International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 14, No.1 (Sep-2023)

http://dx.doi.org/10.12785/ijcds/140180

Comparative Study on Heart Anomalies Early Detection Using Phonocardiography (PCG) Signals

Abbas Hasan* **and Zouhir Bahri****

Electrical and Electronics Engineering Dept., University of Bahrain, Isa Town, Bahrain

Received 05 Apr, 2023, Revised 20 Sep. 2023, Accepted 22 Sep. 2023, Published 24 Sep. 2023

Abstract: In this paper, a Phonocardiography (PCG)- based comparative study for cardiovascular anomalies' early detection system is proposed. Some of the main signal processing and deep learning methods applied in the literature to PCG signals are contrasted to differentiate normal heartbeats from abnormal ones and classify five of the most common murmurs, the "lub" and "dub" of "lub-dub" are formed by combining the typical heartbeat sound S1 and S2. The results of this comparative study show an average of 92.14 % for heart anomaly detection and 71.02 % for classification rates. This is achieved by using Deep Neural Network (DNN) classification with Hyperbolic Tangent (tanh) activation function, a 5-layer with 100 neurons in each layer. The Discrete Wavelet Transform (DWT) was found to be the best denoising algorithm and the Heart Sound Envelogram (HSE) was the best segmentation method for the PCG signal. Mel Frequency Cepstral Coefficients (MFCC) Features outperformed their Time and Frequency Domain counterparts. This work proved to be useful in the framework of intelligent and preventative health care systems, offering a convenient early warning home-care tool that should help to direct potentially ill individuals to cardiologists for more precise diagnoses.

Keywords: Heart Anomalies, Phonocardiography (PCG), Biometric, Signal Processing, Deep Learning.

1 Introduction

One of the main causes of death among people globally is cardiovascular disease, which has a significant impact on the expense of the global healthcare system [1]. Early identification of heart issues should significantly lower the risks of problems and exorbitant treatment expenses before they worsen [2].

Even though the Electro-Cardiography (ECG) is more effective and accurate in the analysis of heart issues compared to the PCG, however, ECG requires special equipment and trained medical staff, necessitating the patient to be admitted to the hospital [3]. On the other hand, PCG is most dominant among medical staff due to its efficacy and simplicity. Hence, this comparative study intends to provide an efficient tool powered by Machine Learning (ML) techniques making it handy to use for untrained users, providing early detection for possible risks, and directing potentially-ill individuals to look for more precise diagnoses through cardiologists. This should establish a simple, fairly reliable, and handy home-care tool for intelligent and proactive public health systems.

In the past few decades, several studies have been conducted to automatically differentiate between normal PCG and cardiac murmurs. Ismail et al. presented a thorough survey on the subject [4] including a summary of the many signal processing, segmentation, feature extraction, and machine learning approaches used to analyze PCG data to identify and categorize cardiac anomalies. The algorithms were separated into time, frequency, and hybrid techniques. Although time-domain segmentation algorithms may accurately and directly localize heart sounds, they require a lot of signal preprocessing because of how sensitive they are to noise. And even though heart sounds may be more easily recognized in another domain than the temporal one, the majority of common techniques convert the PCG signal from the temporal domain into another domain, where heart sounds may be distinguished more readily.

To denoize PCG data, a variety of signal processing techniques have been used. This comprises classical digital filters like the IIR [5], Butterworth [6] Chebyshev (CHV) [7], and Empirical Mode Decomposition (EMD) [8], in addition to the Discrete Wavelet Transform (DWT) [9].

N. Kouras [10] successfully applied segmentation to the PCG signal utilizing Daubechies db10 DWT and Shannon entropy, one of the primary segmentation approaches documented in the open literature. The Template Matching autocorrelation technique, on the other hand, was created by F. Agrafioti and D. Hatzinakos [3] and is a reliable and accurate biometric tool. It considerably improves the classification performance of the ECG signals while acting as an intruder detector. Varghees et al [11] promoted a new automated robust heart sound activity detection method that applied a Hilbert transform to the Shannon entropy envelope and reported an average sensitivity of 99.43% and positive predictivity of 93.56% over clean and noisy pathological and non-pathological PCG signals with a signal-to-noise ratio of 5 dB.

With regards to the feature extraction, Fasil and Rajesh [12] categorized EEG signals using Time domain features and a proposed exponential energy feature with a classification accuracy of 89% and 99.5%, respectively, for the Ralph Andrzejak EEG dataset. This was done in order to

*E-mail address: *abbasalasadi@gmail.com, **Corresponding Author: zkbahri@uob.edu.bh*

test the various feature extraction methods. Singh J. and Anand R. [13] conducted studies in both the time and frequency domains to find specific connections between the PCG representations in the two domains. The findings demonstrated that the time domain analysis offers information on different time intervals associated with the first and second heart sounds. The frequency domain analysis reveals the primary frequency components linked to different heart sounds. The study also demonstrates the relationship between time and frequency domain features, demonstrating that longer S1 and S2 durations result in much larger dominant frequencies in S1 and S2. To construct a strong feature extraction model, Iqtidar K. and Qamar U. [8] proposed 1D-Adaptive Local Ternary Patterns, fused them with the MFCC features, and attained classification accuracy of 98.3% and 97.2% mean by SVM classification.

The literature reports that several techniques were used by people entrusted with classifying PCG signals. Basak H., and Roy A, [14] performed SVM and ANN classification using PCA and LDA features, with ANN classification using PCA features achieving the best classification performance of 96.67%. A small sample of PCG signals was used by Pavlopoulos S [15] to assess the applicability and utility of several decision tree architectures. The findings indicated that differentiation between classes was highly successful. MFCC characteristics were retrieved by Arslan and Karhan M [16] who then examined how the modes derived by the Hilbert-Huang transform influenced the categorization. The KNN, MLP, SVM, and DNN were used as classifiers. Their recommended Hilbert-Huang transform-based feature extraction approach outperformed competing methods, achieving an overall accuracy of 98.9% when categorizing PCG signals with the DNN classifier.

Safara et al. [17] suggested employing Hankel matrix-based classification with singular value decomposition to discriminate between normal cardiac sounds and murmurs in order to evaluate the discriminating capability of the obtained features. They used the discrete wavelet transform (DWT) for segmentation and the adaptive-neuro fuzzy inference system (ANFIS) to build a classifier that used Shannon entropy as a feature. This method does not reveal information on the accuracy of murmur classification, except for mitral and pulmonary stenosis.

A novel technique for combining Heart Sounds into a single cardiac cycle was presented by Gupta et al. [18] utilizing K-means clustering and homomorphic filtering. Systole and diastole energies, spectral coefficients, and cepstral coefficients were used to identify PCG signals as either normal or falling under the murmurs (systolic and diastolic) group. The coefficient eigenvectors of heart sounds were extracted using the Daubechies-2 wavelet decomposition, and they were subsequently employed as the input for neural networks. The classification results indicate an accuracy of 90.29% when using K-means classification and homomorphic filtering.

With the help of a support vector machine with a fifth-order polynomial kernel function, a wavelet packet with 8 levels of decomposition, and two energy characteristics that were retrieved via SLBS from selected nodes, Choi [19] introduced a technique for classifying valvular heart sounds as problematic. This method helped identify normal heart sounds. Nevertheless, Reed et al. [20] provided a computer-aided diagnosis method based on a Coifman fourth-order wavelet kernel and a seven-level wavelet decomposition that could be applied to five distinct clinical scenarios.

In this paper, we provide an extension to our preliminary study [21] with detailed illustrations for the extensive testing and methods applied. The purpose of this work is to provide comparative research in order to integrate some of the major methodologies mentioned above in a system that is reasonably accurate, easy to use, and cost-effective. Therefore, for PCG denoising, the Discrete Wavelet Transform (DWT) is contrasted with a custom IIR filter, as well as traditional Chebyshev and Butterworth filters. Similar comparisons are made between Heart Sound Envelogram and Shannon Energy Segmentation utilizing Template Matching. Likewise, the Mel Frequency Cepstral Coefficients (MFCC) are in contrast to feature extraction based on the Frequency and Time Domains. Finally, to identify and categorize cardiac abnormalities, the Support Vector Machine, KNN, and Decision Tree are compared to Deep Neural Networks.

The rest of this paper is structured as follows. A summary of PCG signals and data collection is given in Section 2. Section 3 provides an overview of the comparative study phases. Main PCG signal processing techniques are compared in Section 4. Section 5 presents a comparison of some PCG signal segmentation methods. Sections 6 and 7 respectively provide a comparison of the main approaches used for PCG feature extraction and classification. The summary of the testing results and recommendations is provided in Section 8. Finally, Section 9 summarizes and concludes this work.

2 PCG SIGNALS AND DATA SET

PCG signals may include many heartbeat phases. Figure 1 illustrates some of the main possible phases in a PCG signal [22]. The systole (S1) and diastole (S2) are the reference heartbeats considered for examining all other unusual signals, including the third & fourth heart sounds, clicks, knocks, ejection, opening, and snaps. Murmurs can be distinguished from normal heart sounds by mainly four features: timing (systolic or diastolic), duration (such as long or short), intensity (loud or soft), and pitch (low or high frequency). These anomalous noises are occasionally described by terms like "musical," "whoop," or "honk," all of which have considerably higher frequencies than trained medical personnel can hear using a stethoscope.

Frequencies associated with heart murmurs range from 20 to 500 Hz. Mitral Stenosis (MS), the third, as well as fourth heart, sounds, and other murmurs, can occur at low frequencies of less than 100 Hz (S3 and S4). However, some other murmurs like mitral valve prolapse (MVP) have higher pitch noises with prominent frequencies at around 400 Hz. However, in general, by setting the frequency measurement between 100 and 400 Hz [23].

Instead, a better quality digital stethoscope that costs more can be used for better measurement of the PCG signals. The objective of this work is to implement the best outcome of this comparative study on smartphones with a sensitive microphone, which will be easy to use for everyone.

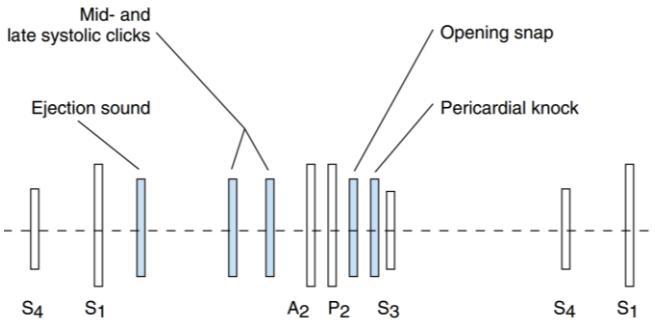


Figure 1: Miscellaneous Heart Sounds [22]

The PCG signal is tainted with the noise surrounding and artifacts all the time due to ambient sounds and stethoscope jitters, therefore, denoising the signal is essential before proceeding to the next phase (segmentation) where a single beat should be separated before extracting its features for the classification process.

At the onset of this work, the intent was to collect heart sounds from people and those with various heart murmurs using a modified stethoscope at hospitals around Bahrain with the help of specialized medical professionals. However, access to patients in hospitals has been severely constrained during the Covid-19 outbreak. As a result, we turned to some recorded data mostly recorded by students on ordinary subjects, in addition to some data we were able to collect for healthy people using a straightforward modified classical stethoscope equipped with a typical computer microphone, as shown in Figure 2.

Figure 2: Stethoscope Built for Data Recording

As a result, we turned to some recorded data that was readily available online [24] [25]. The online data complemented our recorded data with approximately 1000 PCG signals. Nonetheless, integrating these signals into our study presented an additional challenge, as they originated from diverse sources and were recorded using different devices and frequencies. To address this, we developed an application capable of accommodating the varying data quality and frequencies. The total number of the dataset used in this study for the 6 classes is as follows:

(a) Normal heartbeat (601 files)
(b) Third heartbeat (S3) (76 files)
(c) Aortic Stenosis (AS) (227 files)
(d) Mitral Regurgitation (MR) (225 files)
(e) Mitral Stenosis (MS) (222 files)
(f) Mitral Valve Prolapse (MVP) (200 files)

DWT Filter with 20 and 500 Hz selected bands and Daubechies 6 (db6) as a mother wavelet is designed to denoise the PCG signals. Although the ripples can be seen in the frequency domain version of the signal, their magnitude is negligible. The signal was greatly attenuated as shown in Figure 10, and the noise in the PCG signal was successfully filtered. From the clean data for each class, 10 samples were randomly chosen. The average MSE attribute was calculated for the DWT Filter comparing the base clean data to its data after adding White Gaussian Noise (WGN) and artifacts for various classes. The results provide 3.27% for normal, 2.57% for S3, 3.2% for AS, 2.81% for MR, 3.25% for MS, and 2.4% for MVP class.

