International Journal of Reconfigurable and Embedded Systems (IJRES)
Vol. 15, No. 1, March 2026, pp. 183~193
ISSN: 2089-4864, DOI: 10.11591/ijres.v15.i1.pp183-193

Heart disease prediction using hybrid deep learning and medical imaging with wavelet-based feature extraction

Chairmadurai Palanisamy[1], Kavitha Pachamuthu[2], Arun Kumar Ramamoorthy[3]
[1]Department of Computer Science, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS) Pallavaram, Chennai, India
[2]Department of Computer Applications, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS) Pallavaram, Chennai, India
[3]Digital Forensics and Cyber Security, University of South Wales, Treforest, United Kingdom

Article Info

Article history:

Received Nov 14, 2024
Revised Jan 4, 2026
Accepted Feb 17, 2026

Keywords:

Butterworth adaptive 2D
Convolutional neural network
Deep learning algorithms
Heart disease prediction
Long short-term memory
Wavelet filter

ABSTRACT

The process of heart disease prediction is based on patient medical information, which can be addressed in terms of medical image as well as the results of an electrocardiogram (ECG) conducted to determine the risk of developing heart disease. The hybrid deep learning (DL) algorithms are developed using past data that can identify trends related to cardiovascular disease (CVDs). In the current paper, it is possible to offer a new method of heart disease prediction that would combine high-quality image processing and hybrid DL to enhance the effectiveness of predictions and avoid the shortcomings of the modern approaches. First, medical images like ECG images are pre-processed with butterworth adaptive 2D wavelet filter, which ensures maximal noise reduction, followed by maintenance of spatial and frequency information. The Gabor Wavelet-based feature extraction technique is applied to extract meaningful patterns, including both spatial and frequency domain information, which is essential for detecting heart-related anomalies. The resultant features are then categorized, along with both convolutional neural networks (CNN) and long short-term memory (LSTM), to make reliable and precise predictions of heart disease. The performance indicators, including accuracy (92.4%), precision (91.2%), recall (93.5%), and F1-score (91.0%), are utilized. Applying the model yields significant levels of reliability and generalization compared to traditional applications.

Corresponding Author:

Kavitha Pachamuthu
Department of Computer Applications, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS) Pallavaram
Chennai, India
Email: pkavikamal@gmail.com

1. INTRODUCTION

In order to detect and diagnose heart disease early, sophisticated computational methods have been developed because it is one of the world's leading causes of mortality. Deep learning (DL) algorithms have been investigated in recent surveys for their potential to improve prediction accuracy and guide medical decision-making. An example is when the authors suggested a strong heart disease prediction system based on hybrid DL neural networks, whose performance proved to be satisfactory due to the deep integration of features in the system [1]. Evaluated how various feature selection algorithms influence hybrid DL and determined that proper feature selection methods can significantly improve models' efficiency and expert interpretability [2]. Used convolutional neural network (CNN) and long short-term memory (LSTM) in predicting heart diseases and cast light on the performance of CNN and LSTM in dealing with nonlinear data structures [3]. A combination of different classification models and the Boruta feature selection algorithm to further process the input attributes and hence increase the percent of accuracy in the classification results [4]. Moreover, it employed self-measurable attributes that reflect wellness status by using hybrid DL algorithms, and the article highlights the possibility of personal data as the basis of predictive diagnostics. The contributions above are significant in that model selection, feature engineering, and data-driven approaches are highlighted to improve the reliability of methods for predicting heart diseases [5].

Objectives: i) to enhance early diagnosis and treatment planning by creating a reliable heart disease prediction model that utilizes patient data, such as electrocardiogram (ECG) images, ii) to use butterworth adaptive 2D wavelet filtering that would be effective in noise removal without affecting essential parts of the image, iii) to accurately extract heart abnormalities, spatial, and frequency-based features are extracted using Gabor wavelet transforms, and iv) to develop a hybrid DL model that integrates CNN with LSTM to provide an intelligent system that can accurately identify the presence of heart illness.

Contributions of the work: i) an image processing and DL-based hybrid framework is constructed to predict heart diseases, ii) effective noise removal and detail preservation are achieved by the butterworth adaptive 2D wavelet filter, iii) Gabor wavelet identifies abundant spatial and frequency components to improve pattern identification, and iv) CNN and LSTM classifiers enhance the accuracy of prediction, addressing data imbalance and non-linearity.

The remaining portion of the document is divided into significant sections, which are described as follows: section 2 examines the current research efforts in heart disease prediction using hybrid DL and medical imaging with wavelet-based feature extraction used by different authors. Section 3 explains the workflow of the suggested approach in the proposed method. Section 4 presents the findings analysis and performance data. Section 5 presents the conclusion.

2. LITERATURE REVIEW

Bhatt *et al*. [6] demonstrate that this approach may involve a show that can precisely anticipate cardiovascular clutters to decrease the deaths caused by the clutters. This paper proposes utilizing a k-mode clustering strategy to enhance classification accuracy. Models such as decision tree (DT) classifiers, random forests (RF), multilayer perceptrons, and XGBoost (XGB) are employed.

Chandrasekhar and Peddakrishna [7] suggested a hybrid decision support system that might be used to identify issues early on. Multivariate imputation using chained equations has been used to heart disease, depending on the patient's clinical characteristics. An algorithm will handle the missing values. Recursive feature elimination is used to choose relevant features from the given dataset once the genetic algorithm (GA) has been hybridized with the feature selection process [8].

