Application of Signal Processing and Deep Hybrid Learning in Phonocardiogram and Electrocardiogram Signals to Detect Early Stage Heart Diseases

by

Md Tanzil Hoque Chowdhury

A Dissertation

Presented to the Faculty of the Computational Science Program

Middle Tennessee State University

March 2022

In Partial Fulfillment

of the Requirements for the Degree

Doctorate of Philosophy in Computational Science

Dissertation Committee:

Dr. Cen Li, Chair

Dr. Yating Hu

Dr. William Robertson

Dr. Qiang Wu

To my parents, wife and sister.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Prof. Dr. Cen Li, for the continuous support on my Ph.D study and research, her motivation, patience, and immense knowledge. I am also thankful to my previous advisor Dr. Yating Hu for her support and advise on my research. I would like to thank the rest of my thesis committee members: Dr. William Robertson and Dr. Qiang Wu for their insightful comments and encouragement on my dissertation. I would like to thank Middle Tennessee State University and the Computational Science program for funding my graduate studies and research. My sincere thanks also goes to Dr. John F. Wallin who has supported me during each and every academic success and difficult times of my PhD program. I would like to thank Dr. Khem Narayan Poudel for his full support throughout my PhD program. Last but not the least, I would like to thank my parents, wife, and sister for providing unending support and inspiration.

ABSTRACT

This thesis describes a variety of projects on analyzing two common biomedical signals known as Phonocardiogram (PCG) and Electrocardiogram (ECG/EKG) to detect early-stage heart diseases. The projects include the design of prototypes to compress, denoise, segment, and classify PCG and ECG signals accurately. PCG signal is the graphical representation of heart sound which represents the mechanical activities of the human heart. PCG signal contains useful information about the functionality and the condition of the heart. ECG signal represents the electrical activities of the human heart. ECG signal has been widely used in hospitals and clinics to diagnose cardiac diseases. Analysis of PCG and ECG signals is critical in diagnosis of different cardiac diseases as they can provide early indication of potential cardiac abnormalities. Extracting cardiac information from PCG and ECG signals to diagnose heart diseases in the initial stage can play a vital role in remote patient monitoring. In this thesis, we have combined different signal processing techniques, Machine Learning (ML), and Deep Learning (DL) methods to compress, denoise, segment, and classify PCG and ECG signals effectively and accurately. First, PCG signals are compressed and denoised by using a multi-resolution analysis technique based on the Discrete Wavelet Transform (DWT). Then, a segmentation algorithm, based on the Shannon energy envelope and zero crossing is applied to segment the PCG signal into four major parts: the first heart sound (S1), the systole interval, the second heart sound (S2), and the diastole interval. Finally, Mel-scaled power spectrogram and Mel-frequency cepstral coefficients (MFCC) are employed to extract informative features from PCG signals, which are then fed into a classifier to classify each PCG signal into a normal or an abnormal signal. We have combined

traditional ML and DL approaches to develop Deep Hybrid Learning (DHL) models. A Convolutional Neural Network (CNN) is used along with seven traditional ML methods including Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), Decision Tree (DT), Naive Bayes (NB), Support Vector Machine (SVM), and AdaBoost (AB) to build hybrid PCG classification models. Our experimental results have shown that significant improvements in the classification accuracy can be achieved by using DHL models compared to traditional ML and DL models. We have also applied the same methods to analyze ECG signals and got promising results. Besides providing valuable information regarding heart condition, our proposed signal processing and DHL approaches can help cardiologists to take appropriate and reliable steps toward diagnosis if any cardiovascular disorder is found in the initial stage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Heart failure is the one of the leading causes of death worldwide. Every year millions of people in the world die due to heart diseases. Heart diseases are also the main reason of death in the United States. Each year around 6,55,000 person die in the United States due to cardiac diseases which is equal to 1 in every 4 deaths [1]. Many of these deaths could be prevented if it was possible to detect cardiac diseases in the initial stage. Continuous patient monitoring can play a vital role to detect cardiac abnormalities in the initial stage. The traditional health care facilities require to schedule appointments in the clinics or hospitals for preventive health check-ups. Though this model can successfully diagnose diseases in the early stage, it fails to treat any chronic disease which has already progressed to the end stage. Moreover, it is not always possible to detect diseases in the initial stage as accompanying symptoms may be present for a brief amount of time which may lead to end-stage of a disease that cannot be cured. So, the solution of this problem is continuous remote monitoring of ambulatory patients using wearable sensors. This process can continuously collect data from human body and provide a complete information about patient's health. Similarly, to determine the cause of cardiac abnormalities, a patient's heart must be monitored continuously, so his or her physician can diagnose the disorder accurately. This is very important as patients can know the condition of their heart continuously and they can consult with cardiologist instantly if any anomaly occurs.