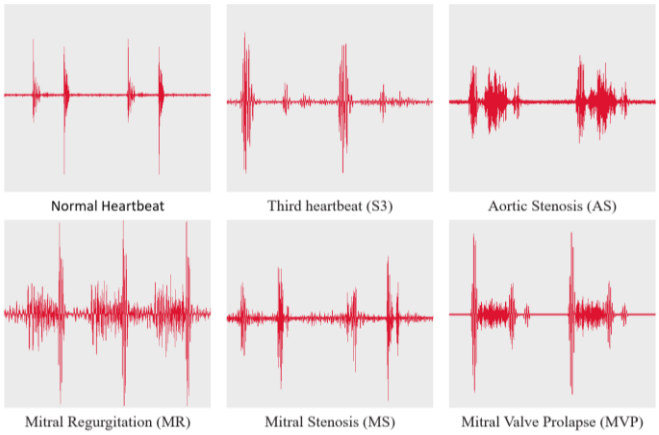


Figure 3: PCG Data Waveform for Different Classes

It is significant to highlight that the data collected were extremely diverse concerning the signal-to-noise ratio (SNR) and sampling rate. While some were of acceptable quality, others (such as those made with the modified stethoscope) were very noisy. This has made the detector/classifier more challenging.

3 Comparative Study Overview

Figure 3 shows the overall process diagram of the main phases of the proposed system. Different techniques shall be performed in each phase in order to select the best method that provides the best performance as depicted in the comparative study block diagram of Figure 4.

For the signal denoising phase, 5 methods of filtering methods were performed: 3 Band Pass Filters configured with a frequency between 20-500 Hz and order 4: Custom-IIR, Butterworth (BWT), and Chebyshev (CHV), as well as Empirical Mode Decomposition (EMD) (with 10 IMFs (Intrinsic Mode Function) with relative energy greater than 25%), in addition to the Discrete Wavelet Transform (DWT) with 20 and 500 Hz selected bands and Daubechies 6 (db6) as a mother wavelet. The Mean Square Error (MSE) attribute was used as a performance measure of the filters and was evaluated by calculating the difference between the original clean data and after adding White Gaussian Noise (WGN) and artifacts across the PCG signals.

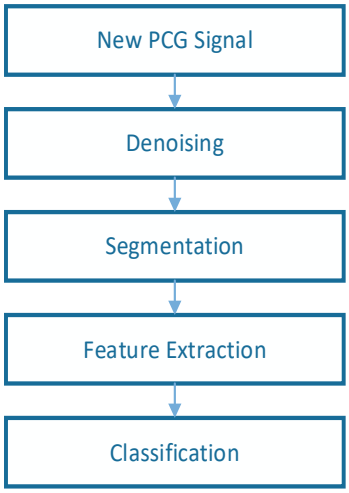


Figure 4: Overall System Diagram

In the segmentation stage, 3 techniques were performed: Segmentation by Wavelet Decomposition using Shannon Energy with Template Correlation (SE-TC), Segmentation based on Heart Sound Envelogram (HSE), and Segmentation by Template Correlation using Correlation Coefficient (CC-TC). To evaluate these techniques' use and compare their effectiveness, a visual comparison was conducted using a stacked comparison graph that was created for all available data to visualize the segmented portion with its reference to the original data, then carried out a visual assessment of all the data.

Next, in the feature extraction phase, 3 different methods were applied to the segmented portions of the PCG signal: Time Domain Features (TDF), Frequency Domain Features (FDF), and Mel Frequency Cepstral Coefficients (MFCC). These features were injected into the training and verification process using 70% of the data for training and 30 % for testing conducted by 4 classification techniques: Support Vector Machine (SVM), Deep Neural Network (DNN), Decision Tree (DT), and K-Nearest Neighbor (KNN).

As MFCC and DNN create the best results in detection and classification rates, an additional process of finetuning was conducted for the DNN parameters: the activation function (tanh, sigmoid, and ReLU), the number of hidden layers (3, 5, and 10), as well as the number of neurons in each hidden layer (10, 50, and 100). The overall process of the detailed functions is shown in Figure 5.

4 PCG SIGNAL PROCESSING

All applications that use noisy data must include signal processing. The PCG signal is recorded using a standard stethoscope that is connected to a computer, and the recorded PCG signal includes lots of noise. Five different types of filtering methods were used in the noise removal phase: In addition to Custom-IIR, Butterworth (BWT), and Chebyshev (CHV) bandpass filters, there is also discrete wavelet transform (DWT) and empirical mode decomposition (EMD) algorithms. To assess the effectiveness of the filters, the Mean Square Error (MSE) attribute was calculated between the clean original data and its noisy version following the addition of the identical White Gaussian Noise (WGN) and artifacts to all signals.

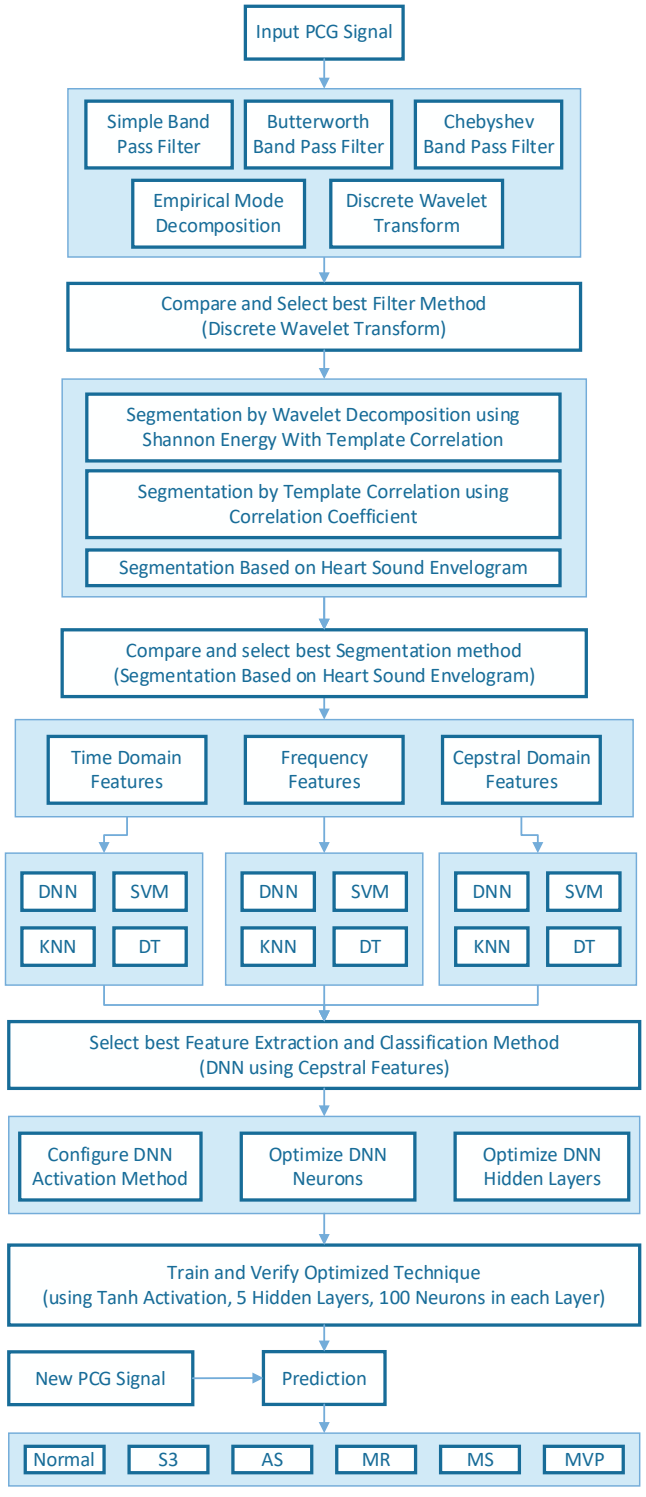


Figure 5: Comparative Study Summary

4.1 Infinite Impulse Response Filter

The current output of an IIR filter depends on the preceding one, hence the names feedback and recursive filters. IIR filters have the advantage of numerous prior outputs (or an infinite). In contrast to other feedback systems, IIR filters don't have a linear phase. IIR filters compute the current output sample utilizing the input sample history as well as the output sample history. The general form is shown in Eq.(1):

$$y(n) = -\sum_{k=1}^{N} a(k).y(n-k) + \sum_{k=0}^{M} b(k).x(m-k) \quad (1)$$

Where a(k) elements are the feedback coefficients and the y(n-k) element is the output. The b(k) elements are the feed-forward elements where the x(m-k) element represents the input data stream. In many IIR filters, the number of feedback coefficients (N) and the number of feed-forward elements (M) are the same (N=M) [5]

The PCG signals were denoised using a custom IIR with a bandwidth of 20-500 Hz and order 4. Figure 6 depicts the PCG signal before and after it has been denoised (filtered). The signal was greatly attenuated and the noise in the PCG signal was successfully filtered. From the clean data for each class, 10 samples were randomly chosen. The average MSE attribute was calculated for the IIR Filter comparing the base clean data to its data after adding White Gaussian Noise (WGN) and artifacts for various classes. The results provide 5.63% for normal, 7.8% for S3, 8.23% for AS, 7.5% for MR, 6.18% for MS, and 6.01% for MVP class.

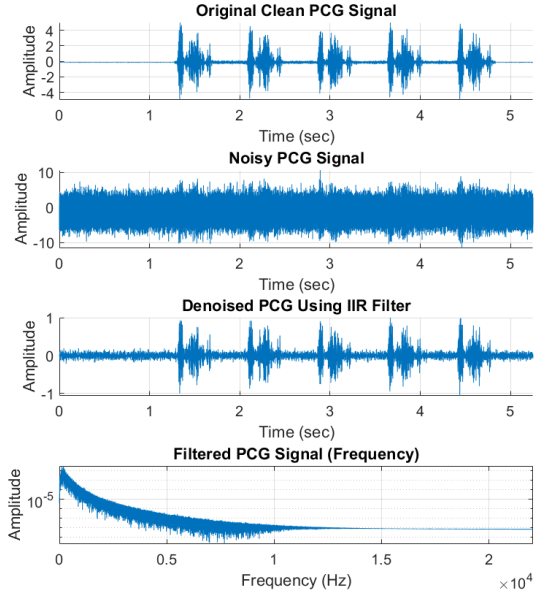


Figure 6: Denoised PCG using IIR Filter

4.2 Wavelet Transform

The most popular technique for converting from the time domain to the frequency domain is Fourier transform-based (FFT) spectral analysis. The Fourier theory states that a signal can be broken down into a variety of sines and cosines. Its disadvantage, though, is that it doesn't offer any time-related information. The wavelet transform, on the other hand, offers an analysis that is comparable to the FFT but uses a different merit function. The primary distinction is that the wavelet transform function, as opposed to the sines and cosines decomposition, is localized in both the time and frequency domains. [26]

The essential idea behind the wavelet transform is to describe any function as a collection of wavelets that create the wavelet transform's basis function. Wavelets, which can be extended and moved to capture features that are confined in time and frequency, are tiny waves that are placed at various moments. The wavelets, also known as daughter wavelets, are scaled and graphically reproduced versions of a finite-length oscillating waveform. The basis for the best wavelet selection is based on the features of the original [9] and the required analysis goal.

