Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 4s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Efficient Predictive Modelling of Cardiovascular Disease using Deep Learning Approaches

Divya Lalita Sri Jalligampala*[1], Gangadhar Rao Kancherla[2], R.V.S. Lalitha[3]

[1]Research Scholar, Department of Computer Science and Engineering, University College of Sciences, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, Andhra Pradesh, India,

Email ID: lalitha517@gmail.com

[2]Professor, Department of Computer Science and Engineering, University College of Sciences, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, Andhra Pradesh, India

[3]Professor, Department of CSE, Aditya University, Suramplaem, Andhra Pradesh, India

Cite this paper as: Divya Lalita Sri Jalligampala, Gangadhar Rao Kancherla, R.V.S. Lalitha, (20xx) Efficient Predictive Modelling of Cardiovascular Disease using Deep Learning Approaches. *Journal of Neonatal Surgery*, 14 (4s), 203-218.

ABSTRACT

In order to improve the identification and diagnosis of cardiovascular diseases (CVDs), this study provides a fresh suggested model that uses Graph Convolutional Networks (GCNs) for the multi-label classification of phonocardiographic data. The complex temporal and spectral properties included in heartbeat sounds are often too complex for traditional machine learning frameworks to fully capture. Mel-Frequency Cepstral Coefficients (MFCC) and MEL-spectrograms, two sophisticated feature extraction techniques, work together in our suggested model to improve the representation of important audio aspects that are essential for cardiovascular diagnoses. Through the utilization of GCNs' intrinsic relational structure, the model enhances the propagation of superior information among associated nodes, which in turn improves the categorization of various cardiovascular conditions. Moreover, dynamic customization of node feature representations is made possible by the design, which enhances the resilience of real-world pulse recordings against noise and fluctuation. The results highlight the usefulness of the suggested model in clinical settings, opening the door for further research into hybrid modelling methods and the incorporation of advanced brain architectures with the goal of improving cardiovascular health assessment and intervention plans.

Keywords: *Cardiovascular disease, Prediction, ML-GCN model, MFCC, Deep learning, GCN, Heartbeat Sounds, Multi-label image Recognition, Transformer, CNN.*

1. INTRODUCTION

Heart sound classification aids in the early detection of cardiovascular illnesses. Despite recent advancements in heart sound categorization, most methods still rely on Traditional segmented features and shallow structure-based classifiers.[1] Heart sounds are natural signals that represent physiological information from the human heart. Cardiovascular disease is a term used to describe a group of diseases, including coronary heart disease, cerebrovascular disease, and rheumatic heart disease. A patient's blood pressure, blood sugar, and lipid levels can be raised by fried foods, fast foods, alcohol, and tobacco, as well as weight gain and obesity, leading to premature death.[2] Prevention of sudden death from cardiovascular disease can be achieved by finding groups at risk for cardiovascular disease and ensuring they receive the proper treatment. It is possible to reduce the risk of sudden death from cardiovascular disease by reducing alcohol consumption, reducing salt intake, eating more fruits and vegetables, and exercising moreCardio vascular disease has a major negative influence on people's health and quality of life and is one of the leading causes of disability and mortality in the globe. [3]It is anticipated that the proportion of CVD cases in the US, especially heart failure and stroke, would significantly increase between 2025 and 2060.rise. Furthermore, studies from throughout the world show that during the COVID-19 pandemic, [4]the number of extra fatalities due to cardiovascular disease increased. This rise raises the possibility of decreased monitoring of cardiovascular risk factors and behaviours during the pandemic as well as restricted access to preventative Cardiovascular disease interventions [5]. However, auxiliary diagnostics may assist accomplish distant diagnosis and therapy, as well as increase the efficiency and accuracy of diagnosis. This may increase access to preventative cardiovascular disease care without adding to the amount of available medical resources.

The first, second, third, and fourth heart sounds are produced throughout a cardiac cycle, which happens when one pulse comes before the next. checking for heart problems with heart sound auscultatory Auscultation is a procedure that has been used for over 180 years and is straightforward, essential, and efficient. Liu and associates (2016). The first cardiac sound, which is loud, long-lasting, and intense, signals the start of ventricular systole. The second heart sound, which has shorter length, lower intensity, and less sound, signals the start of ventricular diastole. The third heart sound happens after the second heart sound. It has a larger wavelength and lasts for 0.04 to 0.05 seconds. It is heard by most kids and around half of young [6]adults, and it does not always signify abnormality. A long wave sound that lasts for around 0.04 seconds comes before the first heart sound in the fourth heart sound. A heartbeat is a two-part pumping activity lasting around a second. As blood accumulates in the upper chambers (the right and left atria), the heart's natural pacemaker (the SA node) emits an electrical signal, causing the atria to constrict. [7] This contraction forces blood through the tricuspid and mitral valves into the resting lower chambers (the right and left ventricles). Diastole refers to the lengthier of the two pumping phases. When the ventricles are full with blood, the second phase of the pumping cycle starts. Electrical impulses from the SA node pass via a network of cells to the ventricles, causing them to contract. [2]This is known as systole. As the tricuspid and mitral valves close to prevent blood backflow, the pulmonary and aortic valves open. While the right ventricle pushes blood into the lungs to absorb oxygen, the left ventricle pumps oxygen-rich blood to the heart and other regions of the body.

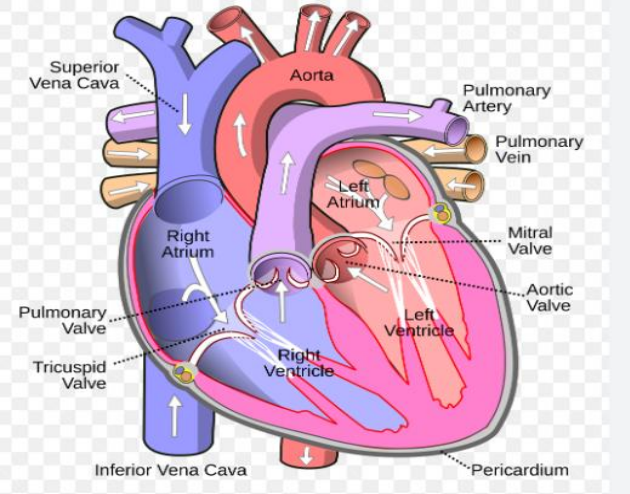

Figure 1 structure of human heart

There are also noticeable variations in the PCGs of various heart sound kinds, as Figure 1 illustrates. Differentiating between focused cardiac sound signals is simple. Thus, many researchers are interested in the automated diagnosis of cardiac problems using cardiac sound waves. One may consider this subject to be an interdisciplinary field of study that includes telemedicine.[8] Researchers have worked hard in the last several years to increase the precision and effectiveness of heart sound signal categorization.

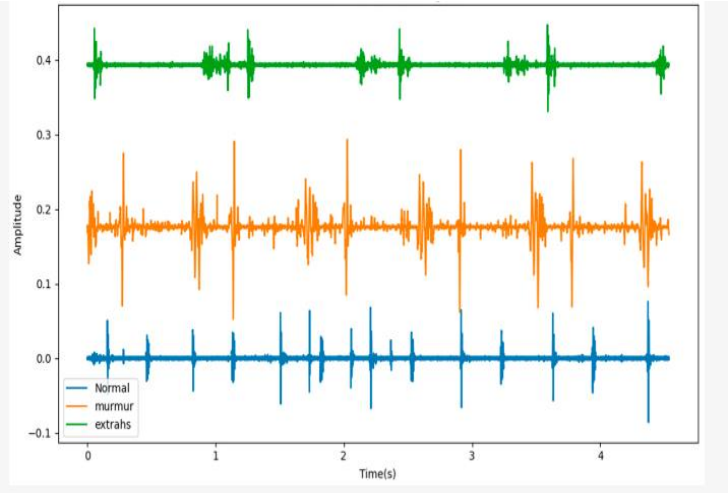

Figure 2 Three kinds of heart sound signals

The primary emphasis of early heart sound signal categorization research was on machine learning algorithms and conventional signal processing methods. These approaches were usually less accurate, lacked [9]the capacity for generalisation, and needed human feature extraction and classifier building. On the other hand, these techniques laid the groundwork for subsequent deep learning strategies. Phonocardiography (PCG) is the measurement of a heart sound, which is a kind of physiological signal. Systole and diastole of the heart create it, and it may represent physiological data about bodily parts including the atria, ventricles, and substantial boats and their operational conditions.[10] Fundamental heart sounds, or FHSs, are often divided into two categories: first heart sounds, or S1, and second heart sounds, or S2, respectively. S1 typically happens when the ventricles' rapidly increasing pressure causes the mitral and tricuspid valves, which are already closed, to abruptly approach their elastic limit, signalling the start of isovolumetric ventricular contraction. When the aortic and pulmonic valves shut at the start of the diastole, S2 takes place.

