SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I4P103

Volume 13 Issue 4, 35-51, April 2026
© 2026 Seventh Sense Research Group®

Original Article

Spatial Attention-Enhanced Bidirectional Transformer Network for Cardiovascular Disease Classification

R. Ratheesh[1], M. Saranya Nair[2], N.V.S. Sree Rathna Lakshmi[3], Blessina Preethi R[4]

[1]*ECE, Agni College of Technology, Chennai, Tamilnadu, India.*
[2]*SENSE, Vellore Institute of Technology, Chennai, Tamilnadu, India.*
[3]*ECE, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamilnadu, India.*
[4]*ECE, Agni College of Technology, Chennai, Tamilnadu, India.*

[2]*Corresponding Author : saranyanair.m@vit.ac.in*

Received: 05 January 2026 Revised: 05 February 2026 Accepted: 04 March 2026 Published: 30 April 2026

Abstract *- Most people were affected by this Cardiovascular Disease (CVD), which is caused by atherosclerosis and some infections. The early detection of CVD has a higher chance of recovering the person from this disease. In tradition, biosensors such as lab–on–a–chip technology were used to detect the CVD. CVD does not accurately detect the disease, and it is more expensive. Even though nanomaterials like nanotubes and nanowires offer special physical and chemical properties, they also have certain limitations. Diagnosing the disease based on just one biometer is not reliable, and it is hard to identify the conditions apart from using a single indicator. In this system, an optimized spatial attention Bidirectional Transformer Network is utilized to detect Cardiovascular Disease. The clinical data is pre-processed by using Missing value imputation and ECG signals by One-hot-encoding algorithms. The pre-processed ECG signals are denoised by using a Narrowband filter (NBF). The features of the pre-processed clinical data are extracted by a Dense-assisted Global Attention-Based Autoencoder (DGA-AE), and features of denoised ECG signals are extracted by using Multiwavelet Transform-Based Feature Decomposition (MT-FD). The data are merged and optimized by the Bidirectional transformer and the POA (Puma Optimization Algorithm) algorithm. The testing of the datasets achieved a high accuracy. In contrast, the ECG dataset from Ptb-XL gives an accuracy of 95.34%, and the second dataset, Cardiovascular Disease detection, gives an accuracy of 97.21%, respectively. The overall performance demonstrated that the proposed model exceeds the baseline models, and it effectively improved the early detection of cardiovascular diseases with high accuracy.*

Keywords *- Cardiovascular Disease Detection, Bidirectional Transformer Network, Missing Value Imputation, One-Hot Encoding, Multiwavelet Transform-Based Feature Decomposition, Puma Optimization.*

1. Introduction

CVDs are among the leading causes of morbidity and mortality worldwide, arising from conditions such as coronary artery obstruction, arrhythmias, and impaired cardiac function [1]. The heart plays a critical role in sustaining life by ensuring continuous blood circulation, and disruptions in its electrical or physiological activity can lead to severe and sometimes fatal outcomes [2]. A major challenge in Cardiovascular care is that early-stage disease often remains asymptomatic, making timely and accurate diagnosis essential for preventing disease progression and reducing long-term health risks [3, 4]. Electrocardiography (ECG) is a widely used, non-invasive, and cost-effective technique for recording the electrical activity of the heart and has become a cornerstone for cardiac diagnosis [5, 6]. Alongside ECG signals, structured clinical attributes such as age, blood pressure, body mass index, and lifestyle factors provide valuable complementary information for disease assessment [7]. However, diagnosis based on a single data modality or isolated biomarkers is often insufficient to capture the complex and multifactorial nature of CVDs, leading to reduced diagnostic reliability and increased false predictions [8, 9].

Recent advances in Machine Learning (ML) and Deep Learning (DL) have enabled automated CVD detection with improved accuracy [10-12]. Traditional ML models rely heavily on handcrafted features and struggle with high-dimensional, noisy, and incomplete medical data. DL models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) offer better feature learning capabilities but suffer from inherent limitations [13, 14]. CNNs are restricted in modeling long-range temporal dependencies in ECG signals, while RNN-based architectures face issues related to vanishing gradients, limited parallelization, and high computation latency [15].

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Although transformer-based models address some of these limitations through self-attention mechanisms, existing approaches often adopt unidirectional processing, inadequately capture bidirectional temporal dependencies, underutilize spatial attention, and lack effective multimodal fusion of ECG and clinical data [16]. Furthermore, many optimization techniques used in these systems suffer from premature convergence and poor robustness to noise and missing values [17]. These limitations reveal a clear research gap, such as the absence of a robust and scalable CVD detection framework that can jointly integrate multimodal data that capture bidirectional temporal dependencies that emphasize diagnostically significant spatial features and maintain reliable performance under noisy and heterogeneous data conditions.

To address these gaps, this paper proposes an optimized spatial attention-enhanced bidirectional transformer network for CVD detection. The proposed approach integrates advanced pre-processing, deep feature extraction from ECG and clinical data, bidirectional transformer-based temporal modelling, and an efficient nature-inspired optimization strategy to improve classification accuracy and robustness, thereby enabling reliable early detection of CVD.

1.1. Motivation

CVD is a group of medical conditions that impact the heart and the network of blood vessels. It covers illnesses like blockage in the coronary arteries, weakening of heart pumping ability, disruptions in the blood flow to the brain, and narrowing of arteries in the limbs. The use of smart technologies in heart care could boost treatment results and increase patient involvement, and revolutionize how healthcare services are provided. In intelligent CAD systems, Machine Learning systems are crucial because of their ability to accurately differentiate between healthy and unhealthy heart patients. These models can easily forecast results using previously available data, which helps in the effective prediction of various diseases. However, the quality of data is poor, and the algorithms should ensure the transparency and privacy of data; it is more difficult to detect the essential features from the pre-processed data. To accurately determine the CVD combination, novel technology is required. For that purpose, an Optimized Spatial and Bidirectional Network with IoT is used, which can precisely detect the CVD disease of the heart.

1.2. Objectives

The objectives of this work are

- To optimize data and balance network-level security requirements of WBAN and IoT, it is implemented.
- To enhance the image segmentation and classification task, optimized spatial attention is used.
- To understand and improve the efficiency of complex data structures, the Bidirectional Transformer Network is used.
- To concatenate features of the Bidirectional transformer on clinical and signal data.

The paper is organized from section 1 introduction and related works in the section 2. The methodology is described in section 3, and its results are given in section 4. The conclusion and future direction of the model are given in Section 5.

2. Related Work

Baghdadi et al. [18] suggested a Gradient Boosting model to forecast the presence of cardiovascular disease and to determine the most significant predictors based on their fuzzy values. Subsequently, various algorithms are applied to examine further and assess cardiovascular disease. A strong and reliable algorithm was able to improve cardiovascular disease in its initial stages. The machine learning model was properly trained after the data had been pre-processed and normalized.

Xtreme Gradient Boosting was used as a supervised learning approach that improved prediction accuracy by combining several decision trees. This Catboost model gave an average precision rate of 90.94%; additionally, it attained a maximum classification performance with higher precision. However, this system needs to enhance the accuracy by using advanced ML algorithms.

Kiliçarslan et al. [19] suggested a combined model to forecast cardiovascular disease by using both optimization techniques and Deep Learning methods. Adjusting the hyperparameters of the deep learning model was a crucial part of the training process. As a result, several research works in the past had focused on optimizing hyperparameters to enhance the effectiveness of deep learning models.

Initially, the adjustment values for the optimization techniques were carefully refined. Thereafter, the models were determined using the optimization strategies. The highest performance for the 1D VGG-16 model was obtained using Particle Swarm Optimization (PSO) and Grey Wolf Optimization (GWO) approach. However, this system leads to data overfitting and is time-consuming.

Elsedimy et al. [20] proposed a heart disease prediction model by incorporating the Quantum-behaved Particle Swarm Optimization (QPSO) with an SVM classifier, termed as QPSO-SVM. The categorical data in the dataset was converted to numerical data, then it was scaled down for normalization. The model optimizes the fitness function and selects key features. Finally, the self-adjust threshold helped fine-tune the parameters. The test outcome proved that the QPSO-SVM model attained 96.31% of accuracy in prediction. However, this system fails to manage the convergence in serious dimensional problems.

Table 1. Prior Research Analysis on Cardiovascular Disease Detection

Author name and reference	Methods	Performance	Disadvantages
Baghdadi et al. [18]	Deep Learning, Machine Learning algorithms, and Xtreme Gradient Boosting.	This Catboost model gave an average precision rate of 90.94 %.	This system needs to enhance the accuracy by using advanced ML algorithms.
Kiliçarslan et al. [19]	CSO, GWO	Average precision rate of 96.23 %.	This system leads to data overfitting and is time-consuming.
Elsedimy et al. [20]	QPSO–SVM.	An accuracy of 96.31% for predicting heart disease.	This system fails to manage the convergence of serious dimensional problems.
Wei et al. [21]	Slap Swarm Algorithm, and Lateral Mutation strategy (LMS).	SOLSSA Catboost achieved a high precision of 90%, respectively.	The Slap Swarm algorithm is data sensitive, and it stops searching for the optimal solution when it finds a good solution.
Yashudas et al. [22]	Bidirectional Gated Recurrent Unit (BiGRU)	Deep cardio approach attained a total precision rate of 99.90%, demonstrating highly effective performance.	This system is not suitable for real-time applications due to lag.
Mohapatra et al. [23]	Stacking Classifier, Machine Learning algorithm	The layered classification model performed better than the other machine learning models, achieving an accuracy rate of 92%.	In this system, a limited amount of data was used to train a machine learning algorithm.
Sumwiza et al. [24]	Random Forest	This model achieves a precision rate of 94% respectively	This system leads to high computational complexity.
Muhammad et al. [25]	K- Nearest Neighbour, RF, LR, GNB, and decision tree	The KNN classifier model achieved a precision rate of 91.8%, respectively	Absence of deep learning techniques and hardware integration had restricted its effectiveness.
Garavand et al. [26]	Artificial Intelligence Classification	Most of these models gave better performance in the feature extraction.	Some additional technologies need to be implemented to enhance the precision of the disease.
Muzammil et al. [27]	Artificial Intelligence and Deep Learning	This system achieved an accuracy of 78% in disease prediction	AI might cause security problems, and it can be easily hacked by attackers