The wavelet transform is regarded as a multi-resolution analysis since it offers a collection of time-frequency properties of the fundamental signal with various resolutions. Wavelet transform methods come in two varieties: Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). The DWT creates a mutually orthogonal set of wavelets; however, the discrete wavelet is more related to the mother wavelet [27] as shown in Eq (2).

$$\psi_{m,k}(t) = \frac{1}{\sqrt{S_0^m}}\psi\left(\frac{t - k\tau_0 S_0^m}{S_0^m}\right) \quad (2)$$

where s_0 is a fixed scaling parameter set for more than 1, m controls wavelet dilation, k controls wavelet translation, τ_0 is the translation value, which must be greater than 0, and ψ is the mother wavelet.

Figure 8 shows the structure of the DWT decomposition process which illustrates the approximation and detail coefficients for a decomposition level 2 where $x(n)$ is the original input signal and the $h(n)$ and $g(n)$ are the high-pass and low-pass filters. The DWT provides a multilevel decomposition with an approximation of the decomposed signal and detail coefficients at each level of decomposition.

The input PCG signal was decomposed into 10 levels. The Daubechies 6 (db6) wavelet template Figure 7 is selected as the mother wavelet because it morphologically resembles the PCG signal.

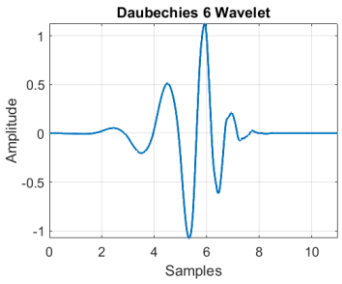


Figure 7: Daubechies 6 Wavelet Function [28]

After the decomposition of PCG up to 10 levels, 11 signals/sub-bands (D1, D2, D3, D4, D5, D6, D7, D9, D10 A10) were obtained, each containing different frequency ranges. Figure 8 illustrates the decomposition of a signal into various levels. among these sub-bands, only the detail coefficients D6, D7, D8, D9, and D10 were chosen since they contain the frequency range of a heart sound and the necessary frequencies for abnormal heartbeats fall between 20 and 500 Hz. The selected bands were added to perform the wavelet reconstruction. The PCG signal was decomposed into 10 levels of wavelet Daubechies Figure 9.

DWT Filter with 20 and 500 Hz selected bands and Daubechies 6 (db6) as a mother wavelet is designed to denoise the PCG signals. Although the ripples can be seen in the frequency domain version of the signal, their magnitude is negligible. The signal was greatly attenuated as shown in Figure 10, and the noise in the PCG signal was successfully filtered. From the clean data for each class, 10 samples were randomly chosen. The average MSE attribute was calculated for the DWT Filter comparing the base clean data to its data after adding White Gaussian Noise (WGN) and artifacts for

various classes. The results provide 3.27% for normal, 2.57% for S3, 3.2% for AS, 2.81% for MR, 3.25% for MS, and 2.4% for MVP class.

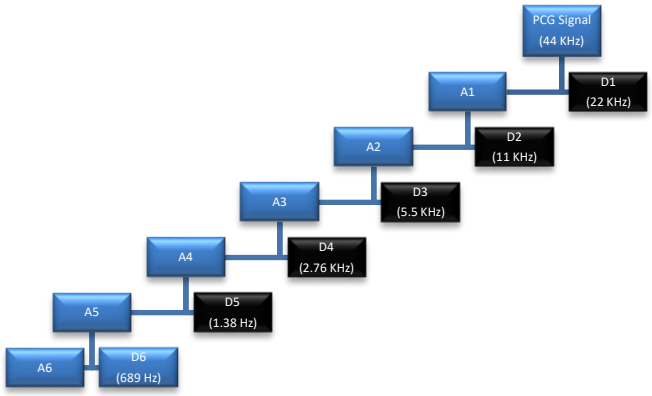


Figure 8: DWT Decomposition for the PCG Signal

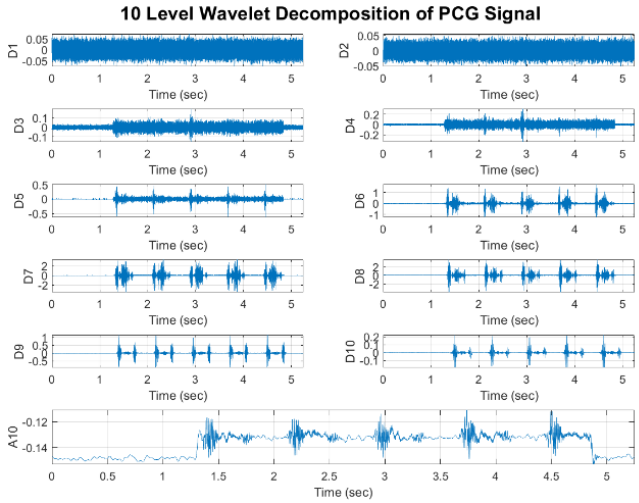


Figure 9: 10 Level DWT Decomposition for the PCG Signal

4.3 Empirical Mode Decomposition

A signal is divided into several intrinsic mode functions (IMF) using the empirical mode decomposition (EMD) technique, coupled with a residue that represents the trend. The frequency information is derived using EMD from non-stationary and non-linear data sets. The nonlinear signal characteristics are extracted using EMD. An IMF has a mean value of zero and only one extreme between zero-crossings [8]. The original PCG signal x(t) is decomposed as:

$$x(t) = \sum_{i=1}^{n} (c_i) + r_n \qquad (3)$$

where c_i represents the extracted IMFs, r_n represents a residual signal, and several functions n in the set depend on the input signal.

10 IMFs with relative energy greater than 25% were selected for reconstructing the denoised version of the PCG signal. The decomposed harmonics disturbance signal is shown in Figure 11. Although Figure 12 below shows that the EMD Filter performed badly with high noise data, it was shown to work effectively with low noise data. From the clean data for each class, 10 samples were randomly chosen. The average MSE attribute was calculated for the EMD Filter comparing the base clean data to its data after adding White Gaussian Noise (WGN) and artifacts for various classes. The results provide 5.98% for normal, 7.58% for S3, 8.69% for AS, 7.42% for MR, 10.1% for MS, and 9.33% for MVP class.

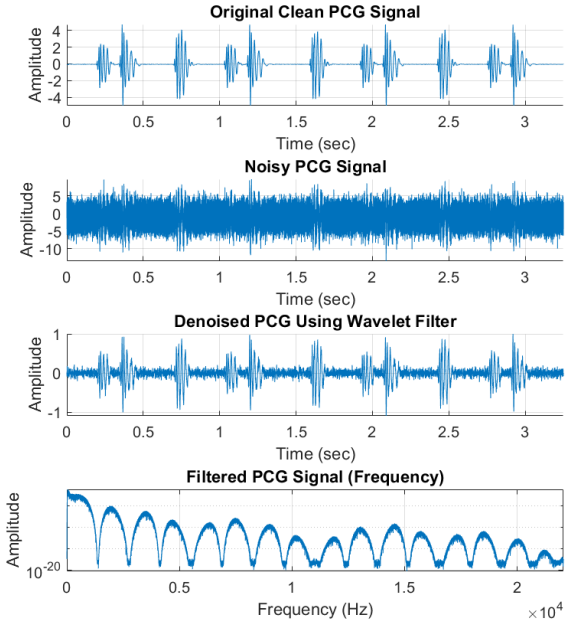


Figure 10: DWT Filter for the PCG Signal

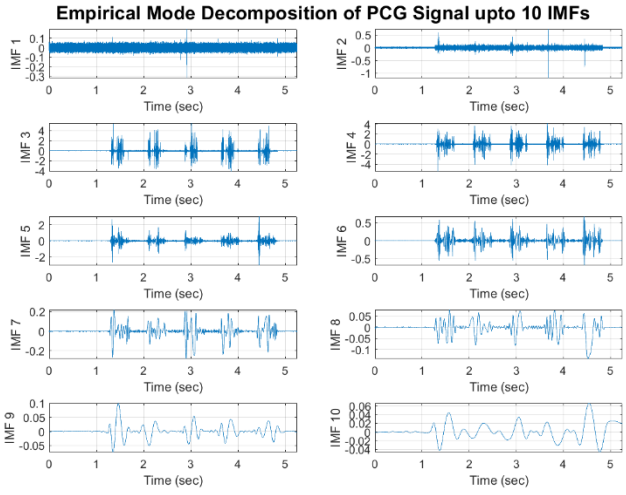


Figure 11: Intrinsic Mode Functions of Noisy PCG Signal

4.4 Butterworth Filter BTW

The Butterworth filter BTW, also known as a maximally flat magnitude filter, is made to have a passband frequency response that is as flat as feasible. Even though it does not produce a very crisp output, the BTW filter is widely employed in many applications that require signal processing to obtain an effectively flat response within the specified frequency limits. The BTW filter seeks to attain its roll-off rate more gradually when compared to other filters (such as Chebyshev), contending that behavior below cut-off

frequency rates is more crucial than at any other frequency [6], [29].

The higher order of the Butterworth filter provides a higher number of cascaded stages; however, the ideal frequency response of BTW is unattainable as it produces an excessive passband ripple. The general form representing the frequency response and the "nth" Order Butterworth filter is given in Eq.**Error! Reference source not found.**:

$$H_{(j\omega)}=\frac{1}{\sqrt{1+\varepsilon^2\left(\frac{\omega}{\omega_c}\right)^{2n}}} \quad (4)$$

Where: n is the filter order, ω is equal to 2πf and ε is the maximum passband gain and ω_c is the cutoff frequency [30].

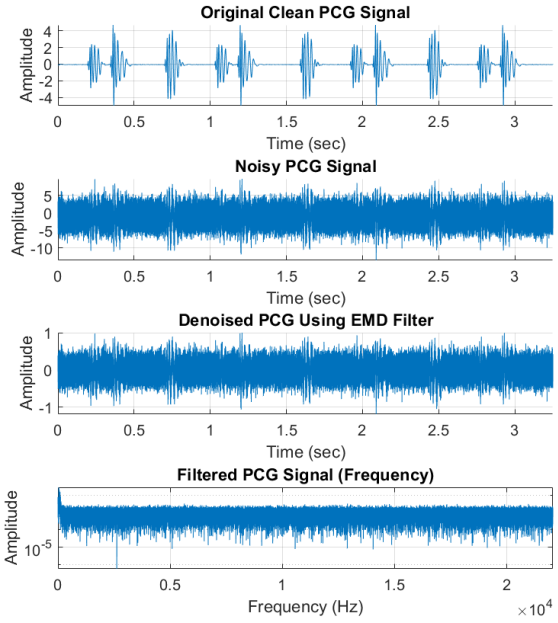


Figure 12: Denoised PCG using EMD Filter [6], [29][30]

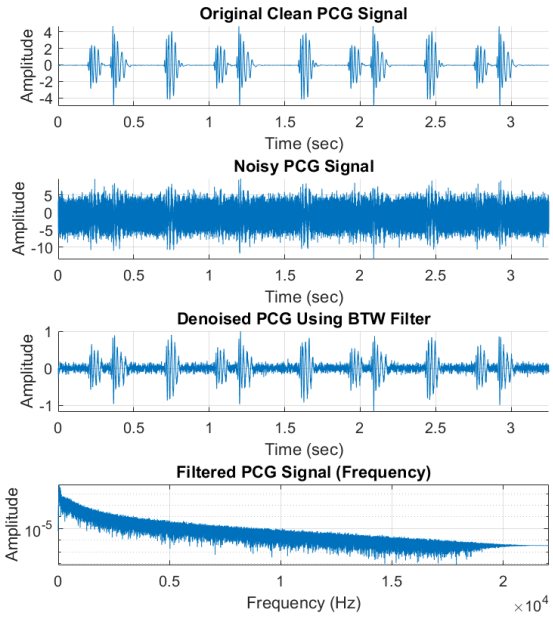


Figure 13: Denoised PCG using BTW Filter.

BTW with a bandwidth of 20-500 Hz and order 4 was designed to denoise the PCG signals. The denoised (filtered) version of the PCG signal given in Figure 13 shows the signal attenuates significantly and the noise in the PCG signal was filtered effectively. From the clean data for each class, 10 samples were randomly chosen. The average MSE attribute was calculated for the BTW Filter comparing the base clean data to its data after adding White Gaussian Noise (WGN) and artifacts for various classes. The results provide 5.63% for normal, 8.27% for S3, 8.29% for AS, 7.5% for MR, 7.43% for MS, and 6.01% for MVP class.