Pathan *et al*. [9] assert that prompt and accurate identification of cardiac disease is essential for preventing further harm to individuals. Artificial intelligence-based medical modalities are examples of non-invasive medical treatments that have been used recently. In particular, machine learning (ML) has employed several widely used algorithms and techniques that are highly adaptable and frequently used to accurately identify cardiac illnesses in a short period.

The contribution of the research of Ahmed and Husein [10] is the critical analysis and use of ensemble learning and other combination machine-learning techniques to forecast cardiac conditions. Ensemble learning techniques are used to data sets and factors used in the cleveland and data port heart disease datasets were age, blood pressure, blood glucose, resting ECG, heart rate, and four types of chest pain.

According to Diwakar *et al*. [11], the diagnosis of infections is the most crucial component of healthcare. An illness can save lives when it is identified sooner than anticipated or when it is more common. ML classification approaches can benefit the healthcare sector by facilitating accurate and timely sickness identification. Due to its difficulty in diagnosis, heart disease is currently one of the most dangerous conditions in the world. It can therefore be advantageous to both patients and specialists.

Rindhe *et al*. [12] reported that in recent decades, heart-related illnesses, often known as cardiovascular diseases (CVDs), have been the leading cause of mortality in the majority of the world's countries. They are lately considered to be the most fatal disease not only in India but also on this planet. Subsequently, a dependable, accurate, and practical framework is required to identify such maladies in development to receive proper treatment. ML calculations and strategies have been connected to assorted restorative information to mechanize the ponder of huge and complex information.

According to Patro *et al*. [13], the category of information in this of intrigued to ponder is the information category. Able to utilize tests and prepare information through categorization to produce a forecast demonstration. With the combination of computer algorithms and numerical methods, a classification calculation generates such information [14]. It develops a versatile show capable of handling detailed information, which will encompass the same information types.

Mehmood *et al*. [15] described Cardio Help, a deep-learning solution that leverages CNN to show the probability of cardiovascular illness in a patient. The specified method proposes a CNN-based heart failure (HF) prediction at the initial level, which encompasses temporal data modeling. We have developed a heart disease dataset, contrasted the results by applying innovative techniques, and achieved decent results.

El-Shafiey *et al*. [16] outlines how the best features are selected to increase the forecast accuracy of heart illness using a hybrid GA that blends RF and particle swarm optimization (PSO) techniques. Initially, the suggested GAPSO-RF uses multivariate statistical analysis to determine the most significant characteristics in the original population. Next, a discriminative mutation process is used by the GA.

According to Dissanayake and Johar [17], the efficiency of classification algorithms used for model development and the relevant characteristics selected by various feature selection methods were examined through an experimental study. Under the feature subset provided based on the backward feature selection strategy, the feature subset with the highest classification rate, 88.52%, precision of 91.30%, and is achieved using the DT classifier.

Nancy *et al*. [18] stated that the internet of things simplifies the connection between people and devices, and when integrated with cloud computing, it enhances quality of life. As new AI and ML technologies proliferate in the medical field, medical predictive analytics holds the potential to transform our proactive approach to healthcare. DL, a branch of ML, has the ground-breaking ability to analyse massive volumes of data precisely at previously unheard-of speeds, yielding incredible results and swiftly and efficiently addressing complex problems.

Ashri *et al*. [19] suggest hybridized classifiers that combine the majority voting approach and the ensemble model to increase prediction accuracy. To improve prediction levels and reduce time consumption, a feature selection and preprocessing strategy based on GA is also demonstrated. A comparison of the study's findings showed that the suggested ensemble classifier model outperformed other pertinent improvements in the classification rate, with the latter proportion increasing by up to 98.18%.

Ramesh *et al*. [20] conducted experiments using the UCI dataset containing 303 instances and 76 features. Among these, 14 features were selected to evaluate the effectiveness of different classification approaches. The separation timberland method normalizes the information utilizing the foremost key highlights and measures of the information, which features a higher degree of exactness. The choice tree classifier incorporates 4 and 18 highlights and the irregular timberland classifiers. The exploratory comes about to demonstrate that K-nearest neighbors (KNN) with eight neighbors is solid sufficient to assess its adequacy, affectability, exactness, and exactness. The comparative analysis is shown in Table 1.

Table 1. Comparative analysis for heart disease prediction

Ref.	Author and year	Dataset used	Algorithm used	Result achieved	Limitations
[21]	Gao *et al*., 2021	UCI heart disease	Ensemble (RF and gradient boosting machine (GBM))	Accuracy of 89%, outperforming single models	Low interpretability and high computational cost
[22]	Asif *et al*., 2021	UCI	SVM, KNN, RF, DT, and logistic regression (LR)	RF achieved ~88% accuracy	Limited to classical ML and no hybrid or DL models
[23]	Akella and Akella, 2021	UCI (Cleveland)	LR, Naïve Bayes (NB), SVM, and KNN	SVM achieved ~85% accuracy	No DL and small benchmark dataset
[24]	Nandy *et al*., 2023	Framingham dataset	Swarm intelligence and artificial neural network	Accuracy above 90%	Small dataset and risk of overfitting
[25]	Yazdani *et al*., 2021	Malaysian and UCI	LR with feature strength score	87% accuracy with improved interpretability	Dataset lacks global diversity
[26]	Ozcan and Peker, 2023	UCI heart dataset	CART (DT)	Good and interpretable performance	Limited capability for complex patterns
[27]	Ishaq *et al*., 2021	HF dataset	SMOTE+data mining	AUC>0.90 after balancing	Risk of overfitting due to oversampling
[28]	Budholiya *et al*., 2022	UCI heart disease	Optimized XGB	Achieved 94.85% accuracy	Black-box nature and tuning complexity
[29]	Chang *et al*., 2022	UCI and Kaggle	SVM, KNN, NB, and DT	SVM achieved 86% accuracy	No feature engineering or ensemble use
[30]	Rubini *et al*., 2021	UCI and Statlog	LR, SVM, DT, RF, and NB	Accuracy of 85%	No real-time or advanced modeling