Wei et al. [21] suggested a Sparrow-Based Algorithm was developed to fine-tune the parameters of the Cat Boost model, and it was applied for assessing the risk of Cardiovascular Disease. The improved SSA called SOLSSA was developed by integrating the elements of Slap Swarm Algorithm, Opposition-Based Algorithm, and Lateral Mutation Strategy (LMS). The SOLSSA (Catboost) model was tested on two CVD datasets from Kaggle to evaluate its performance in boosting prediction accuracy. SOLSSA Catboost achieved a high precision of 90%, respectively. However, the Slap Swarm algorithm is data sensitive, and it stops searching for the optimal solution when it finds a good solution. Yashudas et al. [22] suggested a Cardio Vascular Disease detection based on IoT (Internet of Things) Network (DEEP- Cardio) was implemented by using Bidirectional Gated Recurrent Unit (BiGRU) to offer early diagnosis and medical treatment and dietary advice for heart-related conditions. At the beginning, the patient's heart-related data were gathered remotely through the use of four types of biosensors, such as an ECG sensor, a blood pressure sensor, a pulse sensor, and a glucose sensor. The disease was predicted from the collected data through Adriano Controller. The suggested Deep cardio approach attained a total precision rate of 99.90%, demonstrating highly effective performance. However, this system is not suitable for real-time applications because of the lagging problem.

Mohapatra et al. [23] suggested a prediction system to identify heart diseases by stacking multiple classifiers arranged in two layers, which were the Base layer and Meta layer. Different types of learning algorithms were combined, allowing the model to generate more accurate and robust results. Machine Learning algorithm performed the data analytics, which reduced the processing time, allowing the cardiologist to spend more time planning treatment. The layered classification model performed better than the other machine learning models, achieving an accuracy rate of 92%. However, a limited amount of data was trained using a machine learning algorithm, which limits the scalability.
Sumwiza et al. [24] suggested a powerful ensemble approach using the Random Forest approach to enhance the prediction accuracy by merging several feature selection methods. During the processing, Data pre-processing and feature selection techniques like correlation analysis and data mining-based methods were used to eliminate the outlier values. In the final step, a cardiovascular disease detection model was developed, and its improved accuracy was evaluated using a confusion matrix. The dataset was analysed using the random forest algorithm and its performance was compared against other models such as, K- NN, SVM< Logistic Regression. This model achieves a precision rate of 94%, respectively. However, this system leads to high computational complexity.

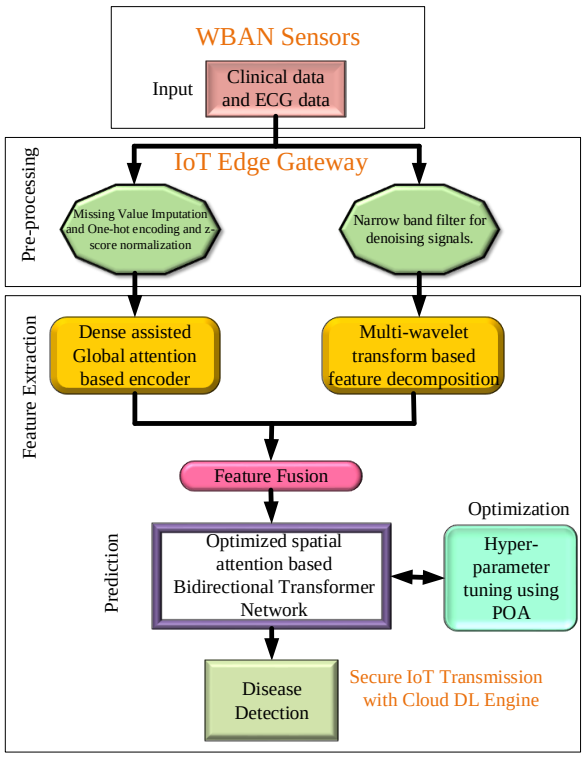

Fig. 1 Architectural diagram for Cardiovascular Disease Detection

Muhammad et al. [25] suggested a system to accurately diagnose ischemic cardiovascular disease using Machine learning algorithms. Six machine learning algorithms, namely K- Nearest Neighbour, Random Forest, Logistic Regression, Gaussian Naïve Bayes, and decision tree, were trained using data related to ischemic heart conditions. The KNN classifier model achieved a precision rate of 91.8% respectively. However, the absence of deep learning techniques and hardware integration had restricted its effectiveness.

Garavand et al. [26] proposed three machine learning approaches, including Machine Learning, Deep Learning/ Neural Network, and Ensemble methods, to diagnose Coronary Artery Disease. Among the techniques utilized, SVM, RF, Linear Regression (LR), and K-Nearest Neighbour were the most commonly implemented and better in feature extraction performance. However, some additional technologies need to be implemented to enhance the precision of the disease.

Muzammil et al. [27] suggested a model using Electrocardiography (ECG), which was enhanced using Artificial Intelligence, emerged as a promising method for accurately diagnosing and treating heart-related diseases. By applying AI, particularly in Deep Learning models like CNN, to analyse single, continuous, and occasional ECG signals, researchers developed fully automated systems that interpreted ECG in a way similar to humans and in some cases even more accurately and reliably. This system achieved an accuracy of 78% in disease prediction. However, AI might cause security problems, and it can be easily hacked by attackers. Table 1 outlines the overview of previous work.

2.1. Background

Several studies summarized in Table 1 [18, 21, 24, 25] have explored traditional machine learning approaches such as logistic regression, k-nearest neighbors, decision trees, and random forest models for cardiovascular disease detection. Baghdadi et al. [18] and Wei et al. [21] report reasonable predictive performance using ensemble-based models; however, these methods rely heavily on handcrafted features and are sensitive to noise and data quality.

Similarly, Sumwiza et al. [24] and Muhammad et al. [25] demonstrate moderate improvements in accuracy but suffer from high computational complexity, limited scalability, and reduced robustness when applied to high-dimensional or heterogeneous datasets. To address these limitations, deep learning-based approaches have been proposed [19, 22, 26, 27], including CNN-and RNN-based architectures. While these models improve automatic feature extraction, studies such as Kiliçarslan et al. [19] and Yashudas et al. [22] report challenges related to overfitting, increased training time, and latency, whereas other works [26, 27] highlight insufficient robustness to noise and lack of scalability in real-world deployments.

More recent research has incorporated optimization-driven and attention-based frameworks to enhance cardiovascular disease prediction performance [19-21]. Elsedimy et al. [20] employ QPSO-SVM to improve classification accuracy but encounter convergence difficulties in high-dimensional feature spaces, while Wei et al. [21] introduce SOLSSA-CatBoost, which remains data-sensitive and prone to premature convergence.

Overall, the comparative analysis in Table 1 [18-27] indicates that existing approaches often suffer from inadequate modeling of long-range and bidirectional temporal dependencies, limited exploitation of spatially relevant ECG features, weak multimodal fusion of ECG and clinical data, and optimization strategies with poor generalization. These limitations collectively motivate the proposed optimized spatial attention-enhanced bidirectional transformer network, which is designed to overcome the shortcomings of prior work by integrating bidirectional temporal modeling, spatial attention, multimodal feature fusion, and efficient optimization within a unified framework.

2.2. Research Gap

Despite significant progress in cardiovascular detection, existing methods exhibit several critical limitations that motivate this research. Traditional machine learning approaches rely on handcrafted features and struggle with high-dimensional, noisy, and heterogeneous clinical and ECG data, resulting in limited generalization and scalability. Deep learning models such as CNNs and RNNs improve feature learning.

However, they are constrained by their inability to effectively capture long-range and bidirectional temporal dependencies, high computational complexity, and latency issues, which hinder real-time deployment. Although recent optimization-driven and attention-based models enhance prediction accuracy, most existing approaches process ECG signals in a unidirectional manner, underutilize spatial attention to emphasize clinically relevant features, and lack a robust multimodal fusion strategy for integrating ECG and clinical data.

Furthermore, commonly used optimization suffers from premature convergence and poor exploration and exploitation balance, reducing model stability under noisy and imbalanced conditions. To handle these gaps, the proposed work introduces an optimized spatial-attention enhanced bi-directional transformer framework that explicitly models both past and future temporal dependencies, prioritizes diagnostically significant spatial features, integrates multimodal data within a unified architecture, and employs an efficient optimization strategy to ensure robustness, scalability, and superior diagnostic performance compared to existing methods.

3. Proposed Methodology

This paper proposes an optimized spatial attention-based bidirectional transformer network for Cardiovascular Disease Detection. In this system, IoT (Internet of Things) is combined with Wireless Body Area Network, which provides an effective method to improve health monitoring and medical diagnosis, especially in the case of heart-related conditions such as Cardiovascular Disease. Clinical and ECG data are preprocessed by Dense-assisted Global Attention-Based Autoencoder (DGA-AE) as given in figure.

3.1. Preprocessing

Initially, the clinical and ECG datasets are preprocessed by applying missing value imputation, followed by one-hot encoding for categorical features and z-score normalization to standardize the data.

3.1.1. Missing Value Imputation

Handling Missing values involves filling absent entries with reasonable estimates to preserve the quality of the data set and support effective analysis. Techniques for filling in missing data generally fall into two types: Single imputation, where one value is used to replace the missing value. Multiple imputation, where multiple values are estimated to reflect the uncertainty. In this proposed methodology, multiple missing value imputation is used to fill the various missing values in the dataset, which gives more accurate results and reduces bias. The missing value approach is illustrated in Figure 2.