4.5 Chebyshev Filter for PCG Signal Denoising

Chebyshev filters have passband ripple (also known as "Chebyshev filters") or stopband ripple (also known as "inverse Chebyshev filters") that provide analog or digital filters with higher roll-off than BTW filters. Chebyshev minimizes the error between the actual filter with ripples and the idealized in the passband [7].

The gain response $G_n(\omega)$ in CHV Filter is calculated as described in Eq.(5) [31]:

$$G_n(\omega) = |H_n(j\omega)| = \frac{1}{\sqrt{1+\varepsilon^2\, C_n^2(\omega/\omega_0)}} \quad (5)$$

Where ω is the angular frequency of the *n*th-order low-pass filter is equal to the absolute value of the transfer function $H_n(s)$ evaluated at s=jω, ε is the ripple factor, ω_0 is the cutoff frequency and C_n is a Chebyshev polynomial of the nth order.

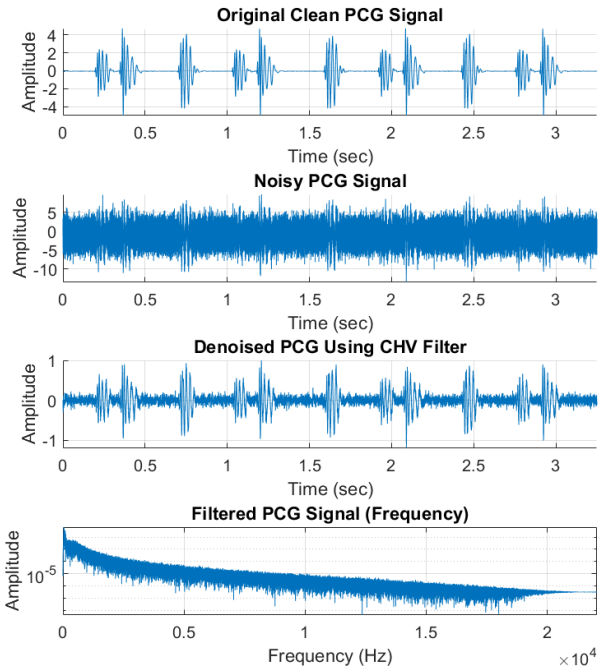


Figure 14: Denoised PCG using Chebyshev

CHV with a bandwidth of 20-500 Hz and order 4 is designed to denoise the PCG signals. The denoised (filtered) version of the PCG signal given in Figure 14 shows the signal attenuates significantly and the noise in the PCG signal was filtered effectively. From the clean data for each class, 10 samples were randomly chosen. The average MSE attribute was calculated for the CHV Filter comparing the base clean data

to its data after adding White Gaussian Noise (WGN) and artifacts for various classes. The results provide 5.11% for normal, 15.67% for S3, 7.01% for AS, 6.65% for MR, 14.51% for MS, and 5.29% for MVP class.

4.6 Filters Performance Comparison

Using the identical White Gaussian Noise (WGN) and artifacts added to all of the signals, the mean square error (MSE) between the noisy data and its original clean data was calculated to evaluate the performance of the filters. Table 1 lists the MSE for each of the used filters and demonstrates that the DWT produces the least MSE.

Table 1. Mean Square Error results for all filter

	Normal	S3	AS	MR	MS	MVP
IIR	5.63%	7.80%	8.23%	7.50%	6.18%	6.01%
DWT	3.28%	2.57%	3.20%	2.81%	3.25%	2.40%
EMD	5.98%	7.58%	8.69%	7.42%	10.10%	9.33%
BTW	5.63%	8.27%	8.29%	7.50%	7.43%	6.01%
CHV	5.11%	15.67%	7.01%	6.65%	14.51%	5.29%

5 PCG SIGNAL SEGMENTATION

Signal segmentation is a very crucial step in the signal processing performed on signals that include patterns. The objective of segmentation is to remove the unwanted segments of the signal and to extract only relevant portions from the PCG waveform.

3 techniques were applied: Segmentation by Wavelet Decomposition using Shannon Energy with Template Correlation (SE-TC), Segmentation based on Heart Sound Envelogram (HSE), and Segmentation by Template Correlation using Correlation Coefficient (CC-TC).

A stacked comparison graph is generated for all accessible data to show the segmented section referred to the original data, and then a visual of all data is undertaken to test the operation of these methods and evaluate their effectiveness.

5.1 Segmentation by Wavelet Decomposition using Shannon Energy With Template Correlation

This method uses the Discrete Wavelet Transform (DWT) in conjunction with Shannon entropy. The entropy of the detailed coefficients will be calculated at each level [10]. The algorithm applied initially decomposes the PCG signal into six wavelet filter levels (Daubechies (db)), where the best level of decomposition is set based on the Shannon entropy calculated values at each level of decomposition. Then the mean value of the Shannon entropy determines the boundary between the main components and the pathological murmur and divides the main signal into several segments. finally, all segments are matched with a template heartbeat signal (Figure 15) using the Sum of Squared Difference (SSD) algorithm [32] using Eq.(6):

$$\text{ssd} = \frac{1}{N}\sum_{k=1}^{N}(x_k - y_k)^2 \qquad (6)$$

Where x_k is the feature of the registered template and y_k is from the authentication template, N represents the total number of features. After calculating similarity, a threshold is used to decide whether either given template is matched or not. SSD is a simple method and gives better results as compared to the correlation used in template matching.

5.2 Segmentation by Template Correlation using Correlation Coefficient

This method uses feature selection based on Autocorrelation (AC) of PCG segments matched with the template signal [3]. The method consists of four stages:

1. *Windowing*, where the preprocessed PCG signal is subjected to segmentation into nonoverlapping windows.
2. *Normalized autocorrelation* computation for every window.
3. *Dimensionality reduction* with the Discrete Cosine Transform (DCT) or Linear Discriminant Analysis (LDA).
4. *Classification* based on features obtained from the DCT or LDA.

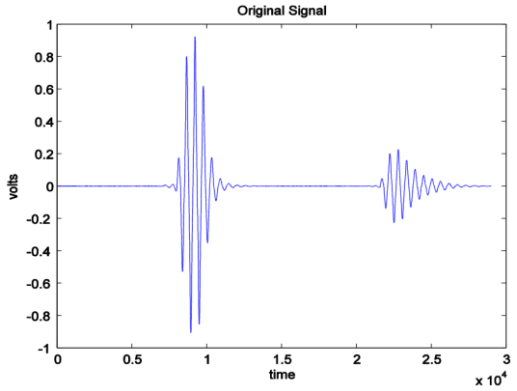


Figure 15: Template used for Template Matching

The PCG is a nonperiodic but highly repetitive signal. The motivation behind AC is that it embeds information about representative and highly distinctive characteristics of signals recorded from different individuals.

Autocorrelation is applied to windowed PCG signals. Windowing can be performed if only the signal is longer than the average heartbeat length to include more than one pulse. On such PCG windows, autocorrelation provides a shift-invariant perspective of similarity features over multiple heartbeat cycles. The normalized autocorrelation coefficients $R_{xx}[m]$ can be computed by Eq.(7):

$$\hat{R}_{xx}[m] = \frac{\sum_{i=0}^{N-|m|-1} x[i]x[i+m]}{\hat{R}_{xx}[0]} \qquad (7)$$

Where x[i] is the windowed PCG for i=0, 1… (N−|m|−1), x[i+m] is the time-shifted version of the windowed PCG with a time lag of m = 0, 1,…(M−1); M<<N, and N is the length of the windowed signal. The dimensionality of autocorrelated signals is considerably high and DCT or LDA are applied for reduction.

5.3 Segmentation based on Heart Sound Envelogram

This method is based on a heart sound envelogram which divides the spectrum into several envelopes. In this method, cardiac sounds will be divided into four parts: the cardiac sound (S1), systolic period, cardiac sound (S2), and diastolic period. The PCG cycle will be segmented into systolic and diastolic parts using the carotid pulse, which provides high accuracy in the characterization and localization of murmurs. The pattern of energy distribution is crucial for the diagnosis of valvular and septal defects, which cause murmurs of characteristic envelope shapes in particular locations of the cycle [33]. The pattern of the energy distribution of the PCG signal is very important in the analysis of valvular and septal defects, which is the reason for murmurs and characteristic envelope shapes in particular locations of the cardiac cycle. We make use of a time-varying energy function calculated as described in Eq.(8)

$$E(n) = \sum_{k=1}^{M} x^2(n-k)\, w(k) \tag{8}$$

where x(n) is the signal and the weighting sequence w(k)= M+I-k with M=32 [34]. This function is computed for the systolic and diastolic segments separately. The presence of murmurs is indicated when significant energy appears after S1 in the systolic segment or after S2 in the diastolic segment. The murmur envelope E(n) is then calculated as given in Eq.5.

5.4 Segmentation algorithms comparison

A stacked comparison graph (Figure 16) is generated for all accessible data to display the segmented section referred to the original data, and then a visual rating for all data is undertaken to assess the operation of these algorithms and compare their performance.

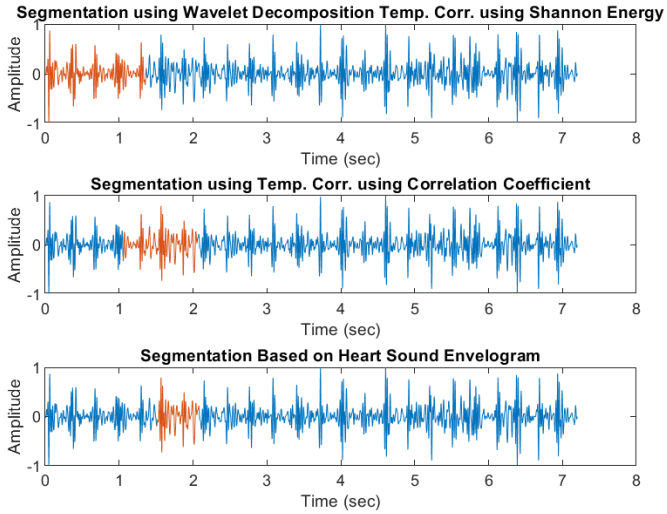


Figure 16: Segmentation stacked graph

The criteria for measuring the segmentation's effectiveness are listed in Table 2. The rating outcomes are displayed in

Table 3 after analyzing the segmentation findings for all data using the stacked graphs.

Table 2. Segmentation evaluation criteria for Segmentation methods

Rating	Grade	Criteria
3	Excellent	detected S1 then S2
2	Good	detected S2 then S1 or long portion
1	Poor	detected short portion or very long portion
0	Not detected	

Table 3. Segmentation Rating results

	Shannon	Corr.Coef	Envelogram
Total	1548	1548	1548
3 scored (Excellent)	61.0%	27.6%	74.6%
2 scored (Good)	26.0%	66.1%	22.5%
1 scored (Bad)	12.7%	6.2%	2.2%
0 scored (Error)	0.3%	0.1%	0.8%
Rate	3814	3390	4192

As can be shown, when compared to other techniques, segmentation based on the heart sound envelogram algorithm produces the best results.

6 PCG Signal Feature Extraction

Feature extraction is an essential part of every machine-learning process. One type of feature is often insufficient to comprehensively represent the PCG signal of all categories, thus it makes the training more challenging. For the feature extraction phase, 3 methods were applied to the PCG signal: Time Domain Features (TDF), Frequency Domain Features (FDF), and Mel Frequency Cepstral Coefficients (MFCC) to evaluate each method and select the best-performing feature extraction method for the training process.