In this thesis, we have examined two common biomedical signals known as PCG and ECG. These two signals play a vital role to detect cardiovascular diseases in the initial stage. PCG is the graphical depiction of the heart sound signal that illustrates the condition of the heart through intensity, frequency, time duration and other valuable information regarding the heart valves [2]. Identifying pathological symptoms by hearing the heart sound through the stethoscope is a very difficult skill and can take a long time to gain and to be proficient in it. To use heart sounds for continuous monitoring of heart, it is very important to track heart sound components repeatedly over a period of time. But it's not possible for doctors to listen heart sounds continuously for 24 hours. Moreover, human ear has the limitation in hearing the sounds. Human ear can hear the sounds within the frequency range of 20 Hz to 20 KHz and can discriminate lower frequencies less than higher frequencies. The frequencies of the heart sound components are from 20 Hz to 500 Hz. Therefore, it's very difficult to hear low frequency components of heart sound having dominant frequencies of less than 100 Hz [3]. So, we need more objective tools like PCG to extract informative characteristics of the heart sound that can't be identified by the human ear. PCG helps to analyze heart sounds and to detect abnormalities in the heart, thereby improves overall diagnosis efficiency [4–7]. On the other hand, ECG is the most widely used and powerful diagnostic tool to analyze the electrical activity of the heart and to detect cardiac abnormalities. ECG signal is a recording of the bio-electric potential produced by rhythmical cardiac activities. ECG has been extensively used for heart disease diagnosis in hospitals, as well as patient monitoring at home, since it can provide valuable information regarding the functional condition of the heart [8]. The functionality of the human heart can be properly monitored, and cardiovascular diseases can be detected through analyzing ECG

signals. Early detection of cardiac arrhythmia is also possible by analyzing ECG heartbeats continuously. Millions of people in the world are suffering from cardiac arrhythmia which refers to irregular heartbeats. Irregular heartbeats can be life-threatening in some cases if it is not detected in the initial stage. So, accurate detection of irregular heartbeats in the primary stage is very important to decrease the mortality rate caused by heart diseases [12]. Thus, PCG and ECG signals have considerable interest in the hospital and health community to improve the cardiac patient survival rate and decrease their hospitalized rate.

In biomedical signal processing, the aim is to extract clinically relevant pathological information from raw signals to enhance the medical diagnosis process. Every organ of the human body delivers pathological signals. These signals can be electrical, mechanical, or chemical. The main task of signal processing in biomedical research is to pull out relevant clinical information from those signals and filter out noise and redundant information. Different signal processing, ML, and DL techniques are available in the scientific literature to extract pathological information from PCG and ECG signals for the proper clinical diagnosis. In continuous remote patient monitoring through PCG and ECG signals, it is required to analyze signals without loosing any physiological and clinical information. Biomedical signals analysis techniques can be mainly categorized into three methods such as direct method, transformation method, and parameter extraction method. The direct methods only analyze signals in the time domain. The examples of this technique are Amplitude Zone Time Epoch Coding (AZTEC), Improved modified AZTEC technique, Coordinate Reduction Time Encoding system (CORTES), Turning point (TP), and so on. These techniques are relatively old and showed poor performance to analyze non-stationary biomedical signals like ECG and PCG. The transformed methods usually convert time domain to frequency domain and

analyze signals based on their frequency. The examples of transformed method are Fourier descriptor, Fourier Transform (FT), Wavelet Transform (WT), Discrete Cosine Transform (DCT), and so on. In biomedical signals, pathological or clinical information are usually hidden inside the frequency domain. So, transformed techniques usually show better performance compared to direct methods to analyze biomedical signals. The parameter extraction method is dependent on the extraction of dominant features from raw signals. Features can be in the time domain, or in the frequency domain, or the combination of both domain [9]. ML and DL techniques have enhanced the importance of research on biomedical signals. Modern ML techniques are dependent on the feature extraction process. Extracting appropriate features from the raw signal is very challenging. DL algorithms can solve this problem by extracting high-quality optimal features through its own neural network and reduce the need for feature engineering or the parameter extraction process. Thus, DL algorithms lead to better performance and higher accuracy compared to ML algorithms [5, 6, 12]. In this thesis we have applied different signal processing, ML, and DL techniques to analyze PCG and ECG signals. Our proposed signal processing techniques provide significant information regarding the heart condition that helps to detect heart diseases in the primary phase.