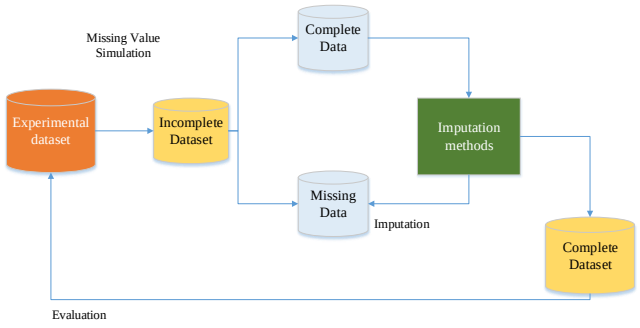

Fig. 2 Workflow of missing value imputation

Existing CVD detection studies predominantly rely on conventional ML models and DL architectures such as CNNs and RNNs, which depend on handcrafted features or suffer from limited ability to capture long-range temporal dependencies and bidirectional contextual information in ECG signals. Recent transformer-based approaches improve temporal modeling but are often unidirectional, lack a spatial attention mechanism, and primarily focus on single modality data, which limits the diagnostic robustness. In contrast, the novelty of this work lies in the unified integration of optimized spatial attention with a bidirectional transformer network that explicitly captures both past and future temporal dependencies in ECG signals while emphasizing clinically

significant feature regions. Unlike prior multimodal approaches that perform simple feature concatenation, the proposed framework jointly optimizes deep features extracted from ECG signals and clinical data using a nature-inspired puma optimization algorithm, which improves convergence stability and generalization. This comprehensive design enables superior performance under noisy and heterogeneous data conditions that demonstrate clear advancement over existing state-of-the-art CVD detection models.

3.1.2. One-Hot Encoding

One-hot encoding is used to convert the categorical features into a numerical data type. A categorical variable with values left, right, and center is processed using one-hot encoding, which transforms it into binary variables for left and right. Each of the new variables takes value of 0 and 1 to indicate whether the data is present or not.

If both left and right variables are 0, then the category is in the center. Also, one-hot encoding increases the dataset's feature count. In this proposed methodology, one-hot encoding is used to process the raw data and categorize the diseased and non-diseased hearts from the input images.

3.1.3. Z-score Normalization

Standardization, often referred to as Z-score Normalization, is used for data pre-processing. The dataset contains different types of continuous variables, which may be related but exist in different units and value ranges. Normalization adjusts these continuous variables to a normal scale, where the mean becomes 0, and the standard deviation becomes 1.

The transformation uses the z-score formula, which converts each value into a standardized score by subtracting the mean and dividing by the standard deviation. In this proposed methodology, Z-score normalization is used to convert all continuous features into a uniform scale or equal-weighted value.

$$N_1 = \frac{Y-\bar{Y}}{\sigma} \quad (1)$$

Here $Y$is the original data value, $\bar{Y}$ the sample mean, and σ the sample standard deviation.

3.1.4. Narrow Band Filter

Narrowband filters are designed to let through only a small, specific range of frequencies while stopping all others. This proposed system includes two matching reflective layers labelled b and c. Those are separated by a specific spacer. Based on the principle of Multi-beam inference, the amount of light that successfully passes through (called transmittance) within the allowed frequency range (passband) can be mathematically described in Equation (1).

$$S = \frac{(1-Q_b)(1-Q_c)}{\left[1-(Q_bQ_c)^{\frac{1}{2}}\right]^2} \bullet \left[\frac{1}{1+E\,sin^2\,\varphi}\right] \quad (2)$$

$$E = \frac{4(Q_bQ_c)^{\frac{1}{2}}}{\left[1-(Q_bQ_c)^{\frac{1}{2}}\right]^2}, \varphi = \frac{2\pi}{\lambda} nd - \frac{\varphi_b+\varphi_c}{2} = v\pi. \quad (3)$$

Where, Q_b,Q_c, $\varphi_b$and $Q_c$are the Refractive Index (RI) and the physical thickness of the spacer, respectively. The incident of light is referred to by$\lambda$and v=0, 1, 2 ...is the order of the space layer.

Typically, the passband width refers to a range of frequencies where the filter allows signals to pass, measured at the point where the signal strength is half of the maximum value, for narrowband pass filters that have mirrored surfaces arranged symmetrically. The bandwidth can be calculated by using the following formula.

$$BW_L(v,y) = \frac{1}{v}\left(\frac{t_K}{t_L}\right)^{2y} \frac{4t_U}{\pi t_L}\lambda_d\left(\frac{t_L-t_K}{t_L-t_K+t_K/v}\right) \quad (4(a))$$

(Height index spacer)

$$BW_K(v,y) = \frac{1}{v}\left(\frac{t_K}{t_L}\right)^{2y} \frac{4t_u}{\pi t_K}\lambda_d\left(\frac{t_L-t_K}{t_L-t_K+t_K/v}\right) \quad (4(b))$$

Wheret_K, $t_L$and t_U are the RI of the high-index material, low-index material, and substrate, the term y represents the layer number of the high refractive index material in the reflective stack, while $\lambda_d$stands for the filter's central wavelength. Because v and y are whole numbers, the bandwidth of the filters cannot be adjusted in a continuous way. This means that with a Fabry-Perot structure. It is not possible to create filters with just any desired bandwidth when the coating materials are already chosen.

3.1.5. Dense-Assisted Global Attention-Based Autoencoder

In this method, a 1D DenseNet architecture was employed to extract deep- level features. The Architectural diagram of DenseNet is given in Figure 3.

The primary reason for choosing DenseNet instead of ResNet is that ResNet can partially solve the vanishing gradient problem. It tends to become computationally heavy as the network goes deeper, leading to an increase in the number of parameters.

On the other hand, DenseNet handles these problems more efficiently by incorporating densely connected convolutional layers, where each layer gets the input from all previous layers. This results in better feature flow and reduced redundancy. Hence, DenseNet was used to provide a more resource-efficient and high-performing approach for predicting stock prices.

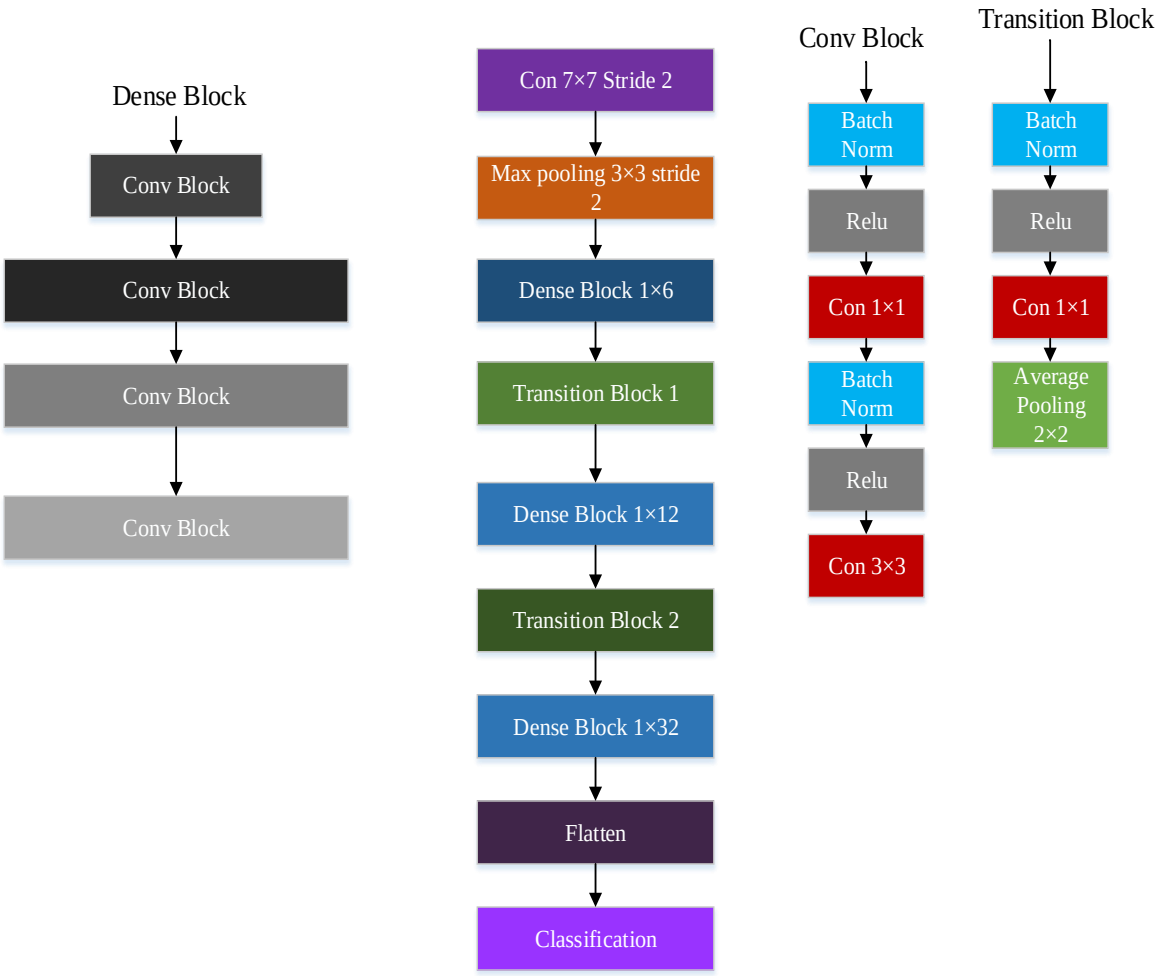

Fig. 3 Architectural diagram of DenseNet

DenseNet was built using a structure that includes dense blocks, along with transition layers and bottleneck layers. As shown in Figure 2, this step is made up of a network with (N-1) layers and a composite function, batch normalization, and convolutional layers. In this system, clinical data from physionet.org was fed into a 1D-DenseNet model to extract features.

