SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V12I6P112

Volume 12 Issue 6, 146-157, June 2025

Original Article

Permutation of Deep Learning Approaches to Perceive Heart Problems Using Audio Signals

K. Vetriselvi[1], G. Karthikeyan[2]

[1,2] Department of Computer Science, Periyar Government Arts College, Tamil Nadu, India.

[1]Corresponding Author : vetriselvik2009@gmail.com

Received: 08 April 2025 Revised: 10 May 2025 Accepted: 09 June 2025 Published: 30 June 2025

Abstract - *Heart disease, which is widely known as the "Silent Killer," kills many people worldwide, and it is now affecting younger people in India. Prevention of fatal outcomes starts with early detection, but traditional machine learning models widely used are inadequate in providing accurate results. So, deep learning techniques are implemented for better solutions. Deep learning models are applied in this research to forecast heart disease by examining heart sound signals. This research uses a machine learning approach with Modified LSTM and a transfer learning technique by feeding spectrograms. Three important steps during detection are normalization, feature extraction, and classification, which are used to analyse heart sounds precisely. Three pre-trained models, Enhanced Inception V3, Elevated CNN and Facelifted ResNet 50, are assessed against a sequence classifier. Facelifted ResNet 50 can detect heart disease by abnormal heartbeat sounds such as artifact, extras, extrasystole and murmur with 98% accuracy. The unique aspect is using transfer learning classifiers based on spectrograms and compared with sequence model Modified LSTM to diagnose patients. This method promotes better detection of cardiovascular diseases and helps introduce intelligent medical technologies.*

Keywords - *Audio signals, Cardiovascular syndrome, Convolutional Neural Network, Inception V3, Long Short-Term Memory, Residual Network 50 model, Silent killer.*

1. Introduction

Globally, heart disease is a major health issue; according to the Indian Council of Medical Research, almost 60% of Indians suffer from disorders including heart attacks and cardiac arrest. Affecting the heart and blood vessels, these disorders usually arise from blocked arteries that restrict blood flow to the heart and brain, so causing major problems, including strokes. The basic problems are irregular heartbeats, malfunctioning heart valves, and blood vessel constriction. While strokes usually follow bleeding or clots in the blood vessels, fat deposits within the blood vessels mostly cause heart attacks and cardiac arrests. Unhealthy lifestyles, including high blood pressure, diabetes, obesity, and bad habits like smoking and too much alcohol intake, count as contributing causes of heart disease.

The growing prevalence of processed and unhealthy foods has elevated the risk of heart failure, even among younger individuals [1]. Because of these lifestyle choices, heart problems are becoming more common in India. Early diagnosis and management are critical in combating heart disease. Conventional techniques such as stethoscope tests detect sounds suggestive of inadequate blood circulation, helping to identify abnormal heart rhythms. These techniques, meanwhile, have difficulties, including outside noise and the need for professional knowledge. As a result, a non-invasive and more precise approach is required. A Novel machine learning-based method aims to develop early diagnosis, enhance treatment outcomes and reduce the reliance on invasive procedures. This method demonstrates the potential of AI-driven healthcare solutions in pediatric cardiology [2]. A combination of Recurrent Neural Networks (RNN) and Bidirectional Long Short Term Memory (Bi-LSTM) networks effectively capture the temporal dynamic heart sounds. A Generative Adversarial Network (GAN) is combined to optimize feature selection and recognise cardiovascular disease patterns [3]. Khan and Khan use residual learning to merge heart sound data into more accurately detecting abnormal cardiac rhythms. To extract better features, integrate MobileNet and DenseNet201, using both approaches' capabilities. The method delivers an exceptional accuracy level of 95.67% on the PhysioNet-2016 Spectrogram dataset.

The dataset has also proven to be strong because validation of the BreakHis results shows an accuracy of 96.55 percent. The study shows how deep learning can be used for cardiac rhythm analysis [4]. The system provides a novel approach for classifying heart sound signals by converting them into image representations. The study uses deep learning models to explore different image transformation techniques to improve classification accuracy. This method improves automated heart sound analysis, aiding in the early diagnosis of cardiovascular diseases. This method gives promising results, highlighting the potential image-based classification in biomedical signal processing [5]. An ensemble-based transfer learning model for predicting imbalanced heart sound signals using spectrogram images. Different frequency representations, such as Mel spectrograms and Short-Time Fourier transform (STFT), are used for better feature extraction. Pre-trained deep learning models such as VGG (Visual Geometry Group) are applied to generate synthetic spectrogram images, addressing class imbalance. This method effectively handles the issue of class imbalance, reducing bias towards majority class samples.

The ensemble-based transfer learning model significantly improves the accuracy of heart sound classification compared to individual deep learning models [6]. A deep learning approach using Convolutional Neural Network (CNN) to predict abnormalities in heartbeat sounds. Before classification, the study takes heart sound data and creates spectrogram images. In diagnostics, CNN allows computers to find better and label normal and abnormal heartbeats. This reveals that CNN-based approaches may be powerful tools for early identification of heart diseases [7]. The new approach to identifying heart sounds using radar signals and modern signal processing methods.

The research examines whether radar-based technology can pick up and analyze heart sounds by staying outside of physical contact. A non-invasive cardiovascular procedure uses machine learning and signal processing to improve how well heart sounds are monitored. The outcome shows high precision, pointing to the potential of radar for both clinical and at-home healthcare [8]. Reviewing audio signals to spot heart disease is much more advanced than electrocardiograms or imaging methods. With this approach, detection happens early with less risk, at a lower cost and is easily available, supporting prompt intervention and good patient outcomes.

So, deep learning developments are creating promising possibilities. Analyzing Mel spectra data using neural networks helps to classify the various heart sounds heard on the recording. MFCC (Mel Frequency Cepstral Coefficient) turns heart sounds from recordings into numbers understood by programs for classification. This recorded audio data is then used to make images on which the Classifier relies. The main goal of this research is to use deep learning on heart sounds captured in recordings to spot heart problems as early as possible. The recorded sounds from the heart are converted into useful features and used to distinguish between normal, artefact, extras, extrasystole and murmur sounds. Heart disease patients are identified by abnormal heart sounds such as artifacts, murmurs, extras and extrasystole sounds and the normal patient is identified by normal heartbeat sounds. As shown in Figure 1, the images in this graph illustrate how heart sounds from different types of subjects help with effective classification and diagnosis.