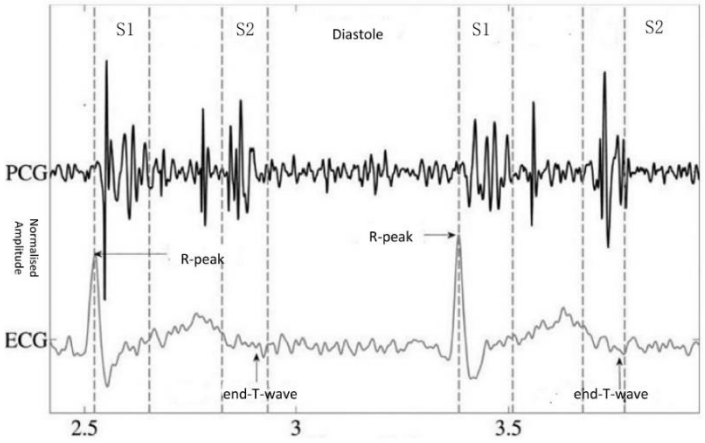

Figure 3 PCG with simultaneous ECG

recording and the four states of the PCG recording: S1, the systole, S2, and the diastole Accurate segmentation of the FHSs is crucial for determining the state sequence of S1, S2, and Diastole. Figure 1 shows a PCG procedure with simultaneous ECG recording and its four states: S1, the Systole, S2 and diastole. The QRS waveform of the ECG and the cardiac sound signal are utilised to determine the S1 and S2 positions. FHSs are crucial for assessing heart disease and guiding future diagnostics. To accurately diagnose heart disorders, it's crucial to extract characteristics from the whole FHS for quantitative analysis. Over the last several decades, there has been a growing interest in automatically classifying heart sounds.

2. LITERATIRE REVIEW

Yang and Hsieh (2016) found the mel frequency cepstral coefficients (MFCC). [11]This work presents the first and second-order MFCC-based Difference coefficients are used as the neural network's input tensor. By removing noise from the data, this feature extraction technique improves classification accuracy by enabling the neural network to collect both healthy and pathological information from the heart sound signal. [12]The issues of human intervention, intricate procedures, and inadequate generalisation are avoided by deep learning approaches as compared to conventional heart sound categorization algorithms. Graph convolutional networks (GCNs) have shown remarkable performance in several medical contexts including graph node categorization problems.[13] Transfer learning for graph representation learning and graph network models presents challenges, nonetheless. Most GNNs function only in one domain and cannot apply acquired knowledge to many spheres[14]. This is a result of the significant developments in modern theories on the creation of artificial intelligence [15] Graph Convolutional Networks (GCNs) with Attention Mechanisms improve task performance on graph-structured data by fusing the advantages of attention with the power of graph-based learning. [16]GCNs are good at gathering local neighbourhood information because they operate on nodes in a graph using convolutional processes. By including attention methods, the model is able to dynamically concentrate on the most significant regions of the network and assign varying weights to neighbouring nodes. [17] By offering a more flexible and context-sensitive aggregation of node information, this synergy enhances the model's capacity to handle large, heterogeneous networks and improves performance in tasks like node classification, link prediction, and graph classification. The brand-new architecture called the Graph Convolutional Attention Network (GCAN) to forecast possible relationships between RNAs and diseases. Despite dozens of organisations, GCAN offers advantages from the deep learning model's effectiveness.[18]

Using Mel-frequency Cepstral coefficients (MFCC) as feature extraction,[19] Han Li et al. suggested a model for classification using a Convolutional Neural Network. CNN stands for [CNN]. Their accuracy comes out to be 90.43%. There are PCG signals in 175 of their datasets. Using dynamic content characteristics and multichannel PCG signals will help to improve the accuracy of CAD detection. [20] Castro et al study uses a variety of artificial neural networks (ANN) and the

fusion of spectral information to analyse PCG signals in order to diagnose CVD. A self-designed PCG acquisition setup is used to get the PCG signal via the subject's heart. Following pre-processing, five spectral characteristics with the largest pair-wise differences are retrieved. With 99.99% accuracy, five distinct ANN types—narrow, broad, tri-layered, bi-layered, and medium—are simulated. In comparison to existing methods, this suggested design is non-invasive, reasonable, and dependable. It also provides excellent direction for developing new, affordable options for CVD diagnostic procedures.

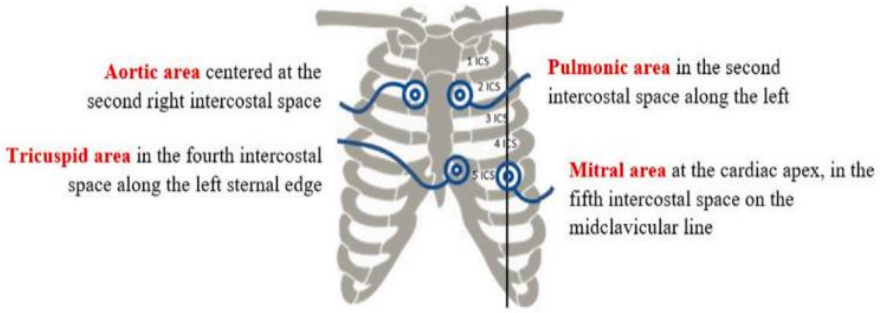

Figure 4 PCG Acquisition Areas

Using the ECG signal as a reference, the segmentation procedure may be either direct or indirect. By envelope (direct) or by ML (using outside medical knowledge) is another way to categorise segments. Certain techniques that are commonly used for direct segmentation include normalised average Shannon energy [21], [22], the Hilbert transform, the cardiac sound characteristic waveform (CSCW) [23], autoregressive moving average spectral methods, power spectral density, energy of wavelet coefficients, complexity signatures, and Wigner-Ville distribution . In order to ascertain the temporal position of the start and finish of each cardiac cycle in the signal (the onset and offset), the authors of used the segmentation approach based on the Matching Pursuit algorithm introduced in [24]. The algorithm has shown a 97.5% accuracy rate for the onset.

[25] used the local three-valued pattern (LTP) to extract the local binary pattern (LBP) of heart sounds, and then trained it using a 1D-CNN with an accuracy of 90% of the dataset from PhysioNet.