At first, the filter count was initiated as 32 and then raised by 16 with every successive range. As a result, the size of the feature vector also grew because of merging the feature maps. To manage this growth and reduce the data size, the transition layer was added after each dense block, which handled the down-sampling process. The output at n^{th} the layer is given in equation (5).

$$M_n = L_n([M_0, M_1, M_2, \ldots, M_{n-1}]) \quad (5)$$

Here $M_n = L_n([M_0, M_1, M_2, \ldots, M_{n-1}])$ are the feature maps.

Autoencoder methods are a feed-forward neural network that includes multiple hidden layers, which has the setup of an equal number of features in the input and output layer, but the in-between layers have fewer features. Since the compression happens at the stage, the encoder section takes an input value Y_i from a space of dimensions d_y . It compresses it into a smaller representation h_i in a space of dimension d_h by applying an activation function $f(y)$, which is defined using the formulas below.

$$f(y) = \frac{1}{1+g^{-rx}} \quad (6)$$

The encoding of data is carried by equation (7) and decoding using equation (8)

$$h_i = f(ry_i) \quad (7)$$

$$y^s i = e(r^s y_i) \quad (8)$$

$$F = \sum_{i=1}^{t} g_i(e, r^s) \quad (9)$$

Here $F$is the overall sample error by summing individual samples as in equation (9).

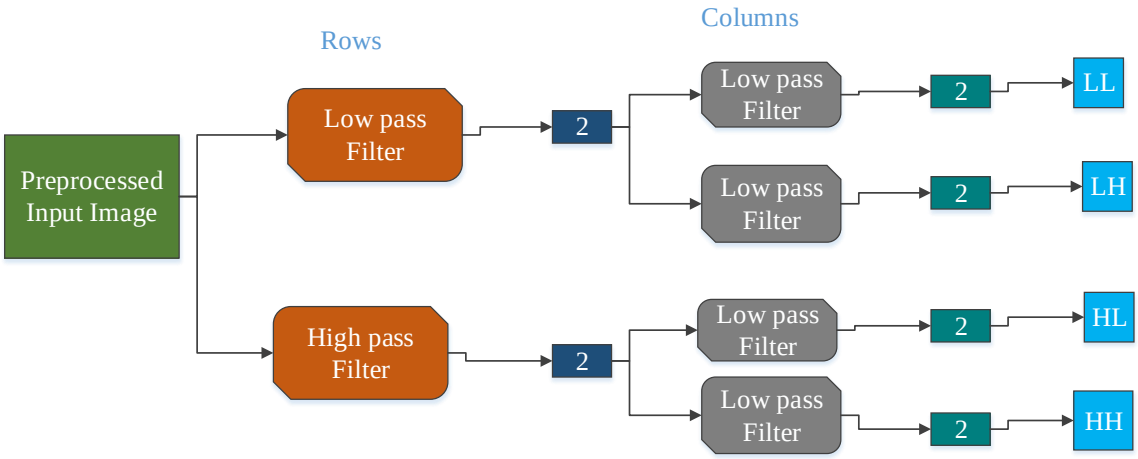

Fig. 4 Multi-Wavelet Transform Based Feature Decomposition

3.1.6. Multiwavelet Transform-Based Feature Decomposition (MT-FD)

In this proposed system, the Wavelet transform can be implemented in either a continuous or discrete format. The Discrete Wavelet Transform (DWT) focuses on a sampled and quantized version of the original continuous transform, which significantly reduces computational complexity. The workflow of the multiwavelet transform-based feature decomposition is given in Figure 4.

This is achieved by employing scaling and wavelet functions that are shifted and dilated by powers of two. These dyadic scales and positions allowed for efficient signal representation. For a particular dilation and translation c, the wavelet coefficient $R_f(b,c)$ for a signal f can be calculated as,

$$R_f(b,c) = \langle f, \psi_{b,c} \rangle = \int f(y)\psi_{b,c}(y)\, dy \qquad (10)$$

In DWD, a signal is passed through a sequence of filters (associated with scaling functions) and a high-pass filter (related to wavelet functions). This process is known as multiresolution analysis, which breaks down the signal into approximation and detail components. The low-pass filter extracts the general trend of the signal, while the high-pass filter captures its rapid fluctuations. Wavelet coefficients represent the information contained in a signal at the corresponding dilation and translation. The inverse transform is applied to reconstruct the original signal.

$$f(y) = \frac{1}{D_r}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{f(b,c)}\psi_{b,c}(y)dc\,\frac{db}{b^2} \qquad (11)$$

After filtering, the signal is down-sampled by a factor of two, effectively reducing the number of samples and focusing on significant information. This process can be replaced by the approximation coefficients to generate multiple levels of decomposition, each offering a more refined resolution of the signal's characteristics, where Dψ is the normalization factor. To obtain a finite number of scaling and wavelet coefficients, a finite number of translations are used in applying multiresolution analysis at a given scale$I$.

$$f(y) = \sum_L D_{IL}\,\varphi_{IL}(y) + \sum_{i=1}^{I}\sum_l d_{il}\,\psi_{il}(y) \qquad (12)$$

Where D_{IL}represents the scaling coefficients and d_{il} denotes wavelet coefficients. The first part of the equation shows a simplified version of the low resolution. DWT makes it particularly useful in applications such as signal compression, noise reduction, and feature extraction. Its ability to localize both time and frequency information ensures a more compact and efficient representation compared to traditional Fourier techniques. The second part captures finer details at various levels of precision, starting from the original signal to the current levels of analysis. Applying the discrete wavelet transform involves using a series of filters, illustrated in the Figure. For a 2D image, the transform can separate the image into four different detail levels: LL (approximation), LH (Horizontal), HL (Vertical), and HH (diagonal components). For deeper analysis, applying N levels of DWT can yield 3N+1 frequency sub bands.

Each stage of decomposition splits the image into high and low frequency components. The low-frequency part can undergo further breakdown until the required detail level is achieved. When decomposition is done multiple times, it is known as multilevel or multiresolution analysis. In practice, especially for image fusion, just one level of breakdown might be enough. This depends on how image resolution is being combined and compared. The standard DWT can be used in two ways, through a decimal method or a non-decimal method. In decimation, the data size is reduced after each step by picking one value in every two columns and rows. This especially shrinks the image into one of its original size with each transformation, decreasing the detail levels.

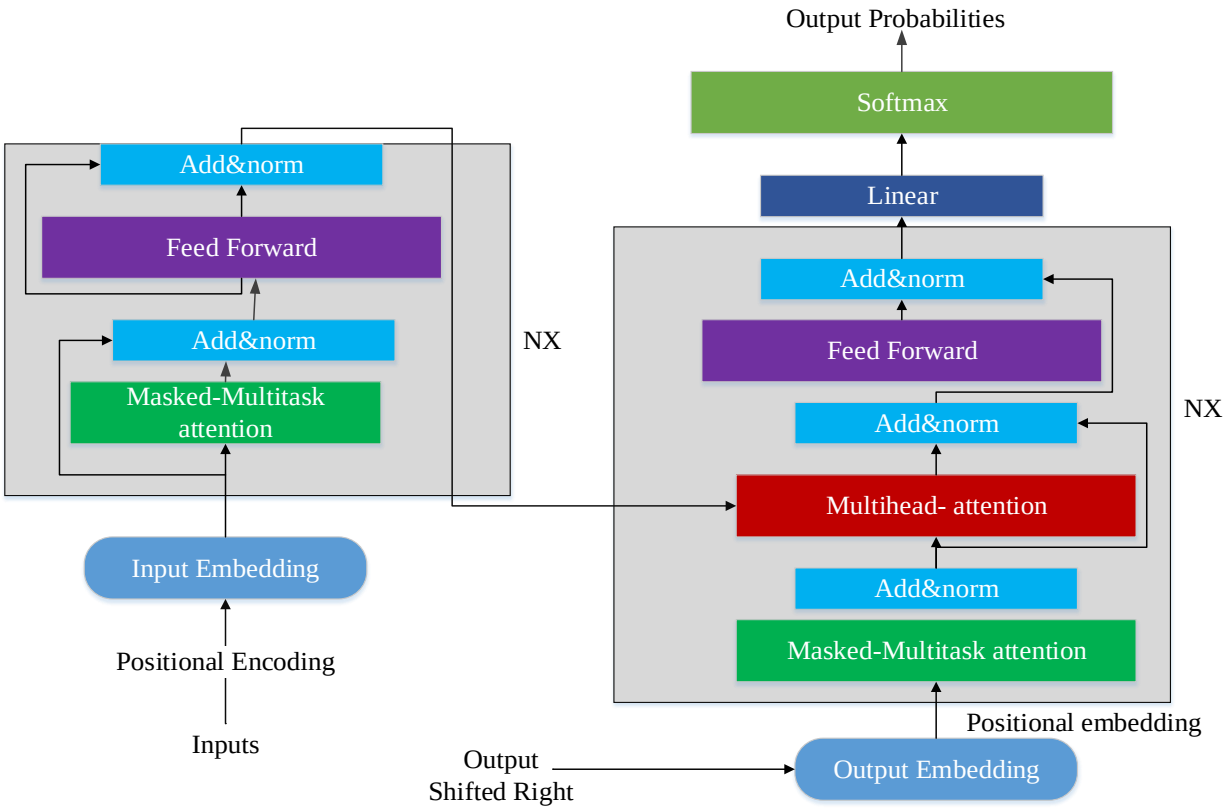

Fig. 5 Bidirectional Transformer Architecture

3.2. Optimized Spatial Attention-Based Bidirectional Transformer Network (OPA-BETN)

The transformer model, which was first developed for translating text in Natural language processing tasks, is naturally good at capturing long-distance relationships in sequence. However, standard Transformers may struggle to understand complex contextual details. To improve this, the proposed methodology extended the transformer architecture and developed a Bidirectional Transformer (BiTrans). It improves how the model learns time-based relationships in the features extracted from the clinical and ECG signals by using the data from both past and future waveforms. Since the input to the BiTrans model is a feature map, which contains abstract spatial features obtained from earlier layers. This Bidirectional transformer is similar in concept to Bidirectional LSTM, which uses a method to reverse the order of the feature map, adding an extra input that captures the sequence in the opposite direction, allowing the model to understand the entire context better. The workflow of the Bidirectional Transformer Architecture is given in Figure 5. This method takes into account both earlier and later parts of the input at the same time while making predictions about a particular point. By using data from both directions (past and future), it builds a deeper understanding of the input, allowing it to detect complex patterns more effectively. This differs from the regular transformer encoder, which typically looks only at the previous parts of the sequence and ignores what comes after the current position. To express it more technically, the self-attention method is implemented by generating three versions of the input features, denoted by e). These are then passed through linear layers to generate the query(R), key (L), and value (M) used in attention processing e_R^y, e_L^y, e_M^yas in equation (13).