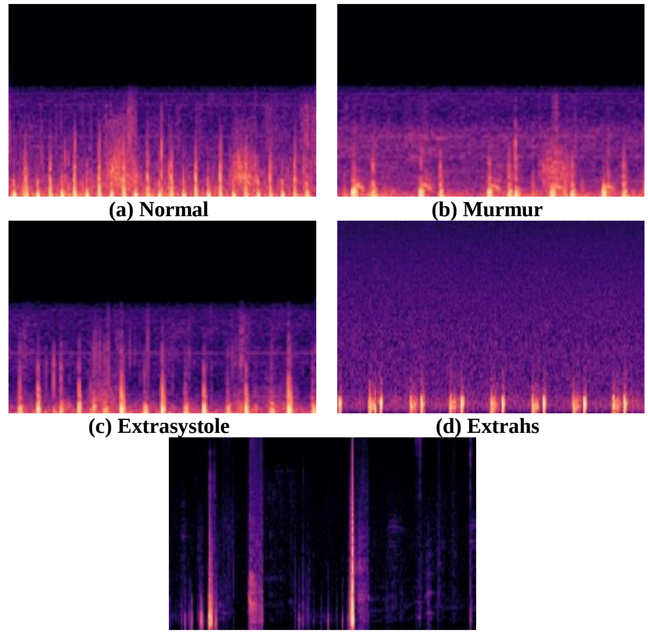
(a) Normal **(b) Murmur** **(c) Extrasystole** **(d) Extrahs** **(e) Artifact**

Fig. 1 Visual representation depicts sample images of heart sounds

2. Related Work

Utilizing wearable devices, deep learning, continuous monitoring, and advanced algorithms is important for discovering and predicting heart disease as soon as possible, as discussed in [9]. Early heart disease prediction accuracy using feature engineering and machine learning can be improved by picking the most useful features [10]. This artificial intelligence method proves that it is possible to find heart disease early with tools that do not involve invasive procedures [11]. Studying different methods for classifying and forecasting heart disease in machine learning and deep learning highlights the advantages and disadvantages [12].

The authors examine the use of deep learning in heart sound analysis, from technical solutions to clinical applications. The experts discuss how heart sound analysis is challenging and how deep learning models help with better results. The paper combines major heart sound collections and describes the most recent research. It encourages doctors to use deep learning methods to improve cardiac auscultation. The research points out ways to enhance the analysis of heart sounds in the future [13]. A hybrid lightweight version of deep learning for cardiac valvular disease classification demonstrates the usefulness of multiple deep learning approaches [14]. Using recent techniques on electrocardiogram data summarizes the latest achievements and future areas of development [15]. Using deep learning to classify heart sounds shows that diagnosis can be made more accurately [16]. Several models are compared by examining machine learning methods' efficiency in predicting heart disease [17]. Improving model performance for heart disease prediction through fusing features in a mixed cascaded kernel extreme learning machine is demonstrated in [18]. Zhao et al. present a Convolution-Transformer neural network called the HCTNN to detect CHD using PCG data. CNN and Vision Transformer modules are combined to capture both local and global information. A method using neural networks achieves better accuracy, reaching 94.24% over CNNs used alone.

The study demonstrates that Transformer-based approaches can effectively detect heart disease. With this research, doctors can use deep learning to improve non-invasive heart diagnostic tools [19]. Machine learning for better heart attack prediction was based on a study conducted at Jordan University Hospital [20]. The authors suggest an original way to anticipate heart risks by examining retinal fundus photographs the deep learning approach to uncover vascular problems associated with cardiovascular diseases. The research suggests that viewing the retina can be a safe early indicator for cancer risk assessment.

The model improves its ability to predict cardiovascular diseases when using CNNs. The research helps advance precision medicine for diagnostics in cardiovascular care [21]. The efficiency of using trees for decision-making is seen in a heart disease regression tree and classification technique [22]. To compare various supervised classification models, these model abilities to detect heart disease are analyzed [23]. The use of a model that makes heart disease classification easier to understand using 12-lead ECG data underlines how interpretability is helpful for medical models. The Performance of SVM and KNN algorithms is studied to see how a hybrid machine-learning technique helps predict heart disease [25]. The usefulness of reliable prognosis tools is made clear by studies of the anatomic-physiologic classification for adults with congenital heart disease [26].

The ability to use optical ECGs for non-invasive screening is proven by a study that predicts heart disease from ECG readings processed by hybrid learning [14]. Combining a CNN and an LSTM in a hybrid model successfully predicts heart disease [27]. According to Singh et al., class imbalance in heart sound signal classification can be resolved by borrowing CNN technology from other datasets using an ensemble approach. Using spectrogram images and STFT to obtain features improves the method's accuracy. AlexNet, SqueezeNet and VGG19 as an ensemble attained an accuracy of 99.20%. The research found deep learning to work effectively with heart sound data that is not balanced. The research leads to better cardiac diagnoses provided by AI [6].

3. Materials and Methods

3.1. Problem Declaration

The current problem in the medical field is to diagnose heart issues without any symptoms or in the early stages of the ailment. For this purpose, deep learning develops the model using audio signals to classify the heartbeat sound into five classes: normal, artifact, extras, extrasystole and murmur. The audio signals are converted into images for processing. Images should be pre-processed for further enhancement, and a feature extraction method is used to retrieve arithmetic features.

These features are used in classifying normal and abnormal individuals. Deep learning uses two types of classifications: sequence and transfer learning methods for accuracy. The sequence method employs the LSTM scheme, while Transfer learning consists of Elevated CNN with Heart Rate Image Classification model, Enhanced Inception V3 for the Heart Rate Classification model and Facelifted ResNet 50 for the Heart Rate Prediction model for classification. Finally, the output is determined for all four models and accuracy is calculated and compared. The architecture depicting the above-mentioned techniques is described and given in Figure 2.

3.2. Sources of Dataset

For the proposed methodology of heart sound classification, datasets are retrieved from prominently available databases. The data is categorized into groups based on heart sound types and includes labelled and unlabelled audio files. The audio recordings vary from 1 to 30 seconds and are available in WAV and AIF formats. A segmentation file provides the locations of key heart sounds (S1 and S2) in a subset of the recordings. The dataset consists of two primary sources: Dataset A, collected from the general public using the stethoscope Pro iPhone app and Dataset B, which was collected during clinical trials in hospitals using the DigiScope digital stethoscope.

The dataset is divided into training and test sets for both Dataset A and Dataset B. Each set includes labelled recordings for specific heart sound categories. Nearly 585 audio patterns, both normal and abnormal, are processed from the Peterjbentley Heart Challenge dataset [28]. The sound patterns are recorded using a digital stethoscope or mobile devices as audio signals and stored in WAV format for further processing.

This research employs data augmentation techniques to increase the complexity of data. Different changes can be applied to audio data through augmentation techniques such as changing tempo or pitch, adjusting volume and time-stretching procedures. Augmentation is used to get 1000 heartbeat audio signals. The main idea is to transform 1D audio signals into 2D images for feature extraction and classification. The audio signals are first pre-processed using filtering and normalizing techniques suitable for further handling. The dataset contains 800 images for training, and 200 images are used to test the proposed models.