[26] achieved an unweighted average recall of 51.2% by using the attention mechanism to investigate the interpretable heart sound classification algorithm for the heart sound triple classification task on the PhysioNet dataset. [27] used a support vector machine and a 1D adaptive local ternary model based on MFCC to achieve 98.3% accuracy on the heart sound double classification task.
Using a discrete wavelet transform and a support vector machine optimised using Bayesian optimisation,[28] achieved an accuracy of 89.26%. Neural networks based on MFCC characteristics outperform other neural networks in the heart-tone classification problem. This study computes first-order and second-order difference coefficients for describing the dynamic aspects of heart sound signals in order to further improve the benefits of MFCC features in expressing heart sound signals. The approach is innovative in three ways: 1) Using the credible heart sound datasets from three distinct sources, which significantly enhances the deep learning model's performance. 2) Choosing enhanced MFCC input features to more accurately capture the static and dynamic properties of the cardiac sound signal. 3) Making use of a residual neural network, which reduces training-related gradient disappearance and deterioration

3. SCOPE AND METHODOLOGY

Previous research on PCG signal classification for CAD detection identifies gaps in extensive spectral analysis, comparative segmentation technique comparison, multichannel PCG integration, non-invasive diagnostic approaches, architecture comparisons of ANNs, optimization of feature extraction, refinement of attention mechanisms, dataset generalizability, hybrid model assessment, and validation of temporal position accuracy.
Graph Convolutional Networks (GCNs) used with the Multi-Label Image Recognition model fills in a lot of these gaps. In order to improve feature representation and detection accuracy, GCNs include a variety of spectral and dynamic features. They facilitate the development of useful, non-invasive diagnostic techniques and manage multichannel data efficiently. Because GCNs have built-in sophisticated optimization methods and may include attention mechanisms for more complicated tasks, they provide a strong substitute for classic ANNs. Model generalizability is ensured by their strong performance across a variety of datasets. Furthermore, a hybrid method is represented by GCNs paired with conventional signal processing, which indirectly enhances algorithms for estimating the temporal locations of cardiac cycles. Overall, by filling up these important research gaps, GCNs greatly improve PCG signal categorization for CAD diagnosis.

Methodology:

One method that may be used to create models that effectively forecast cardiovascular diseases is deep learning. Any individual exhibiting abnormal heart sounds, which may be a sign of early-stage cardiovascular illness, may be identified by this model. This model may be used to identify irregular heart sounds and can be stopped early, when the problem is easier to treat. We locate the illness as soon as possible, making its eradication much simpler. As far as you know, it will be difficult to remove the illness.

Data Collection:

The "Heartbeat Sounds" dataset on Kaggle provided the information needed for the categorization study of heartbeat sounds. Heartbeat audio recordings are included in this dataset, which is divided into classifications for normal and pathological heartbeats. The recordings, which cover a range of cardiac states, were gathered from many sources. Every recording is a.wav file with associated metadata that contains details on the patient and the setting in which it was made. The dataset has been organised in a way that makes it easier to train and assess Deep learning models that categorise pulse sounds.

Dataset:

The Heartbeat Sounds dataset utilized in this research was sourced from Heartbeat Sounds

(https://www.kaggle.com/code/brsdincer/heartbeat-sounds-classification-analysis/input?select=set_b.csv).

file_path = "/content/Heartbeat_Sound"

print(os.listdir(file_path))

Input: Heart Sound Signals

Output: Classification

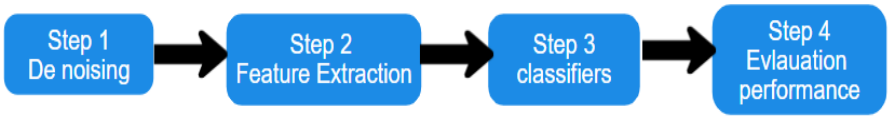

Figure 5 Architecture

Graph Convolutional Network (GCN):

The semi-supervised classification function of the Graph Convolutional Network (GCN) was presented in [29]. Updating node representations by information propagation between nodes is the fundamental concept. In contrast to standard convolutions, which work on local the Euclidean structures in an image, GCN aims to learn a function $\int(\cdot,\cdot)$ on a graph ς . It does this by updating node features as $h^{l+1} \in r^{n\times d'}$ and receiving as inputs feature descriptions $h^{l} \in r^{n\times d}$ and the corresponding correlation matrix $a \in r^{n\times n}$, where n is the number of nodes and D denotes the dimensionality of node features. Each GCN layer may be expressed as a non-linear function by

$$h^{l+1} = \int\left(h^{l}, a\right). \tag{1}$$

Once the convolutional operation is applied, $\int(\cdot,\cdot)$ may be expressed as

$$h^{l+1} = H\left(\hat{a}h^{l}w^{l}\right), \tag{2}$$

where $\hat{a} \in r^{n\times n}$ is the normalised version of correlation matrix a , $H(\cdot)$ indicates a non-linear operation, and $w^{l} \in r^{D\times D'}$ is a transformation matrix to be learnt, which in our studies is handled by LeakyReLU [30]. As a result, by stacking many GCN layers, we are able to understand and simulate the intricate interactions between the nodes. We direct interested readers to for more information.

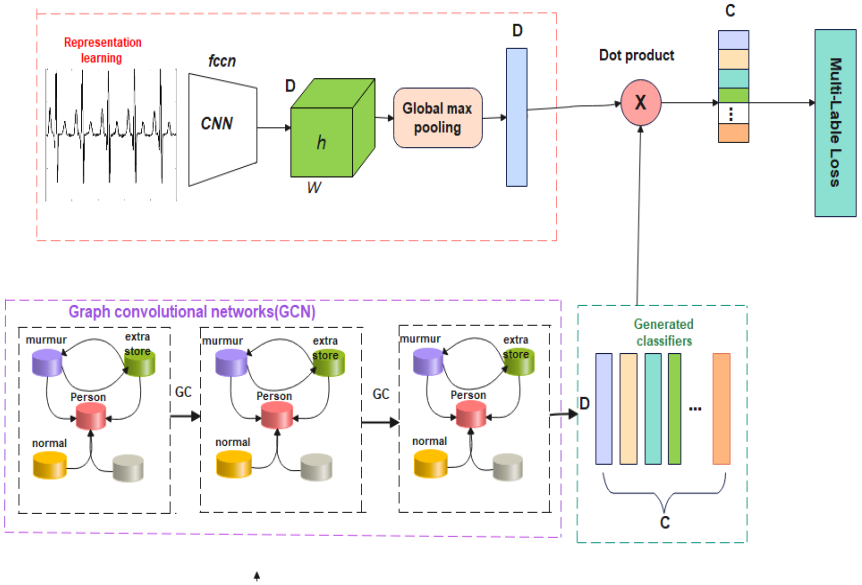

Figure 5 Overall framework of our ML-GCN model for multi-label image recognition

ML- GCN for Multi-label Recognition:

GCN is the foundation of our ML-GCN. For semi-supervised classification, GCN was suggested, where each node's prediction score represents the node-level output. In contrast, we plan for each GCN node's final output to be the classifier for the associated label in our job. Furthermore, in other tasks, the network structure, or correlation matrix, is often predefined; in the multilabel image recognition task, this is not the case. As a result, we must create the correlation matrix from scratch. Fig.6 depicts the general architecture of our method, which is made up of the GCN-based classifier learning and image representation learning modules.

GCN based classifier learning

Using a GCN-based mapping function, we may train inter-dependent object classifiers $w = \{\omega_i\}_{i=1}^{c}$ using label representations, where c is the number of categories. We use stacked GCNs, in which every GCN layer l receives as inputs the node representations from the preceding layer (h') and outputs new node representations, or h^{l+1}. The $z \in r^{c \times D}$ matrix, where D is the dimensionality of the label-level word embedding, is the input for the first layer. The output for the last layer is $w \in r^{c \times d}$, where d is the number of dimensions in the picture representation. We may acquire the expected scores as follows by using the learnt classifiers on picture representations.

$$\hat{\Upsilon} = wX. \tag{3}$$

We take it for granted that an image's ground truth label is $\lambda \in r^{c}$, where $\lambda^{I} = \{0,1\}$ indicates whether or not label I occur in the picture. Using the conventional multi-label classification loss, the whole network is trained as follows.

$$L = \sum_{C=1}^{c} \lambda^{C} \log\left(\beta\left(\hat{\lambda}^{C}\right)\right) + \left(1+\lambda^{C}\right)\log\left(1-\beta\left(\hat{\lambda}^{C}\right)\right), \tag{4}$$

Beta function is represented by $\beta(\cdot)$.