$$e_R^y = T_R^y \cdot e, e_L^y = T_L^y \cdot e, e_M^y \cdot e, \tag{13}$$

There T_R^y, T_L^y, T_M^y are learnable projection matrices for the y^{th} head. To keep things simple, ignore the specific layer notation. The spatial attention mechanism for each head is then applied as mentioned below:

$$B^y = Attention\left(e_R^y, e_L^y, e_M^y\right) = soft\,max\left(\frac{e_R^y (e_L^y)^W}{\sqrt{d_L}}\right) e_M^y \tag{14}$$

Where d_L is the dimension e_L^y for the y-th head?

The multihead attention system gathers outputs from several separate attention heads. After combining all these outputs, it applies a final layer to generate the complete attention result used by the model.

$$MHA(e,e,e) = Concat(B^1, \ldots, B^Y) \cdot T_{o,} \tag{15}$$

Where $MHA(e,e,e)$represents the multihead self-

attention output and $T_{o,}$is the output projection matrix for the final multihead attention output.

The main advancement in the Bidirectional Transformer Network is the use of attention in both forward and backward directions. To enable this, the input feature map is also processed in reverse order, allowing the model to consider both earlier and later elements (tokens) in the sequence when making decisions.

$$e_{rev}[p,i] = z[p,d-i-1] \quad (16)$$

For p=0, 1,..,.n-1 and i=0,1,…..,d-1. Apply the reversed input to equations (14) and (15), the output of the original feature map and its reversed counterpart, as follows.

$$e_{out} = MHA(e,e,e) + MHA\left(e_{rev,}e_{rev,}e_{rev}\right) \quad (17)$$

Each layer within the Bidirectional Transformer passes its output to the next layer in the sequence, enabling the model to learn and capture more complex patterns progressively. The final result of the Bidirectional transformer Network, referred to as BiTrans, is produced by repeatedly applying each layer on the original input e. Essentially, this process can be described as a step-by-step transformation of the input through all layers of the network.

$$BiTrans(e) = Layer_c(Layer_{c-1}(\ldots(Layer_1(e))\ldots)\ ) \quad (18)$$

3.2.1. Puma Optimization Algorithm (POA)

The optimized solution from the Bidirectional transformer is used in this Puma optimization. In this proposed methodology, the Puma Optimization (PO) algorithm introduces a unique strategy for shifting between exploration and exploitation phases, which is described as both innovative and intentional and is reportedly introduced in this algorithm. In both exploration and exploitation phases, distinct techniques are used to carry out the optimization process. Here is how it works with nature-inspired metaphors. The workflow of the puma optimization algorithm is given in Figure 6.

- The best solution is to identify the particular affected area in the heart
- The entire search identifies the disease that affects the heart
- Other potential research has identified that the heart is unhealthy.

In each iteration, the algorithm intelligently and deliberately decides whether a solution should enter exploration or exploitation based on a phase switching mechanism. This decision is not random but follows a calculated logic. During exploration, two specialized strategies are used, indicating the symptoms of heart disease and the affected area in the heart.

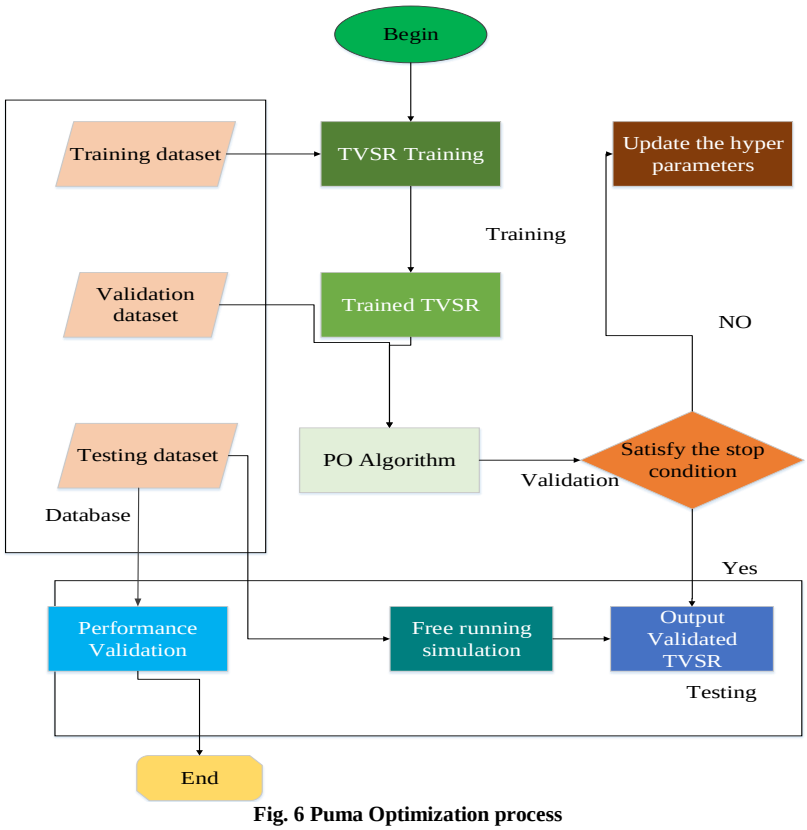

Fig. 6 Puma Optimization process

$$g1_{Explor} = PF_1 \cdot \left(\frac{Seq^1_{CostExplore}}{Seq_{Time}}\right) \quad (19)$$

$$g1_{Exploit} = PF_1 \cdot \left(\frac{Seq^1_{CostExploit}}{Seq_{Time}}\right) \quad (20)$$

$$g2_{Explor} = PF_2 \cdot \left(\frac{Seq^1_{CostExplore}+Seq^2_{CostExplore}+Seq^3_{Cost\,explore}}{Seq^1_{Time}+Seq^2_{Time}+Seq^3_{Time}}\right) \quad (21)$$

$$g^2_{Exploit} = PF_2 \cdot \left(\frac{Seq^1_{CostExploit}+Seq^2_{Cost\,explore}+Seq^3_{Cost\,explore}}{Seq^1_{Time}+Seq^2_{Time}+Seq^3_{Time}}\right) \quad (22)$$

Where Seq_{cost} is determined from equations (19-22), Seq_{time} is set to 1. PF_1 and PF_2 are prioritized the $g^1(f)$ and $g^2(f)$.

$$Seq^1_{CostExplore} = \left|Cost^{Initial}_{Best} - Cost^1_{Explore}\right| \quad (23)$$

$$Seq^2_{CostExplore} = \left|Cost^2_{Explore} - Cost^1_{Explore}\right| \quad (24)$$

$$Seq^3_{CostExplore} = \left|Cost^3_{Explore} - Cost^2_{Explore}\right| \quad (25)$$

$$Seq^1_{CostExploit} = \left|Cost^{Initial}_{Best} - Cost^1_{Exploit}\right| \quad (26)$$

$$Seq^2_{CostExploit} = \left|Cost^2_{Exploit} - Cost^1_{Exploit}\right| \quad (27)$$

$$Seq^3_{CostExploit} = \left|Cost^3_{Exploit} - Cost^2_{Exploit}\right| \quad (28)$$

In Equations (23) and (26), $Cost^{Initial}_{Best}$,$Cost^1_{Explore}$, $Cost^2_{Explore}$, $Cost^3_{Explore}$, $Cost^1_{Exploit}$, $Cost^2_{Exploit}$, $Cost^3_{Exploit}$ are the cost functions with the exploitation and exploration as given in equations (29) and (30).

$$Score_{Explore} = \left(PF_1 \cdot g1_{Explore}\right) + \left(PF_2 \cdot g2_{Explore}\right) \quad (29)$$

$$Score_{Exploit} = \left(PF_1 \cdot g1_{Exploit}\right) + \left(PF_2 \cdot g2_{Exploit}\right) \quad (30)$$

The values $Score_{Explore}$ and $Score_{Exploit}$ decide the exploration or exploitation at the 3rd iteration; the solutions were produced independently in each step.

4. Results and Discussion

The implementation of this proposed system is done in a Python program. The materials used for training and testing ML classification models are SVM, LR, and Adaboost. The experimental setup was carried out on a system running Windows 10 Pro, equipped with an Intel® Core™ i5-3570 processor operating at 3.40 GHz.

The machine was configured with 8 GB of RAM and utilized a 64-bit x64-based architecture, providing a stable and efficient computing environment for executing the proposed model and related simulations.

4.1. Dataset Description

The proposed model utilizes two types of datasets, namely, ECG (Electrocardiography) and Cardiovascular Disease datasets obtained from the Kaggle repository [28, 29].