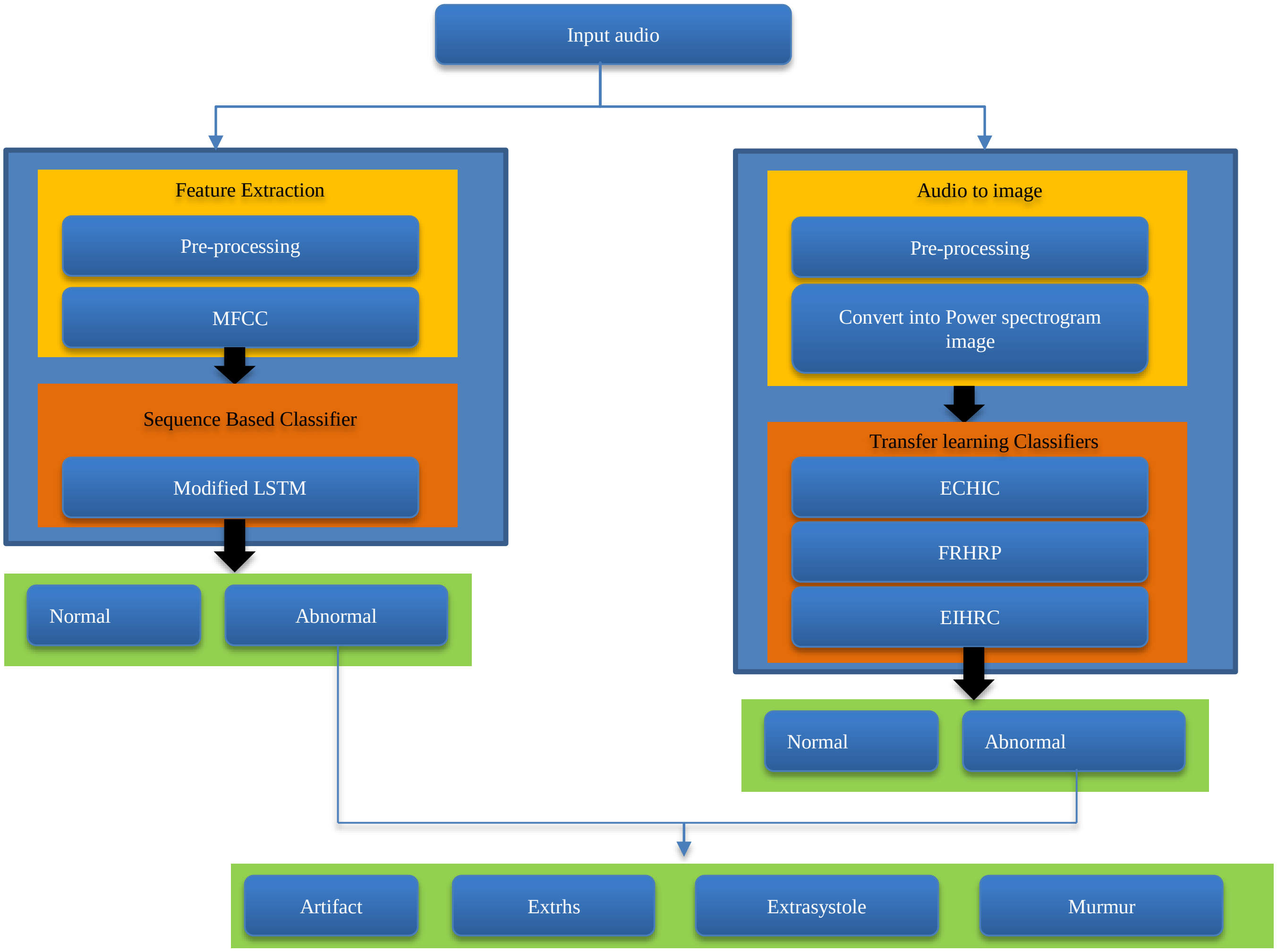

Fig. 2 Proposed architecture of heart disease prediction model

3.3. Pre-Processing Step

There are two stages in the pre-processing method. First, the audio signals are normalized and filtered for processing. Next, the audio waves are converted into power spectrogram images and resized to a standard scale for the next investigation.

3.3.1. Normalization

One crucial step is removing systematic noise in the audio signal and receiving a large audio volume. Normalization changes the amplitude so the signal's degree reaches the standard level.

3.3.2. Audio Conversion

The next step is converting the audio signals into images. A power spectrogram is the visual conversion of sound frequencies varying with time or other factors. The frequencies are the features that shape the form of sound recorded. The audio files needed are converted into images using log transformation with Fourier series, and only mono files are used. The frequencies are decomposed, and the amplitude of each frequency is displayed as an image.

3.3.3. Resizing of Images

The images of varying sizes may not be suitable for feature extraction. So, the images are resized to standard size 200 x 200 for easy application. Resizing changes the dimensions of the image without content loss. The interpolation method is used for resizing, and the type is an inter-area method for shrinking the images. Pixel area relation is used to scale the images to the standard size. One sample image is shown in Figure 3 for rescaling to regular size.

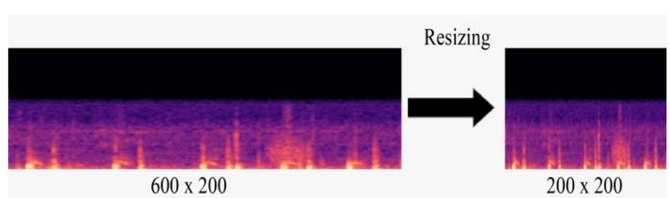

Fig. 3 Resizing of images

3.4. Feature Extraction

Feature extraction is the method of mining and processing the numerical features renewed from the raw data, preserving the information without any loss. This step is needed to reduce the redundant data, thereby increasing the simplification speed of the model that will be constructed for classification.

3.4.1. MFCC Features

The method used for extracting features from audio signals recorded is the MFCC (Mel Frequency Cepstral Coefficients), representing the power spectrum of the audio signals based on the Mel scale of frequency. Following are the steps used. The essential steps are,

Pre-Emphasis

Pre-emphasis is the filtering technique to suppress noise so that high-frequency sound waves can be recorded. All other disturbances, such as emotional background blasts, can be eliminated in this step.

Frame Blocking and Windowing

To receive a better resolution of spectral sounds, the window is used along with short segments to examine the frequency. The temporal features of the audio are recorded using a 20 MS window. This technique is applied on each sound frame to spill the signal towards the borders. This step is used to level the signal for an accurate DFT function while deriving a spectral representation of the signal.

DFT and Mel Spectrum

Discrete Fourier Transform (DFT) calculates the frequency spectrum of any signal tracked. Using DFT, every sound frame is changed into a magnitude spectrum [10]. Mel spectrum is derived using a Fourier signal with triangular filters called Mel filter, a measure of physical frequency. It is represented as

$$Mel(f) = 2595 \log_{10}(1 + f/700) \quad (1)$$

Where f represents physical frequency in hertz. The filter banks with Mel frequency warping calculate the Mel spectrum by multiplying filters with the magnitude spectrum.

Discrete Cosine Transform

The Mel frequency coefficients, which are converted, create cepstral features by applying the DCT function. Log scale is used to indicate Mel spectrum and audio content is provided by a few first MFCC coefficients discarding higher order DCT components.

Dynamic MFCC features

Mel coefficients contain only vital information about each sound frame, which is called static features. So, the first and second-order cepstral coefficients are calculated to provide additional evidence of the sound waves.

3.5. Classification Models

After feature extraction, the next step in processing audio signals is the classification methods for finding normal and abnormal persons from the heart sound. From the abnormal sound, four categories are finalized. Two types of classification methods are used. One is a sequence-based classifier, and the other is a transfer learning classifier. Four models are implemented for classification: LSTM, which belongs to the sequence method, and CNN, Inception and Residual network mode, which belongs to the transfer learning method.

3.5.1. Sequence-Based Classifier

In classification, where early diagnosis is important, the sequence method is applied, especially in health care observation. A sequence classifier predicts the result with an arrangement of input data collected over time or space.

The result is to categorize the sequence into various classes. One of the sequence classifiers is the LSTM method, which belongs to the recurrent neural network model used to classify data oriented with time.