6.1 Time-domain features

The preprocessed signal's time domain properties, including mean, standard deviation, skewness, kurtosis, peak amplitude, root mean square, shape factor, impulse factor, and clearance factor, are computed to assess their capacity to identify the existence of murmur. Additionally, the relevance of their statistical significance is assessed [35].
The standard deviation (SD) of the preprocessed PCG signal is given as:

$$SD = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(Xp_t - \mu(Xp_t))^2} \tag{9}$$

Where $\mu(Xp_t)$ represents the mean value of the preprocessed signal Xp_t. the Mean is calculated by,

$$\mu(Xp_t) = \frac{1}{N}\sum_{t=1}^{N} Xp_t \tag{10}$$

The Peak Amplitude (X_{peak}) is calculated by:

$$X_{Peak} = Max(|Xp_t|) \tag{11}$$

The Root Mean Square (RMS) value is calculated as described in Eq.(12)

$$\mathrm{RMS} = \sqrt{\frac{1}{\mathrm{N}}\left(\sum_{\mathrm{t}=1}^{\mathrm{N}}\left(\mathrm{Xp_t}^2\right)\right)} \tag{12}$$

The crest factor (CRF) is calculated as described in Eq.(13)

$$\mathrm{CRF} = \frac{\mathrm{Y_{Peak}}}{\mathrm{RMS}} \tag{13}$$

The Impulse Factor (IF) is calculated as described in Eq.(14)

$$\mathrm{IF} = \frac{\mathrm{Y_{Peak}}}{\mu(\mathrm{Xp_t})} \tag{14}$$

The Shape Factor (SHF) is calculated as described in Eq.(15)

$$\mathrm{SHF} = \frac{\mathrm{RMS}}{\mu(\mathrm{Xp_t})} \tag{15}$$

The Clearance Factor (CLF) is calculated as described in Eq.(16)

$$\mathrm{CLF} = \frac{\mathrm{X_{Peak}}}{\mathrm{E}} \tag{16}$$

The Skewness (Sk) is calculated as described in Eq.(17)

$$\mathrm{Sk} = \frac{\frac{1}{\mathrm{N}}\sum_{\mathrm{t}=1}^{\mathrm{N}}(\mathrm{Xp_t} - \mu(\mathrm{Xp_t}))^3}{\left[\frac{1}{\mathrm{N}}\sum_{\mathrm{t}=1}^{\mathrm{N}}(\mathrm{Xp_t} - \mu(\mathrm{Xp_t}))^2\right]^{3/2}} \tag{17}$$

The Kurtosis (Kurt) is calculated as described in Eq.(18)

$$\mathrm{Kurt} = \frac{\frac{1}{\mathrm{N}}\sum_{\mathrm{t}=1}^{\mathrm{N}}(\mathrm{Xp_t} - \mu(\mathrm{Xp_t}))^4}{\left[\frac{1}{\mathrm{N}}\sum_{\mathrm{t}=1}^{\mathrm{N}}(\mathrm{Xp_t} - \mu(\mathrm{Xp_t}))^2\right]^2} - 3 \tag{18}$$

6.2 Frequency domain features

Frequency domain (spectral-domain) features are used to extract the frequency-related relationship of the signals of various categories. These features are extracted by first transforming the time domain signal to the frequency domain using Fast Fourier Transform (FFT) algorithm. From the FFT spectrum, 10 features are extracted, such as Spectral Flux, Spectral Crest, Spectral Flatness, Spectral Skewness, Spectral Centroid, Spectral Kurtosis, Spectral Spread, Spectral Roll Off, Spectral Slope, and Spectral Decrease [13], [36].
The spectral flux is calculated as described in Eq.(19)

$$\mathrm{flux(t)} = \left(\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} |\mathrm{S_k(t)} - \mathrm{S_k(t-1)}|^{\mathrm{P}}\right)^{1/\mathrm{P}} \tag{19}$$

Where S_k is the spectral value at bin k. b_1 and b_2 are the band edges, in bins.
The spectral crest is calculated as described in Eq.(20)

$$\mathrm{crest} = \frac{\max\left(\mathrm{S_k} \in [\mathrm{b_1}, \mathrm{b_2}]\right)}{\frac{1}{\mathrm{b_2} - \mathrm{b_1}}\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} \mathrm{S_k}} \tag{20}$$

The spectral flatness is calculated as described in Eq.(21)

$$\mathrm{flatness} = \frac{\left(\prod_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} \mathrm{S_k}\right)^{\frac{1}{\mathrm{b_2}-\mathrm{b_1}}}}{\frac{1}{\mathrm{b_2} - \mathrm{b_1}}\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} \mathrm{S_k}} \tag{21}$$

The spectral skewness is calculated as shown in Eq.(22)

$$\mathrm{skewness} = \frac{\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}}(\mathrm{f_k} - \mu_1)^3\,\mathrm{S_k}}{(\mu_2)^3\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} \mathrm{S_k}} \tag{22}$$

Where f_k is the frequency in Hz corresponding to bin k. μ_1 is the spectral centroid, calculated as described by the spectral Centroid function. μ_2 is the spectral spread, calculated as described by the spectral Spread function.
The spectral centroid is calculated as described in Eq.(23)

$$\mathrm{centroid} = \frac{\sum_{k=b_1}^{b_2} f_k S_k}{\sum_{k=b_1}^{b_2} S_k} \tag{23}$$

The spectral kurtosis is calculated as described in Eq.(24)

$$\mathrm{kurtosis} = \frac{\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}}(\mathrm{f_k} - \mu_1)^4\,\mathrm{S_k}}{(\mu_2)^4\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} \mathrm{S_k}} \tag{24}$$

The spectral spread is calculated as described in Eq. (25)

$$\mathrm{Spread} = \sqrt{\frac{\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}}(\mathrm{f_k} - \mu_1)^2\,\mathrm{S_k}}{\sum_{\mathrm{k}=\mathrm{b_1}}^{\mathrm{b_2}} \mathrm{S_k}}} \tag{25}$$

The spectral roll-off point (i) is calculated as described in Eq. (26)

$$\sum_{k=b_1}^{i} S_k = \kappa \sum_{k=b_1}^{b_2} S_k \tag{26}$$

Where κ is the percentage of the total energy contained between b_1 and i.
The spectral slope is calculated as described in (27)

$$slope = \frac{\sum_{k=b_1}^{b_2}(f_k - \mu_f)(S_k - \mu_S)}{\sum_{k=b_1}^{b_2}(f_k - \mu_f)^2} \tag{27}$$

Where μ_s is the mean spectral value.
The spectral decrease is calculated as described in (28)

$$decrease = \frac{\sum_{k=b_1+1}^{b_2} \frac{S_k - S_{b_1}}{k-1}}{\sum_{k=b_1+1}^{b_2} S_k} \quad (28)$$

6.3 Mel-frequency Cepstral Coefficients (MFCC) features

A representation of a sound's short-term power spectrum used in sound processing is called a Mel-Frequency Cepstrum (MFC), which is based on a nonlinear Mel scale frequency log power spectrum linear cosine transform.
An MFC is made up of a number of coefficients known as Mel-Frequency Cepstral Coefficients (MFCCs) [37]. They were created using an audio clip's cepstral representation (a nonlinear "spectrum-of-a-spectrum"). When used in audio compression, for instance, this frequency warping can improve the representation of sound and potentially lower the transmission bandwidth and storage needs of audio signals.
MFCC is used extensively recently for feature extraction and classification of heart sounds [8], [38]. MEL frequency is related to linear frequency in Eq.(29).

$$\text{Mel(f)} = 2595 \log_{10}\left(1 + \frac{\text{f}}{700}\right) \quad (29)$$

As shown in Figure 17, the signal-to-noise ratio will be improved as the PCG signal is pre-weighted. The segmented PCG signals will be divided into frames using an overlapping window of around 30 ms considering 20 ms for the frame-blocking stage. To avoid parasitic spectral leakage, a hamming window of the length of 1323 samples was chosen by setting the sampling frequency to 44.1 kHz.

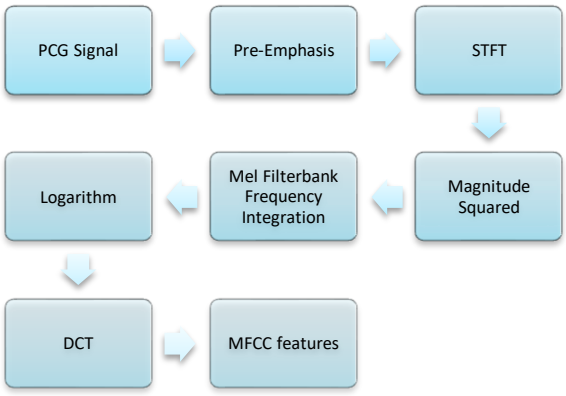


Figure 17: The process of MFCC feature extraction.

Fast Fourier transform (FFT) is applied to the data to transform each segment to its frequency domain version. After the PCG signal is segmented, the selected best segment is transformed into the Mel inverse spectrum domain after being filtered by a set of band-pass Mel triangular filters. The higher band of the PCG signal is compressed then by the algorithm of the Mel spectrum coefficient extracted from each Mel triangular filter, then DCT transforms the logarithmic Mel spectrum coefficients as illustrated in Eq.(30) [38].

$$c[n] = \sum_{M=0}^{N-1} S[m] \cos\left(\frac{\tau n}{M}\left(m - \frac{1}{2}\right)\right), n = 1,2,..,M \quad (30)$$

Where C is the number of MFCCs, c[n] is the number of cepstral coefficients, and M is the filter bank. Typically, 8 to 13 cepstral coefficients are used in conventional MFCC systems. Since the zeroth coefficient indicates the average log energy of the input signal, which conveys very little speaker-specific information, it is frequently removed. 13 MFCCs were selected for this study.

7 PCG Signal Classification

The last stage of the process is the classification phase which categorizes the features extracted from the PCG signal into 6 classes (Normal, third heart sound S3, Aortic valve stenosis AS, Mitral valve regurgitation MR, Mitral valve stenosis MS, Mitral valve prolapse MVP). The training and verification process uses a split of 70% for training and 30% for verification. This ratio was selected based on multiple studies that demonstrated optimal results for the training sets [8]. 3 kinds of features were extracted from the PCG signal (Time Domain Features (TDF), Frequency Domain Features (FDF), and Mel Frequency Cepstral Coefficients (MFCC)). These were incorporated into the training and verification phase using the following classification techniques: Deep Neural Network (DNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Decision Tree (DT) to measure their accuracy and select the best performing method.
The accuracy of the classification is presented as confusion matrixes which is a performance evaluation method for the success of the machine learning classification. These measures are identified from the classifier's true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values as shown in Figure 18.

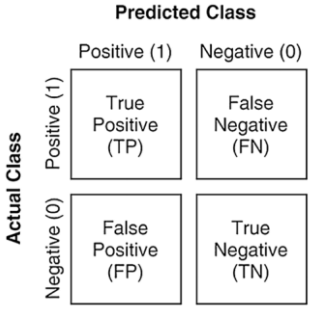


Figure 18: Confusion matrix

7.1 K-nearest neighbors

The idea of the nearest neighbor algorithm is to trace the nearest predetermined parameter of the training samples to a new point end and estimate its label. A constant (k) can be defined in K-nearest neighbor learning KNN, or it can be related to its local density of points.
KNN is a non-generalized machine learning which is calculated by a simple majority vote of the other nearest neighbors and points, where a query point is assigned the class with the most representative of the nearest neighbors. [39]. If the feature representation values go on vastly different scales, then the normalization of the training data improves its accuracy dramatically [16]. For optimum performance, a K value of 1 was selected. All other features extracted from beats were standardized. The result confusion matrices for time, frequency, and cepstral domains are shown in Figure 19, Figure 20, And Figure 21 respectively.