The ECG data is obtained from the PTB-XL repository, which contains. PTB-XL ECG has clinical 12- lead ECG recordings of 21, 799, which comprises 52% male recordings and 48% female patient recordings for every 10 seconds with different ages. Each recording contains a raw waveform of ECG data, which is reviewed by one or more cardiologists who assigned one or more interpretations for each ECG.

The Cardiovascular Disease dataset has 70,000 instances with 11 features of patients' data, which comprises the patient age, height, weight, and gender, followed by the 4 examination features. The dataset also has 3 habit-related features to support the disease prediction, and the last is the binary label for the detection of cardiovascular disease.

4.1.1. Performance Analysis of the ECG Dataset

The metrics used for the evaluation are described as follows,

$$A_C = \frac{D_{TN}+D_{TP}}{D_{TN}+D_{TP}+D_{FN}+D_{FP}} \quad (31)$$

$$P_C = \frac{D_{TP}}{D_{TP}+D_{FP}} \quad (32)$$

$$R_C = \frac{D_{TP}}{D_{TP}+D_{FN}} \quad (33)$$

$$F_{1\_score} = 2 * \frac{P_C * R_C}{P_C + R_C} \quad (34)$$

Here A_C is the accuracy for true positives and true negatives in the data D_{TN} and D_{TP}. D_{FN} and D_{FP} It is the false positives and negatives. P_C is the precision, R_C is the recall, and F_{1_score} is F1score. To access the proposed model's performance, a K-fold cross-validation is carried out. Table 2 and Figure 7 illustrate the down-sampling results of ECG data.

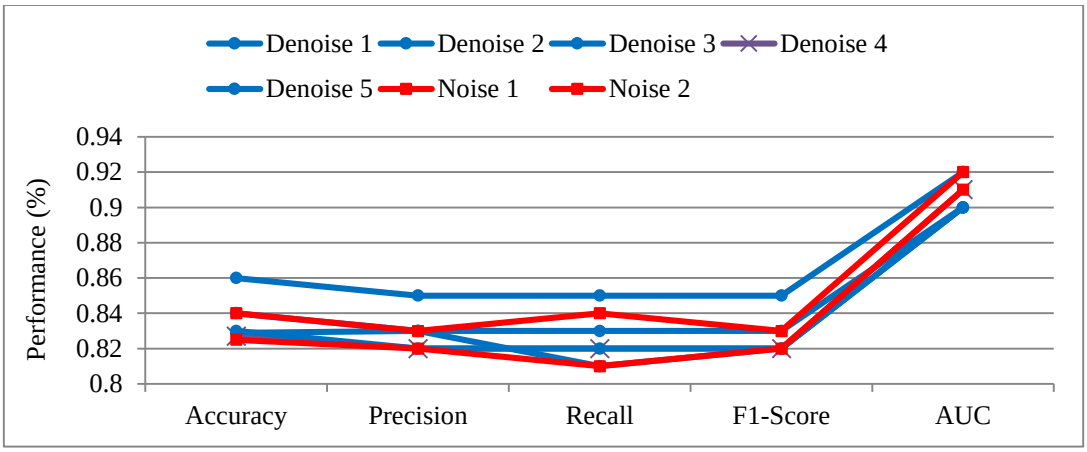

Fig. 7 Down-sampling results of the proposed model

Table 2 presents the performance evaluation of the proposed model under both denoised and noisy ECG signal conditions using multiple statistical metrics. For the denoised signals, the classification accuracy ranges approximately between 0.82 and 0.86, with precision and recall values remaining consistently above 0.82, indicating stable and reliable detection capability. The corresponding F1-scores, which lie in the range of 0.82 to 0.85, reflect a balanced trade-off between precision and recall. Additionally, the AUC values, varying from 0.90 to 0.92, demonstrate the model's strong ability to discriminate between classes when noise is reduced. In the presence of noise, the model continues to maintain competitive performance, achieving an accuracy of up to 0.84. The precision, recall, and F1-score values remain close to 0.82–0.84, suggesting that the proposed approach exhibits robustness even under degraded signal conditions. The high AUC scores of 0.91 and 0.92 further confirm the model's effectiveness in distinguishing pathological ECG patterns despite noise interference. Overall, the results indicate consistent and reliable performance across varying signal qualities.

Table 2. Down-sampling results

Signals	Accuracy A_C	P_C	R_C	$F_{1\ score}$	AUC
Denoise	0.86	0.85	0.85	0.85	0.92
	0.829	0.83	0.81	0.82	0.90
	0.84	0.83	0.83	0.83	0.90
	0.827	0.82	0.82	0.82	0.91
	0.83	0.82	0.82	0.82	0.90
Noise	0.84	0.83	0.84	0.83	0.92
	0.825	0.82	0.81	0.82	0.91

Table 3. All signals results

Signals	A_c	P_C	R_C	$F_{1\ score}$	AUC
Denoise	0.85	0.84	0.95	0.89	0.90
	0.86	0.87	0.83	0.84	0.93
	0.85	0.86	0.81	0.83	0.91
Noise	0.87	0.87	0.85	0.86	0.93
	0.86	0.86	0.83	0.84	0.92
	0.86	0.86	0.83	0.84	0.92

The results are summarized in Table 3 and Figure 8, showing the performance of the proposed model for both denoised and noisy ECG signals across standard evaluation metrics. Under the denoised condition, the model achieves accuracies in the range of 0.85 to 0.86, with precision values remaining consistently high. The recall varies across experiments, reaching a peak value of 0.95, which indicates the model's strong ability to identify relevant cardiac events correctly. The corresponding F1-scores, ranging from 0.83 to 0.89, reflect a balanced classification performance, while AUC values between 0.90 and 0.93 demonstrate reliable class discrimination. The model continues to perform effectively, recording accuracies between 0.86 and 0.87. Precision and recall values remain stable, leading to F1-scores of up to 0.86, which suggests robustness against noise interference. The consistently high AUC scores (0.92–0.93) further confirm the model's capability to distinguish between classes even in challenging signal conditions.

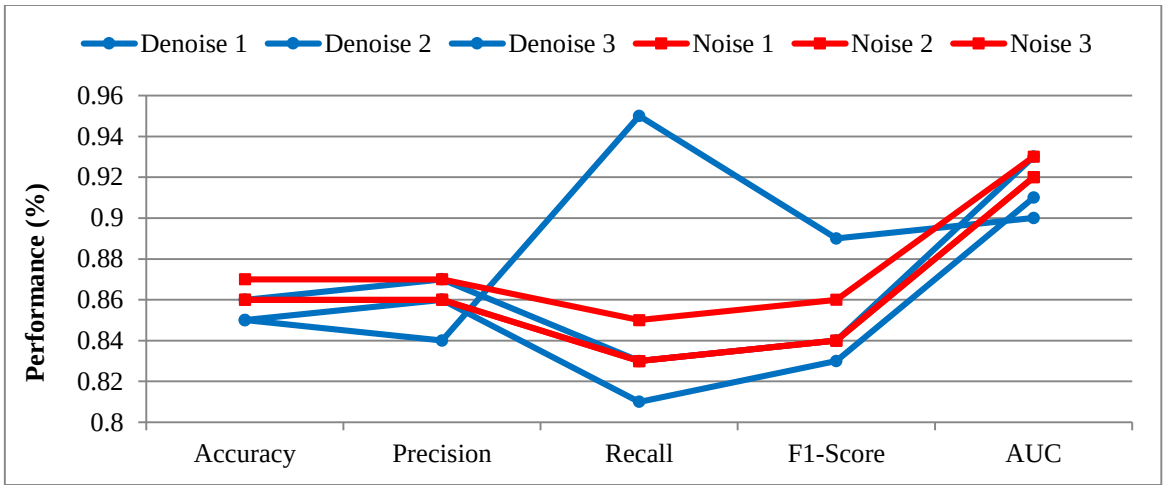

Fig. 8 Performance evaluation of all the signals

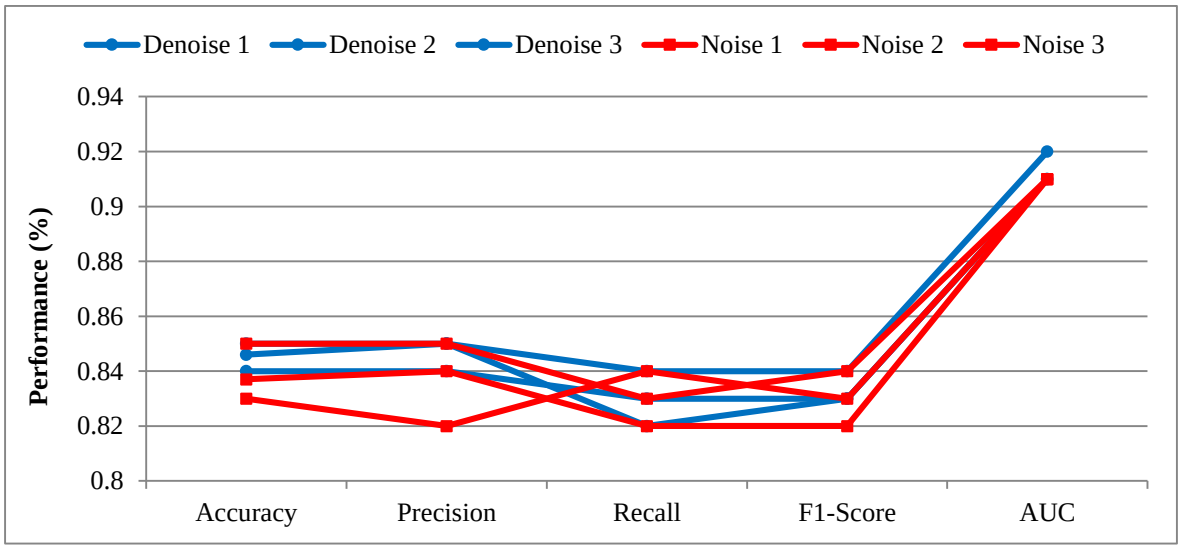

Fig. 9 Upsampling results of the proposed model

Table 4. Upsampling results

Signals	A_c	P_C	R_C	$F_{1\ score}$	AUC
Denoise	0.84	0.84	0.83	0.83	0.91
	0.85	0.85	0.84	0.84	0.92
	0.846	0.85	0.82	0.83	0.91
Noise	0.83	0.82	0.84	0.83	0.91
	0.85	0.85	0.83	0.84	0.91
	0.837	0.84	0.82	0.82	0.91