Table 1 lists 10 current studies that use different machine-learning techniques to predict cardiac disease [21]–[30]. In order to compare the types of contributions made by each study, the table displays the findings, methods, restrictions, and datasets. The UCI dataset and other publicly accessible datasets were used in the majority of the studies and employed well-known classifiers, including RF, support vector machine (SVM), and XGB. The accuracy was expressed in percentages, ranging from 85 to 94.85%, with the ensemble models and hybrid ones performing better. However, in general, it was limited by the absence of DL applications, the threat of overfitting, and the difficulties associated with the interpretability of models and the diversity of the datasets.

3. METHOD

The suggested algorithm of heart disease forecasting combines the current high-end image processing and DL practices to enhance the accuracy of diagnosis. To start with, medical images of patients, ECG scans of a patient, or cardiac imagery data will first be preprocessed using a butterworth adaptive 2D wavelet filter, which reduces noise content in an image while retaining some of its edges and frequency content in an efficient manner. The Gabor wavelet transform has been used to extract features, as it provides both spatial and frequency domain information, which are helpful in identifying meaningful patterns that are significant to heart disease. To enable the precise and dependable prediction of the characteristics displayed in Figure 1, they are then input into two classifiers, specifically CNN and LSTM. To achieve robustness, generalization, and efficiency in clinical decision-making, the system is evaluated using measures including accuracy, precision, recall, F1-score, and area under the curve-receiver operating characteristic (AUC-ROC).

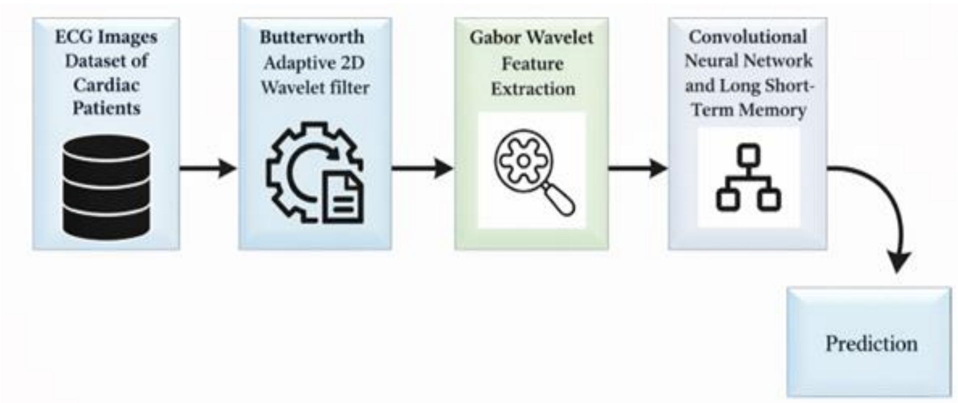

Figure 1. Overall flow diagram of proposed work

3.1. Electrocardiogram images dataset of cardiac patients

The dataset illustrates the four primary categories into which the dataset is divided, each of which represents a distinct heart disease. Myocardial infarction (MI) patients' ECG images, patients' ECG images with irregular heartbeats, patients with a history of MI and normal people's ECG images are among them.

Table 2 patients with MI, sometimes referred to as a heart attack, are shown in these pictures. The ECG patterns commonly associated with this serious state are visible in the photos. ECG pictures from patients with irregular heartbeat patterns fall under this group. These patterns may indicate a range of arrhythmias or other cardiac conditions, providing valuable information for research and diagnosis.

Table 2. ECG image dataset overview

Dataset overview	No. of images	Total dimensions
Myocardial infarction patients ECG images	240	240×12 (total of 2880 images)
Patients with abnormal heartbeat ECG images	233	233×12 (total of 2796 images)
ECG images of patients with a history of MI	172	172×12 (total of 2064 images)
Normal person ECG images	284	284×12 (total of 3408 images)

3.2. Preprocessing using butterworth adaptive 2D wavelet

To improve the quality of data for DL models and minimize noise, the butterworth adaptive 2D wavelet preprocessing of heart disease prediction is an excellent method. In medical datasets, particularly those containing signal data or images such as ECGs or heart scans, noise and irrelevant variance can reduce the accuracy of prediction. To eliminate this, the butterworth filter, known for its smooth frequency response, is applied to filter out high-frequency noise while retaining the key features of the signal. This is integrated with the adaptive 2D wavelet transform, which can further break down the data of the signal or image into various resolutions, allowing for better noise spotting and feature extraction.

$$Mean\ (\mu) = \sum_T X_{m,n} \tag{1}$$

Where μ mean of the input signal/image or data matrix, $X_{m,n}$ is intensity or value at the pixel, T is total number of elements in the data.