Modified LSTM Method

The technique popularly known as the Long Short-Term Memory method is an improved version of deep RNN suitable for solving the vanishing gradient problem. The basic building blocks of LSTM are called gates, and there are three types of gates that have different functions. The first one is the Forget gate, which selects the relevant data from the previous layer to be remembered.

The second gate is the input gate, which is used to acquire new information from the input signal, and the last gate is the output gate, which trains the corrected information and passes it to the next layer. The architecture designed for the proposed method is given in Figure 4. For classification purposes, features are extracted from the MFCC method described above, and 40 features are selected as input for the Classifier. The features prescribe the slight variations in the heartbeat sound of normal and abnormal persons, sorting them into five classes, including normal identity. LSTM classifier decides the result based on the previous sequence of audio signals received over time.

The data are collected until the function return sequence is true, and the units called as cells represent the layers holding the hidden state vector appended with the input layer. When the return sequence becomes false, automatically, the model learns the class by reducing the features using the dense layer. Thereby, classification is done based on the information retrieved from the output gate, and the softmax classifier is used to finalize the classification into five required classes.

3.5.2. Transfer Learning Classifier

The modern technique used for classification is transfer learning-based models, whose accuracy is comparatively precise. For classification purposes, three models are discussed using audio signals to group the individuals into normal and abnormal categories.

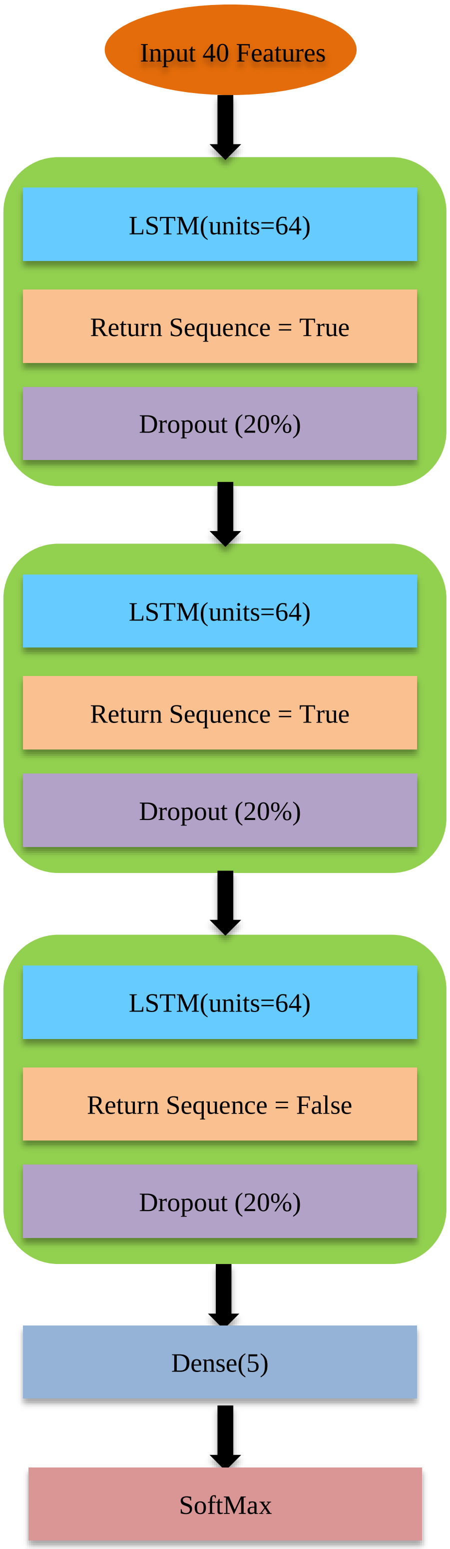

Fig. 4 Modified LSTM architecture

Elevated CNN with Heart Rate Image Classification (ECHIC)

The new architecture built for the classification of audio signals is discussed in detail below. The architecture used for this model is given in Figure 5.

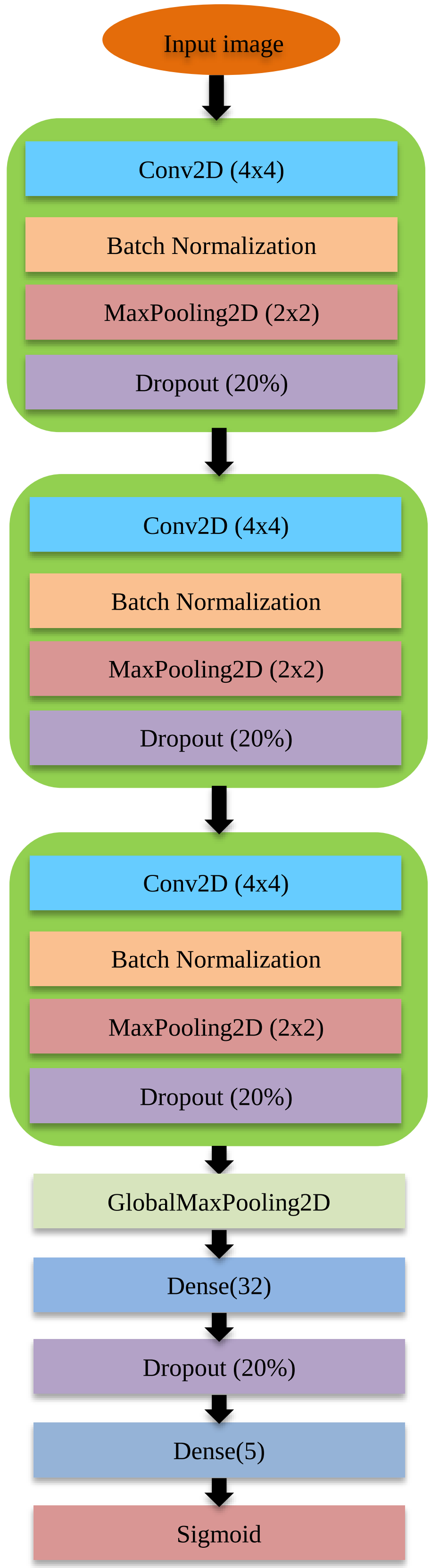

Fig. 5 Architecture of elevated CNN with heart rate image classification model

The steps are,

- First, the input image is fed to the model
- The image is passed to the first convolutional layer, where 4 x 4 filters are used to convolve the image, producing feature maps of the same size as the filter size.
- Next, the Max pooling layer is applied with stride 2 x2 to reduce the parameters, decreasing the image size while the depth remains constant.
- Batch normalization stabilises the model and trains faster by rescaling the image.
- Dropout is used to avoid the overfitting problem. The proposed model averages 10% of dropouts to average the network.
- The above layers are repeated twice in the architecture described above, followed by the Global Max pooling layer. This layer derives the maximum features of the input image, and the values are flattened to convert the output into different classes required.
- Dense layers act as final layers to predict the classes. Dropout is used to reduce the attributes that prescribe the classes.
- Sigmoid is used to classify images, thereby predicting the class. The image belongs as one and other classes as 0.

The audio signals are classified using the proposed Elevated CNN with Heart Rate Image Classification model using the automatic features extracted from the model. When the heart sound is abnormal, it is categorized as a murmur, artifact, extras, or extrasystole.