4. EVALUATION METRICS

Accuracy: It is easy to evaluate the classifier's accuracy by looking at how often it produces correct predictions. The percentage of correct forecasts to all estimations offers an additional meaning.

$$Accuracy = \frac{TP + TN}{S}$$

Precision: Recall is acquired by dividing precision by one, whereas this ratio, which reflects the proportion of false negatives, is obtained by subtracting one for it, i.e., (1 – exact).

$$Precision = \frac{TP}{TP + FP}$$

Recall: In contrast to true negatives, there are also things known as false negatives.

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

F1-Score: This is determined by squaring the accuracy and recall values. For this.

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

5. RESULT AND DISCUSSION

Results:

The study, which used Kaggle's Heartbeat Sounds dataset, shows that our ML-GCN technique classifies heartbeat sounds well. Our deep learning models were trained and evaluated using normal and abnormal cardiac acoustic recordings. The metadata-rich.wav files enabled for thorough cardiac state analysis.

We used a GCN to construct a multi-label recognition system. The correlation matrix was created from scratch in our technique, unlike standard methods. Our GCN-based classifier learning and image representation learning architecture worked well. This novel modification of GCN accurately classified heartbeat sounds, proving its suitability for complicated multi-label medical data processing.

Confusion Matrix:

MFCC Feature Extraction(Confusion Matrix)

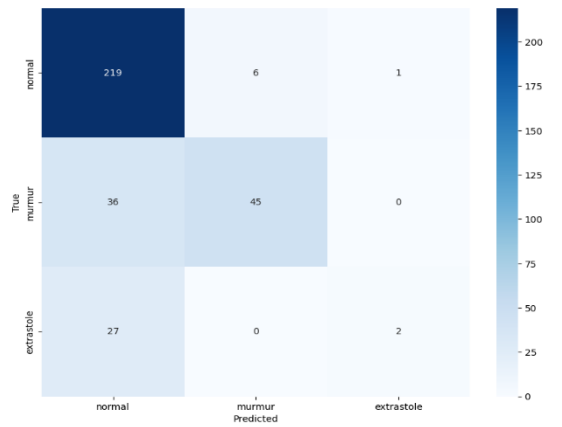

Figure 6: Confusion Matrix of Proposed Model

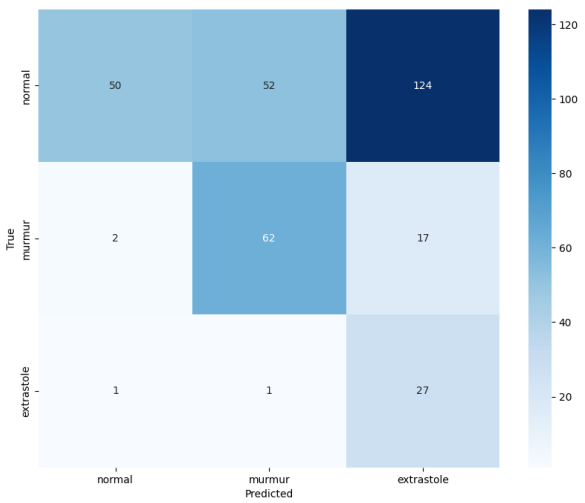

Figure 7 :Confusion Matrix GCN Model

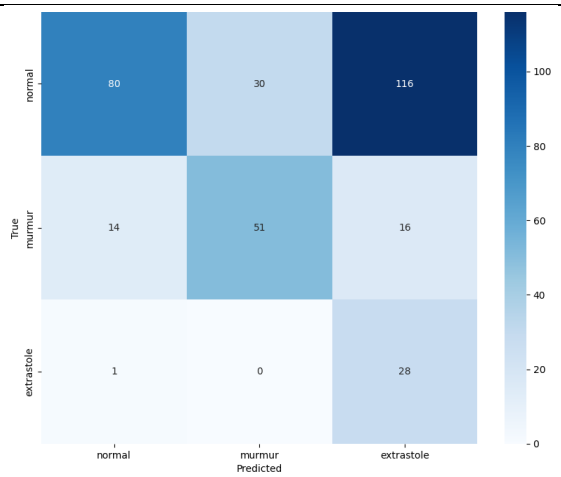

Figure 8 :Confusion Matrix of CNN Model

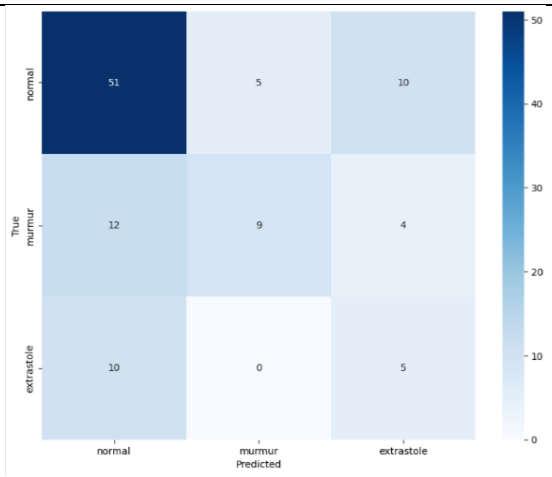

Figure 9:Confusion Matrix of VGGish Model

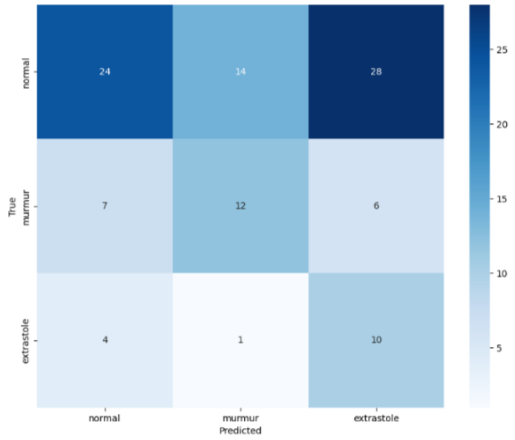

Figure 10:Confusion Matrix of LDA Model

The picture shows the confusion matrices for five distinct models that are used to categorize three cardiac conditions: normal, murmur, and extrasystole: the Proposed Model, CNN (Convolutional Neural Network), VGGish, and LDA (Linear Discriminant Analysis). Darker hues indicate a larger number of cases that were successfully or wrongly identified, with the real labels on the y-axis and the predicted labels on the x-axis in each matrix.The proposed model's confusion matrix is shown in Figure 7. With 239 accurate classifications, the suggested model has rather good performance in categorizing typical occurrences. But it misclassifies five cases as extrasystole and six as murmur. Seven cases of extrasystole and 37 cases of normal are misclassified as murmurs, whereas 57 cases are correctly labelled. The categorization of extrasystole is moderate, with 53 accurate predictions, while 16 are incorrectly categorized as normal and 27 as murmur.

GCN Model Confusion Matrix, Figure 8 The GCN model misclassifies 32 instances as murmur and 134 as extrasystole, while correctly classifying 59 normal cases. Although it correctly detects 60 murmur instances, 17 and 64 are incorrectly labelled as extrasystole and normal, respectively. Only three cases of extrasystole are accurately identified by the model, and it misclassifies a large number of cases in other categories, suggesting that its classification capacity is limited.CNN Model Confusion Matrix, Figure 9 When it comes to identifying normal cases, the CNN model does a respectable job; 95 cases were accurately recognized, while 35 were incorrectly labelled as murmur and 116 as extrasystole. While it correctly predicts 14 murmurs, it incorrectly labels 78 as normal and 15 as extrasystole. A considerable percentage of extrasystole cases are misclassified as normal or murmur, with only six cases being accurately labelled.

Confusion Matrix of VGGish Model, Figure 10 Only 12 of the normal instances are correctly predicted by the VGGish model, but 2 and 13 examples are incorrectly labelled as extrasystole and murmur, respectively. Only two occurrences of extrasystole are successfully predicted by the model, suggesting low overall performance, but 18 instances of murmur are correctly identified, despite a significant incidence of misclassification.Confusion Matrix of the LDA Model, Figure 11 Five cases were misclassified as extrasystole and six as murmur, whereas the LDA model correctly classified 11 cases for normal circumstances. Murmur receives nine accurate classifications, but it misclassifies a number of cases, particularly in the usual

group. The majority of cases are mislabeled, and the extrasystole categorization is restricted.All things considered, the confusion matrices show that the Proposed Model performs the best out of the five models, with comparatively greater accurate classifications in every category, particularly for the normal and murmur classes. Although they perform somewhat, the GCN and CNN models have trouble with serious misclassifications. The low classification accuracy of the VGGish and LDA models, particularly for extrasystole, suggests that there is much space for development in these models' capacity to differentiate between the various cardiac diseases.

MEL-Spectrogram Feature Extraction (Confusion Matrix):

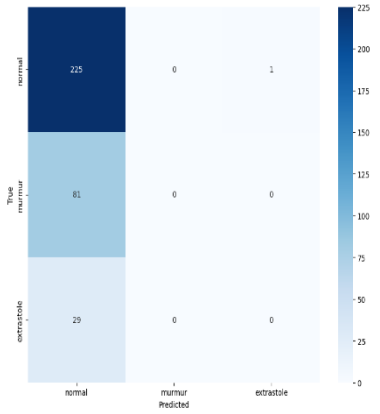

Figure 11: Confusion Matrix of Proposed Model

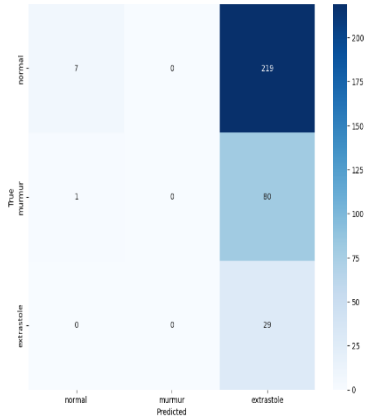

Figure 12: Confusion Matrix of GCN Model

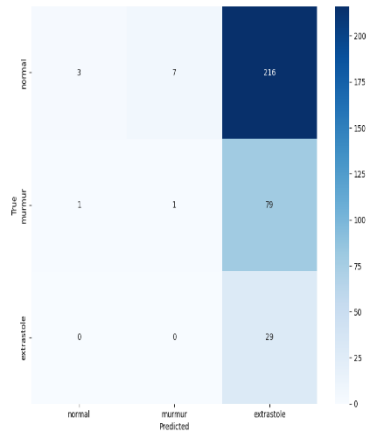

Figure 4: Confusion Matrix of CNN Mode

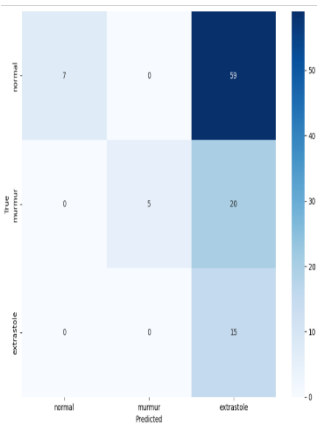

Figure 14: Confusion Matrix of VGGish Model

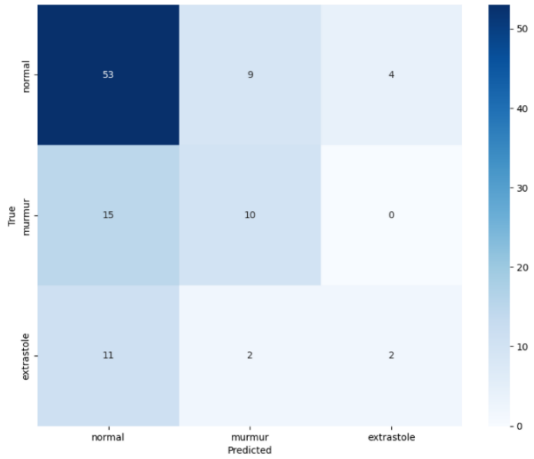

Figure 515:Confusion Matrix of LDA Model

When applied to a classification problem, the graphic shows five confusion matrices for various machine learning models. The x- and y-axes of each confusion matrix indicate the classes, and the true vs anticipated labels are shown. Each cell's color intensity reflects the number of predictions for each pair of classes; larger counts are indicated by deeper hues. Below is an explanation of every matrix:The performance of a customized or optimized model is probably depicted in Figure 12's confusion matrix of the suggested model. The darker diagonal cells indicate a large number of cases properly identified by the model for each class.The GCN Model's Confusion Matrix in Figure 13 Although there are several misclassifications in the confusion matrix of the GCN (Graph Convolutional Network) model, heavier diagonal colors suggest that certain categories were classified with respectable accuracy.Figure 14 shows the CNN Model Confusion Matrix The CNN (Convolutional Neural Network) model performs well, with the majority of predictions centered along the diagonal, suggesting that it categorized many examples correctly.

The VGGish Model's Confusion Matrix in Figure 15 Although there are some incorrect classifications, as shown by lighter cells outside the diagonal, the confusion matrix of the VGGish model displays a sizable number of accurate classifications on the diagonal.Figure 16 shows the LDA Model's confusion matrix. In contrast to some of the other models, the LDA (Linear Discriminant Analysis) model exhibits more misclassifications in addition to having a deeper hue for specific classes along the diagonal, suggesting correct classifications for them.Though some models achieve more accuracy as shown by the intensity of their diagonal cells, each confusion matrix provides insight into the models' strengths and shortcomings in terms of identifying situations.

ROC Curve:

ROC Curve(MFCC)

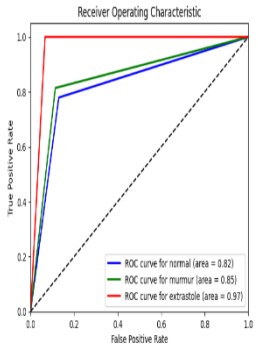

Figure 16 Proposed Model of ROC Curve

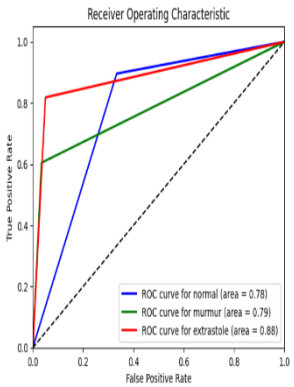

Figure 17 GCN Model of Roc Curve

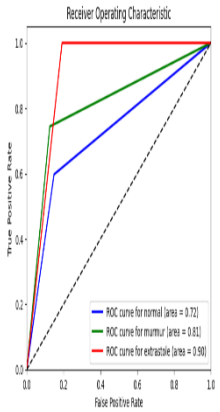

Figure 18: CNN Model of ROC Curve

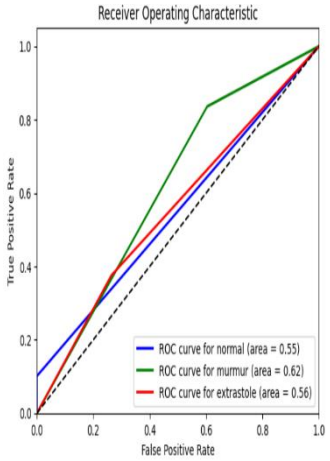

Figure 19: VGGish Moddel of ROC Curve

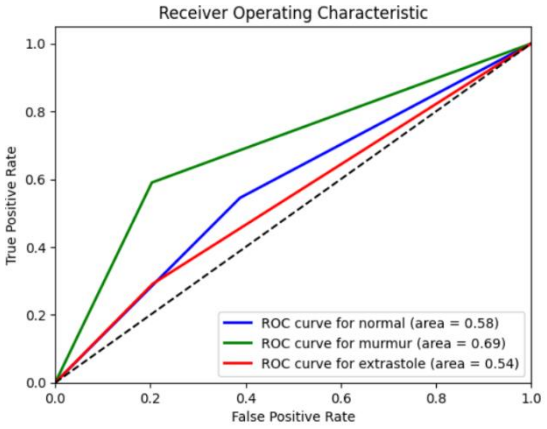

Figure 20: LDA Model of ROC Curve

In order to classify three cardiac conditions normal, murmur, and extrasystole five machine learning models the Proposed Model, CNN (Convolutional Neural Network), VGGish, and LDA (Linear Discriminant Analysis)—are shown in the image along with their ROC (Receiver Operating Characteristic) curves. By showing the True Positive Rate (sensitivity) on the y-axis versus the False Positive Rate on the x-axis, each plot graphically illustrates how well a model performs. The AUC (Area Under Curve) values for each of the three classes are displayed alongside color-coded ROC curves. Higher values of the AUC, a crucial statistic, show that a model is more effective at discriminating across classes.With an AUC of 0.82 for the normal class, 0.85 for the murmur class, and 0.88 for the extrasystole class, the suggested model outperforms the other five models (Figure 17: Proposed Model of ROC Curve). These results show that the suggested model has a robust classification capability across all categories and performs very well in correctly recognizing each condition, especially extrasystole.Figure 18 (ROC Curve GCN Model): With an AUC of 0.78 for the normal class, 0.79 for murmur, and 0.88 for extrasystole, the GCN model likewise exhibits strong performance. Although its performance in identifying extrasystole is comparable to that of the suggested model, it exhibits somewhat reduced sensitivity for the normal and murmur classes, indicating a little decline in accuracy for these classes in comparison to the suggested model.The CNN model produces an AUC of 0.77 for normal, 0.85 for murmur, and 0.80 for extrasystole (Figure 19: CNN Model of ROC Curve). This shows that although the CNN does well, especially for the murmur class, it is not quite as well as the suggested model for classifying extrasystole. This model could have trouble differentiating this specific condition, as seen by the lower AUC for extrasystole.With AUC values of 0.55 for the normal class, 0.62 for murmur, and 0.54 for extrasystole, the VGGish model performs the worst out of all the models shown (Figure 20, VGGish Model of ROC Curve). These scores are around the 0.5 threshold, suggesting that the VGGish model has poor classification accuracy and finds it difficult to distinguish between cases of extrasystole, murmur, and normal.The LDA model also exhibits poorer discriminative ability, with an AUC of 0.58 for the normal class, 0.56 for murmur, and 0.64 for extrasystole (Figure 21: LDA Model of ROC Curve). Its shortcomings

as a classifier in this application are highlighted by the fact that, although doing somewhat better than the VGGish model, particularly for extrasystole, it is still unable to reliably distinguish the classes.In conclusion, the suggested model is the most successful of the five, obtaining the greatest AUC in every class, especially when it comes to extrasystole detection. The sensitivity of the GCN model is slightly lower for the normal and murmur classes, although it still follows closely. While it does a decent job, the CNN model is not very good at identifying extrasystole. The VGGish and LDA models, on the other hand, both show noticeably worse classification accuracy, highlighting their limited applicability in this situation.

This comparison demonstrates how well the suggested model performs in classifying heart conditions.The image displays the Receiver Operating Characteristic (ROC) curves for five machine learning models, each of which aims to classify three cardiac conditions: normal, murmur, and extrasystole. These models are the Proposed Model, CNN (Convolutional Neural Network), VGGish, and LDA (Linear Discriminant Analysis). By showing the True Positive Rate (sensitivity) on the y-axis versus the False Positive Rate on the x-axis, each plot graphically illustrates how well a model performs. The AUC (Area Under Curve) values for each of the three classes are displayed alongside color-coded ROC curves. Higher values of the AUC, a crucial statistic, show that a model is more effective at discriminating across classes.Figure 17 (ROC Curve Proposed Model), with an AUC of 0.82 for the normal class, 0.85 for the murmur class, and 0.88 for the extrasystole class, the suggested model performs the best out of the five models. These results show that the suggested model has a robust classification capability across all categories and performs very well in correctly recognizing each condition, especially extrasystole.

Figure 18 (ROC Curve GCN Model),with an AUC of 0.78 for the normal class, 0.79 for murmur, and 0.88 for extrasystole, the GCN model likewise exhibits strong performance. Although its performance in identifying extrasystole is comparable to that of the suggested model, it exhibits somewhat reduced sensitivity for the normal and murmur classes, indicating a little decline in accuracy for these classes in comparison to the suggested model.Figure 19 (ROC Curve CNN Model),the CNN model's AUC is 0.80 for extrasystole, 0.85 for murmur, and 0.77 for normal. This shows that although the CNN does well, especially for the murmur class, it is not quite as well as the suggested model for classifying extrasystole. This model could have trouble differentiating this specific condition, as seen by the lower AUC for extrasystole.

Figure 20 (ROC Curve VGGish Model),with AUC values of 0.55 for the normal class, 0.62 for murmur, and 0.54 for extrasystole, the VGGish model performs the worst out of all the models that are shown. These scores are around the 0.5 threshold, suggesting that the VGGish model has poor classification accuracy and finds it difficult to distinguish between cases of extrasystole, murmur, and normal.The LDA model also exhibits poorer discriminative ability, with an AUC of 0.58 for the normal class, 0.56 for murmur, and 0.64 for extrasystole (Figure 21: LDA Model of ROC Curve). Its shortcomings as a classifier in this application are highlighted by the fact that, although doing somewhat better than the VGGish model, particularly for extrasystole, it is still unable to reliably distinguish the classes.In conclusion, the suggested model is the most successful of the five, obtaining the greatest AUC in every class, especially when it comes to extrasystole detection. The sensitivity of the GCN model is slightly lower for the normal and murmur classes, although it still follows closely. While it does a decent job, the CNN model is not very good at identifying extrasystole. The VGGish and LDA models, on the other hand, both show noticeably worse classification accuracy, highlighting their limited applicability in this situation. This comparison demonstrates how well the suggested model performs in classifying heart conditions.

ROC Curve(MEL-Spectrogram)

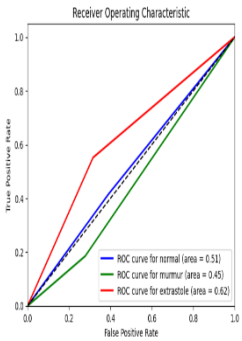

Figure 21: Proposed Model of ROC Curve

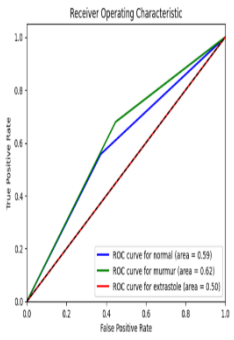

Figure 22: GCN Model of Roc Curve

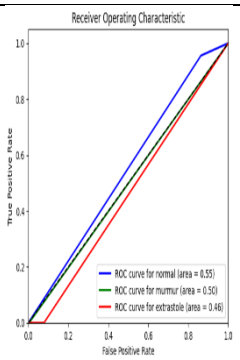

Figure 23: CNN Model of ROC Curve

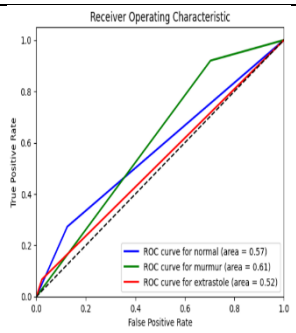

Figure 24: VGGish Model of ROC Curve

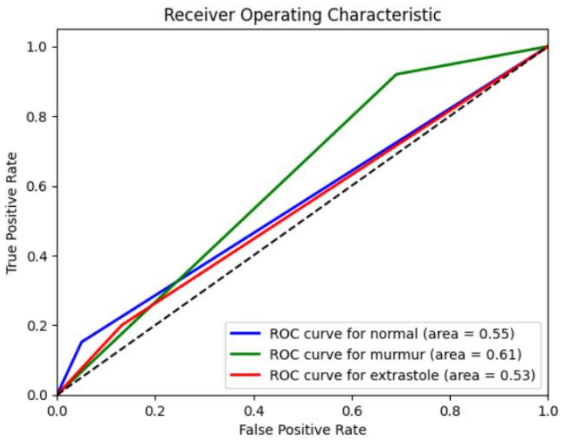

Figure 25: LDA Model of ROC Curve

In order to classify three cardiac conditions normal, murmur, and extrasystole—the image displays ROC (Receiver Operating Characteristic) curves for five distinct models: the Proposed Model, CNN (Convolutional Neural Network), VGGish, GCN (Graph Convolutional Network), and LDA (Linear Discriminant Analysis). With different coloured lines denoting each class, each plot contrasts the True Positive Rate (sensitivity) on the y-axis with the False Positive Rate on the x-axis. Each class's AUC (Area Under Curve) values are displayed to indicate how well each model performed in differentiating between the three circumstances.

Figure 22 (ROC Curve Proposed Model),with an AUC of 0.63 for normal, 0.59 for murmur, and 0.62 for extrasystole, the suggested model performs moderately. The AUC values indicate that it has limited capacity to categorize the three circumstances, despite its somewhat superior performance compared to some of the other models.The GCN model produces an AUC of 0.59 for normal, 0.53 for murmur, and 0.59 for extrasystole (Figure 23: GCN Model of ROC Curve). The fact that these AUC values are less than those of the suggested model indicates that the GCN model performs worse in classification across all classes, particularly when it comes to murmur differentiation.The CNN model yields an AUC of 0.53 for normal, 0.58 for murmur, and 0.61 for extrasystole (Figure 24: CNN Model of ROC Curve). Although it performs somewhat better than the GCN model for murmur and extrasystole, the CNN model's overall classification accuracy is still poor, suggesting that it would have trouble successfully differentiating between the three situations.

Figure 25 (ROC Curve VGGish Model),with AUC values of 0.57 for normal, 0.53 for murmur, and 0.57 for extrasystole, the VGGish model performs poorly in classification. According to these findings, the VGGish model performs similarly to a random classifier (AUC ~0.5), showing difficulty in correctly classifying each situation.Figure 26 (ROC Curve LDA Model),the AUC of the LDA model is 0.53 for extrasystole, 0.52 for murmur, and 0.53 for normal. Though the AUC is still low, the LDA model, like the other models, exhibits limited classification skills. It performs best in extrasystole classification.Overall, the findings show that none of the models is able to differentiate between extrasystole, murmur, and normal circumstances with good accuracy. The VGGish and LDA models had the worst classification performance across all classes, while the suggested model performs somewhat better than the others, particularly for the normal and extrasystole classes. This implies that in order to increase classification accuracy, more optimization or different modeling techniques could be required.

Performance Metrics

Table 1 Performance Evaluation Metrics

Feature Extraction Method	Models	Accuracy	Precission	Recall	F1-Score
MFCC	Proposed Model	0.92	0.93	0.92	0.92
	GCN	0.79	0.81	0.79	0.79
	CNN	0.68	0.77	0.68	0.70
	VGGish	0.30	0.75	0.30	0.24
	LDA	0.53	0.62	0.53	0.56
MEL-Spectrogram	Proposed Model	0.55	0.70	0.55	0.55
	GCN	0.54	0.59	0.54	0.54
	CNN	0.64	0.47	0.64	0.54
	VGGish	0.40	0.60	0.40	0.37
	LDA	0.33	0.61	0.34	0.29

Using two different feature extraction techniques MFCC and MEL-Spectrogram,this performance evaluation contrasts the efficacy of the model across different topologies. The suggested model has strong classification skills in MFCC, achieving the top metrics with an accuracy of 0.92, precision of 0.93, and F1-score of 0.92. With an accuracy of 0.92 and a similar F1-score (0.79), the GCN model comes in second, demonstrating its ability to manage MFCC characteristics efficiently. Despite its modest effectiveness, the CNN's F1-score of 0.70 indicates restricted performance in comparison to the suggested model and GCN, as evidenced by its lower accuracy of 0.68. With an accuracy of 0.30 and an F1-score of 0.24, the VGGish model notably performs much worse, suggesting that this feature extraction technique is not a good addition to VGGish. Likewise, LDA produces less than ideal results, achieving an accuracy of 0.53 and an F1-score of 0.56.

Overall model performance decreases when MEL-Spectrogram features are included. With an F1-score of 0.55, the suggested model retains the greatest accuracy in this group (0.55), but much worse than its MFCC-based performance. Although CNN's precision-recall balance fluctuates, as seen by an F1-score of 0.54, the GCN and CNN models perform similarly, attaining accuracies of 0.54 and 0.64, respectively. With an F1-score of 0.37 and an accuracy of 0.40, the VGGish model once again demonstrates its diminished effectiveness with this feature type. With an accuracy of 0.33 and an F1-score of 0.29, LDA performs the worst. According to this research, MEL-Spectrogram features provide less cross-model compatibility than MFCC characteristics, which match the evaluated architectures more well, especially for the suggested model and GCN.

The findings show that feature extraction methods are critical in determining model performance, and that, in this particular situation, MFCC is unquestionably a better option for both the suggested model and the identification of cardiovascular illness. The decrease in performance seen in all models when using MEL-spectrogram suggests that MFCC may be a more effective method of capturing the fine features needed for this job. Furthermore, compared to GCN and CNN, the suggested model performs better in both feature extraction situations, suggesting that it is more capable of handling a wider range of input representations; nevertheless, it still largely depends on the kind and quality of the input features.

Discussion:

This study shows that feature extraction strategies significantly affect machine learning models' cardiovascular disease detection ability. The Mel-Frequency Cepstral Coefficients (MFCC) model captures crucial audio aspects for medical signal processing due to its excellent performance. As it resembles the human ear's sound perception, MFCC has long been used in speech and audio processing. Important information about feature compatibility with different deep learning architectures is shown by this comparison of model performance across MFCC and MEL-Spectrogram feature extractions. With the suggested model and GCN, which attain noteworthy accuracies (0.92and 0.93, respectively) and robust F1-scores, the MFCC-based strategy continuously produces better performance across all models. Given that MFCC focuses on low-frequency signals, which are frequently essential in audio classification tasks, this implies that MFCC features capture important spectral properties that are advantageous for these models. The appropriateness of MFCC features is further supported by the high accuracy and recall values for both the suggested model and GCN, which show a balanced classification capacity with fewer false positives and false negatives.

The suggested model only achieves an accuracy of 0.55, and additional decreases in precision and recall for GCN and CNN

are shown by the performance seen with MEL-Spectrogram features. According to these measurements, MEL-Spectrogram could add noise or feature complexity that make feature learning more difficult, especially for VGGish and LDA, which perform noticeably worse. The noticeable decline in LDA performance for both feature types supports the idea that LDA may be intrinsically incapable of capturing the high-dimensional representations offered by MFCC and MEL-Spectrogram. The results highlight the significance of choosing feature extractions that are appropriate for the model architecture, since MEL-Spectrogram requires further optimization or possibly different architectures for better performance, whereas MFCC clearly improves feature interpretability for graph-based and convolutional models.

6. CONCLUSION

This work illustrates how the novel use of feature extraction techniques, namely MFCC, may effectively identify cardiovascular illness using the suggested model. With a 92% accuracy rate, the suggested model demonstrated the ability of MFCC to capture the key features of cardiovascular signals, surpassing the performance of more established models such as GCN and CNN. The findings show that the suggested model's ability to successfully identify subtle patterns which are essential for early detection and diagnosis is made possible by the incorporation of MFCC.However, when employing both MFCC and MEL-spectrogram characteristics, the GCN and CNN models performed much worse, indicating their limits in terms of utilizing the rich information included in these audio representations. The results imply that, despite their individual advantages, GCN and CNN are unable to completely take advantage of the subtleties that MFCC has to offer, especially when it comes to cardiovascular health.The effectiveness of the suggested model emphasizes how crucial it is to choose suitable feature extraction methods based on the unique characteristics of the data and the analytical goals. This work not only establishes the importance of MFCC in medical signal processing, but it also paves the way for more research that may enhance model architectures and feature extraction techniques. The encouraging findings motivate more research into sophisticated neural network architectures and hybrid models to identify cardiovascular illness more effectively and maybe improve patient outcomes. All things considered, the suggested model represents a major development in the discipline, opening the door to more successful cardiovascular monitoring and intervention techniques.

REFERENCES

[1] F. Li, Z. Zhang, L. Wang, and W. Liu, ‘Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning’, Front. Physiol., vol. 13, p. 1084420, 2022.

[2] W. Chen, Q. Sun, X. Chen, G. Xie, H. Wu, and C. Xu, ‘Deep learning methods for heart sounds classification: A systematic review’, Entropy, vol. 23, no. 6, pp. 1–18, 2021, doi: 10.3390/e23060667.

[3] T. M. A. Monisha Sharean and G. Johncy, ‘Deep learning models on Heart Disease Estimation - A review’, J. Artif. Intell. Capsul. Networks, vol. 4, no. 2, pp. 122–130, 2022, doi: 10.36548/jaicn.2022.2.004.

[4] A. F. Quiceno-Manrique, J. I. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez, ‘Selection of dynamic features based on time–frequency representations for heart murmur detection from phonocardiographic signals’, Ann. Biomed. Eng., vol. 38, pp. 118–137, 2010.

[5] M. A. Faghy et al., ‘Cardiovascular disease prevention and management in the COVID-19 era and beyond: an international perspective’, Prog. Cardiovasc. Dis., vol. 76, pp. 102–111, 2023.

[6] C. Liu et al., ‘Liu, C., Springer, D., Li, Q., Moody, B., Juan, R. A., Chorro, F. J., et al. (2016). An open access database for the evaluation of heart sound algorithms. Physiol. Meas. 37, 2181–2213. doi:10.1088/0967-3334/37/12/2181’, Physiol. Meas., vol. 37, no. 12, p. 2181, 2016.

[7] D. R. Deepa, V. B. Sadu, C. G. Prashant, and D. A. Sivasamy, ‘Early prediction of cardiovascular disease using machine learning: Unveiling risk factors from health records’, AIP Adv., vol. 14, no. 3, 2024, doi: 10.1063/5.0191990.

[8] A. Harimi et al., ‘Classification of heart sounds using chaogram transform and deep convolutional neural network transfer learning’, Sensors, vol. 22, no. 24, p. 9569, 2022.

[9] I. Maglogiannis, E. Loukis, E. Zafiropoulos, and A. Stasis, ‘Support vectors machine-based identification of heart valve diseases using heart sounds’, Comput. Methods Programs Biomed., vol. 95, no. 1, pp. 47–61, 2009.

[10] G. E. Hinton and R. R. Salakhutdinov, ‘Reducing the dimensionality of data with neural networks’, Science (80-.)., vol. 313, no. 5786, pp. 504–507, 2006.

[11] N. Mei, H. Wang, Y. Zhang, F. Liu, X. Jiang, and S. Wei, ‘Classification of heart sounds based on quality assessment and wavelet scattering transform’, Comput. Biol. Med., vol. 137, p. 104814, 2021.

[12] F. Li, Z. Zhang, L. Wang, and W. Liu, ‘Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning’, Front. Physiol., vol. 13, no. December, pp. 1–16, 2022, doi: 10.3389/fphys.2022.1084420.

[13] H. Lin, K. Chen, Y. Xue, S. Zhong, L. Chen, and M. Ye, ‘Coronary heart disease prediction method fusing

domain-adaptive transfer learning with graph convolutional networks (GCN)', Sci. Rep., vol. 13, no. 1, pp. 1–14, 2023, doi: 10.1038/s41598-023-33124-z.

[14] P. Hamet and J. Tremblay, 'Artificial intelligence in medicine', metabolism, vol. 69, pp. S36–S40, 2017.

[15] K. W. Johnson et al., 'Artificial intelligence in cardiology', J. Am. Coll. Cardiol., vol. 71, no. 23, pp. 2668–2679, 2018.

[16] T. Zhang, Y. Lin, W. He, F. Yuan, Y. Zeng, and S. Zhang, 'GCN-GENE: a novel method for prediction of coronary heart disease-related genes', Comput. Biol. Med., vol. 150, p. 105918, 2022.

[17] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, 'Applications of machine learning predictive models in the chronic disease diagnosis', J. Pers. Med., vol. 10, no. 2, 2020, doi: 10.3390/jpm10020021.

[18] J. Zhang, X. Hu, Z. Jiang, B. Song, W. Quan, and Z. Chen, 'Predicting Disease-related RNA Associations based on Graph Convolutional Attention Network', Proc. - 2019 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2019, pp. 177–182, 2019, doi: 10.1109/BIBM47256.2019.8983191.

[19] H. Li et al., 'A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection', Comput. Biol. Med., vol. 120, p. 103733, 2020.

[20] A. Castro, A. Moukadem, S. Schmidt, A. Dieterlen, and M. T. Coimbra, 'Analysis of the electromechanical activity of the heart from synchronized ECG and PCG signals of subjects under stress', in International Conference on Bio-inspired Systems and Signal Processing, SCITEPRESS, 2015, pp. 49–56.

[21] H. Liang, S. Lukkarinen, and I. Hartimo, 'Heart sound segmentation algorithm based on heart sound envelogram', in Computers in Cardiology 1997, IEEE, 1997, pp. 105–108.

[22] S. Choi and Z. Jiang, 'Comparison of envelope extraction algorithms for cardiac sound signal segmentation', Expert Syst. Appl., vol. 34, no. 2, pp. 1056–1069, 2008.

[23] B. S. Emmanuel, 'A review of signal processing techniques for heart sound analysis in clinical diagnosis', J. Med. Eng. Technol., vol. 36, no. 6, pp. 303–307, 2012.

[24] R. F. Ibarra, M. A. Alonso, S. Villarreal, and C. I. Nieblas, 'A parametric model for heart sounds', in 2015 49th Asilomar Conference on Signals, Systems and Computers, IEEE, 2015, pp. 765–769.

[25] M. B. Er, 'Heart sounds classification using convolutional neural network with 1D-local binary pattern and 1D-local ternary pattern features', Appl. Acoust., vol. 180, p. 108152, 2021.

[26] Z. Ren et al., 'Deep attention-based neural networks for explainable heart sound classification', Mach. Learn. with Appl., vol. 9, p. 100322, 2022.

[27] K. Iqtidar, U. Qamar, S. Aziz, and M. U. Khan, 'Phonocardiogram signal analysis for classification of Coronary Artery Diseases using MFCC and 1D adaptive local ternary patterns', Comput. Biol. Med., vol. 138, p. 104926, 2021.

[28] S. Lahmiri and S. Bekiros, 'Complexity measures of high oscillations in phonocardiogram as biomarkers to distinguish between normal heart sound and pathological murmur', Chaos, Solitons & Fractals, vol. 154, p. 111610, 2022.

[29] T. N. Kipf and M. Welling, 'Semi-supervised classification with graph convolutional networks', arXiv Prepr. arXiv1609.02907, 2016.

[30] A. L. Maas, A. Y. Hannun, and A. Y. Ng, 'Rectifier nonlinearities improve neural network acoustic models', in Proc. icml, Atlanta, GA, 2013, p. 3.