$$\sigma = \sqrt{\frac{1}{T}\sum_T (X_{m,n} - \mu)^2} \tag{2}$$

Where σ is standard deviation of the data, $X_{m,n}$ is original data value, and μ is mean of the data.

$$N_f(m,n) = m\left(\mu_{y1} - \mu\right)^2 + n\left(\mu_{y2} - \mu\right)^2 \tag{3}$$

Where $N_f(m,n)$ is frequency energy or feature emphasis at location, μ_{y1}, μ_{y2} are local means in horizontal and vertical directions, m,n pixel or coefficient positions, and μ is global mean.

$$E_e = \sum_T (\mu_{mn} - \mu) \tag{4}$$

Where E_e is energy difference or total deviation energy in the transformed domain, μ is global mean, and μ_{mn} is local mean at each position,

$$N_q(i,j) = \frac{N_f(m,n) - \mu_y}{\sigma_y} \tag{5}$$

Where $N_q(i,j)$ is normalized frequency, $N_f(m,n)$ is frequency features a m,n, μ_y, and σ_y is mean and standard deviation of the local or directional components,

$$C_v = \sum_{m,n=0}^{T-1} N_{q(m,n)}(i-j)^2 \tag{6}$$

Where C_v is final adaptive variation or contrast measure across data, $N_{q(m,n)}$ is normalized values at each point, i,j is position indices or feature dimensions, and T is total number of elements.

3.3. Gabor wavelet feature extraction

Gabor wavelet feature extraction is a process of forecasting heart disease by analyzing medical data, primarily in the form of imaging or signal-based data, such as ECGs, to identify patterns related to cardiovascular conditions. Gabor wavelets are useful feature extractors because they can extract the spatial and frequency characteristics of specific signals or images, which would be helpful in emphasizing the texture and slight differences in biomedical signals or images. Under this strategy, preprocessing of the input data is necessary, followed by the application of Gabor filters of various orientations and scales to obtain meaningful features that may then be used to describe the nature of the heart signal or image.

$$\varphi_{u,v}(hm) = \frac{\left|\left|k_{u,v}\right|\right|^2}{\sigma^2} exp(\frac{|k_{u,v}|\text{|^2||}z\text{||^2}}{2\sigma^2} \times [exp\left(ik_{u,v}z\right) - exp\left(-\frac{\sigma^2}{2}\right)] \tag{7}$$

Where $\varphi_{u,v}(hm)$ is the Gabor wavelet function for scale u and orientation v, applied to input hm, $k_{u,v}$ is wave vector that defines the frequency and orientation of the Gabor wavelet is $\left|\left|k_{u,v}\right|\right|^2$, magnitude of the wave vector, controlling the central frequency of the wavelet, and z is spatial position:

$$A_{u,v}(z) = \sqrt{Re\left(o_{\mu,v}(z)\right)^2 + Im\left(o_{\mu,v}(z)\right)^2} \tag{8}$$

Where $A_{u,v}(z)$ is the magnitude (amplitude) of the Gabor filter response at position z, for a given scale u and orientation v, $o_{\mu,v}(z)$ is the complex output (response) of the Gabor filter at z, and $\left(o_{\mu,v}(z)\right)$ is imaginary part of the Gabor filter response.

$$\theta_{u,v}(z) = \frac{archtan\ (ImIm\left(o_{\mu,v}(z)\right)}{Re\left(o_{\mu,v}(z)\right)} \tag{9}$$

Where $\theta_{u,v}(z)$ is phase angle of the Gabor response at position z indicating the phase shift of the texture or signal at that point and $archtan$ is inverse tangent function used to compute the phase angle from the imaginary and real parts of the complex response.

3.4. Classification: convolutional neural network and long short-term memory

Heart disease may be accurately predicted using a hybrid CNN-LSTM model that combines temporal sequence modeling (LSTM) with spatial prediction (CNN), as shown in Figure 2. CNN extracts local and regional patterns and characteristics from medical records, including organized health data, image patterns, and ECG signals. The popular network for classifying heart diseases is the CNN. This is due to training on ECG or medical scan images. Automatically, CNN predetermines the crucial features, which include the pattern of waves, abnormalities, or textures in the input image. The network can learn local and global patterns associated with heart disease detection using convolution and pooling, which capture the regional and global patterns, respectively.

$$Z_{i,j}^{(k)} = \left(\mathrm{X} * \mathrm{W}^{(\mathrm{k})}\right)i,j + b^{(k)} \tag{10}$$

Where X is input ECG image, $\mathrm{W}^{(\mathrm{k})}$ is filter (kernel) for the k^{th} feature map, $b^{(k)}$ is bias, $*$ convolution operation, and $Z_{i,j}^{(k)}$ is output at position (i, j) in feature map k.

$$A_{i,j}^{(k)} = ReLU\ (Z_{i,j}^{(k)} = \max(0, Z_{i,j}^{(k)}) \tag{11}$$

$$y = \sigma(W_{fc}.x + b_{fc})$$

Where x is flattened feature vector, W_{fc}, b_{fc} are weights, and bias of FC layer, and σ is SoftMax or sigmoid function depending on binary/multiclass classification.

$$L = -\left[y \log(\widehat{y}) + (1-y)\log(1- \hat{y})\right] \tag{12}$$

LSTM is used to model the temporal correlations, which is useful for time-series data like heart rate and ECG readings. LSTM networks are a compelling method for classifying heart disease when analyzing a sequence, such as ECG signals. It is also capable of learning long-term dependencies and patterns, such as intervals of heartbeat.