Enhanced Inception V3 for Heart Rate Classification (EIHRC)

Inception V3 is the transfer learning method used by modifying the standard method for classifying audio signals. The basic model contains three blocks

- Block A - Factorization of smaller convolutions.
- Block B - Factorization of 7 x 7 convolutions into factors (1 and 7).
- Block C - Factorization of asymmetric convolutions.
- Final factorization replaces 7 x 7 into a series of 3 x 3 convolutions.

The proposed architecture is depicted in Figure 6, and this. The proposed model modifies the standard model by updating the last layer for classification by freezing the bottom layers in the given model to produce accurate results. An input image of standard size is used in the given model, and convolution and pooling layers are performed.

Module A is performed 5 times, module B is performed 4 times, and module C is performed 2 times. An Auxiliary Classifier is used for classification, and the final output is 8 x 8 x 2048 with 2048 features used for processing.

The added layers contain dense layers with a sigmoid Classifier for categorizing normal and abnormal classes. The abnormal classes identified by the heart sound include artifact, murmur, extras and extrasystole.

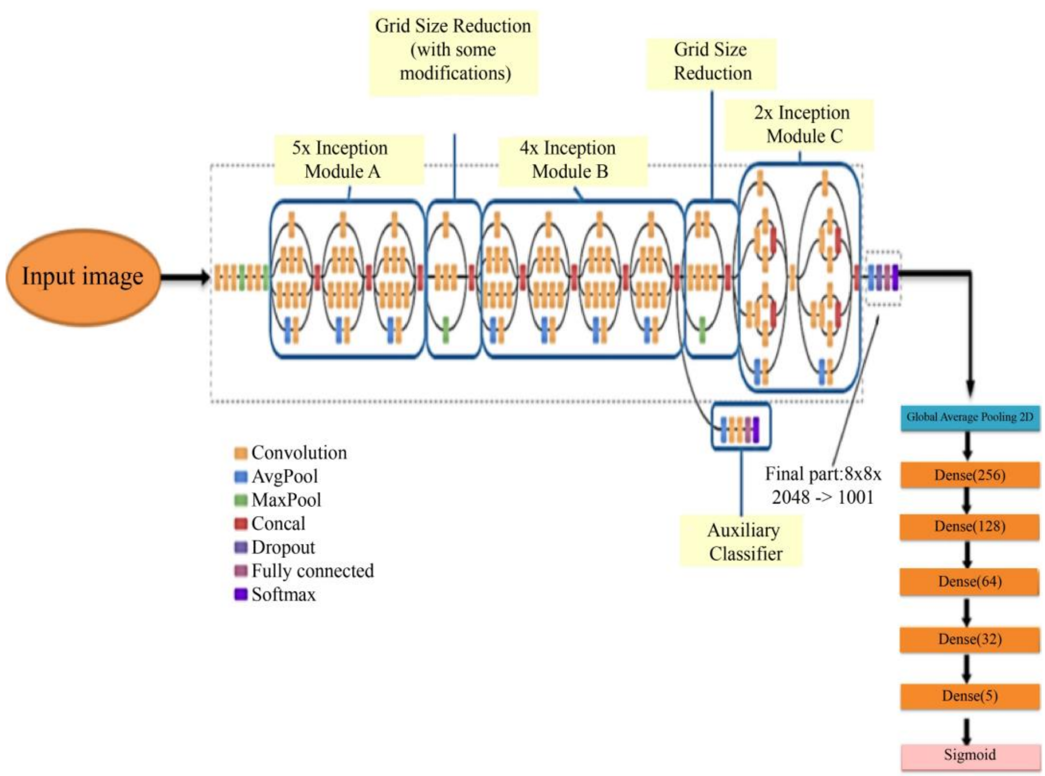

Fig. 6 Enhanced Inception V3 for Heart Rate Classification (EIHRC)

Facelifted ResNet 50 for Heart Rate Prediction (FRHRP)

- The standard convolution layer with stride 3 and filter 7x7 is followed by the max pooling layer with 3x3 filter and stride 2 -Batch normalization and activation function is used for the proper classification process.
- The proposed architecture contains four convolution and four identity block layers performing one after another.
- Next, the global max pooling layer is performed to flatten the values fed to the dense layers.
- The proposed layers appended in the given model contain dense layers followed by a sigmoid Classifier used to classify five classes.
- The architecture contains two additional blocks, namely the convolutional block, which contains a shortcut path to add another convolutional layer to make the input image size equal to the output image.
- The identity block is another module with 3 convolutional layers, and filter sizes 1x1 and 3x3 are used.
- The residual network 50 model architecture is updated for the classification of the given dataset and is depicted in Figure 7.