1.2 Objective

The main objective of this thesis is to develop a novel framework to continuously monitor the heart. We have combined signal processing, ML, and DL approaches to compress, denoise, encrypt, segment, and classify PCG and ECG signals. Until now, very little research has been done on analyzing PCG and ECG signals which covers all five major techniques while

continuous monitoring of the heart. Besides providing valuable information regarding heart condition, this signal processing approach can help cardiologists take appropriate and reliable steps toward diagnosis if any cardiovascular disorder is found in the initial stage. The major contributions of this thesis are described below.

1.2.1 Compression algorithm

Extracting cardiac information from PCG and ECG signals to detect cardiac abnormalities can play a vital role to diagnose heart diseases in the primary stage. This process can significantly decrease the mortality rate caused by cardiovascular diseases. Therefore, continuous PCG and ECG monitoring are of great interests for remote patient monitoring. However, 24 hours of online monitoring generates a large amount of data to be transferred and stored at healthcare facilities. It is also a very difficult task to get the proper information from PCG and ECG signals, as these signals are usually mixed with different kinds of noise, murmurs, and unnecessary information that may lead to an inaccurate clinical diagnosis. So, before segmentation, it is required to remove these redundant pieces of information from PCG and ECG signals to evaluate the proper functionality of the heart. Moreover, an end-to-end encryption is also required to share the data without compromising privacy or security. So, compression, denoising, and encryption are necessary for the continuous monitoring of PCG and ECG signals [7, 17, 18]. FT and short-time Fourier transform (STFT) are commonly used tools for examining stationary signals. But they show limited performance in examining non-stationary signals as they can't provide simultaneous time and frequency localization [11, 17, 19]. Non-stationary signals like PCG and ECG can be analyzed properly by using DWT, as it provides very good time-frequency localization [18, 20–22]. By using a

multi-resolution analysis technique based on DWT, PCG and ECG signals can be decomposed into different sub-bands having different frequency range. The required sub-bands which contain the valuable clinical information can be picked up for further analysis. The other sub-bands which contain murmurs, noise or other unnecessary information can be discarded [24, 26–28]. Furthermore, these compressed and denoised signals will be encrypted while transferring, so that only the physician or doctor can decode the encrypted signal. An average PCG compression of 93.67% and an average ECG compression of 74.57% were achieved using our proposed compression technique. We compared our proposed compression technique with other compression methods and our compression technique has outperformed most of those techniques by a large margin.

1.2.2 Segmentation algorithm

The duration of heart sounds, systole, and diastole intervals, cardiac cycle, and the number of heart beats per minute are important parameters to determine the cardiac function of a person [26]. To extract these important characteristics from PCG signals, it is required to segment PCG signals properly [29, 31, 58]. For the segmentation, normalized average Shannon energy is used in our research to detect the components of the PCG signal by calculating the envelope of its energy [32, 33]. The zero-crossing algorithm is used to detect the starting and stopping points of the heart sounds [40]. Based on this information we can obtain the duration and amplitude of each heart sound, systole interval, diastole interval, one cardiac cycle, and heartbeat of a person [5]. Thus, it will continuously provide the values of different important cardiac parameters of the heart. This will help patients to know the condition of their heart continuously and they can consult with doctors if they find any

cardiac abnormality during the very early stage of a disease.