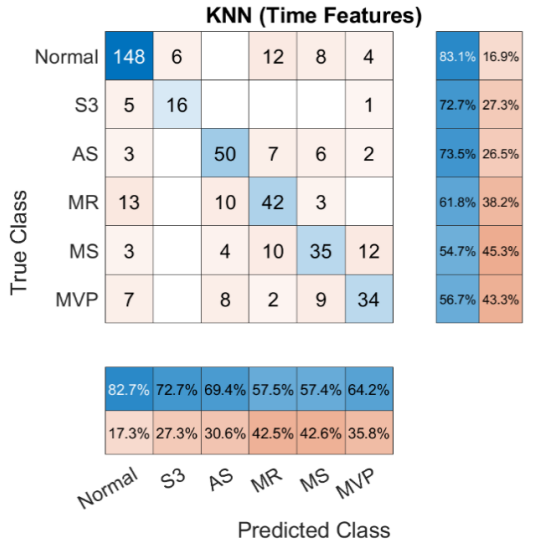


Figure 19: KNN classification with Time domain features

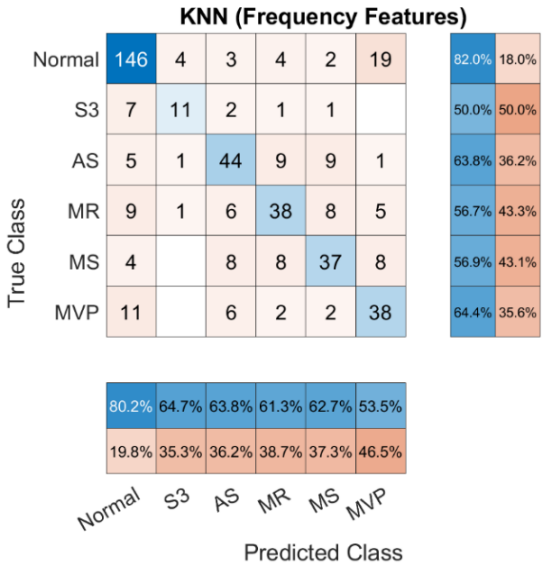


Figure 20: KNN classification with frequency domain features

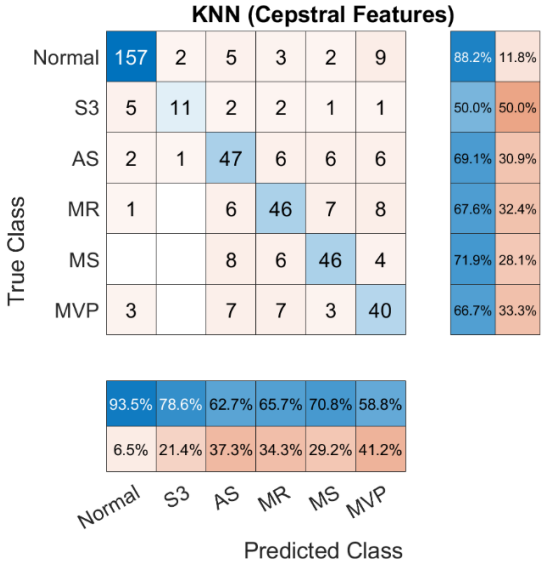


Figure 21: KNN classification with cepstral domain features

7.2 Support vector machines

Support Vector Machine (SVM) is considered one of the main methods used in myoelectric signals. The function of the SVM is to categorize an n-dimensional hyperplane and categorize a set of input features into classes. SVM works by data mapping to a high-dimensional feature space so that it categorizes the data points, even if the data is not linearly separable. The data will be changed if a separator between the categories is found so that the separator may be drawn as a hyperplane [14]. A linear kernel function was used in SVM to distinguish between multiple classes of beats extracted from PCG signals. One vs One technique for multiclass SVM was employed. The result confusion matrices for time, frequency, and cepstral domains are shown in Figure 22, Figure 23, and Figure 24 respectively.

7.3 Decision Tree

A Decision Tree (DT) is a method for supervised machine learning that is used for both regression and classification. It divides the dataset into smaller subsets while associated DT is incrementally developed.

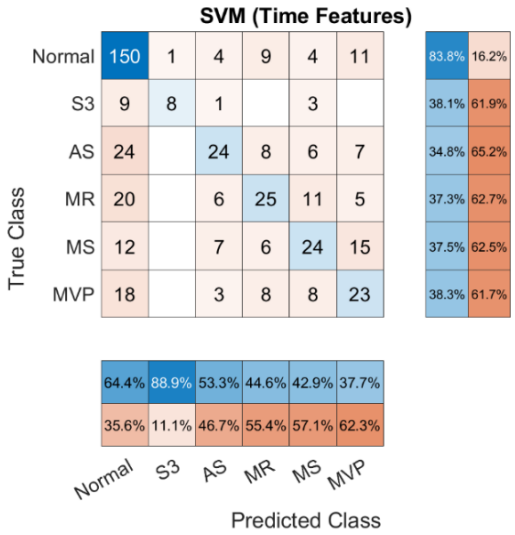


Figure 22: SVM classification with Time domain features

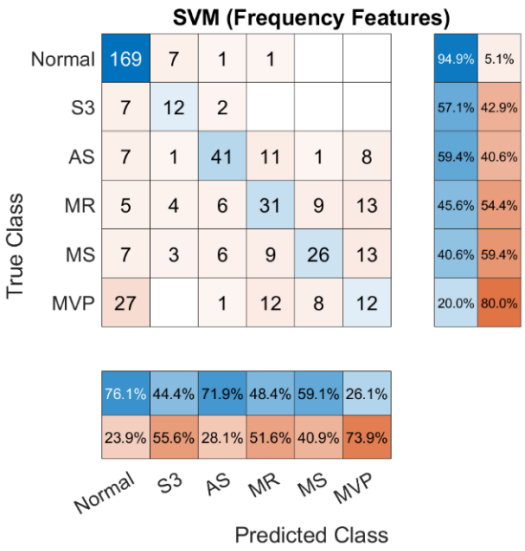


Figure 23: SVM classification with frequency domain features

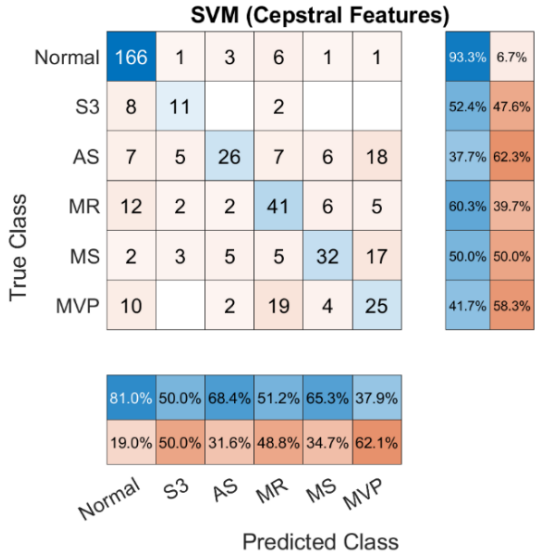


Figure 24: SVM classification with cepstral domain features

In the construction of the decision tree algorithm, we used the features extracted from the training dataset, which we know their feature attributes (independent variables) and the classifying attribute (dependent variable). The best attribute and condition for splitting the dataset are determined, taking into consideration the classifying attribute to get the highest classification accuracy [15]. The maximum number of splits was set to 100 while developing the decision trees. The result confusion matrices for time, frequency, and cepstral domains are shown in Figure 25, Figure 26, and Figure 27 respectively.

7.4 Deep Neural Network

Deep Neural Network (DNN) learns by communicating with other nodes (neurons) in multi-layered structures. We can categorize the data, spot trends, and predict upcoming events thanks to DNN. DNN separates the input data into layers of abstraction with several levels and varying weights for each element. These weights can be automatically changed throughout training until the DNN successfully completes the required aim [14].

A nonlinear mapping is carried out by the neurons in the neural network's layers one after the other. With the use of backpropagation algorithms and a weighted sum of neurons, this mapping enables network learning [25].

Activation functions are used to learn non-linear and complex functional mappings between the inputs and the response variable in an artificial neural network. activation function should be differentiable so that backpropagation optimization can be performed in the network while computing gradients of error (loss) concerning weights, to optimize weights and reduce errors. [40]

Many types of activation functions are available in the DNN, however, in this work, we consider three of the most popular non-linear activation functions that are Sigmoid, Tangent Hyperbolic (tanh), and Rectified Linear Units (ReLU).

The DNN was configured to 5 layers considering 100 Neurons within each layer, and the activation function of Tanh was employed. The total number of iterations limit is set to 1000, and features were standardized before processing. The result confusion matrices for time, frequency, and cepstral domains are shown in Figure 28, Figure 29, and Figure 30 respectively.

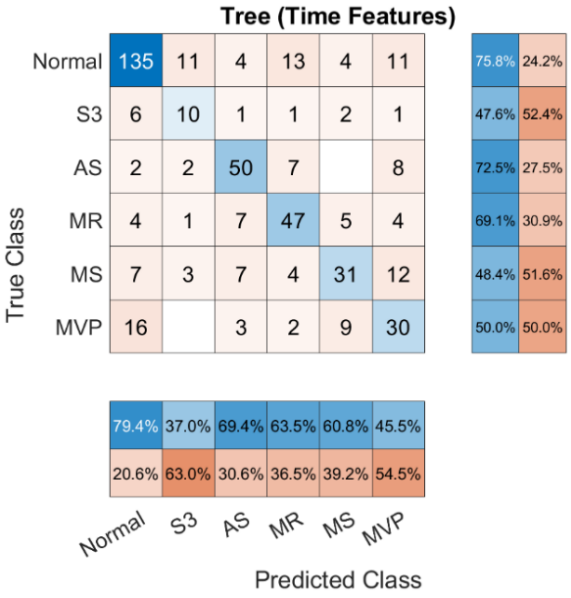


Figure 25: Decision Tree with Time domain features

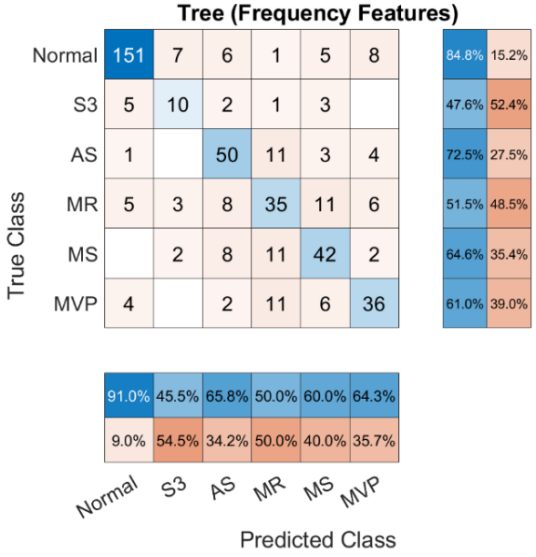


Figure 26: Decision Tree with frequency domain features

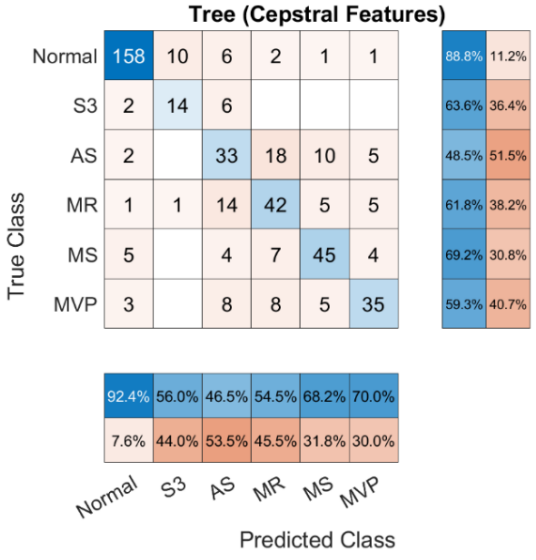


Figure 27: Decision Tree with cepstral domain features

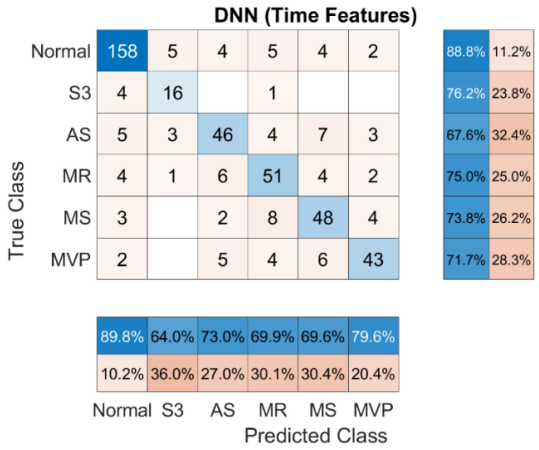


Figure 28: ANN classification with Time domain features

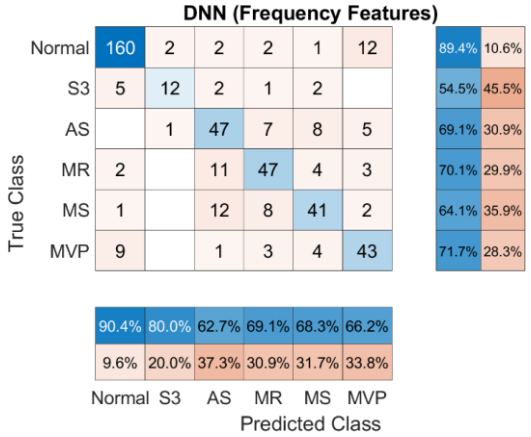


Figure 29: ANN classification with frequency domain features

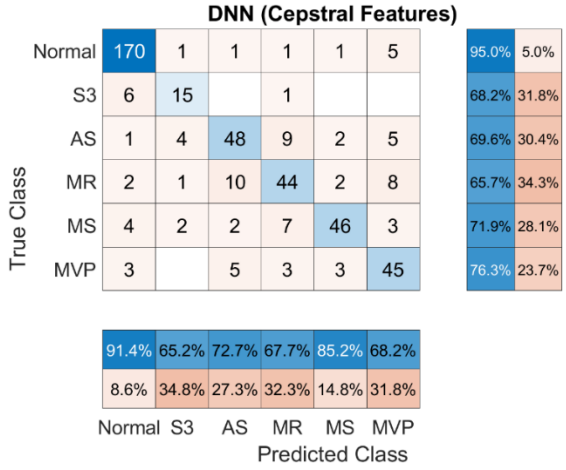


Figure 30: ANN classification with cepstral domain features

7.5 Classification Results

There are two objectives for this work:

1. To detect anomalies (differentiate normal from abnormal heartbeat)
2. To classify the anomalies

Considering these objectives, on the one hand, the accuracy of anomaly detection can be calculated by averaging the true positive and false negative results for the normal class. On the other hand, anomaly classification accuracy can be calculated by averaging the true positive and false negative for other classes.

The summary table (Table 4) shows that the anomaly detection accuracy and the anomaly classification accuracy for the DNN using Cepstral features are the highest, with 93.2% accuracy for anomaly detection and 71.07% accuracy for anomaly classification.

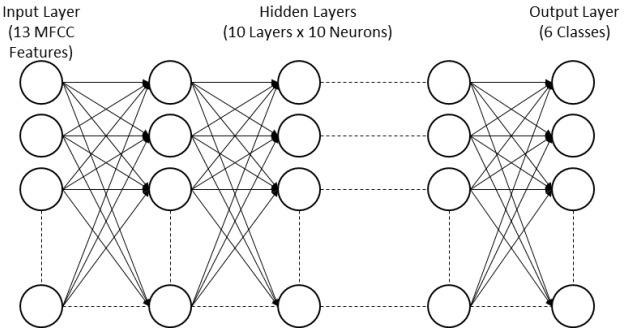


Figure 31: DNN Architecture

The findings indicate that, when compared to alternative classification and feature methods, DNN employing Cepstral features provides the highest accuracy for anomaly detection and anomaly classification. The DNN's following parameters were adjusted further by changing them: The activation function, the number of hidden layers (3, 5, and 10), the number of neurons in each hidden layer (10, 50, and 100), and (tanh, sigmoid, and ReLU). Figure 32 illustrates the process of the proposed system.

7 RESULTS SUMMARY

To evaluate the system's effectiveness, extensive system testing and training were carried out, using 70% of the data for training and 30% for testing. The DNN was modified as indicated below:

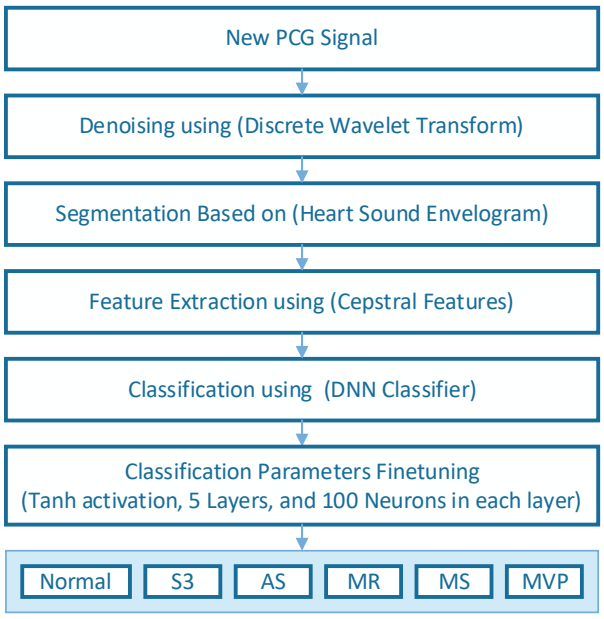


Figure 32: Final Proposed Methodology

Stage 1: Select the ideal activation function from (Sigmoid, Tanh, and ReLU) using preset parameters for the number of DNN hidden layers and neurons in each layer (5 layers and 100 neurons within each layer)

Stage 2: Select the number of neurons with fixed activation functions that work well (Tanh performed best in stage 1).
In stage 3, pick the number of hidden layers using the stage 1 and stage 2 parameters that perform best. (100 Neurons and Tanh's activation function)
The Tanh activation function, 5 hidden layers, and 100 Neurons within each layer are the DNN settings that gave the greatest performance for anomaly detection and classification as given in Table 6. Additionally, 10 simulations were run with this arrangement, and the results of the multiple tests were averaged for better accuracy as given in
Table 5. In terms of anomaly detection and classification rates, the results showed average performances of over 92.14% and 71.02%, respectively.

8 CONCLUSION

Five common heart murmurs, the third heart sound (S3), aortic stenosis (AS), mitral regurgitation (MR), mitral stenosis (MS), and mitral valve prolapse (MVP), were the focus of an efficient Phonocardiography (PCG) based system that was designed and tested for the early detection and classification of cardiovascular anomalies.
The preliminary findings point to average rates for detection and classification of heart anomalies of approximately 92% and 71%, respectively. This was accomplished by utilizing a variety of filtering, segmentation, feature extraction, and deep learning approaches in comparison to the PCG signals.
It was found that better outcomes were obtained utilizing a 5-layer Deep Neural Network (DNN) with 100 neurons per layer and the Hyperbolic Tangent (tanh) as the activation function.
The best methods for segmenting and denoising the PCG signal, respectively, were the Discrete Wavelet Transform (DWT) and Heart Sound Envelogram (HSE). The Mel Frequency Cepstral Coefficients (MFCC) Features were utilized to train and execute the DNN since they performed better than their Frequency and Time Domain equivalents.
While a more expensive option, such as high-quality cellphones or digital electronics stethoscopes, can be utilized to provide higher-quality PCG signals, a low-cost but noisy modified conventional stethoscope was employed as the input device in this study. As the COVID limits are altered, more data about a larger variety of heart murmurs could be collected thanks to easier access to heart patients in hospitals. The suggested system is anticipated to perform better with extra data and further modification of the DNN parameters, to classify more heart murmurs with higher accuracy.
The proposed system's success in creating early warnings and preliminary indications to potentially ill people and directing them to specialized cardiologists for more precise diagnostics—despite the fact that it was not intended to be a primary tool for the diagnosis—was demonstrated by its relatively high detection rate. In the context of preventive public health systems, this should be a useful home-care tool.

Table 4. Classification Results Summary

	True Positive						True Negative							
	Normal	S3	AS	MR	MS	MVP	Normal	S3	AS	MR	MS	MVP	Anomaly Detection Accuracy	Anomaly Classification Accuracy
KNN Time	83.1	72.7	73.5	61.8	54.7	56.7	82.7	72.7	69.4	57.5	57.4	64.2	82.90	64.06
KNN Frequency	82.0	50.0	63.8	56.7	56.9	64.4	80.2	64.7	63.8	61.3	62.7	53.5	81.10	59.78
KNN Cepstral	88.2	50.0	69.1	67.6	71.9	66.7	93.5	78.6	62.7	65.7	70.8	58.8	90.85	66.19
SVM Time	83.8	38.1	34.8	37.3	37.5	38.3	64.4	88.9	53.3	44.6	42.9	37.7	74.10	45.34
SVM Frequency	94.9	57.1	59.4	45.6	40.6	20.0	76.1	44.4	71.9	48.4	59.1	26.1	85.50	47.26
SVM Cepstral	93.3	52.4	37.7	60.3	50.0	41.7	81.0	50.0	68.4	51.2	65.3	37.9	87.15	51.49
DT Time	75.8	47.6	72.5	69.1	48.4	50.0	79.4	37.0	69.4	63.5	60.8	45.5	77.60	56.38
DT Frequency	84.8	47.6	72.5	51.5	64.6	61.0	91.0	45.5	65.8	50.0	60.0	64.3	87.90	58.28
DT Cepstral	88.8	63.6	48.5	61.8	69.2	59.3	92.4	56.0	46.5	54.5	68.2	70.0	90.60	59.76
DNN Time	88.8	76.2	67.6	75.0	73.8	71.7	89.8	64.0	73.0	69.9	69.6	79.6	89.30	72.04
DNN Frequency	89.4	54.5	69.1	70.1	64.1	71.7	90.4	80.0	62.7	69.1	68.3	66.2	89.90	67.58
DNN Cepstral	95.0	68.2	69.6	65.7	71.9	76.3	91.4	65.2	72.7	67.7	85.2	68.2	93.20	71.07

Table 5. Multiple Testing using Best-Performing DNN Parameters

					True Positive						True Negative					
			Anomaly Detection Accuracy	Anomaly Classification Accuracy	Normal	S3	AS	MR	MS	MVP	Normal	S3	AS	MR	MS	MVP
Layers	5	Test 01	92.30%	72.32%	93.3%	52.4%	76.5%	63.2%	84.6%	64.4%	91.3%	100.0%	68.4%	70.5%	75.3%	67.9%
Neurons	100	Test 02	92.30%	70.95%	91.0%	77.3%	70.6%	59.7%	80.0%	73.3%	93.6%	65.4%	69.6%	70.2%	67.5%	75.9%
Activation	Tanh	Test 03	93.30%	72.14%	93.3%	59.1%	69.6%	70.1%	83.1%	72.9%	93.3%	76.5%	71.6%	69.1%	74.0%	75.4%
Features	Cepstral	Test 04	91.75%	71.46%	90.4%	81.8%	72.1%	61.2%	83.1%	66.7%	93.1%	62.1%	68.1%	82.0%	74.0%	63.5%
		Test 05	93.20%	71.07%	95.0%	68.2%	69.6%	65.7%	71.9%	76.3%	91.4%	65.2%	72.7%	67.7%	85.2%	68.2%
		Test 06	91.15%	70.61%	93.9%	68.2%	65.2%	67.2%	78.1%	62.7%	88.4%	83.3%	66.2%	72.6%	71.4%	71.2%
		Test 07	91.80%	69.36%	90.5%	52.4%	72.1%	67.6%	81.5%	74.6%	93.1%	57.9%	71.0%	68.7%	76.8%	71.0%
		Test 08	93.25%	71.15%	92.7%	76.2%	60.3%	63.2%	84.6%	76.7%	93.8%	66.7%	68.3%	69.4%	76.4%	69.7%
		Test 09	91.20%	70.51%	92.7%	63.6%	66.7%	76.1%	73.4%	67.8%	89.7%	66.7%	67.6%	70.8%	79.7%	72.7%
		Test 10	91.10%	70.62%	92.1%	63.6%	62.3%	74.6%	73.8%	71.2%	90.1%	77.8%	61.4%	73.5%	71.6%	76.4%
		Average	92.14%	71.02%	92.5%	66.3%	68.5%	66.9%	79.4%	70.7%	91.8%	72.2%	68.5%	71.5%	75.2%	71.2%