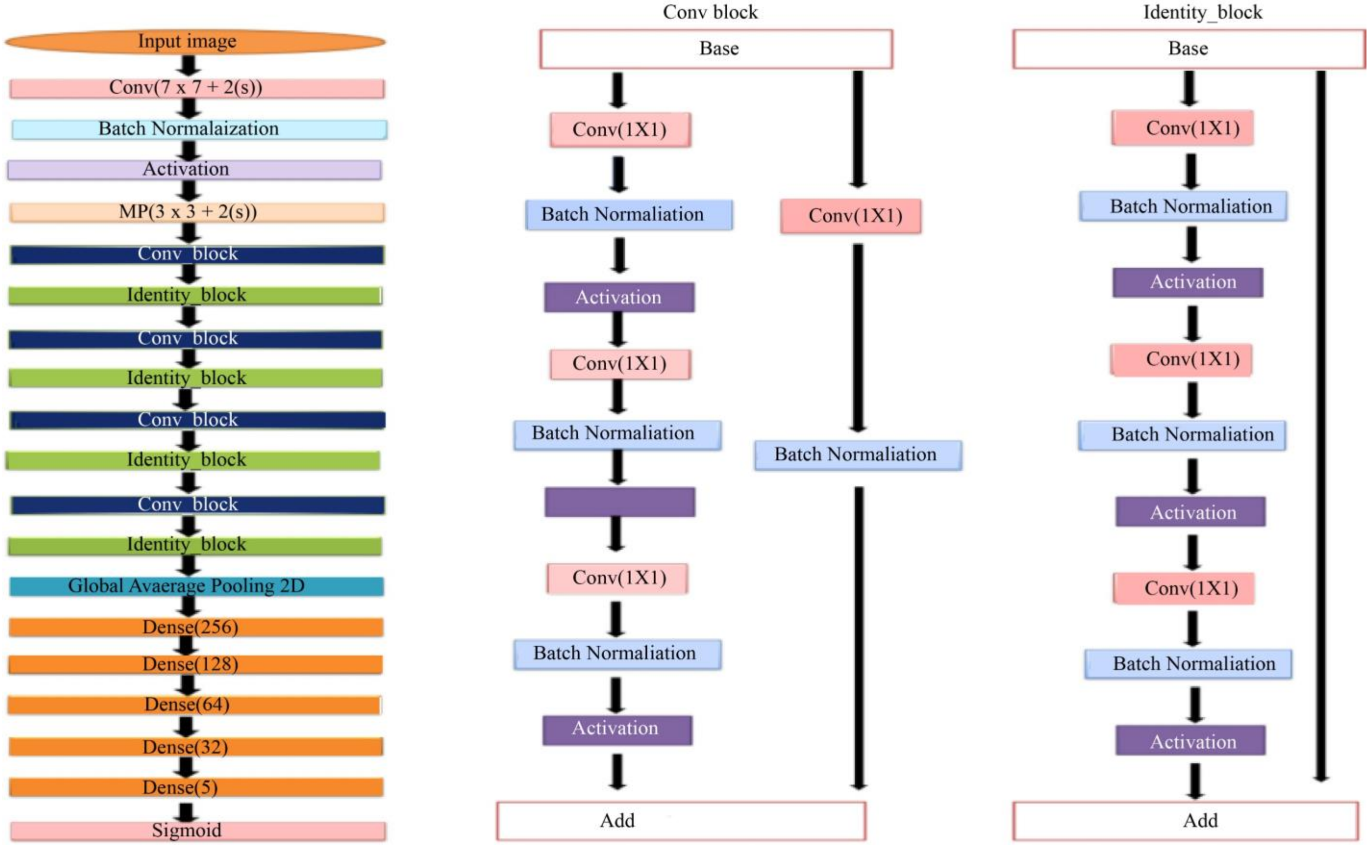

Fig. 7 Facelifted ResNet 50 for Heart Rate Prediction (FRHRP)

4. Results and Discussion

All four classification models are processed using the given dataset to segment normal candidates from abnormal candidates using audio signals of heartbeat sound. LSTM provides a good accuracy of 92%, and the performance measures of this model are given in Table 1.

Table 1. Performance measures of the Modified LSTM model

Categories of HeartBeat	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Normal	99.10	87.02	93.50	92
Murmur	84.05	95.12	89.54	
Extrasystole	94.12	89.32	92.23	
Extrahs	90.45	95.23	92.32	
Artifact	95.01	95.30	95.00	

Elevated CNN with Heart Rate Image Classification model provides an accuracy of up to 90%, and the performance metrics of this method are given in Table 2.

Table 2. Elevated CNN with heart rate image classification model performance measures

Categories of HeartBeat	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Normal	92.09	87.02	89.50	90
Murmur	81.05	92.12	86.54	
Extrasystole	99.12	87.32	93.02	
Extrahs	90.45	92.23	91.32	
Artifact	90.01	92.30	91.00	

The next model that proves good accuracy is the Enhanced Inception V3 for Heart Rate Classification model, and the accuracy is proven to be 95% in sorting normal and abnormal persons. The performance measures of this model are given in Table 3 as follows.

Table 3. Performance measures of enhanced inception V3 for heart rate classification model

Categories of HeartBeat	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Normal	99.10	92.02	96.50	95
Murmur	90.05	97.12	94.54	
Extrasystole	95.12	95.32	95.02	
Extrahs	93.45	97.23	95.32	
Artifact	99.01	95.30	97.00	

All the models are discussed here; the Facelifted ResNet 50 for Heart Rate Prediction model shows the highest accuracy of 98% in categorizing the sound, whether the dataset belongs to normal or abnormal.

The performance measures are detailed in Table 4, including precision, recall, F1 score, and overall accuracy.

Table 4. Performance measure of facelifted ResNet 50 for heart rate prediction model

Categories of Heart Beat	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Normal	99.09	99.02	99.50	98
Murmur	97.05	97.12	97.54	
Extrasystole	95.12	99.32	97.02	
Extrahs	97.45	97.23	97.32	
Artifact	99.01	95.30	97.00	