Table 4 and Figure 9 present the performance outcomes of the proposed model under up-sampling conditions for both denoised and noisy ECG signals. In the denoised scenario, the model achieves accuracies ranging from 0.84 to 0.85, with precision and recall values remaining idle, indicating consistent classification behaviour. The resulting F1-scores are between 0.83 and 0.84, which reflect a balanced trade-off between sensitivity and precision. Moreover, the AUC values of 0.91 to 0.92 demonstrate the model's reliable discriminative capability when up-sampled, denoised signals are used. Table 5 and Figure 10 indicate that using a 1D convolutional layer for training took about 995 seconds on average to complete 5 training cycles (Epochs). In contrast, the 2D Convolutional layer required more time, taking around 1428 seconds on average for the same number of epochs, indicating that the 2D model was more computationally intensive. Ultimately, a 1D convolutional layer was selected due to its fast processing and low memory demand, as well as the 5-fold cross-validation performance of the trained model across different training epochs.

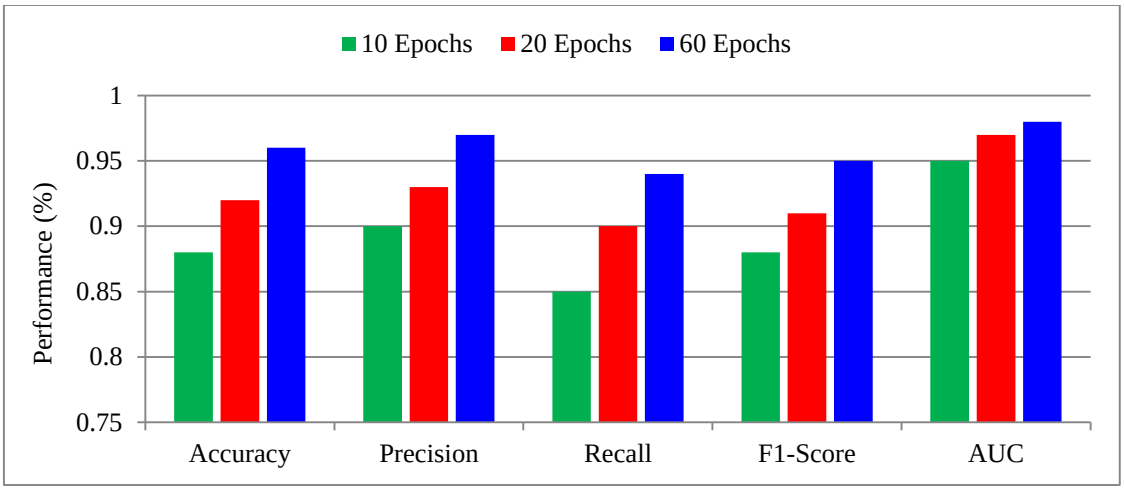

Fig. 10 5-fold cross-validation results of the proposed model

Table 5. 5-fold cross-validation results

Epochs of the Trained Model	A_c	P_C	R_C	F_{1_score}	AUC
10	0.88	0.90	0.85	0.88	0.95
20	0.92	0.93	0.90	0.91	0.97
60	0.96	0.97	0.94	0.95	0.98

At 10 epochs, the model achieves an accuracy of 0.88, with a precision of 0.90 and a recall of 0.85, resulting in an F1-score of 0.88 and an AUC of 0.95, indicating promising early learning behavior. When the number of epochs is increased to 20, the model shows notable improvement, reaching an accuracy of 0.92, along with enhanced precision (0.93) and recall (0.90). This leads to an F1-score of 0.91 and an AUC of 0.97, reflecting better generalization. The comparative analysis of the proposed approach against existing state-of-the-art models was evaluated on the PTB-XL dataset. Conventional convolutional neural network (CNN)–based methods that incorporate entropy-driven features have reported a maximum classification accuracy of 89%. Similarly, ResNet-based architectures utilizing visibility graph representations combined with ResNet or Inception frameworks have achieved an accuracy of 89.71%, along with an F-score of 79.61% and an AUC of 3.46%. Hybrid architectures that integrate beat segmentation with CNN and attention mechanisms followed by a Bi-LSTM classifier have demonstrated a mean accuracy of 88.85%, accompanied by a macro ROC-AUC of 0.9216 and an F1-score of 0.8057. In comparison, the proposed Bi-directional Transformer Network significantly outperforms these existing models, attaining a superior classification accuracy of 95.34%. This improvement highlights the effectiveness of the transformer-based architecture in capturing complex temporal dependencies within ECG signals, leading to enhanced diagnostic performance.

4.1.1. Performance Analysis of Cardiovascular Disease dataset

Here is a performance analysis of the CVD dataset from the Kaggle repository, presented in Table 6 and Figure 11. In this proposed model, major techniques were used to identify cardiovascular disease. These contribute by identifying a proposed model, which provides the best performing technique that can predict cardiovascular disease with better accuracy of early detection.

Table 6. Performance analysis on Cardiovascular Disease dataset

Algorithm	R_C	F_{1_score}	P_C	A_c
Decision Tree	64.40	63.94	63.42	63.69%
K-NN	61.46	67.02	73.68	69.87%
LR	67.99	71.13	74.58	72.36
NB	32.30	44.43	71.11	59.44%
SVM	64.21	70.17	77.35	72.66%
Proposed Model	**72.22**	**80.12**	**90.32**	**97.21%**

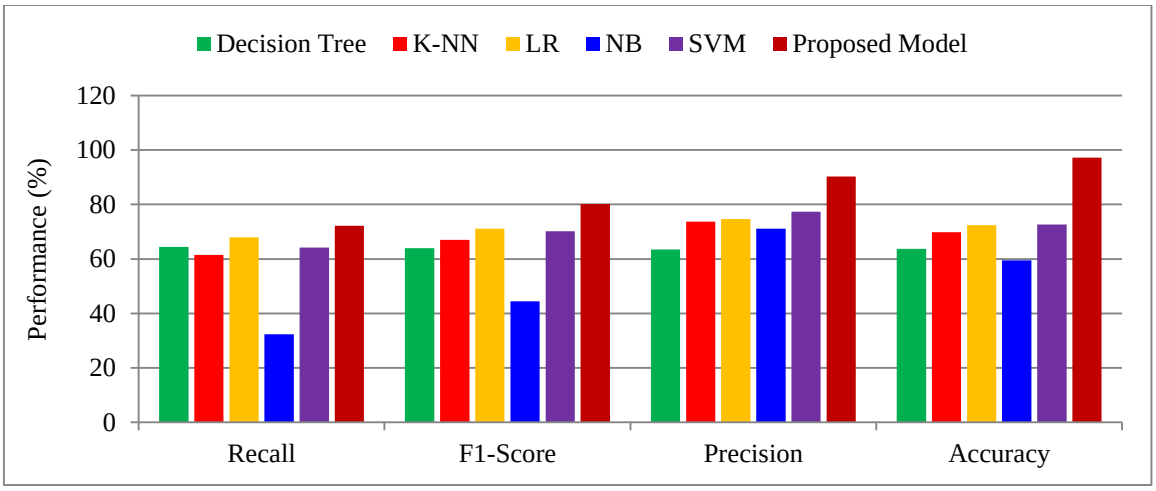

Fig. 11 Performance analysis on Cardiovascular Disease dataset for the proposed model

4.2. Discussion

This proposed model effectively overcomes the limitations given in the previous research. The Gradient Boosting model [18] provides less accuracy due to the usage of existing modules, so this proposed model utilizes the advanced classification algorithms to attain maximum accuracy. The cardiovascular disease forecast model [19] leads to data overfitting and is more time-consuming. To overcome this problem, the proposed methodology used ECG signals and utilized the DenseNet to overcome overfitting and reduce time complexity. Further, QPSO - SVM [20] fails to manage the convergence in high-dimensional problems, so this proposed system overcomes the convergence of high dimensional by utilizing the DenseNet, which improves the gradient flow and feature reuse through dense connections. Batch Normalization helps to stabilize the convergence. Likewise, a Sparrow-Based Algorithm [21] is data sensitive, so this proposed methodology used the optimized spatial-based bi-directional transformer-based network to find the optimal solution effectively. Bidirectional Gated Recurrent Unit (BiGRU) [22] is not suitable for real-time applications because of the lagging problem, so this proposed system used the Puma Optimization algorithm to optimize the model's performance and ensure high accuracy. Likewise, the Heart Disease prediction system [23] possesses limited data to train the model, so this proposed system used two datasets, including a large dataset to train the models. K- NN, SVM, Logistic Regression [24] leads to high computational complexity, which was overcome by a 1D CNN tailored for ECG time series data, which was lighter than a 2D CNN. Multiple machine learning models [25] enhance the accuracy. However, they do not give better performance as compared to transformer-based models, so this proposed system aims to overcome the drawback by using advanced deep learning algorithms. Likewise, a Coronary Artery Disease (CAD) and diagnosis model was developed by Garavand et al. [26] by utilizing the AI technology. Most of the AI models gave better performance in the feature extraction, but they failed to enhance the accuracy rate. So this proposed model used the DenseNet and two optimization algorithms to enhance the accuracy. Further, the Cardiovascular Disease detection model was developed by Muzammil et al. [27], which emerged as a promising method for accurately diagnosing and treating heart-related diseases. Additionally, developed fully automated systems that interpreted ECG in a way similar to humans and in some cases even more accurately and reliably, which failed to secure the data of patients. To overcome the insecurity of data, this proposed methodology uses secure data encryption and transmission protocols to ensure the security of data.