Table 6. DNN Parameters Selection Test Results

Testing Different Activation Methods

Layers	5
Neurons	100

	Training Time (s)	True Positive						True Negative						Anomaly Detection Accuracy	Anomaly Classification Accuracy
		Normal	S3	AS	MR	MS	MVP	Normal	S3	AS	MR	MS	MVP		
Relu	112	92.2%	63.6%	58.8%	65.7%	71.9%	60.0%	92.7%	56.0%	58.0%	62.9%	69.7%	69.2%	92.45%	63.58%
Sigmoid	155	90.5%	66.7%	60.9%	70.1%	75.0%	70.0%	89.5%	60.9%	72.4%	58.0%	77.4%	76.4%	90.00%	68.78%
Tanh	83	93.3%	59.1%	69.6%	70.1%	83.1%	72.9%	93.3%	76.5%	71.6%	69.1%	74.0%	75.4%	93.30%	72.14%

Testing Different Number of Neurons

Layers	5
Activation	Tanh

	Training Time (s)	True Positive						True Negative						Anomaly Detection Accuracy	Anomaly Classification Accuracy
		Normal	S3	AS	MR	MS	MVP	Normal	S3	AS	MR	MS	MVP		
10 Neurons	35	92.1%	63.6%	73.5%	73.5%	75.4%	66.1%	94.8%	60.9%	60.2%	68.5%	86.0%	76.5%	93.45%	70.42%
50 Neurons	52	90.5%	54.5%	72.1%	68.7%	68.8%	58.3%	92.0%	52.2%	61.3%	63.9%	72.1%	72.9%	91.25%	64.48%
100 Neurons	83	93.3%	59.1%	69.6%	70.1%	83.1%	72.9%	93.3%	76.5%	71.6%	69.1%	74.0%	75.4%	93.30%	72.14%

Testing Different Number of Hidden Layers with 100 Neurons

Neurons	100
Activation	Tanh

	Training Time (s)	True Positive						True Negative						Anomaly Detection Accuracy	Anomaly Classification Accuracy
		Normal	S3	AS	MR	MS	MVP	Normal	S3	AS	MR	MS	MVP		
3 Layers	43	91.0%	50.0%	69.6%	54.4%	73.4%	69.5%	87.6%	52.4%	58.5%	71.2%	75.8%	70.7%	89.30%	64.55%
5 Layers	83	93.3%	59.1%	69.6%	70.1%	83.1%	72.9%	93.3%	76.5%	71.6%	69.1%	74.0%	75.4%	93.30%	72.14%
10 Layers	244	89.3%	61.9%	63.2%	72.1%	81.5%	65.0%	88.8%	81.2%	68.3%	64.5%	71.6%	75.0%	89.05%	70.43%

REFERENCES

[1] P. Hult, T. Fjällbrant, K. Hildén, U. Dahlström, B. Wranne, and P. Ask, "Detection of the third heart sound using a tailored wavelet approach: Method verification," *Med Biol Eng Comput*, vol. 43, no. 2, 2005, doi: 10.1007/BF02345957.

[2] M. E. Tavel, "Cardiac auscultation: A glorious past - But does it have a future?," *Circulation*, vol. 93, no. 6. 1996. doi: 10.1161/01.CIR.93.6.1250.

[3] F. Agrafioti and D. Hatzinakos, "ECG based recognition using second order statistics," in *Proceedings of the 6th Annual Communication Networks and Services Research Conference, CNSR 2008*, 2008. doi: 10.1109/CNSR.2008.38.

[4] S. Ismail, I. Siddiqi, and U. Akram, "Localization and classification of heart beats in phonocardiography signals —a comprehensive review," *Eurasip Journal on Advances in Signal Processing*, vol. 2018, no. 1. 2018. doi: 10.1186/s13634-018-0545-9.

[5] R. Oshana, "Overview of Digital Signal Processing Algorithms," in *DSP Software Development Techniques for Embedded and Real-Time Systems*, 2006. doi: 10.1016/b978-075067759-2/50006-5.

[6] O. Deperlioglu, "Segmentation of Heart Sounds by Re-Sampled Signal Energy Method," *Brain (Bacau)*, vol. 9, no. 1, 2018.

[7] L. H. Cherif, M. Mostafi, and S. M. Debbal, "Digitals filters in heart sound analysis," *International Journal of clinical Medicine Research*, vol. 1, no. 3, 2014.

[8] 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, 2021, doi: 10.1016/j.compbiomed.2021.104926.

[9] A. W. Galli, G. T. Heydt, and P. F. Ribeiro, "Exploring the power of wavelet analysis," *IEEE Computer Applications in Power*, vol. 9, no. 4, 1996, doi: 10.1109/67.539845.

[10] N. Kouras, D. Boutana, and M. Benidir, "Wavelet based segmentation and time-frequency caracterisation of some abnormal heart sound signals," in *Proceedings of the International Conference on Microelectronics, ICM*, 2012. doi: 10.1109/ICM.2012.6471392.

[11] V. N. Varghees and K. I. Ramachandran, "A novel heart sound activity detection framework for automated heart sound analysis," *Biomed Signal Process Control*, vol. 13, no. 1, 2014, doi: 10.1016/j.bspc.2014.05.002.

[12] O. K. Fasil and R. Rajesh, "Time-domain exponential energy for epileptic EEG signal classification," *Neurosci Lett*, vol. 694, 2019, doi: 10.1016/j.neulet.2018.10.062.

[13] J. Singh and R. S. Anand, "Computer aided analysis of phonocardiogram," *J Med Eng Technol*, vol. 31, no. 5, 2007, doi: 10.1080/03091900500282772.

[14] H. Basak, A. Roy, J. B. Lahiri, S. Bose, and S. Patra, "SVM and ANN based Classification of EMG signals by using PCA and LDA," Oct. 2021, [Online]. Available: http://arxiv.org/abs/2110.15279

[15] S. A. Pavlopoulos, A. C. H. Stasis, and E. N. Loukis, "A decision tree - Based method for the differential diagnosis of Aortic Stenosis from Mitral Regurgitation using heart sounds," *Biomed Eng Online*, vol. 3, 2004, doi: 10.1186/1475-925X-3-21.

[16] Ö. Arslan and M. Karhan, "Effect of Hilbert-Huang transform on classification of PCG signals using machine learning," *Journal of King Saud University - Computer and Information Sciences*, 2022, doi: 10.1016/j.jksuci.2021.12.019.

[17] F. Safara, S. Doraisamy, A. Azman, A. Jantan, and S. Ranga, "Wavelet packet entropy for heart murmurs classification," *Adv Bioinformatics*, vol. 2012, 2012, doi: 10.1155/2012/327269.

[18] C. N. Gupta, R. Palaniappan, S. Swaminathan, and S. M. Krishnan, "Neural network classification of homomorphic segmented heart sounds," *Applied Soft Computing Journal*, vol. 7, no. 1, 2007, doi: 10.1016/j.asoc.2005.06.006.

[19] S. Choi, "Detection of valvular heart disorders using wavelet packet decomposition and support vector machine," *Expert Syst Appl*, vol. 35, no. 4, 2008, doi: 10.1016/j.eswa.2007.08.078.

[20] T. R. Reed, N. E. Reed, and P. Fritzson, "Heart sound analysis for symptom detection and computer-aided diagnosis," in *Simulation Modelling Practice and Theory*, 2004. doi: 10.1016/j.simpat.2003.11.005.

[21] A. Hasan and Z. Bahrai, "Phonocardiography (PCG)-Based Early Detection of Heart Anomalies Using Deep Learning ," Isa Town: EDAS IET Smart Cities Symposium 6-8 Dec. 2022, Dec. 2022, pp. 1–3.

[22] R. Pechetty and L. Nemani, "Additional Heart Sounds—Part 1 (Third and Fourth Heart Sounds)," *Indian J Cardiovasc Dis Women WINCARS*, vol. 5, no. 02, 2020, doi: 10.1055/s-0040-1713828.

[23] S. McGee, *Evidence-Based Physical Diagnosis: Fourth Edition*. 2017.

[24] P. Bentley, "http://www.peterjbentley.com/heartchallenge/."

[25] Y. Khan, "https://github.com/yaseen21khan/Classification-of-Heart-Sound-Signal-Using-Multiple-Features-."

[26] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering*. 2019. doi: 10.1017/9781108380690.

[27] P. S. Addison, *The illustrated wavelet transform handbook*. 2017.

[28] "https://www.mathworks.com/help/wavelet/gs/choose-a-wavelet.html."

[29] P. Oktivasari, F. Haryanto, S. A. Hamidah, R. Rian-Dini, and S. Suprijadi, "A real-time heart rate signal detection using an electronic stethoscope with labview," *J Biomed Phys Eng*, vol. 10, no. 3, 2020, doi: 10.31661/jbpe.v0i0.1183.

[30] "https://www.electronics-tutorials.ws/filter/filter_8.html."

[31] L. Weinberg and P. Slepian, "Takahasi's Results on Tchebycheff and Butterworth Ladder Networks," *IRE Transactions on Circuit Theory*, vol. 7, no. 2, 1960, doi: 10.1109/TCT.1960.1086643.

[32] A. Rehman, N. A. Saqib, S. M. Danial, and S. H. Ahmed, "ECG based authentication for remote patient monitoring in IoT by wavelets and template matching," in *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, 2018. doi: 10.1109/ICSESS.2017.8342871.

[33] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," in *Computers in Cardiology*, 1997. doi: 10.1109/cic.1997.647841.

[34] R. J. Lehner and R. M. Rangayyan, "A Three–Channel Microcomputer System for Segmentation and Characterization of the Phonocardiogram," *IEEE Trans Biomed Eng*, vol. BME-34, no. 6, 1987, doi: 10.1109/TBME.1987.326060.

[35] P. Careena, M. Mary Synthuja Jain Preetha, and P. Arun, "Detection of murmur from non-stationarity of heart sounds," *Biomedical and Pharmacology Journal*, vol. 12, no. 2, 2019, doi: 10.13005/bpj/1721.

[36] M. U. Khan, S. Aziz, K. Iqtidar, G. F. Zaher, S. Alghamdi, and M. Gull, "A two-stage classification model integrating feature fusion for coronary artery disease detection and classification," *Multimed Tools Appl*, vol. 81, no. 10, 2022, doi: 10.1007/s11042-021-10805-3.

[37] Y. M. R. Huang *et al.*, "Advances in Multimedia Information Processing - PCM 2008: Preface," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5353 LNCS. 2008.

[38] S. Aziz, M. U. Khan, M. Alhaisoni, T. Akram, and M. Altaf, "Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features," *Sensors (Switzerland)*, vol. 20, no. 13, 2020, doi: 10.3390/s20133790.

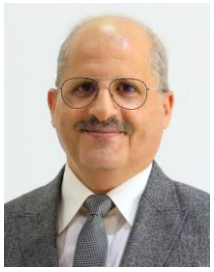
[39] B. Kamgar-Parsi and L. N. Kanal, "An improved branch and bound algorithm for computing k-nearest neighbors," *Pattern Recognit Lett*, vol. 3, no. 1, 1985, doi: 10.1016/0167-8655(85)90036-4.

[40] S. Gupta, R. A. Ansari, and D. Sarkar, *Deep learning with R cookbook*.

BIOGRAPHIES

Abbas Hasan received his B.Sc. degree in Electronics Engineering from the University of Bahrain in 2007. Building upon a rich industrial career with 10 years in ALBA and 10 years in Pipeline Technology (KSA) and served as project manager for several oil and gas pipeline leak detection systems.

Abbas is currently pursuing his MSc degree in Artificial Intelligence at the University of Bahrain. His research interests include Artificial Intelligence and Machine Learning.



Zouhir Bahri Completed his BSc in Electrical Engineering from the University of Pittsburgh (1985) and his MSc and PhD from Carnegie Mellon University (CMU) in Electrical and Computer Engineering (1986 and 1989) with Summa Cum Laude distinction.

Since 1989 he has been with the Electrical and Electronics Engineering Department at the University of Bahrain. His current research interests include Signal Processing and Artificial Intelligence for Communication and Biomedical Engineering and is a reviewer for several international journals.

Dr. Bahri has pioneered the establishment of UOB's Electrical & Electronics Engineering Dept. which he is currently chairing. He has been on the technical and organizational committees of several local and international confer