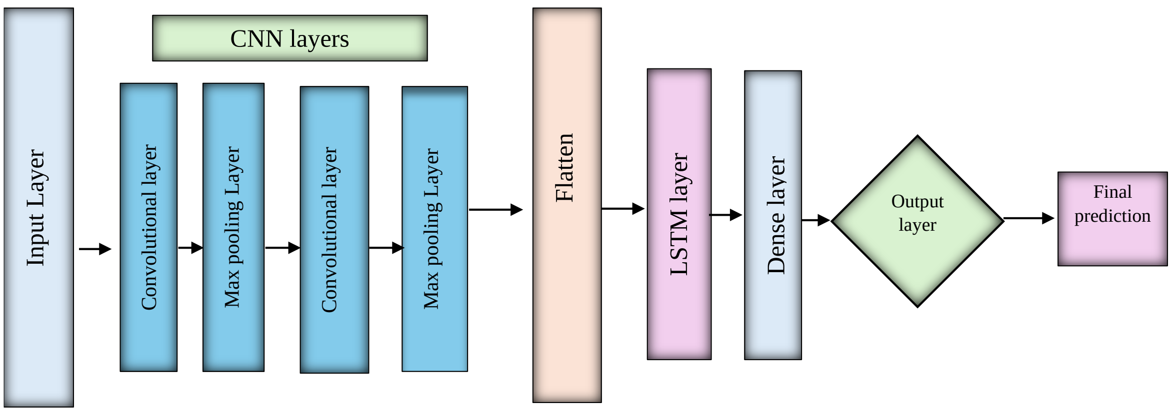

Figure 2. Structure of proposed hybrid CNN-LSTM model

4. RESULTS AND DISCUSSION

The standard execution measures were utilized to test the proposed half-breed show based on its viability in anticipating heart maladies. Its results illustrated way better precision, accuracy, review, and AUC-ROC in comparison with conventional models. Figure 3 illustrates the raw, unprocessed ECG recording. The signal appears blurred and has low contrast, making it difficult to distinguish essential waveform features such as P waves, QRS complexes, and T waves. The filtered image is shown on the right side of Figure 3. The image denoising and enhancement using a hybrid filtering approach that combines the butterworth adaptive wavelet filtering approach.

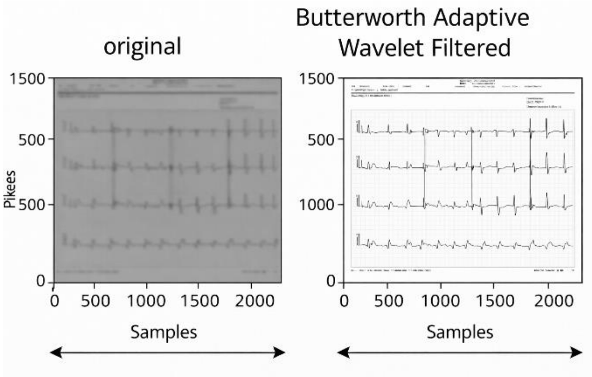

Figure 3. Preprocessed ECG signal image using butterworth adaptive wavelet filtered

Table 2 presents a comparison of the performance of various models used to classify heart diseases. The proposed CNN+LSTM hybrid achieved the best accuracy of 92.4%, along with excellent levels of precision (91.2%), recall (91.2%), specificity (93.5%), and F1-score (91.0%), making it a powerful model. The classic models, such as KNN, ANN, and DT, exhibited lower scores to a lesser degree. ANN managed to prove itself better than the KNN and DT methods; however, the hybrid model CNN+LSTM outdid all the rest in all measures of evaluation.

Table 2 Performance comparison

Models	Accuracy	Precision	Recall	F1-score
KNN	85.3	82.1	83.5	82.8
ANN	88.9	87.0	87.0	86.7
DT	83.2	80.5	80.5	80.7
Proposed CNN+LSTM	92.4	91.2	91.2	91.0

Figure 4 illustrates the comparison of classification accuracy among four deep-learning models in detecting heart diseases. The accuracy of the proposed CNN+LSTM model was 92.4%, indicating that it is better suited to learn both temporal and spatial characteristics of ECG data. Although it has a respectable accuracy of 88.9%, the ANN model lacks temporal memory. With respective accuracies of 85.3% and 83.2%, the KNN and DT models showed a moderate level of efficacy. A comparison of the accuracy scores of several categorisation models used to predict heart disease is displayed in Figure 5. With relatively few false positives, the proposed CNN+LSTM hybrid model also obtains the maximum precision of 91.2%, demonstrating its capacity to effectively identify positive instances of heart disease. Next is the ANN model, which outperforms the conventional models with an accuracy of 87.0%. Lower accuracy values of 82.1% and 80.5% are attained by KNN and DT models, respectively. The relative recall performance of several DL models in the classification of heart disease is shown in Figure 6.

Figure 7 presents a comparison of specificity values for four models of heart disease classification prediction. The proposed CNN+LSTM has shown a specificity of 93.5%, indicating that it is better at predicting negative (non-heart disease) cases. The ANN model achieves 90.2% accuracy and indicates good classification ability. The KNN has a specificity of 86.7%, which is higher than that of the DT at 84.6%. This analogy highlights that deep learning-based models, particularly CNN+LSTM, are more accurate than conventional hybrid deep learning strategies in detecting healthy individuals.