1.2.3 Classification algorithm

Due to the significance of PCG and ECG signals to detect cardiac abnormalities, different automatic PCG and ECG classification models have been developed using ML and DL approaches. Recent advancements in ML and DL have played a vital role to detect abnormalities in hearts through PCG and ECG signals. Different ML algorithms such as NB, DT, LR, RF, KNN, SV, and AB are usually used for PCG and heartbeat classification. All of these classification models use feature extraction methods to remove redundant features and to increase the classification accuracy. However, one of the major limitations of ML is the feature extraction. Extracting appropriate features from the raw signal is very challenging. DL algorithms can solve this problem by extracting high-quality optimal features through its own neural network and reduce the need for feature engineering. Thus, DL algorithms lead to better performance and high accuracy compared to ML algorithms [5,6,12]. But DL algorithms also suffer from over fitting problem when the model is fed with unstructured or less data. We have combined the advantages of ML and DL to develop DHL algorithms. Our proposed DHL algorithms can overcome the limitation of traditional ML and DL algorithms. DHL algorithms have less time and computational complexity compared to DL algorithms. DHL algorithms also don't require feature engineering technique like ML algorithms. Mel-scaled power spectrogram and MFCC are used to extract informative features from the PCG signals which are then fed into ML, DL, and DHL classifiers to classify each PCG signal into a normal or an abnormal signal. To classify heartbeats from ECG signals, we used each ECG beat as input to ML, DL, and DHL classification models. Our proposed DHL algorithms

outperformed ML and DL algorithms by a large margin to classify PCG signals and ECG beats. For the classification of PCG signals, our proposed CNN-RF classifier outperformed other classifier by giving an accuracy of 94.30%. For the ECG heartbeat classification, among all the ML, DL, and DHL models, an excellent testing accuracy of 98.60% is achieved using our proposed CNN-RF model. The overall performance of the proposed PCG and ECG classification models are compared with other recent methods on PCG and ECG classification, and we got better result compared to most of the classification techniques in terms of overall accuracy.

1.3 Thesis outline

The organization of the thesis is as follows. Chapter 2 explains the background of heart, PCG, and ECG signals. Chapter 3 presents the proposed compression technique. Chapter 4 demonstrates the proposed segmentation technique. Chapter 5 examines different ML and DL algorithms and also describes the proposed DHL classification algorithms. Chapter 6 describes and analyzes the result of this research. Conclusion and future works are discussed in chapter 7.

CHAPTER 2

HEART, PCG, AND ECG

PCG and ECG are two most vital biomedical signals to evaluate the current condition of the heart. PCG annotates the timing, duration and amplitude of different heart sounds and can detect structural defects of the heart valves by analyzing heart sound signals. ECG represents the electrical activity of a heart. ECG is used to monitor the cyclical contraction and relaxation of the human heart muscles [8]. In this chapter, the structure of the human heart and the characteristics of PCG and ECG signals are discussed.

2.1 Biological overview of the heart

2.1.1 Anatomy and physiology of the heart

Heart is one of the most important organs of the human body. The primary function of the heart is to pump adequate blood to the entire body through the network of arteries and veins. This process is known as the cardiovascular system. The surface of the heart is surrounded by the coronary arteries which provide oxygen and nutrients rich blood to the heart muscles and take way the waste products from the heart. This helps to maintain the normal metabolism rate [13–16]. Fig. 1 shows the basic anatomy of the human heart.

The heart has four chambers: the right atrium, the right ventricle, the left atrium, and the left ventricle. The right atrium receives blood from the veins and pumps it to the right ventricle. The right ventricle receives blood from the right atrium and pumps it to the lungs. Blood is loaded with oxygen in the lungs. The left atrium receives oxygenated blood from the lungs and pumps it to the left ventricle. The left ventricle pumps oxygen-rich blood to

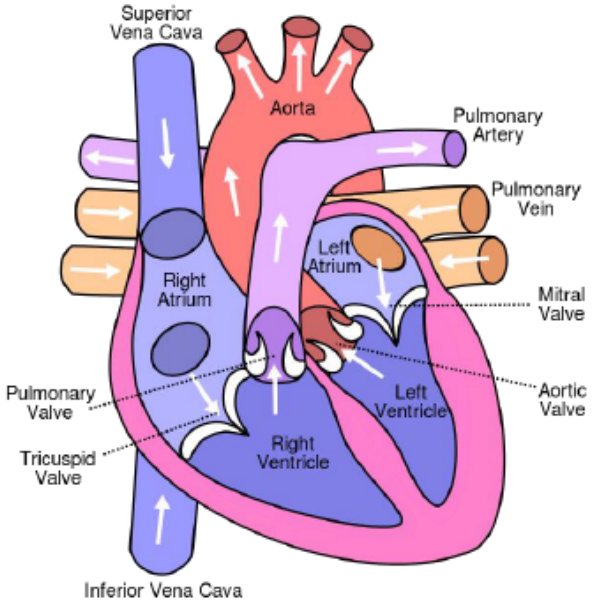

Figure 1: The basic anatomy of the human heart [15]

the rest of the body [13–16].