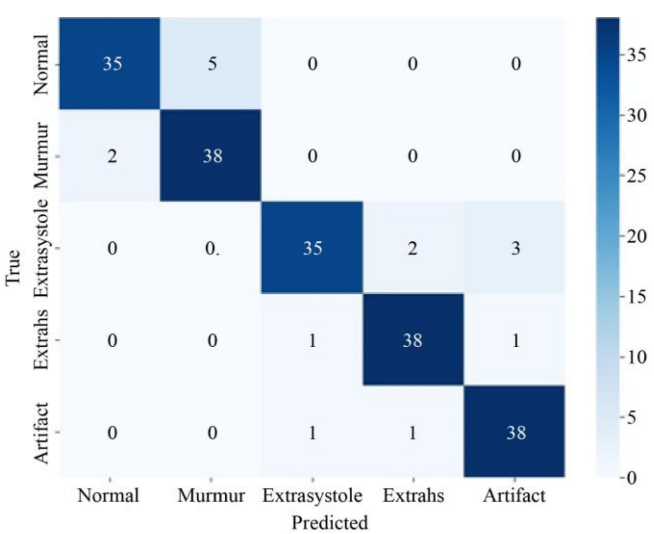

Fig. 8(a) Confusion matrix of modified LSTM model

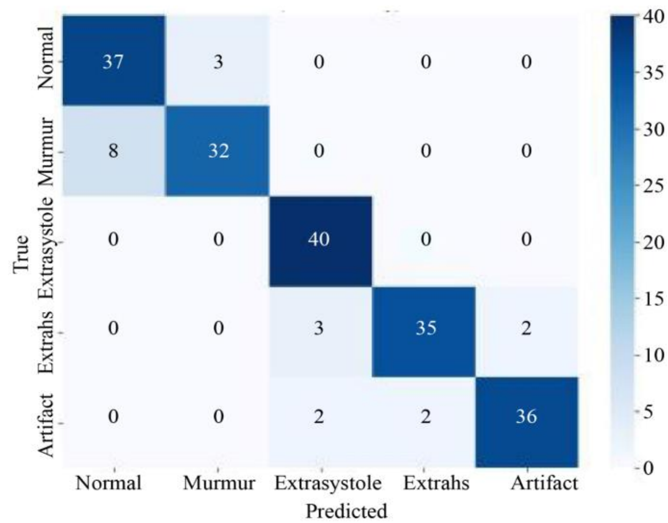

Fig. 8(b) Confusion matrix of ECHIC model

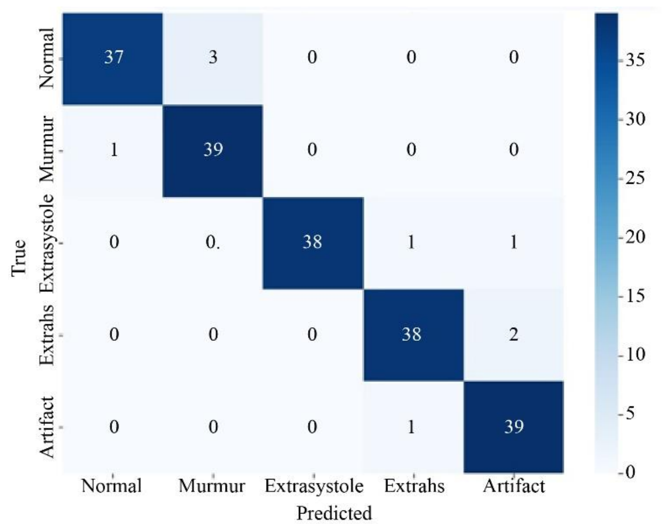

Fig. 8(c) Confusion matrix of EIHRC model

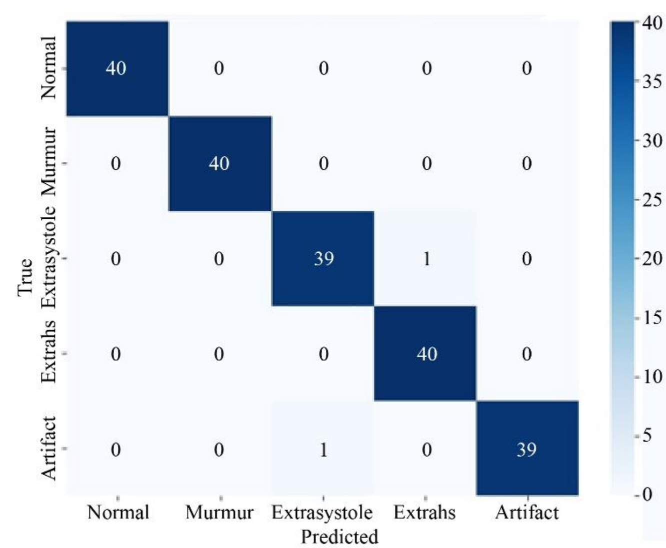

Fig. 8(d) Confusion matrix of FRHRP model

The confusion matrix of the Modified LSTM model depicted in Figure 8 (a) presents a classification of efficacy across various categories of heart disease. There are five classes: normal, extras, extrasystole, murmur and artifact. The normal class for finding not heart disease patients, and the other four classes (artifact, murmur, extrahs and extrasystole) for abnormal heart disease patients. The confusion matrix provides an analysis of the prediction of the model for each class, presenting the quantities of true positives, false negatives, false positives, and false negatives and also identifying the instances of misclassification, and gaining insights into the model's capabilities and limitations in classifying various forms of heart disease. The diagonal elements of the confusion matrix represent the correctly classified instances, while all other elements indicate instances misclassified by the model. The modified LSTM model performs well only for extras and artifacts, but misclassification occurs within other classes, such as 5 normal predicted as a murmur and 3 extrasystoles predicted as an artifact. The confusion matrix in Figure 8 (b) measures the Elevated CNN model performance. This ECHIC model accurately classifies the extrasystole class of abnormal heartbeat sounds but has higher confusion between other classes. The confusion matrix of the Enhanced Inception V3 (EIHRC) model is shown in Figure 8 (c). This EIHRC model provides consistent Performance for all categories, but minor misclassification occurs, such as 1 murmur predicted as 1 normal and 1 extra as 1 artefact. The confusion matrix of the FRHRP model is shown in Figure 8(d), which presents nearly perfect classification across all categories of classes and only 2 misclassifications, as shown in Figure 8(d). So, Modified LSTM and ECHIC provide lower accuracy of 92% and 90%, respectively, and there is frequent confusion between classes. The EIHRC model is also strong, with 95%. Finally, the FRHRP model outperforms all others with 98% accuracy. So, the ResNet 50 FRHRP model performs well for heart sounds and finds heart disease patients easily.

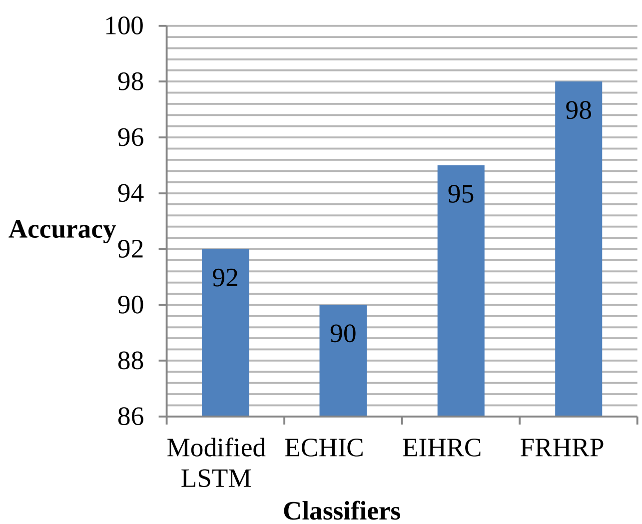

Fig. 9 Bar chart of four classification models for heart disease prediction

Facelifted ResNet 50 for Heart Rate Prediction model outperforms other models primarily due to its robust architecture, which includes deep residual learning to ease the training of networks. Additionally, data augmentation techniques enhance the model's ability to generalize from limited data, further improving its Performance. The Performance of all the four models is given in Figure 9.

Table 5. Comparison with existing methodologies

Year	Author & Methodology	Accuracy (%)
2023	[29] Keikhosrokiani et al. (Adaptive Neuro-Fuzzy Inference System (ANFIS) + Artificial Bee Colony (ABC))	**93.0**
2023	[27] Sudha & Kumar (Hybrid CNN-LSTM)	**89**
2024	[30] Muthaiah et al. (ResNet-50 vs. Particle Swarm Optimization (PSO))	**96.40**
2024	[3] Vinay et al. (RNN-BiLSTM with GAN)	**96.01**
2024	[11] Abbas et al. (Multilayer Perceptron (MLP) with Feature Ensemble)	**95.65**
2025	[19] Zhao et al. [19] (Hybrid Vision Transformer)	**94.24**
2025	[3] Khan et al. (Feature fusion + spectrograms)	**96.55**
2025	Proposed FRHRP Model	**98**

In Table 5 above, research has turned to deep learning to estimate better heart disease, and models from 2023–2025 are now much more accurate. Using FRHRP with spectrogram preprocessing, the model records 98% accuracy, higher than the usual ECG methods. During performance benchmarking, FRHRP exceeds the results of sophisticated models such as Vision Transformers (94.24%) and RNN-BiLSTM (96.01%), while models that unify features (Khan 2025 - 96.55%) have promise but are not as successful at artefact classification. Because of its high reliability, FRHRP helps prevent many errors when diagnosing severe cardiac diseases. Later advancements should combine wearable device data with contactless technology (like that found in the Farooq study, 2024) to improve the usefulness in real-world situations and help find issues early.

4.1. Limitations

Despite their impressive Performance, each model has its limitations. The FRHRP model, while extremely accurate, requires significant computational power and large datasets for training. This may be impractical in resource-limited environments [30]. It is possible that the EIHRC model will not perform well on new, previously unseen data unless thoroughly trained [31]. The temporal dependencies captured by Modified LSTM models are time-consuming to train and susceptible to data noise [15]. The simplicities and speed of the ECHIC models may hinder the ability to discern intricate patterns [16].

5. Conclusion

The FRHRP strategy provides a remarkable improvement in organizing heart disease diagnoses and reaches a 98% level of accuracy. It establishes a new standard for predictive healthcare by being superior to modified LSTM, ECHIC, and EIHRC. By using deep learning and processing images as spectrograms, FRHRP refines how heart noise is analyzed and makes medical diagnostics easier. It will be important for future research in the field to increase data variance and expand model flexibility so that Performance is not affected by diverse patient populations. To enhance live monitoring using wearable devices, AI diagnostics can be used sooner for more targeted healthcare. As a result, using AI for cardiovascular health will change how diseases are managed, so AI-assisted diagnosis becomes a key tool in medicine. Optimizing how accurate predictive AI is, increasing how it can be used and including instant health tracking will help FRHRP and similar technologies transform healthcare into one where early intervention leads to better patient results and helps save lives.

