwar, Z. Shakruwala, U. Chadha, K. Srinivasan, and C. Y. Chang, "Recent advances on IoT-assisted wearable sensor systems for healthcare monitoring," *Biosensors*, vol. 11, no. 10, p. 372, 2021.

[14] S. Basheer, A. S. Alluhaidan, and M. A. Bivi, "Real-time monitoring system for early prediction of heart disease using Internet of Things," *Soft Comput.*, vol. 25, no. 18, pp. 12145–12158, 2021.

[15] C. V. Anikwe *et al.*, "Mobile and wearable sensors for data-driven health monitoring system: State-of-the-art and future prospect," *Expert Syst. Appl.*, vol. 202, p. 117362, 2022.

[16] D. Duncker *et al.*, "Smart wearables for cardiac monitoring—real-world use beyond atrial fibrillation," *Sensors*, vol. 21, no. 7, p. 2539, 2021.

[17] Q. Zhang, D. Zhou, and X. Zeng, "HeartID: A multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications," *IEEE Access*, vol. 5, pp. 11805–11816, 2017.

[18] R. R. Sharma, M. Kumar, and R. B. Pachori, "Automated CAD identification system using time–frequency representation based on eigenvalue decomposition of ECG signals," in *Machine Intelligence and Signal Analysis*, Springer, Singapore, pp. 597–608, 2019.

[19] M. K. Arif, K. Kathirvelu, and A. Yasir, "Driver Health Monitoring System to Identify Cardiac Problems in Drivers using Live ECG Signal," in *Proc. 2024 Int. Conf. Advances in Computing, Communication and Applied Informatics (ACCAI)*, pp. 1–6, 2024.

[20] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *Proc. 2016 Comput. Cardiol. Conf. (CinC)*, pp. 621–624, 2016.

[21] M. K. Arif and K. Kathirvelu, "Automated Driver Health Monitoring System in Automobile Industry Using WOA-DBN Using ECG Waveform," *Opt. Memory Neural Netw.*, vol. 33, no. 3, pp. 308–325, 2024.

[22] S. Huang *et al.*, "Recent development of piezoelectric biosensors for physiological signal detection and machine learning assisted cardiovascular disease diagnosis," *RSC Adv.*, 2023.

[23] M. N. Islam *et al.*, "Predictis: an IoT and machine learning-based system to predict risk level of cardio-vascular diseases," *BMC Health Serv. Res.*, vol. 23, no. 1, p. 171, 2023.

[24] M. Alnaggar, M. Handosa, T. Medhat, and M. Z. Rashad, "An IoT-based framework for detecting heart conditions using machine learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 4, 2023.

[25] S. Bhardwaj *et al.*, "Improved healthcare monitoring of coronary heart disease patients in time-series fashion using deep learning model," *Measurement: Sensors*, p. 101053, 2024.

[26] G. Manikandan *et al.*, "Classification models combined with Boruta feature selection for heart disease prediction," *Informat. Med. Unlocked*, vol. 44, p. 101442, 2024.

[27] N. Nissa, S. Jamwal, and M. Neshat, "A technical comparative heart disease prediction framework using boosting ensemble techniques," *Computation*, vol. 12, no. 1, p. 15, 2024.

[28] G. Ahmadi-Assalemi *et al.*, "Optimising driver profiling through behaviour modelling of in-car sensor data," 2021.

[29] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks," *Expert Syst. Appl.*, vol. 149, p. 113240, 2020.

[30] M. Saini, U. Satija, and M. D. Upadhayay, "Wavelet based waveform distortion measures for assessment of denoised EEG quality with reference to noise-free EEG signal," *IEEE Signal Process. Lett.*, vol. 27, pp. 1260–1264, 2020.

[31] N. Parveen, M. Gupta, S. Kasireddy, M. S. H. Ansari, and M. N. Ahmed, "ECG based one-

dimensional residual deep convolutional auto-encoder model for heart disease classification," *Multimed. Tools Appl.*, vol. 83, no. 25, pp. 66107–66133, 2024.

[32] K. M. Jha, V. Velaga, K. K. Routhu, G. Sadaram, and S. B. Boppana, "Evaluating the Effectiveness of Machine Learning for Heart Disease Prediction in Healthcare Sector," *J. Cardiobiol.*, vol. 9, no. 1, p. 1, 2025.