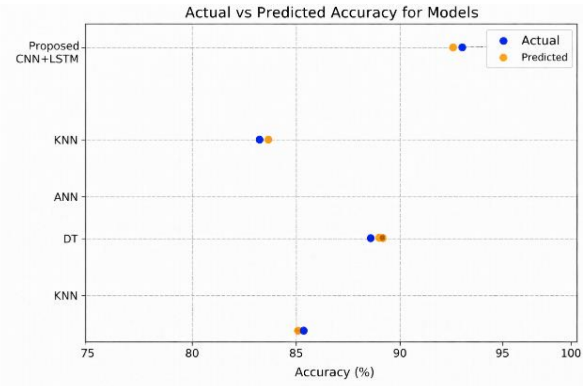

Figure 4. Comparison of actual and predicted accuracy

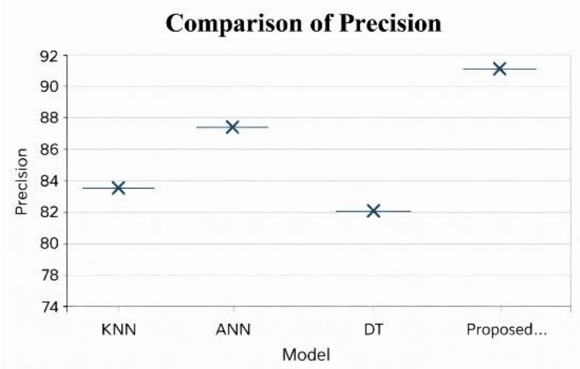

Figure 5. Comparison of precision

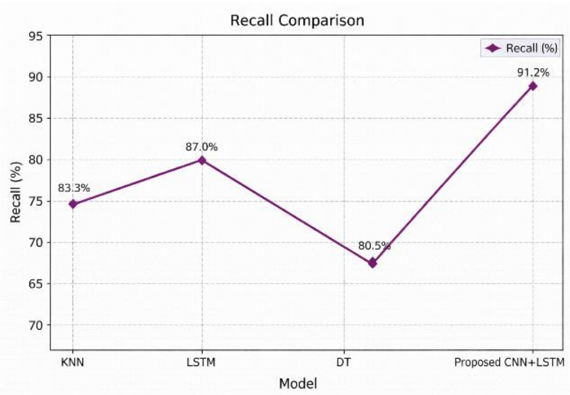

Figure 6. Comparison of recall

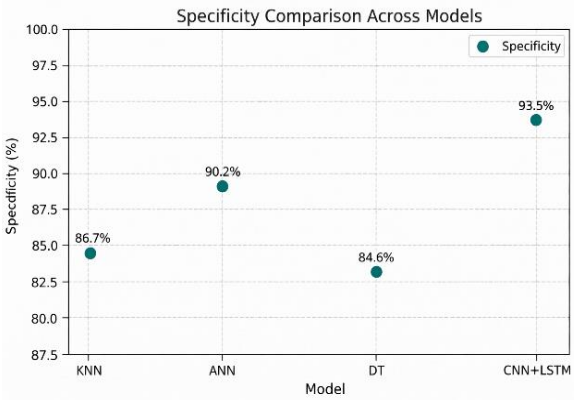

Figure 7. Comparison of specificity scores

A comparative examination of the F1-score of four hybrid deep-learning models that identified heart disease is shown graphically in Figure 8. With an F1-score of 91.0%, the suggested CNN+LSTM model produces the greatest results and strikes the ideal mix between recall and accuracy. This shows that it performed effectively in the identification of positive and negative cases of heart diseases. The ANN model continues and generates a percent score of 86.7, and the performance is somewhat effective yet holding back. In comparison, traditional models such as KNN and DT have lower values of F1-scores, which are 82.8% and 80.7%, demonstrating better F1-score performance in deep learning compared with traditional methods. The ROC curve for the suggested CNN and LSTM hybrid model is shown in Figure 9.

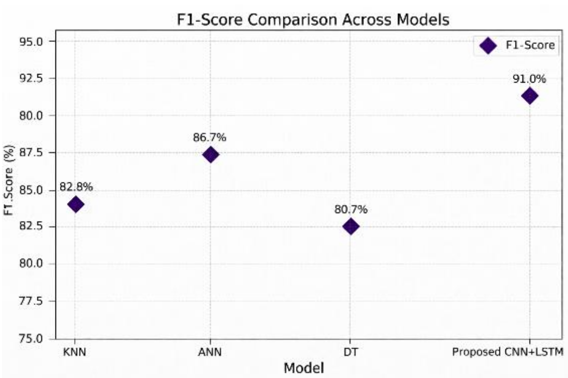

Figure 8. Comparison F1-score

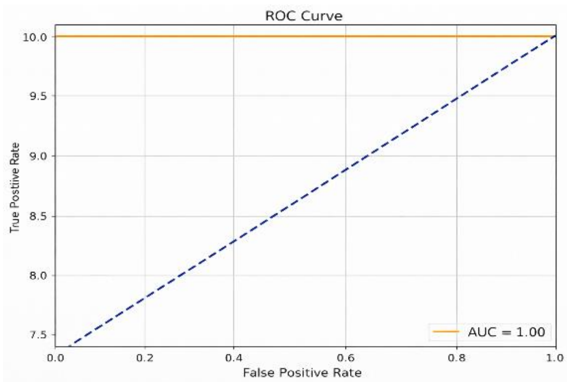

Figure 9. ROC curve for proposed CNN and LSTM

5. CONCLUSION

The combination of practical image processing and hybrid deep learning methods is recommended as a superior technique for predicting heart diseases. Incorporating the butterworth adaptive 2D wavelet filtering method to provide noise-free image preprocessing. The Gabor wavelet-based feature extraction method is used to extract pattern recognition features in the system, thus ensuring the quality of the input data. Based on CNN and LSTM models, it is possible to more easily classify heart conditions due to the non-linearity and complex data variations. The method used works better than conventional models to control such problems as the imbalance of data and dependence on the irrelevance of the features. Such an extensive application of evaluation metrics such as accuracy 92.4%, precision 91.2%, recall 93.5, and F1-score 91.0%, as well as indicates the credibility and generalizing performance of the model. The end product of this hybrid is an optimistic prospect of early detection, better clinical judgment, better treatment planning, and, therefore, the prevention of the mortalities associated with heart diseases and overall better patient treatment by developing better technology.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Chairmadurai Palanisamy	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Kavitha Pachamuthu		✓				✓		✓	✓	✓	✓	✓		
Arun Kumar Ramamoorthy	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] M. S. A. Reshan, S. Amin, M. A. Zeb, A. Sulaiman, H. Alshahrani, and A. Shaikh, "A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks," in *IEEE Access*, vol. 11, pp. 121574-121591, 2023, doi: 10.1109/ACCESS.2023.3328909.

[2] Z. Noroozi, A. Orooji, and L. Erfannia, “Analyzing the impact of feature selection methods on machine learning algorithms for heart disease prediction,” *Scientific Reports*, vol. 13, no. 1, p. 22588, 2023, doi: 10.1038/s41598-023-49962-w.

[3] R. Hoque, M. Billah, A. Debnath, S. S. Hossain, and N. B. Sharif, “Heart disease prediction using SVM,” *International Journal of Science and Research Archive*, vol. 11, no. 2, pp. 412-420, 2024, doi: 10.30574/ijsra.2024.11.2.0435.

[4] G. Manikandan, B. Pragadeesh, V. Manojkumar, A. L. Karthikeyan, R. Manikandan, and A. H. Gandomi, “Classification models combined with Boruta feature selection for heart disease prediction,” *Informatics in Medicine Unlocked*, vol. 44, p. 101442, 2024, doi: 10.1016/j.imu.2023.101442.

[5] H. Sun and J. Pan, “Heart disease prediction using machine learning algorithms with self-measurable physical condition indicators,” *Journal of data analysis and information processing*, vol. 11, no. 1, pp. 1-10, 2023.

[6] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, “Effective heart disease prediction using machine learning techniques,” *Algorithms*, vol. 16, no. 2, p. 88, 2023, doi: 10.3390/a16020088.

[7] N. Chandrasekhar and S. Peddakrishna, "Enhancing heart disease prediction accuracy through machine learning techniques and optimization," *Processes*, vol. 11, no. 4, p. 1210, 2023, doi: 10.3390/pr11041210.

[8] P. Rani, R. Kumar, N. M. O., S. Ahmed, and A. Jain, "A decision support system for heart disease prediction based upon machine learning," *Journal of Reliable Intelligent Environments*, vol. 7, no. 3, pp. 263–275, 2021, doi: 10.1007/s40860-021-00133-6.

[9] M. S. Pathan, A. Nag, M. M. Pathan, and S. Dev, "Analyzing the impact of feature selection on the accuracy of heart disease prediction," *Healthcare Analytics*, vol. 2, p. 100060, 2022, doi: 10.1016/j.health.2022.100060.

[10] M. Ahmed and I. Husein, "Heart disease prediction using hybrid machine learning: A brief review," *Journal of Robotics and Control (JRC)*, vol. 5, no. 3, pp. 884-892, 2024, doi: 10.18196/jrc.v5i3.21606.

[11] M. Diwakar, A. Tripathi, K. Joshi, M. Memoria, and A. Singh, "Latest trends on heart disease prediction using machine learning and image fusion," *Materials today: proceedings*, vol. 37, pp. 3213-3218, 2021, doi: 10.1016/j.matpr.2020.09.078.

[12] B. U. Rindhe, N. Ahire, R. Patil, S. Gagare, and M. Darade, "Heart disease prediction using machine learning," *Heart Disease*, vol. 5, no. 1, 2021, doi: 10.48175/IJARSCT-1131.

[13] S. P. Patro, G. S. Nayak, and N. Padhy, "Heart disease prediction by using novel optimization algorithm: A supervised learning prospective," *Informatics in Medicine Unlocked*, vol. 26, p. 100696, 2021, doi: 10.1016/j.imu.2021.100696.

[14] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, 2021, doi: 10.1155/2021/8387680.

[15] M. I. Mehmood, Z. Mehmood, A. Irtaza, M. Nawaz, T. Nazir, and M. Masood, "Prediction of heart disease using deep convolutional neural networks," *Arabian Journal for Science and Engineering*, vol. 46, no. 4, pp. 3409–3422, 2021, doi: 10.1007/s13369-020-05105-1.

[16] M. G. El-Shafiey, A. Hagag, E.-S. A. El-Dahshan, and M. A. Ismail, "A hybrid GA and PSO optimized approach for heart-disease prediction based on random forest," *Multimedia Tools and Applications*, vol. 81, no. 13, pp. 18155–18179, 2022, doi: 10.1007/s11042-022-12425-x.