5.1. Further Enhancement

1. Model selection: The right models should be chosen depending on the activity, enough resources and a mix of accuracy and processing speed.
2. Future investigation is needed to identify the weaknesses of these models and strengthen the Performance in clinical care. Plans for new heart disease diagnostics involve continuous monitoring, connecting devices and possibly using the same methods for other diseases.
3. Propose expanding the uses of these methods to discover various health issues, taking advantage of deep learning in many healthcare scenarios.

AI can improve how heart disease is diagnosed and help create new healthcare treatments.

References

[1] Paolo Palatini et al., "Healthy Overweight and Obesity in the Young: Prevalence and Risk of Major Adverse Cardiovascular Events," *Nutrition, Metabolism and Cardiovascular Diseases*, vol. 34, no. 3, pp. 783-791, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Prabhu Pachiyannan et al., "A Novel Machine Learning-Based Prediction Method for Early Detection and Diagnosis of Congenital Heart Disease Using ECG Signal Processing," *Technologies*, vol. 12, no. 1, pp. 1-23, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] N.A. Vinay et al., "An RNN-Bi LSTM Based Multi Decision GAN Approach for the Recognition of Cardiovascular Disease (CVD) from Heart Beat Sound: A Feature Optimization Process," *IEEE Access*, vol. 12, pp. 65482-65502, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Saif Ur Rehman Khan, and Zia Khan, "Detection of Abnormal Cardiac Rhythms Using Feature Fusion Technique with Heart Sound Spectrograms," *Journal of Bionic Engineering*, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[5] Erqiang Deng et al., "Heart Sound Signals Classification with Image Conversion Employed," *Electronics*, vol. 13, no. 7, pp. 1-25, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Sinam Ajitkumar Singh et al., "An Ensemble-Based Transfer Learning Model for Predicting the Imbalance Heart Sound Signal Using Spectrogram Images," *Multimedia Tools and Applications*, vol. 83, no. 13, pp. 39923-39942, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] P. Thendral et al., "Prediction of Abnormalities in Heart Beat Sounds Using Convolutional Neural Networks," *International Journal of Health Sciences*, vol. 6, no. S4, pp. 9844-9855, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Muhammad Farooq et al., "Contactless Heart Sound Detection Using Advanced Signal Processing Exploiting Radar Signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 2, pp. 1009-1020, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] S. Sivasubramaniam, and S.P. Balamurugan, "Early Detection and Prediction of Heart Disease Using Wearable Devices and Deep Learning Algorithms," *Multimedia Tools and Applications*, vol. 84, pp. 6187-6201, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[10] Mohammed Amine Bouqentara et al., "Early Heart Disease Prediction Using Feature Engineering and Machine Learning Algorithms," *Heliyon*, vol. 10, no. 19, pp. 1-23, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Sidra Abbas et al., "Artificial Intelligence Framework for Heart Disease Classification From Audio Signals," *Scientific Reports*, vol. 14, no. 1, pp. 1-21, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Pooja Rani et al., "An Extensive Review of Machine Learning and Deep Learning Techniques on Heart Disease Classification and Prediction," *Archives of Computational Methods in Engineering*, vol. 31, no. 6, pp. 3331-3349, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Qinghao Zhao et al., "Deep Learning in Heart Sound Analysis: From Techniques to Clinical Applications," *Health Data Science*, vol. 4, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Avinash L. Golande, and T. Pavankumar, "Optical Electrocardiogram Based Heart Disease Prediction Using Hybrid Deep Learning," *Journal of Big Data*, vol. 10, no. 1, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Georgios Petmezas et al., "State-of-the-Art Deep Learning Methods on Electrocardiogram Data: Systematic Review," *JMIR Medical Informatics*, vol. 10, no. 8, pp. 1-29, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Wei Chen et al., “Deep Learning Methods for Heart Sounds Classification: A Systematic Review,” *Entropy*, vol. 23, no. 6, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Halah A. Al-Alshaikh et al., “Comprehensive Evaluation and Performance Analysis of Machine Learning in Heart Disease Prediction,” *Scientific Reports*, vol. 14, no. 1, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Sumit Kumar et al., “Optimized Feature Fusion-Based Modified Cascaded Kernel Extreme Learning Machine for Heart Disease Prediction in E-Healthcare,” *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 27, no. 8, pp. 980-993, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Wenhao Zhao et al., “Detection of Coronary Heart Disease Based on Heart Sound and Hybrid Vision Transformer,” *Applied Acoustics*, vol. 230, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[20] Mohammad Alshraideh et al., “Enhancing Heart Attack Prediction with Machine Learning: A Study at Jordan University Hospital,” *Applied Computational Intelligence and Soft Computing*, vol. 2024, no. 1, pp. 1-16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] A. Jeba Sheela, and M. Krishnamurthy, “Revolutionizing Cardiovascular Risk Prediction: A Novel Image-Based Approach Using Fundus Analysis and Deep Learning,” *Biomedical Signal Processing and Control*, vol. 90, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Mert Ozcan, and Serhat Peker, “A Classification and Regression Tree Algorithm for Heart Disease Modeling and Prediction,” *Healthcare Analytics*, vol. 3, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Ezekiel Adebayo Ogundepo, and Waheed Babatunde Yahya, “Performance Analysis of Supervised Classification Models on Heart Disease Prediction,” *Innovations in Systems and Software Engineering*, vol. 19, no. 1, pp. 129-144, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Yehualashet Megersa Ayano et al., “Interpretable Hybrid Multichannel Deep Learning Model for Heart Disease Classification Using 12-Lead ECG Signal,” *IEEE Access*, vol. 12, pp. 94055-94080, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Rehan Ahmed, Maria Bibi, and Sibtain Syed, “Improving Heart Disease Prediction Accuracy Using a Hybrid Machine Learning Approach: A Comparative Study of SVM and KNN Algorithms,” *International Journal of Computations, Information and Manufacturing (IJCIM)*, vol. 3, no. 1, pp. 49-54, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Alexander C. Egbe et al., “Prognostic Value of the Anatomic-Physiologic Classification in Adults with Congenital Heart Disease,” *Circulation: Heart Failure*, vol. 16, no. 9, pp. 809-817, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[27] V.K. Sudha, and D. Kumar, “Hybrid CNN and LSTM Network for Heart Disease Prediction,” *SN Computer Science*, vol. 4, no. 2, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Peter Bentley et al., “Classifying Heart Sounds Challenge,” PASCAL, 2011. [Google Scholar] [Publisher Link]

[29] Pantea Keikhosrokiani et al., “Heartbeat Sound Classification Using A Hybrid Adaptive Neuro-Fuzzy Inference System (ANFIS) and Artificial Bee Colony,” *Digital Health*, vol. 9, pp. 1-22, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[30] M. Muthaiah et al., “Efficient Approach to Predict the Accuracy of Heart Disease by Generating Heartbeat-Based Audio Signal Using ResNet-50 Compared with Particle Swarm Optimization Classifier,” *AIP Conference Proceedings*, vol. 3161, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[31] Jinhee Bae, Minwoo Kim, and Joon S. Lim, “Feature Extraction Model Based on Inception V3 to Distinguish Normal Heart Sound from Systolic Murmur,” *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea (South), pp. 460-463, 2020. [CrossRef] [Google Scholar] [Publisher Link]