The superior performance of the proposed optimized spatial attention-enhanced bidirectional transformer network can be attributed to several key design choices that directly address the limitations of existing state-of-the-art methods. First, unlike conventional CNN-based approaches that primarily capture local temporal patterns, the proposed bidirectional transformer explicitly models long-range dependencies by processing ECG features in both forward and backward directions. This enables the network to leverage complete contextual information from past and future signal segments, which is critical for accurately identifying subtle and transient cardiac abnormalities. Second, the incorporation of spatial attention allows the model to dynamically focus on diagnostically significant feature regions while suppressing irrelevant or noisy components, leading to more discriminative feature representations compared to attention-free or uniform-attention models reported in prior studies.

In addition, the proposed framework outperforms existing methods due to its effective multimodal feature integration strategy. The suggested model takes advantage of a unified learning framework to optimally fuse deep representations from both modalities, as opposed to previous

research that relies on clinical data or ECG signals alone or on simple feature concatenation. This complementary use of physiological signals and clinical attributes enhances robustness and generalization, particularly in heterogeneous and real-world datasets. Furthermore, the use of a DenseNet-based autoencoder improves feature reuse and gradient flow, reducing overfitting and stabilizing convergence compared to deeper CNN architectures reported in the literature. Another critical factor contributing to improved performance is the integration of the Puma Optimization Algorithm, which provides a balanced exploration-exploitation mechanism during optimization. In contrast to commonly used metaheuristic techniques that suffer from premature convergence or high sensitivity to data distribution, the proposed optimization strategy enables stable convergence in high- dimensional feature spaces. It improves hyperparameter tuning and feature selection efficiency. This results in consistent performance gains across different datasets and signal conditions, including noisy and imbalanced scenarios. Collectively, these architectural and algorithmic enhancements explain why the proposed method achieves higher accuracy and robustness than previously reported state-of-the-art cardiovascular disease detection techniques. Table 7 illustrates the Comparison of this proposed system with existing systems. Types of datasets show the accuracy and precision in Table 7. Cleveland Heart Disease and Kaggle are popular datasets for identifying cardiovascular disease.

Table 7. Comparison of this proposed system with existing systems

Author Name	Dataset	Accuracy
Baghdadi et al. [18]	Heart condition data	0.70
Kiliçarslan et al. [19]	Kaggle (Heart Disease)	0.87
Elsedimy et al. [20]	Cleveland Heart Disease database	0.96
Wei et al. [21]	CVD (Kaggle)	0.81
Yashudas et al. [22]	Framingham's and stat log heart disease dataset	0.86
Mohapatra et al. [23]	UCI (Heart Disease)	0.92
Sumwiza et al. [24]	Kaggle (Heart Disease)	0.96
Muhammad et al. [25]	UCI repository (Heart Disease)	0.92
Garavand et al. [26]	Kaggle (Heart Disease)	0.92
Muzammil et al. [27]	Kaggle (Heart Disease)	0.93
Proposed	**ECG ptb-XL**	**95.34**
	Kaggle (Cardiovascular Disease)	**97.21**

5. Conclusion

The proposed model is capable of accurately identifying multiple cardiovascular diseases. The models were developed using the large data sets, which made the detection process more efficient. To support the diagnosis of the early detection system of heart diagnosis data, data were sourced from Kaggle and Ptb-XL. In this system, 70,000 patient records were used. The effective model is used in this method to identify the ECG signals. A powerful Machine Learning algorithm based on logistic regression is used to predict cardiovascular disease. In this proposed methodology, Kaggle (Cardiovascular Disease dataset) was used to predict cardiovascular diseases, which gives a High accuracy of 97.21% respectively by using Multiwavelet transform, Machine Learning, and Logistic regression. The second dataset, ECG ptb-XL, achieves an accuracy of 95.34% in measuring the electrical activity of the heart. This proposed model effectively detects the disease in the heart earlier with higher accuracy of detection. Future work will explore this model with AI robotics to detect heart disease.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Funding Statement

The authors have not received any funding for this project

References

[1] Abdullah Alqahtani et al., "Cardiovascular Disease Detection using Ensemble Learning," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Yuansheng Gao, *Architecture of the Blood Vessels*, Biology of Vascular Smooth Muscle: Vasoconstriction and Dilatation, pp. 3-17, Springer, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jay P. Rabadia et al., *Cardiovascular System, its Functions and Disorders*, Cardioprotective Plants, pp. 1-34, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Christiana Wuche, "The Cardiovascular System and Associated Disorders," *British Journal of Nursing*, vol. 31, no. 17, pp. 886-892, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Muhammad Hamid et al., "Fine Tuned CatBoost Machine Learning Approach for Early Detection of Cardiovascular Disease through Predictive Modeling," *Scientific Reports*, vol. 15, pp. 1-15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[6] Mohammed B. Abubaker, and Bilal Babayiğit, "Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 2, pp. 373-382, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Vivek Pandey, Umesh Kumar Lilhore, and Ranjan Walia, "A Systematic Review on Cardiovascular Disease Detection and Classification," *Biomedical Signal Processing and Control*, vol. 102, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mohammad Moshawrab et al., "Reviewing Multimodal Machine Learning and Its Use in Cardiovascular Diseases Detection," *Electronics*, vol. 12, no. 7, pp. 1-30, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Adedayo Ogunpola et al., "Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases," *Diagnostics*, vol. 14, no. 2, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Pedro A. Moreno-Sánchez et al., "ECG-based Data-driven Solutions for Diagnosis and Prognosis of Cardiovascular Diseases: A Systematic Review," *Computers in Biology and Medicine*, vol. 172, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jiachuan Long et al., "A Comprehensive Review of Signal Processing and Machine Learning Technologies for UHF PD Detection and Diagnosis (II): Pattern Recognition Approaches," *IEEE Access*, vol. 12, pp. 29850-29890, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Tahseen Ullah et al., "Machine Learning-based Cardiovascular Disease Detection using Optimal Feature Selection," *IEEE Access*, vol. 12, pp. 16431-16446, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Awad Bin Naeem et al., "Heart Disease Detection Using Feature Extraction and Artificial Neural Networks: A Sensor-Based Approach," *IEEE Access*, vol. 12, pp. 37349-37362, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Omneya Attallah, "CerCan·Net: Cervical Cancer Classification Model Via Multi-layer Feature Ensembles of Lightweight CNNs and Transfer Learning," *Expert Systems with Applications*, vol. 229, pp. 1-40, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Yoon Sang Cho, and Seoung Bum Kim, "Supervised Contrastive Learning for Multisensor Signals Classification in Automobile Engine Manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 5, pp. 7764-7776, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Lin Liu et al., "Camera-Based Dual-Wavelength Defocused Speckle Imaging for Multi-Point Seismocardiographic Motion Measurement," *IEEE Journal of Biomedical and Health Informatics*, pp. 1-13, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[17] D. M. G. Preethichandra et al., "Wireless Body Area Networks and Their Applications—A Review," *IEEE Access*, vol. 11, pp. 9202-9220, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Nadiah A. Baghdadi et al., "Advanced Machine Learning Techniques for Cardiovascular Disease Early Detection and Diagnosis," *Journal of Big Data*, vol. 10, pp. 1-29, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Serhat Kiliçarslan, "PSO + GWO: A Hybrid Particle Swarm Optimization and Grey Wolf Optimization based Algorithm for Fine-Tuning Hyper-Parameters of Convolutional Neural Networks for Cardiovascular Disease Detection," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 87-97, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] E. I. Elsedimy, Sara M. M. AboHashish, and Fahad Algarni, "New Cardiovascular Disease Prediction Approach using Support Vector Machine and Quantum-behaved Particle Swarm Optimization," *Multimedia Tools and Applications*, vol. 83, pp. 23901-23928, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Xi Wei et al., "Risk Assessment of Cardiovascular Disease based on SOLSSA-CatBoost Model," *Expert Systems with Applications*, vol. 219, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] A. Yashudas et al., "DEEP-CARDIO: Recommendation System for Cardiovascular Disease Prediction Using IoT Network," *IEEE Sensors Journal*, vol. 24, no. 9, pp. 14539-14547, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Subasish Mohapatra et al., "A Stacking Classifiers Model for Detecting Heart Irregularities and Predicting Cardiovascular Disease," *Healthcare Analytics*, vol. 3, pp. 1-10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Kellen Sumwiza et al., "Enhanced Cardiovascular Disease Prediction Model using Random Forest Algorithm," *Informatics in Medicine Unlocked*, vol. 41, pp. 1-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] Ghulam Muhammad et al., "Enhancing Prognosis Accuracy for Ischemic Cardiovascular Disease Using K Nearest Neighbor Algorithm: A Robust Approach," *IEEE Access*, vol. 11, pp. 97879-97895, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Ali Garavand et al., "Towards Diagnostic Aided Systems in Coronary Artery Disease Detection: A Comprehensive Multiview Survey of the State of the Art," *International Journal of Intelligent Systems*, vol. 2023, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[27] Muhammad Ali Muzammil et al., "Artificial Intelligence-enhanced Electrocardiography for Accurate Diagnosis and Management of Cardiovascular Diseases," *Journal of Electrocardiology*, vol. 83, pp. 30-40, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[28] PTB-XL, A Large Publicly Available Electrocardiography Dataset, PhysioNet, 2022. [Online]. Available: https://physionet.org/content/ptb-xl/1.0.3/

[29] Cardiovascular Disease Dataset, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset