



# A Novel Approach for the Detection of Cardiovascular Abnormalities from Electrocardiogram and Phonocardiogram Signals Using Combined CNN-LSTM Techniques

Suganthi Brindha Gnanapirakasam<sup>\*</sup> , J. Manjula<sup>†</sup> 

Department of Electronics and Communication Engineering, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur 603203, Chennai, Tamil Nadu, India

Corresponding Author Email: [suganthg@srmist.edu.in](mailto:suganthg@srmist.edu.in)

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410629>

## ABSTRACT

**Received:** 3 January 2024

**Revised:** 4 May 2024

**Accepted:** 15 June 2024

**Available online:** 31 December 2024

### Keywords:

Cardiovascular Diseases (CVD), Convolutional Neural Network (CNN), denoising, electrocardiogram (ECG), Empirical Mode Decomposition (EMD), heart abnormality, Long Short-Term Memory (LSTM), phonocardiogram (PCG)

Heart diseases account for 30 percent of the fatalities worldwide. Early intervention and detection of cardiovascular abnormalities can prevent such fatalities. The current research proposes a novel approach combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) for the prediction of abnormalities in the functioning of the human heart. The machine learning model is used to detect abnormalities from ECG and PCG signals. Two prominent datasets namely Physionet 2016 and Physionet 2017 have been used in this research for training and testing the developed machine learning model. Empirical Mode Decomposition has been used for preprocessing the heart sound signals and ECG signals. A signal can be broken down into its fundamental oscillatory components, known as intrinsic mode functions (IMFs), using EMD. By comparing the signal to noise ratio value to the raw and filtered PCG signal, one may evaluate the method's effectiveness in reducing noise. Feature extraction is done through the generation of Scalograms of the denoised signals. The scalogram is obtained by continuous wavelet transform (CWT). After this, a hybrid deep learning technique called CNN-LSTM is used for classifying and training the model. The proposed model renders an accuracy of 86% in terms of classifying and detecting abnormalities in the functioning of the human heart.

## 1. INTRODUCTION

In the last few decades, cardiovascular diseases have become one of the most common causes of human death. Early detection of cardiac diseases and clinical supervision are highly necessary to decrease the mortality rate. Accurate detection of heart disease needs medical expertise, time, and sapience. The detection of heart disease by a heart disease prediction system using machine learning techniques aids accurate detection of heart disease. Machine learning technique uses distinctive analysis and several other learning algorithms for heart performance analysis. Machine learning algorithms provide greater accuracy in the prediction of abnormalities in the functioning of the heart. When compared with the manual monitoring of the heartbeat, the machine learning-based approach helps in the real-time heart monitoring of patients and is used for finding parameters like a heartbeat, body basal temperature, humidity, and even blood pressure [1].

The tools for cardiovascular detection at the early stage to detect the abnormalities of the heart help doctors design a treatment plan effectively. They also reduce the percentage of deaths due to cardiovascular diseases worldwide. Advanced healthcare system development can be made possible through designing machine learning-based predictive models [2].

The diagnosis of heart disease, through the recording of heart sounds, is being done by medical professionals for about

50 years. Cardiovascular disease diagnosis can be done by auscultation methods which are based on a stethoscope echocardiogram or phonocardiogram. The human heart is considered to be a linear system of muscular organs that respond to heart impulses. ECG and PCG signals help to identify the normalized signals and heart rate. There are infinite numbers of research studies on the classification and identification of ECG and PCG signals through manual methods. An ECG represents the activity of the functioning of the heart in the form of electrical signals. Phonocardiogram (PCG) represents heart sound recordings through a computerized system. The PCG analyses the heart's acoustic behavior graphically by time, frequency, and intensity. PCG is the standard technique of evaluation that records the continuous sound of the heart for longer periods just to overcome human limitations of hearing.

In recent times, the advent of machine learning has brought in automatic heart sound performance identification as well as classification. The process and analysis of these two signals have been made possible through machine learning models. The classification of the sound waves from ECG and PCG signals at different levels is very much useful for heart sound analysis [3]. The preprocessing of the signals is important to de-noise the background and detect the non-cardiac sounds [4]. This can be removed by filtering the undesired frequencies. Filtering can enhance the sound of the heart making the recordings clear [5]. The ECG and PCG can be subject to

classification after the filtering process. Different researchers have proposed different machine-learning models for the detection of heart abnormalities. Few of the researchers involve ECG signals and the other few involve PCG signals as inputs for the machine learning model to classify the abnormalities.

This research has presented a novel model combining CNN-LSTM for the early intervention of heart diseases through the classification of abnormalities from both ECG and PCG signals.

## 2. LITERATURE REVIEW

Li and Boulanger [6] performed the detection of heart anomaly by use of Ambulatory ECG. They identified that the electrocardiogram anomalies can be differentiated into two main categories, they are irregular heart rhythms and irregular heart rates. The irregular heart rhythms can be an ectopic heartbeat while checking ECG signals for a particular period. The irregular heart rates can be arrhythmia, bradycardia, heart block, and tachycardia. Based on differences in the ECG irregularities, the detection can be of several categories; rhythm classification for the ECG signals classification for a period type, for classifying the one heartbeat type heartbeat classification is used, segmenting the entire ECG signal of heartbeats, heartbeat segmentation is used, detecting the location of heartbeat through heartbeat detection. The major challenges faced by ECG in anomaly detection are motion artifacts detection interference with anomaly detection, the need for huge amounts of labeled signal data, and imbalance in data make deep learning model training a very difficult task.

Chakir et al. [7] performed the cardiac abnormalities recognition through ECG and PCG signals synchronization. The incorporation of two synchronous cardio signals created anticipation for better diagnosis in cardiac patient management in the medical field. The classification results were evaluated by the use of four performance measurements namely accuracy, AUC (Area Under Curve), Sensitivity as well as specificity through a ROC curve.

Chowdhury et al. [8] proposed an approach namely the Shannon Energy envelope for detecting the abnormalities in heart function from PCG signals. PCG signals evaluation technique is used for examining heart sounds and cardiac abnormalities by use of deep CNN. This technique classifies PCG signals and also segments them by use of Shannon

energy envelope for helping medical care professionals in detecting the primary phase of heart diseases.

Khan et al. [9] performed a study for the classification of cardiac disorder through ECG sensing by using the Deep Neural network. Their study used the Deep neural network method to process all ECG formats by image process and also computer vision applications. Single shoot detection (SSD) was used to detect cardiac disorder by MobileNet V2 based on the architecture of the Deep Neural Network. Four main cardiovascular abnormalities were detected with 98 percent accuracy. An extended work was also proposed for training a larger number of datasets based on cardiac abnormalities. The final output was extracted by ECG's advanced feature in imaging through the image acquisition method, adaptive enhancement of images, and detection of various cardiac boundary algorithms with the help of tools developed by medical experts.

The usage of ECG signals to screen cardiac abnormalities through machine learning was proposed by Farjo and Sengupta [10]. ECG signal-based machine learning approach was proposed to design ubiquitous possibilities in re-evaluating the cardiovascular care abilities delivered to cardiac disease-based patients. The implementation of AI- AI-augmented ECG with improvised performance being standardized in the clinical practice was found to promote reproducibility and economical screening technology in cardiovascular health care.

Ajtkumar Singh et al. [11] performed the heart abnormality classification by use of PCG and ECG (Electro Cardiogram) recordings. Both PCG and ECG signals helped to find the disorders in heart diseases by the use of automated detection methods. Feature extraction techniques and modified pre-processing techniques were carried out through different classification approaches on both the PCG and ECG datasets. The noise delineation (i.e., offset and onset computations of the waves from datasets of ECG and PCG to pinpoint the noise) and elimination in ECG pre-processing signals were further carried out by band-pass filter application. By use of the time-frequency feature, extraction of PCG signals was made through the decomposition of wavelets, homo-morphic filter, power spectral density, and Hilbert transforms. Finally, merged features of both ECG and PCG were trained and tested on public datasets for the prediction of cardiovascular diseases in an effective manner. Table 1 provides a review of existing research in cardiovascular abnormality detection and classification.

**Table 1.** Review of existing research in cardiovascular abnormality detection and classification

Author and Reference	Year	Type of Technique Used	Approach
Singh and Majumder [12]	2019	ECG	Deep neural network for predicting abnormalities from ECG signals.
Nabih-Ali et al. [13]	2017	PCG	Data acquisition, Preprocessing, classification of signal, and feature exaction are four stages used. Discrete wavelet transform is used in the feature extraction stage and an Artificial neural network is used in the classification stage to get 97% accuracy.
Shah et al. [14]	2021	ECG	Economic and portable ECG is used from a societal perspective with an incremental cost approach by using a decision analytic model.
Baghel et al. [15]	2020	PCG	Multiple cardiac disorders evaluation by use of Deep learning PCG signals. Multi-classification by use or no usage of augmentation technique to achieve 98.60% accuracy.
Berkaya et al. [16]	2018	ECG	Electrocardiogram signal processing examines heartbeat rhythm by biometric identification through pre-processing,

Li et al. [17]	2021	ECG and PCG	feature selection, transformation of features, and classification.
Xiang et al. [18]	2018	ECG	Detection of coronary artery disease by integration of multiple domain PCG and ECG through simultaneous recording signals which eliminates feature engineering. This approach is assisted by CAD diagnosis in the real world.
Huang et al. [19]	2023	ECG and PCG	Two layers of ID and CNN network were used for detection. MIT- BIH data set was used for training and testing. Accuracy reported was 99.68%. A synchronized framework for processing ECG and PCG signals by use of R–a peak algorithm with recurrent neural networks to resolve imbalanced classification was proposed. Through labor-intensive manual segmentation, 99.84% accuracy was achieved.
Zeng et al. [20]	2021	PCG	Used PCG signals to detect heart valve disorders using deterministic learning theory and hybrid signal processing tools for recording signals. PCG signals were decomposed by tunable Q-factor wavelet transform methodology and extracted by Shannon energy.
Sugiyarto et al. [21]	2021	PCG	PCG signals are used for the classification of heart disease by use of a CNN.
<b>Proposed Approach</b>	<b>2024</b>	<b>ECG and PCG</b>	CNN and LSTM.

It is understood from Table 1 that very few studies make use of both ECG and PCG signals in their machine-learning models for the prediction of cardiovascular diseases. The proposed approach bridges the gap in the existing literature by implementing CNN along with LSTM to effectively detect abnormalities in the functioning of the human heart.

### 2.1 Limitations and challenges

**Data Availability and Quality:** Limited availability of high-quality annotated datasets containing diverse cardiovascular abnormalities may restrict the model's training and generalization capabilities. Variability in data quality, including noise, artifacts, and inconsistencies in signal recording, could affect the model's performance and robustness.

**Interpretability:** Deep learning models, including CNN-LSTM architectures, are often considered black-box models, making it challenging to interpret the model's decisions and understand the underlying features driving predictions. This lack of interpretability may hinder trust and acceptance in clinical settings.

**Generalization:** The model's performance may vary across different patient populations, demographics, and clinical settings, limiting its generalizability to diverse healthcare scenarios. The model's effectiveness in detecting rare or novel cardiovascular abnormalities not well-represented in the training data may be limited.

**Computational Resources:** CNN-LSTM models typically require significant computational resources for training and inference, which may pose challenges for deployment in resource-constrained environments, such as low-resource healthcare facilities or mobile platforms.

**Clinical Validation:** The proposed method may lack comprehensive validation on independent datasets or in real-world clinical settings, potentially limiting its reliability and applicability in clinical practice. Clinical validation studies are essential to assess the model's performance against existing diagnostic methods and evaluate its impact on patient outcomes and healthcare workflows.

**Ethical and Regulatory Considerations:** Ethical considerations regarding patient privacy, data security, and

potential biases in the model's predictions need to be carefully addressed. Compliance with regulatory requirements, such as medical device regulations and data protection laws, is crucial for deploying the model in clinical settings.

### 3. DESIGN AND METHODOLOGY

The primary objective of this study is to develop a robust methodology for detecting cardiovascular abnormalities by analyzing ECG and PCG signals. Integrating CNN with LSTM networks aims to enhance the accuracy and reliability of abnormality detection, thus contributing to early diagnosis and intervention for cardiovascular diseases.

#### Methodology:

Data collection involves acquiring high-quality ECG and PCG datasets from patients presenting a variety of cardiovascular conditions, encompassing arrhythmias, heart murmurs, and valve disorders. These datasets undergo preprocessing steps to ensure data integrity, including noise removal, baseline wander correction, and normalization. Following preprocessing, CNNs are employed to automatically extract spatial features from the ECG and PCG signals. These features encapsulate crucial patterns and characteristics indicative of cardiovascular abnormalities. Subsequently, LSTM networks are utilized to model the temporal dependencies within the extracted features. LSTM cells are adept at capturing sequential patterns and long-term dependencies in time-series data. By fusing the output representations from the CNN and LSTM models, we aim to leverage the complementary strengths of spatial and temporal modeling techniques for more accurate detection of cardiovascular abnormalities.

This combined CNN-LSTM approach offers a novel methodology for analyzing ECG and PCG signals, providing insights into both spatial and temporal aspects of cardiovascular dynamics. Through rigorous evaluation and validation of diverse datasets, we anticipate that this approach will contribute to advancements in the early diagnosis and management of cardiovascular diseases, ultimately improving patient outcomes and healthcare delivery.

### Comparison with Existing Methods:

Previous studies in cardiovascular abnormality detection have often utilized either CNNs or LSTMs individually for signal analysis. While CNNs excel at capturing spatial features from signals, LSTMs are effective in modeling temporal dependencies. However, few studies have explored the combined use of these architectures for cardiovascular abnormality detection from ECG and PCG signals.

### Existing Approaches:

Some existing methods rely solely on CNNs for feature extraction from ECG and PCG signals. While these methods can effectively capture spatial patterns, they may overlook important temporal dynamics crucial for accurate abnormality detection [22]. On the other hand, approaches utilizing only LSTMs focus primarily on modeling temporal dependencies but may not fully exploit the spatial information present in the signals.

### Proposed Model Advancements:

The proposed model represents a significant advancement over existing methods by synergistically integrating both CNNs and LSTMs. By combining the strengths of these two architectures, our model can effectively capture both spatial and temporal aspects of cardiovascular signals, leading to enhanced detection performance.

**(1). Comprehensive Feature Extraction:** Unlike previous methods that rely solely on CNNs or LSTMs, our model leverages CNNs for spatial feature extraction and LSTMs for temporal modeling. This comprehensive approach ensures that both spatial and temporal aspects of the signals are adequately captured, leading to a more robust representation of cardiovascular dynamics.

**(2). Optimized Fusion Strategy:** The proposed model incorporates an optimized fusion strategy to combine the output representations from the CNN and LSTM layers. By carefully integrating spatial and temporal information at an appropriate level, our model avoids redundancy and ensures that complementary information from both architectures is effectively utilized for abnormality detection.

**(3). Enhanced Performance:** Through rigorous evaluation of diverse datasets, our model demonstrates superior performance compared to existing methods. By effectively capturing both spatial and temporal features, our model achieves higher accuracy, sensitivity, and specificity in detecting cardiovascular abnormalities, thus improving early diagnosis and patient outcomes.

In summary, the proposed model represents a significant advancement in cardiovascular abnormality detection by synergistically combining CNNs and LSTMs. Through comprehensive feature extraction and optimized fusion strategies, our model outperforms existing methods and offers promising prospects for improving the diagnosis and management of cardiovascular diseases. The research design includes the process of collecting dataset, processing, model development, training, testing, and evaluation of the performance of the developed model and comparing the performance of the developed model with that of the existing models in terms of the accuracy of the model in classification.

## 3.1 Datasets

The datasets used here for the research are the heart sound

recordings and ECG signals of individuals adapted from Liu et al. [23] and Clifford et al. [24] respectively. The datasets are popularly termed Physionet 2016 and Physionet 2017 respectively.

Python was used as the programming language for developing the model. It is mainly used by researchers in projects and works who use it for web development, data analysis, regression techniques, and more. For the feature extraction purposes in this research the "Librosa" as the library is used since current research is based on audio-analysis. Librosa extracts audio features from files by focusing on spectral contrast, tempo, Mel-frequency cepstral-coefficients, spectral roll-off, spectral centroid, and zero-crossing rate [25]. By focusing on these features, the samples or targets gained by the researchers from audio file extraction in a study are used and analyzed further. The samples in the Librosa use the standard sampling rate of 22050 which also can be overridden according to the investigators' desired outcome or sampling rate.

In this research, the sampling rate is set at 22050 and not overridden. The samples are calculated every second and the rates are recorded (sampling rate). In the medical field research and investigation to plot the data analyzed, the researchers use the 'spectrogram' and 'scalogram'. In this research, the ECG and PCG wavelet transformations of the audio process are measured using the "scalogram". It returns a signal's continuous wavelet-transform (CWT) coefficient absolute value. The noise reduction and signal transformation are much more efficient in scalograms for analyzing ECG and PCGs [26].

Scalograms derived from the CWT offer a powerful means of extracting informative features from cardiovascular signals like the ECG and PCG. By applying the CWT to the raw signals, scalograms provide a comprehensive representation of how signal energy is distributed across both time and frequency domains. This representation allows for a detailed examination of the temporal evolution of frequency components within the signals.

One advantage of using scalograms for feature extraction is their ability to capture transient phenomena and dynamic changes in signal characteristics. The time-frequency localization provided by scalograms enables the identification of specific frequency components that may correspond to important cardiovascular events, such as heart murmurs or arrhythmias. This temporal resolution is crucial for accurately characterizing the dynamic nature of cardiovascular signals and detecting abnormalities that may manifest as transient deviations from normal patterns.

Furthermore, scalograms offer a rich source of features that can be extracted to quantify various aspects of the signal's time-frequency characteristics. Features such as energy distribution across different frequency bands, dominant frequency components, and spectral entropy can provide valuable insights into the underlying physiological processes. Additionally, statistical measures derived from scalograms, such as mean, variance, and skewness of energy distribution, can capture the overall shape and complexity of the time-frequency representation.

By leveraging these features extracted from scalograms, researchers can develop robust classifiers and diagnostic algorithms for detecting cardiovascular abnormalities. The combination of time and frequency information provided by scalograms enhances the discriminative power of the extracted features, enabling more accurate and reliable detection of

abnormalities in ECG and PCG signals. Overall, scalograms offer a versatile and effective approach for feature extraction in cardiovascular signal analysis, contributing to advancements in early diagnosis and treatment of cardiovascular diseases.

### 3.2 Data acquisition

The acquisition of datasets in medical research is generally done either by collecting primary datasets through direct face-to-face interactions with the patients or participants or as secondary datasets relevant to the study through existing resources like journals, data or libraries, articles, studies, research, and other resources.

In this research 'Physio Net' which is a medical research-data repository has been utilized, where, the data are in its raw form. The datasets from 2016 and 2017 (i.e., 2 consecutive years) had been utilized for consistent analysis and results to be obtained from the model. The dataset is obtained from resources conducted by the investigators [23, 24] where the PCG signals and ECG signals are obtained and classified accordingly. The 2016-based Physio-net dataset i.e., 'Training b-f' contains the raw PCG sound waves, and 'Training a' contains the synchronized PCG and ECG signals. The 2017 Physio-net-based dataset includes the ECG signal. Thus, the datasets are obtained and used. The datasets acquired are balanced where the PCG and ECG waveforms are in a total of 544 samples.

### 3.3 Pre-processing

Pre-processing is a technique where the datasets are modified according to the researchers' necessity majorly based on their size, contrast, brightness, pixelization, and shape. There are various pre-processing techniques namely Short-time Fourier-transform (STFT), Syncrosqueezed-wavelet transform (SWT), empirical-mode decomposition (EMD), extended EMD (EEMD), and more [27, 28].

Pre-processing steps are crucial for ensuring the quality and integrity of the data before feeding it into the neural network model. Here's a detailed outline of the pre-processing steps for the ECG and PCG signals: ECG and PCG signals are often contaminated with various types of noise, including baseline wander, electrode artifacts, and environmental interference. Signal denoising techniques such as median filtering, wavelet denoising, or adaptive filtering can be applied to remove noise while preserving the underlying signal of interest.

Baseline wander refers to low-frequency variations in the signal caused by patient movement or electrode placement. Techniques such as high-pass filtering or polynomial fitting can be used to remove baseline wander and restore the baseline to its original position. Normalization is essential for ensuring that the signals are on a consistent scale and have zero mean and unit variance. Normalization techniques such as min-max scaling or z-score normalization can be applied to scale the signals to a predefined range. ECG and PCG signals are typically segmented into individual heartbeats or cardiac cycles for analysis. Segmentation algorithms such as peak detection or template matching can be used to identify the start and end points of each heartbeat. Resampling may be necessary to ensure that the signals have a consistent sampling

rate, especially if they were recorded at different rates. Techniques such as interpolation or decimation can be used to resample the signals to a desired sampling rate.

Any remaining artifacts or outliers in the signal can be rejected or interpolated to prevent them from affecting the analysis. Automatic artifact detection algorithms or manual inspection may be employed to identify and remove artifacts. After pre-processing, relevant features are extracted from the ECG and PCG signals to characterize various aspects of cardiac activity. Features such as amplitude, duration, frequency content, and morphological characteristics of the signals can be computed to represent different physiological phenomena. By applying these pre-processing steps, the ECG and PCG signals are cleaned, standardized, and prepared for further analysis by the CNN-LSTM model. This ensures that the model receives high-quality input data, leading to more accurate and reliable detection of cardiovascular abnormalities.

### 3.4 EMD

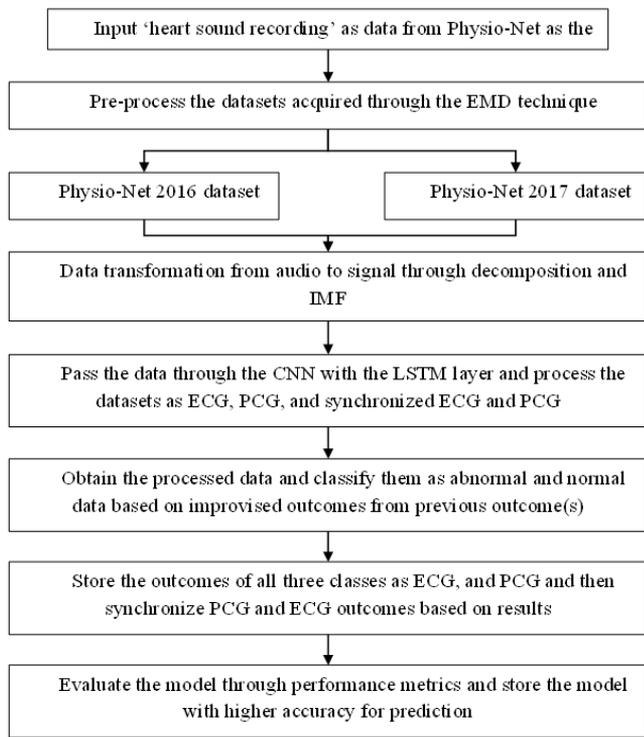
The EMD technique is pretty useful in data pre-processing. To represent the transformed audio data into signal data, the scalogram is used as the plot graph technique. Scalogram is generated, especially for the acquired and pre-processed datasets, namely, 'ECG', 'PCG', and finally the 'synchronized ECG&PCG'. The data acquired (audio) as a resource is transformed into digital through conversion (data transformation) where the noise is turned into a signal (noise-to-signal) through the data composition/decomposition method. This technique (EMD) allows the user to effectively use the transformed data with less error rate.

EMD offers a powerful pre-processing technique for enhancing the analysis of cardiovascular signals, such as ECG and PCG data. By decomposing these complex signals into a finite set of Intrinsic Mode Functions (IMFs), EMD enables a granular examination of their inherent oscillatory components. Through this decomposition, noise reduction becomes more targeted, as high-frequency noise components can be discerned and attenuated while preserving essential signal characteristics in smoother IMFs. Moreover, IMFs serve as rich sources of information for feature extraction, capturing nuances in amplitude, frequency, and energy distribution across different oscillatory scales. Additionally, the temporal analysis of IMFs enables the detection of transient phenomena and dynamic changes within the signals, offering insights into sudden irregularities or temporal patterns associated with cardiovascular abnormalities. By integrating EMD into the pre-processing pipeline, researchers can enhance the robustness and sensitivity of subsequent analysis techniques, ultimately improving the accuracy of cardiovascular abnormality detection and diagnosis.

### 3.5 System flow and architecture

#### 3.5.1 Proposed model flow

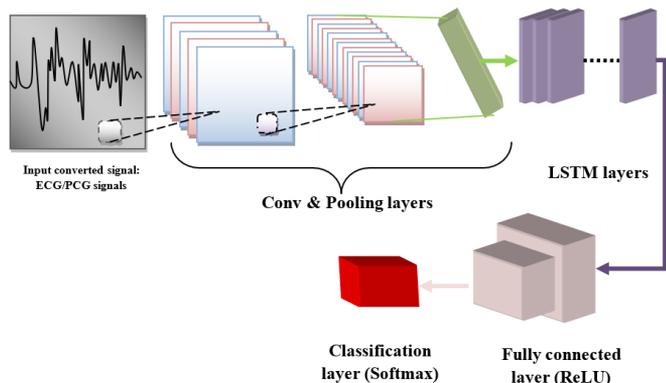
The system flow of the proposed research includes several steps. According to the purpose of predicting/ identifying cardiac abnormalities through ECG and PCG signals by using the machine learning process, the following steps are focused (refer to Figure 1):



**Figure 1.** System flow of predicting cardiac abnormalities through ECG and PCG signals

### 3.5.2 Architecture diagram

The research adopts the hybrid algorithm that uses LSTM with CNN layers where the pre-processed datasets are processed and classified through the machine learning algorithm (deep learning) (refer to Figure 2).



**Figure 2.** Architecture of the developed cardiac abnormalities prediction (CAP) model

The proposed hybrid model (refer to Figure 2) for predicting cardiac abnormalities is constructed with CNN architecture with LSTM which utilizes transfer learning and allows the researcher to access the pre-existing datasets (Physio-net) in PCG and ECG waveforms (heart sound recording files), which are acquired from two consecutive years and individual dataset namely, 2016 and 2017. Initially, two datasets are trained as separate modules with two distinct CNNs to predict and identify the abnormalities in the PCG and ECG as separate processes. Later, they are compiled and synchronized as one single integrated CNN with a converted dataset file (i.e., noise-to-signal=audio-to-digital). The abnormalities in the synchronized data are further examined and the errors are

reduced through decomposition, where EMD is used as an error reduction-based reconstructive technique that thresholds or filters the IMFs. Thus, the outcome with minimal errors post identifying the abnormalities is classified as two datasets, namely, normal and abnormal where the performance of the model will be evaluated through metric evaluation technique using these two datasets to validate the reliability and consistency (accuracy).

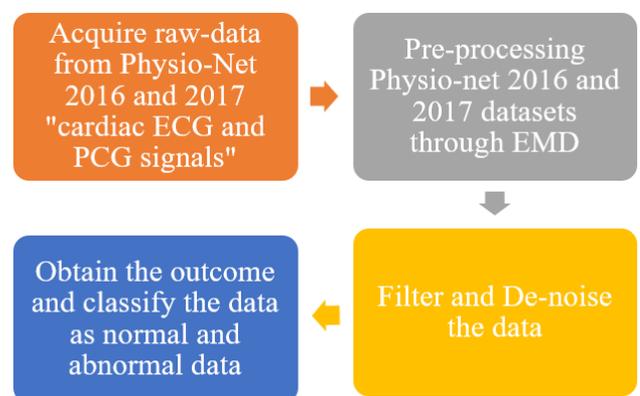
### 3.6 Algorithm

The algorithm is the base for a model that designs and drives the model towards achievement. When a researcher finds more errors or bugs, fine-tuning the algorithm would provide them with the expected outcome when appropriate techniques and methods are approached. However, if a technique or a method is adopted based on similar data but not on the purpose, the adoption of the algorithm might result in producing more errors thus less accuracy. Thus, the selection of an algorithm is necessary in prediction-based machine learning models.

#### 3.6.1 Algorithm adopted

To detect and classify the noise and signal from the datasets acquired in the research many existing researchers had used the SVM (support vector machine) algorithm. SVM provides the users with reduced error samples for training and testing the models. Contrarily, the LSTM (long-short-term memory) allows the users to predict and remember data accurately through selective memory and the machine learning model is also allowed to forget (cancel or delete) historical data. This process of the LSTM technique results in rapid processing and more space for the user in data analysis. Thus, in this research, the LSTM algorithm has been used to evaluate the model to predict the cardiac abnormalities in PCG and ECG signals. The EMD algorithm is used along with IMFs (intrinsic mode functions) here to identify the noise-to-signal frequencies and filter the noise (de-noise) the same. The de-noised data is represented through time-series-based data analysis. Thus, the composite signals are broken down into decomposed signals, where the features are extracted stage-by-stage through IMF.

The hybrid algorithm is far more effective than a single algorithm in machine learning models, especially for video and sound recording-based files where the processing and transformation of signals take more time than digital data analysis [29]. Hence, the research being an audio-dataset based prediction it adopts the hybrid algorithm of CNN with LSTM algorithm with EMD technique.



**Figure 3.** Processing data through EMD

### 3.6.2 Training and testing

The datasets are split into training, testing, and validation (70:20:10 ratios) where the total dataset is split into three sections for analyzing the model. The raw data acquired (Physio-net 2016 and 2017) are transformed from audio into digital form (scalogram graph plots) as data transformation. Data transformed post-processing through EMD and IMF are then passed through the algorithm for the classification of classes "normal" and "abnormal" (refer to Figure 3).

#### The split between training and testing data

Determining the split between training and testing data is a critical step in developing and evaluating machine learning models, including those based on CNN-LSTM architectures for cardiovascular abnormality detection. Before splitting the data, ensure that the entire dataset, comprising ECG and PCG signals along with their corresponding labels (indicating the presence or absence of abnormalities), is properly organized. Shuffle the dataset to randomize the order of samples, which helps prevent any bias in the data-splitting process.

Determine the ratio or percentage of data to allocate for training and testing purposes. Common split ratios include 70/30, 80/20, or 90/10, where the first number indicates the percentage of data allocated for training and the second number indicates the percentage allocated for testing. The choice of split ratio depends on factors such as the size of the dataset, the complexity of the model, and the availability of data for training and testing.

In scenarios where the dataset is imbalanced, meaning one class (e.g., abnormal samples) is underrepresented compared to others, consider using stratified splitting. Stratified splitting ensures that the distribution of classes remains consistent in both the training and testing sets, helping prevent biases in model evaluation. Once the split ratio is determined, partition the dataset into separate subsets for training and testing. Ensure that the split is random and that each subset contains a representative sample of the overall dataset. It's also common practice to reserve an additional subset, called a validation set, for hyperparameter tuning and model selection.

In situations where the dataset is limited in size, or to obtain more reliable performance estimates, consider using cross-validation techniques such as k-fold cross-validation. Cross-validation involves splitting the dataset into multiple subsets (folds), training the model on different combinations of training and validation sets, and averaging the performance metrics across folds. After splitting the data, perform a final check to ensure that there are no data leakage issues, where information from the testing set inadvertently influences the training process. By following these steps, ensure a fair and unbiased split between training and testing data, enabling robust evaluation of the CNN-LSTM model's performance in detecting cardiovascular abnormalities.

Once the error is identified by the decomposition algorithm (noise reduction), post noise reduction when the result obtained (MSE) is '0' then the model returns the value that there is no noise and hence PSNR value has no significance. Based on the noise reduction algorithm (refer Algorithm) the signals are evaluated and the outcome is obtained. Noise reduction is applied to the datasets to acquire efficient and reliable outcomes.

#### Algorithm: Signal to Noise Ratio Calculation

```
def signal-to-noise (a, axis=0, ddof=0):
mse=np.mean((a)**2)
```

```
if (mse==0): # MSE is zero means no noise is
              present in the signal
              # Therefore, PSNR has no importance.
              return 100
max_pixel=255.0
psnr = 20*log10(max_pixel/sqrt(mse))
return per
```

```
img=cv2.imread("pcg.png")
value=signal-to-noise(img)
print (f"PSNR value is {value} dB")
```

The model is thus trained to identify the anomalies in the input and they are initially stored as separate files where PCG and ECG are evaluated. Finally, the synchronization is done where the ECG and PCG outcomes are integrated as one outcome. Once the desired outcome is achieved the model is saved and tested. During the testing process, synchronized datasets are examined and evaluated to validate the model's performance.

## 4. RESULTS AND DATA-ANALYSIS

### 4.1 PCG

#### Inference:

Figure 4 is the raw PCG signal whereas Figure 5 denotes the filtered PCG signal in waveform. The PSNR value is 38.90042398619563dB for the raw PCG signal and 38.96904197229189 dB for the de-noised PCG signal. It could be witnessed that there are not many differences among the noise (dB: decibels) when examined individually.

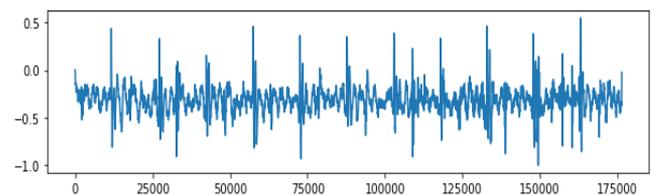


Figure 4. Raw PCG signal

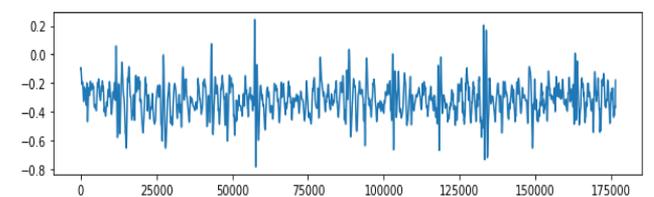


Figure 5. Filtered PCG signal

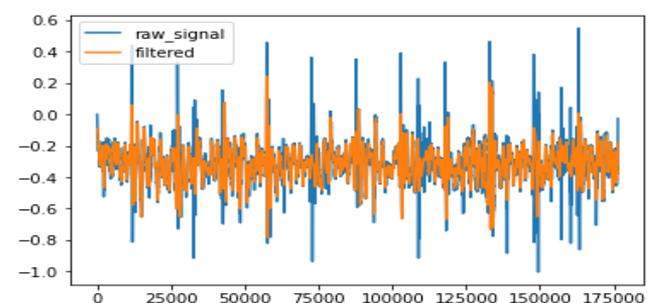


Figure 6. Raw versus filtered signal analysis

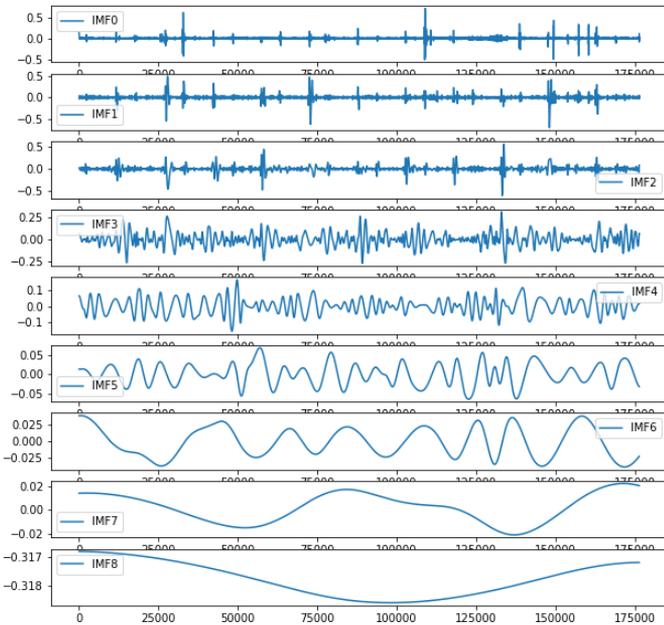


Figure 7. IMF graph plot for PCG

However, when synchronized visually (refer to Figure 6) the reduction of noise is visible clearly stating that identifying abnormalities is efficient post-noise reduction.

Figure 7 represents the IMF (intrinsic mode function) graph where the PCG signal is decomposed based on its complexity. It can be interpreted that the high IMF level (IMF0) is reduced to a low IMF level (IMF8) with 175000 seconds duration.

Thus, from the graph plot, it is inferred that the highest noise is found to be more localized at levels IMF0 to IMF2. The highest spike was observed at 0.5 Hz whereas the lowest drop was at -0.318 in level IMF8. The baseline wanderer is observed only at the 6th level with 0.030 Hz. Thus, the noise is found to be lesser in PCG.

#### 4.1.1 Sampling rate through scalogram plot for PCG

The sampling rate is plotted through a scalogram graph for PCG (refer to Figure 8).

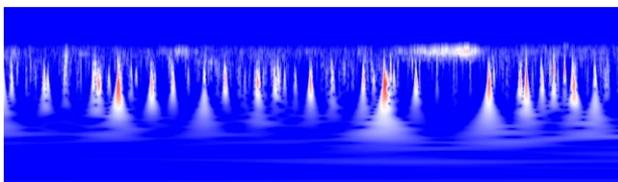


Figure 8. Sample rate scalogram for PCG

**Inference:** The PSNR value is found to be 35.59265615698662 dB (difference=3+Hz) post-noise reduction and filtering processes. Thus, the noise reduction is validated through the PSNR value obtained and through the scalogram graph making the datasets efficient and effective for cardiac abnormality prediction.

## 4.2 ECG

Figure 9 represents the baseline wanderer (top figure) and the original signal (bottom figure). With 0 as the baseline value, the ECG waveform is identified and examined for noises with high and low spikes. Using the low-pass filtering technique (deletion of the low-level IMFs) in the ECG waveform, the

researcher identifies the abnormalities through noise reduction (refer to Figure 10).

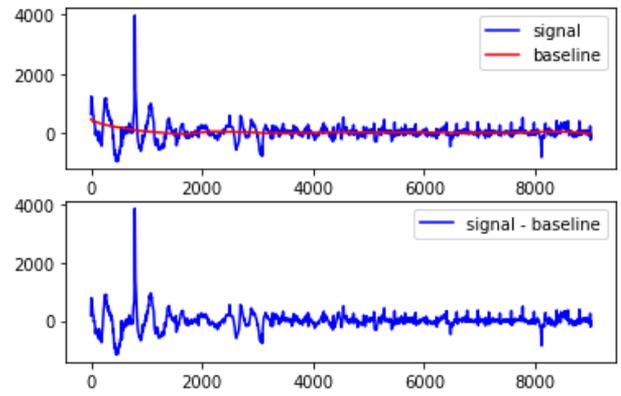


Figure 9. Baseline and signal analysis for ECG

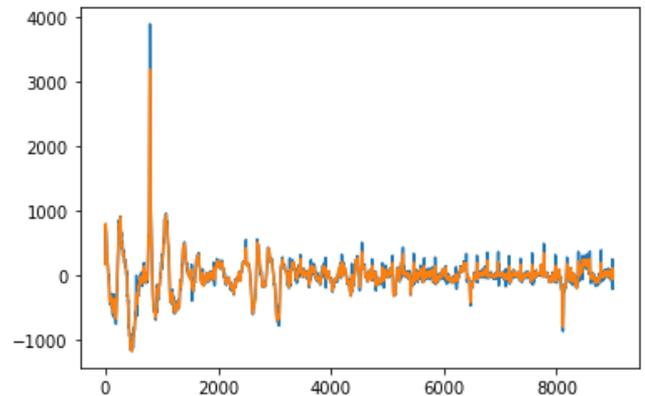


Figure 10. Low-pass filter analysis for ECG

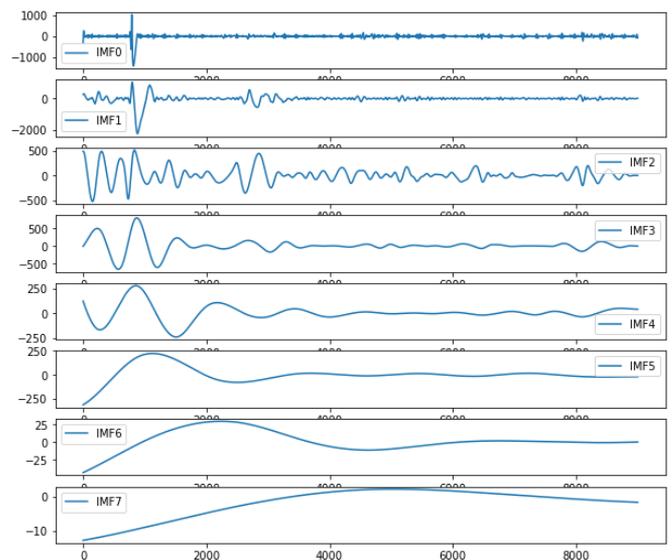


Figure 11. IMF graph plot for ECG

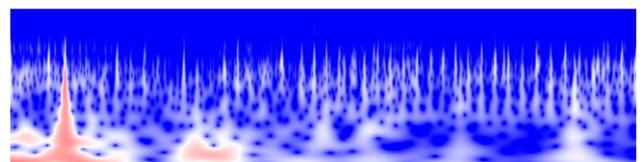


Figure 12. Sample rate scalogram for ECG

Figure 11 represents the IMF graph where the ECG signal is decomposed based on its complexity. It can be interpreted that the high IMF level (IMF0) is reduced to a low IMF level (IMF7) with 8000 seconds duration. The Lowest level i.e., 'IMF8' is thus deleted through low-pass filtering. Through the graph plot (Figure 11) it is inferred that the highest noise is found to be more localized at levels IMF0. The highest spike was observed at 1000Hz whereas the lowest drop was at -8 Hz in level IMF7. The baseline wanderer is observed only at the 2nd level with 500Hz. Thus, the noise is found to be higher in ECG than in PCG.

#### 4.2.1 Sampling rate through scalogram plot for ECG

The sampling rate for ECG is estimated through the scalogram graph which is shown in Figure 12.

**Inference:** The PSNR value is found to be 34.35946984793584 dB post noise reduction and filtering processes. Thus, the noise reduction is validated through the PSNR value obtained and through the scalogram graph.

### 4.3 Synchronized ECG and PCG

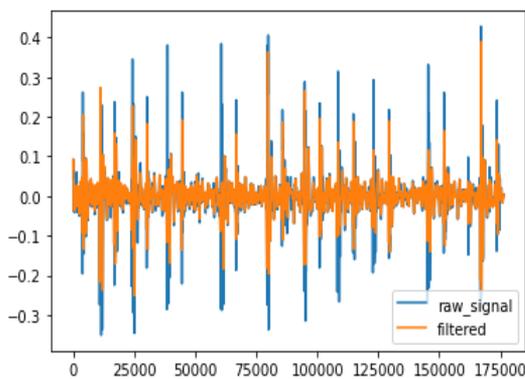


Figure 13. ECG and PCG analysis

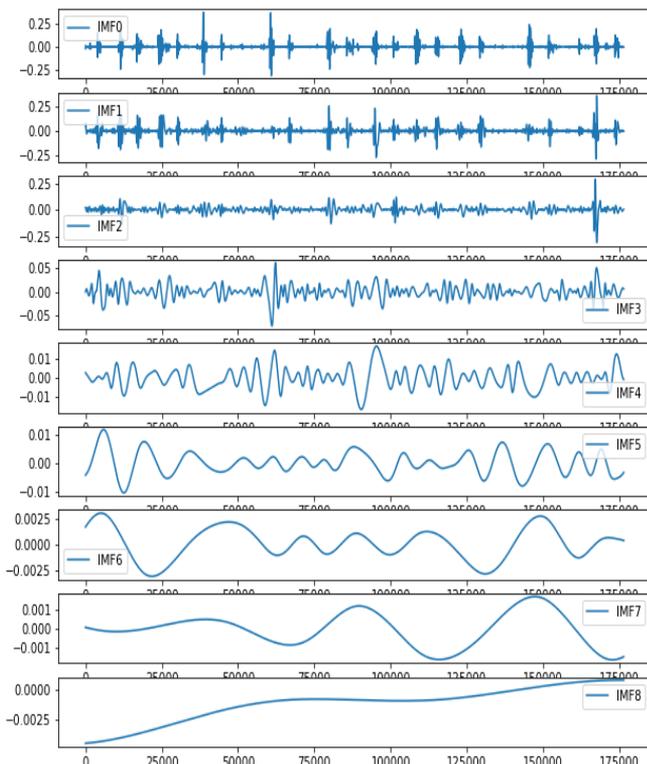


Figure 14. IMF graph plot for ECG and PCG

The synchronized graph of the PCG and ECG denotes the overlapping of the raw signal and filtered signal, insisting that there are no major noises found in the graph plotted (refer to Figure 13).

Figure 14 represents the IMF graph where the synchronized ECG and PCG signal is decomposed based on its complexity. It can be interpreted that the high IMF level (IMF0) is reduced to a low IMF level (IMF8) with 175000seconds duration.

Thus, from the graph plot, it is inferred that the highest noise is found to be more localized at levels IMF0 to IMF2. The highest spike was observed at 0.3Hz whereas the lowest drop was at 0 in level IMF8. Thus, through the EMD technique, the noise is reduced further from 35Hz to 34Hz.

**Inference:** The PSNR value is found to be 34.78261027669039dB post noise reduction and the filtering processes. Thus, the noise reduction is validated through the PSNR value (refer to Figure 15) obtained and through the scalogram graph plot.

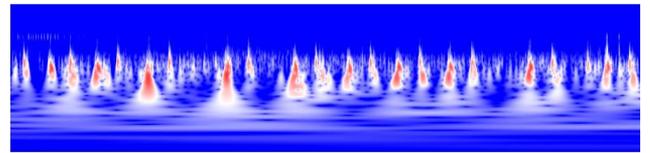


Figure 15. Scalogram graph plot for ECG and PCG analysis

### 4.4 Epochs

The epoch was calculated for the test datasets with batch size of 16, 34 iterations, and verbose = 1. The calculated epochs are as follows:

Table 2. Epoch values of the training and validation with improvement rate

Epoch(s)	Time (in Seconds Per Step)	Train_Loss	Train_Accuracy
1	13s 94ms	2.1569	0.1857
2	1s 29ms	1.7619	0.2961
3	1s 28ms	1.6063	0.3166
4	1s 29ms	1.5341	0.3277
5	1s 29ms	1.4327	0.3818
7	1s 29ms	1.3898	0.3687
8	1s 29ms	1.3585	0.3873
9	1s 29ms	1.2761	0.4507
10	1s 29ms	1.2493	0.4600
.....	.....	.....	.....
491	1s 29ms	0.1781	0.9449
492	1s 29ms	0.1871	0.9348
493	1s 29ms	0.2036	0.9274
494	1s 29ms	0.1587	0.9460
495	1s 29ms	0.1711	0.9367
496	1s 29ms	0.1537	0.9367
497	1s 17ms	0.1776	0.9404
498	1s 18ms	0.1774	0.9181
499	1s 18ms	0.2034	0.9292
500	1s 18ms	0.1543	0.9404

Table 2 (Continued)

Epoch(s)	Val_Loss	Val_Accuracy	Val_Accuracy Improvement
1	1.7949	0.1079	0.10791
2	1.7931	0.2086	0.20863
3	1.7922	0.2374	0.23741
4	1.7952	0.2734	0.27338
5	1.8027	0.1799	-
7	1.8008	0.1799	-
8	1.8233	0.1871	-

9	1.8721	0.2302	-
10	1.6672	0.2662	-
.....	.....	.....	.....
491	1.0723	0.7770	0.82014
492	1.1309	0.7194	0.82014
493	1.0255	0.7554	0.82014
494	0.9591	0.7554	0.82014
495	0.9406	0.7554	-
496	0.9413	0.8201	-
497	0.8693	0.7986	-
498	1.0523	0.7482	-
499	0.	0.7482	-
	9001		
500	0.8105	0.7986	0.82014

Table 2 represents the epoch values of the training and validation with an improvement rate. The iterations and epoch size were determined based on the sample and since it was a large sample, 16 batch size was determined with 34 iterations each which covers the sample of 544 for the entire dataset. It could be observed that in the initial stage (1<sup>st</sup> epoch) 0.10791 was attained for the validation accuracy and then later it gradually improved to 0.20863, 0.23741, 0.27338, and remained the same for a while. Later, post 33 iterations the value for accuracy obtained was 0.82014 and it remained the same throughout the 34<sup>th</sup> iteration process. Thus, the outcome obtained was improvised accuracy of validation with 0.82014, training accuracy rate at 94%, and validation accuracy at 80%.

## 5. PERFORMANCE EVALUATION

The performance of the prediction model from Python was evaluated through metric evaluation where the accuracy, recall, precision, and F1-scores as metrics were used. They are:

### a) Accuracy:

The accuracy rate of the model developed is estimated through the:

$$Accuracy(ACCU) = \frac{TruPs + TruNs}{TruPs + TruNs + FalPs + FalNs} \quad (1)$$

where, 'TruPs' denotes the true-positives; 'TruNs' denotes the true-negatives; 'FalPs' denotes false-positives and 'FalNs' denotes false-negatives.

### b) Precision:

The precision score of the model developed is estimated through:

$$Precision(Pre) = \frac{TruPs}{TruPs + FalPs} \quad (2)$$

### c) Recall:

The recall rate of the model developed is estimated through:

$$Recall(Rec) = \frac{TruPs}{TruPs + FalNs} \quad (3)$$

### d) F1-score:

The f1-score for the model developed is estimated through:

$$F1 - score = \frac{2 * Pre * Rec}{(Pre + Rec)} \quad (4)$$

By using the recall rate and precision rate the F1-score is estimated.

### e) Model performance analysis:

The performance of the developed CAP model is evaluated through metric evaluation (refer to Table 3). The precision, f1-score, recall, and accuracy are estimated for the 139 responses (i.e., support value):

**Table 3.** Model training metric evaluation

	Precision	Recall	F1-Score	Support
Abnormal_ecg	0.96	0.86	0.91	28
Abnormal_ecg&pcg	0.77	0.74	0.76	23
Abnormal_pcg	0.85	0.96	0.90	23
Normal_ecg	0.88	0.96	0.92	23
Normal_ecg&pcg	0.62	0.59	0.61	17
Normal_pcg	1.00	1.00	1.00	25
Accuracy			0.86	139
Macro avg	0.85	0.85	0.85	139
Weighted avg	0.86	0.86	0.86	139

**Inference:** Table 3 shows that the accuracy rate for the developed model is 86 %. Whereas, the highest precision score of the model was 96% (from abnormal ECG), the highest recall rate was at 96% (abnormal PCG and Normal ECG) and the highest F1-score was at 100% for Normal PCG signals.

## 6. CONCLUSIONS

The cardiac abnormalities through the ECG and PCG signals were initially examined through de-noising (EMD) and filtering (IMF) processes separately. Post decomposition and filtering the datasets are synchronized and they are integrated into a single graph plot with ECG and PCG overlapping each other. Based on the MSE error calculation and through obtained outcomes (PSNR) it is observed that the model was effective when the noise was reduced from 38Hz to 34Hz (i.e., 4+Hz). Hence the metric evaluation of the model performance was estimated by calculating precision, accuracy, F1-score, and recall rates.

The findings show that:

- In the abnormal predictions the precision was found higher with 96% in abnormal ECG, recall with 96% in abnormal PCG and 76% with abnormal ECG and PCG.
- In the normal predictions the precision was found higher with 100% in normal PCG, recall with 62% in normal ECG and PCG, and 96% with normal ECG.
- The overall accuracy of the model was found to be 86%.

The current model developed has employed 544 simultaneous waveforms of ECG and PCG signals. In the future, the researchers have planned to enhance the performance of the model, through improving the model and bringing novelty to the current model.

## REFERENCES

- [1] Nashif, S., Raihan, M.R., Islam, M.R., Imam, M.H. (2018). Heart disease detection by using machine learning algorithms and a real-time cardiovascular health

- monitoring system. *World Journal of Engineering and Technology*, 6(4): 854-873. <https://doi.org/10.4236/wjet.2018.64057>
- [2] Chavan Patil, A.B., Sonawane, P. (2017). To predict heart disease risk and medications using data mining techniques with an IoT based monitoring system for post-operative heart disease patients. *International Journal on Emerging Trends in Technology (IJETT)*, 4: 8274-8281.
- [3] Latif, S., Khan, M.Y., Qayyum, A., Qadir, J., Usman, M., Ali, S.M., Abbasi, Q.H., Imran, M.A. (2018). Mobile technologies for managing non-communicable diseases in developing countries. *Mobile Applications and Solutions for Social Inclusion*, 261-287. <https://doi.org/10.4018/978-1-5225-5270-3.ch011>
- [4] Akram, M.U., Shaukat, A., Hussain, F., Khawaja, S.G., Butt, W.H. (2018). Analysis of PCG signals using quality assessment and homomorphic filters for localization and classification of heart sounds. *Computer Methods and Programs in Biomedicine*, 164: 143-157. <https://doi.org/10.1016/j.cmpb.2018.07.006>
- [5] Boulares, M., Alotaibi, R., AlMansour, A., Barnawi, A. (2021). Cardiovascular disease recognition based on heartbeat segmentation and selection process. *International Journal of Environmental Research and Public Health*, 18(20): 10952. <https://doi.org/10.3390/ijerph182010952>
- [6] Li, H., Boulanger, P. (2020). A survey of heart anomaly detection using ambulatory electrocardiogram (ECG). *Sensors*, 20(5): 1461. <https://doi.org/10.3390/s20051461>
- [7] Chakir, F., Jilbab, A., Nacir, C., Hammouch, A. (2020). Recognition of cardiac abnormalities from synchronized ECG and PCG signals. *Physical and Engineering Sciences in Medicine*, 43: 673-677. <https://doi.org/10.1007/s13246-020-00875-2>
- [8] Chowdhury, M., Poudel, K., Hu, Y. (2020). Detecting abnormal PCG signals and extracting cardiac information employing deep learning and the shannon energy envelope. In *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, USA, pp. 1-4. <https://doi.org/10.1109/SPMB50085.2020.9353624>
- [9] Khan, A.H., Hussain, M., Malik, M.K. (2021). Cardiac disorder classification by electrocardiogram sensing using deep neural network. *Complexity*, 2021(1): 5512243. <https://doi.org/10.1155/2021/5512243>
- [10] Farjo, P.D., Sengupta, P.P. (2021). ECG for screening cardiac abnormalities: The premise and promise of machine learning. *Circulation: Cardiovascular Imaging*, 14(6): e012837. <https://doi.org/10.1161/CIRCIMAGING.121.012837>
- [11] Ajitkumar Singh, S., Ashinikumar Singh, S., Dinita Devi, N., Majumder, S. (2021). Heart abnormality classification using PCG and ECG recordings. *Computación y Sistemas*, 25(2): 381-391. <https://doi.org/10.13053/cys-25-2-3447>
- [12] Singh, S.A., Majumder, S. (2019). A novel approach OSA detection using single-lead ECG scalogram based on deep neural network. *Journal of Mechanics in Medicine and Biology*, 19(4): 1950026. <https://doi.org/10.1142/S021951941950026X>
- [13] Nabih-Ali, M., El-Dahshan, E.S.A., Yahia, A.S. (2017). Heart diseases diagnosis using intelligent algorithm based on PCG signal analysis. *Circuits and Systems*, 8(7): 184-190. <https://doi.org/10.4236/cs.2017.87012>
- [14] Shah, K., Pandya, A., Kotwani, P., Saha, S., Saxena, D., Gaidhane, S. (2021). Cost-effectiveness of portable electrocardiogram for screening cardiovascular diseases at a primary health center in Ahmedabad District, India. *Frontiers in Public Health*, 9: 753443. <https://doi.org/10.3389/fpubh.2021.753443>
- [15] Baghel, N., Dutta, M.K., Burget, R. (2020). Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. *Computer Methods and Programs in Biomedicine*, 197: 105750. <https://doi.org/10.1016/j.cmpb.2020.105750>
- [16] Berkaya, S.K., Uysal, A.K., Gunal, E.S., Ergin, S., Gunal, S., Gulmezoglu, M.B. (2018). A survey on ECG analysis. *Biomedical Signal Processing and Control*, 43: 216-235. <https://doi.org/10.1016/j.bspc.2018.03.003>
- [17] Li, H., Wang, X., Liu, C., Li, P., Jiao, Y. (2021). Integrating multi-domain deep features of electrocardiogram and phonocardiogram for coronary artery disease detection. *Computers in Biology and Medicine*, 138: 104914. <https://doi.org/10.1016/j.combiomed.2021.104914>
- [18] Xiang, Y., Lin, Z., Meng, J. (2018). Automatic QRS complex detection using two-level convolutional neuralnetwork. *Biomedical Engineering Online*, 17: 1-17. <https://doi.org/10.1186/s12938-018-0441-4>
- [19] Huang, Q., Yang, H., Zeng, E., Chen, Y. (2023). A deep-learning-based multi-modal ECG and PCG processing framework for cardiac analysis. *Authorea Preprints*, 20(24): 1-11.
- [20] Zeng, W., Lin, Z., Yuan, C., Wang, Q., Liu, F., Wang, Y. (2021). Detection of heart valve disorders from PCG signals using TQWT, FA-MVEMD, Shannon energy envelope and deterministic learning. *Artificial Intelligence Review*, 54: 6063-6100. <https://doi.org/10.1007/s10462-021-09969-z>
- [21] Sugiyarto, A.W., Abadi, A.M., Sumarna, S. (2021). Classification of heart disease based on PCG signal using CNN. *Telkommnika (Telecommunication Computing Electronics and Control)*, 19(5): 1697-1706. <http://doi.org/10.12928/telkommnika.v19i5.20486>
- [22] Tariq, Z., Shah, S.K., Lee, Y. (2022). Feature-based fusion using CNN for lung and heart sound classification. *Sensors*, 22(4): 1521. <https://doi.org/10.3390/s22041521>
- [23] Liu, C., Springer, D., Li, Q., Moody, B., Juan, R.A., Chorro, F.J., Castells, F., Roig, J.M., Silva, I., Johnson, A.E.W., Syed, Z., Schmidt, S.E., Papadaniil, C.D., Hadjileontiadis, L., Naseri, H., Moukadem, A., Dieterlen, A., Brandt, C., Tang, H., Samieinasab, M., Samieinasab, M.R., Sameni, R., Mark, R.G., Clifford, G.D. (2016). An open access database for the evaluation of heart sound algorithms. *Physiological Measurement*, 37(12): 2181-2213. <https://doi.org/10.1088/0967-3334/37/12/2181>
- [24] Clifford, G.D., Liu, C., Moody, B., Lehman, L.H., Silva, I., Li, Q., Johnson, A.E., Mark, R.G. (2017). AF classification from a short single lead ECG recording: The PhysioNet/computing in cardiology challenge 2017. In *2017 Computing in Cardiology (CinC)*, Rennes, France, pp. 1-4. <https://doi.org/10.22489/CinC.2017.065-469>
- [25] Byeon, Y.H., Pan, S.B., Kwak, K.C. (2019). Intelligent deep models based on scalograms of electrocardiogram signals for biometrics. *Sensors*, 19(4): 935. <https://doi.org/10.3390/s19040935>
- [26] Veerappan, P., Palanisamy, M.M., Depha, M. (2022). A

- review on ECG-Signal classification of scalogram snapshots the use of convolutional neural network and continuous wavelet transform. *International Research Journal of Engineering and Technology (IRJET)*, 9(2): 908-911.
- [27] Acharya, U.R., Fujita, H., Sudarshan, V.K., Lih Oh, S., Muhammad, A., Koh, J.E., W, Tan, J.H., Chua, C.K., Chua, K.P., San Tan, R. (2017). Application of empirical mode decomposition (EMD) for automated identification of congestive heart failure using heart rate signals. *Neural Computing and Applications*, 28: 3073-3094.
- <https://doi.org/10.1007/s00521-016-2612-1>
- [28] Stallone, A., Cicone, A., Materassi, M. (2020). New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms. *Scientific Reports*, 10(1): 15161. <https://doi.org/10.1038/s41598-020-72193-2>
- [29] Ozaltin, O., Yeniay, O. (2023). A novel proposed CNN-SVM architecture for ECG scalograms classification. *Soft Computing*, 27(8): 4639-4658. <https://doi.org/10.1007/s00500-022-07729-x>