[17] Dissanayake and M. G. M. Johar, "Comparative study on heart disease prediction using feature selection techniques on classification algorithms," *Applied Computational Intelligence and Soft Computing*, vol. 2021, no. 1, pp. 1-17, 2021, doi: 10.1155/2021/5581806.

[18] A. A. Nancy, P. M. D. R. Vincent, K. Srinivasan, and D. G. Reina, "IoT-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning," *Electronics*, vol. 11, no. 15, p. 2292, 2022, doi: 10.3390/electronics11152292.

[19] S. E. A. Ashri, M. M. El-Gayar and E. M. El-Daydamony, "HDPF: Heart Disease Prediction Framework Based on Hybrid Classifiers and Genetic Algorithm," in *IEEE Access*, vol. 9, pp. 146797-146809, 2021, doi: 10.1109/ACCESS.2021.3122789.

[20] T. R. Ramesh, M. Poongodi, S. Simaiya, A. Kaur, and M. Hamdi, "Predictive analysis of heart diseases with machine learning approaches," *Malaysian Journal of Computer Science*, pp. 132-148, 2022, doi: 10.22452/mjcs.sp2022no1.10.

[21] X.-Y. Gao, A. A. Ali, H. S. Hassan, and E. M. Anwar, "Improving the accuracy for analyzing heart diseases prediction based on the ensemble method," *Complexity*, vol. 2021, no. 1, p. 6663455, 2021, doi: 10.1155/2021/6663455.

[22] M. M. N. Asif, F. Faisal, R. R. Dip, M. H. Udoy, M. F. Shikder, and R. Ahsan, "Performance evaluation and comparative analysis of different machine learning algorithms in predicting cardiovascular disease," *Engineering Letters*, vol. 29, no. 2, 2021.

[23] S. Akella and S. Akella, "Machine learning algorithms for predicting coronary artery disease: Efforts toward an open-source solution," *Future Science OA*, vol. 7, no. 6, 2021, doi: 10.2144/fsoa-2020-0206.

[24] S. Nandy, V. Balasubramanian, X. Li, and M. Zakaria, "An intelligent heart disease prediction system based on swarm-artificial neural network," *Neural Computing and Applications*, vol. 35, no. 20, pp. 14723–14737, 2023, doi: 10.1007/s00521-021-06124-1.

[25] A. Yazdani, Y. K. Chiam, A. W. Malik, and W. A. W. Ahmad, "A novel approach for heart disease prediction using strength scores with significant predictors," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, p. 194, 2021, doi: 10.1186/s12911-021-01527-5.

[26] M. Ozcan and S. Peker, "A classification and regression tree algorithm for heart disease modeling and prediction," *Healthcare Analytics*, vol. 3, p.100130, 2023, doi: 10.1016/j.health.2022.100130.

[27] A. Ishaq *et al*., "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques," in *IEEE Access*, vol. 9, pp. 39707-39716, 2021, doi: 10.1109/ACCESS.2021.3064084.

[28] S. K. Budholiya, S. K. Shrivastava, and V. Sharma, "An optimized XGBoost based diagnostic system for effective prediction of heart disease," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4514-4523, 2022, doi: 10.1016/j.jksuci.2020.10.013.

[29] V. Chang, V. R. Bhavani, A. Q. Xu, and M. A. Hossain, "An artificial intelligence model for heart disease detection using machine learning algorithms," *Healthcare Analytics*, vol. 2, 2022, doi: 10.1016/j.health.2022.100016.

[30] P. E. Rubini, C. A. Subasini, A. V. Katharine, V. Kumaresan, S. G. Kumar, and T. M. Nithya, "A cardiovascular disease prediction using machine learning algorithms," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 2, pp. 904-912, 2021.

BIOGRAPHIES OF AUTHORS

Chairmadurai Palanisamy is a Ph.D. research scholar in Vels Institute of Science, Technology, and Advanced Studies, Chennai, Tamil nadu, India. He has hands-on experience in developing and evaluating predictive models using advanced algorithms and is actively engaged in academic research and experimentation in these domains. His work aims to contribute to the advancement of intelligent computing applications across real-world problem settings. He can be contacted at email: chairma17@gmail.com.

Dr. Kavitha Pachamuthu is working as Assistant Professor in the Department of Computer Applications, Vels Institute of Science, Technology, and Advanced Studies, Chennai India. She has extensive experience in the teaching field and has been actively involved in academic instruction and curriculum development. Her research interests include emerging areas of computer science, and she has published several research papers in reputed national and international conferences and journals. She can be contacted at email: pkavikamal@gmail.com.

Dr. Arun Kumar Ramamoorthy is a distinguished Lecturer in Digital Forensics and Cyber Security at the University of South Wales, Treforest, United Kingdom. He holds a B.Eng. in Computer Science and Engineering from Karunya University, Coimbatore, an M.Tech. in Information Technology, and a Ph.D. from Anna University, Chennai. With 16 years of academic experience, he has made significant contributions to the field of Information Technology, authoring over 25 research articles in various domains. His areas of expertise include information security, digital forensics, mobile application security, artificial intelligence, machine learning, deep learning, and network security. He is an inventor with two Indian patents in neural networks and machine learning, as well as an Australian patent in cyber security. He is actively involved in professional organizations, being a member of the Computer Society of India (CSI) and the Indian Society for Technical Education (ISTE). He can be contacted at email: arun.kumar@southwales.ac.uk.