2.1.2 Heart valves

The heart is composed of four valves in which blood goes by before lifting each chamber of the heart. The valves keep away from the blood back flow. These valves are located on each end of the two ventricles. They act as one-way entrance of blood on one side of a ventricle and one-way entrance of blood on the other side of a ventricle. Normal valves have three flaps except the mitral valve. The mitral valve has two flaps. The name of the four heart valves are given below:

- **Tricuspid valve:** It is located between the right atrium and the right ventricle.
- **Pulmonary valve:** It is located between the right ventricle and the pulmonary artery.
- **Mitral valve:** It is located between the left atrium and the left ventricle.

- **Aortic valve:** It is located between the left ventricle and the aorta.

The valves between the atria and the ventricles are called atrioventricular (AV) valves. The left one is the mitral valve and the right one is the tricuspid valve. The valve between the right ventricle and pulmonary trunk is the pulmonary semilunar valve. The valve between the left ventricle and the aorta is the aortic semilunar valve. The pulmonary valve controls the blood from the right ventricle into the pulmonary artery and the aortic valve regulates blood from the left ventricle into the aorta [13–16].

2.1.3 Cardiac cycle

A cardiac cycle is defined as a complete heartbeat. It consists of a complete relaxation and contraction of both the atria and ventricles. It defines the electrical and mechanical activities of the heart throughout the systole and diastole interval. Systole interval is the duration of the cardiac contraction and diastole interval is the duration of the cardiac relaxation. The average duration of a cardiac cycle is around 0.8 second. Fig. 2 shows the heart systole and diastole.

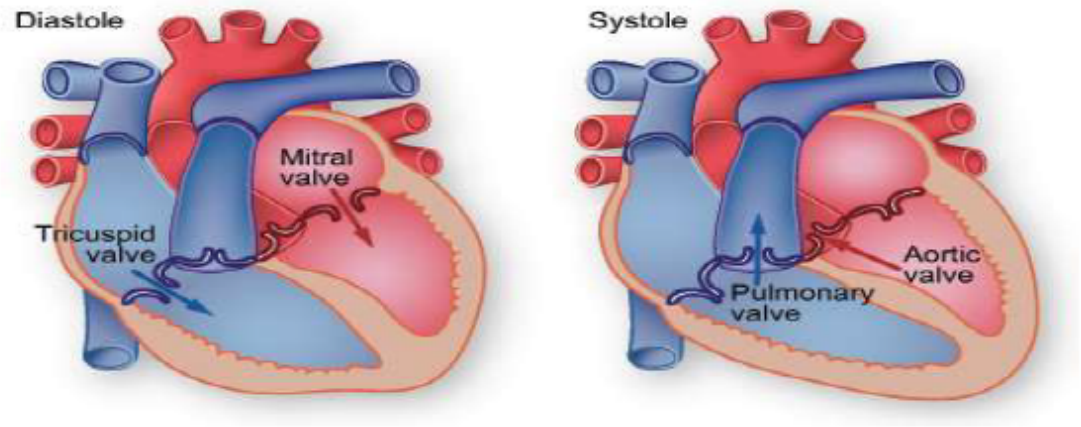

Figure 2: Heart diastole and systole [14]

In a cardiac cycle, atria and ventricles contract and relax mutually to pump the blood to move from the atria to the ventricles and then into the pulmonary. In the interval of diastole, it allows filling up the chambers to get ready for the next contraction. Events that occur in left chambers of heart is same as right chambers of heart [13–16] . For each chamber, the cardiac cycle has seven phases which are explained below:

- **Atrial contraction:** The first stage is called the atrial contraction. In this stage the atrium contracts and blood moves from the atrium to the ventricle.

- **Isovolumetric contraction:** The second stage is known as the isovolumetric contraction. This is the primary stage of the ventricular systole. In this stage the ventricles start to contract. During this ventricular contraction both valves stay remain closed and there is no change in the volume and size of the ventricle.

- **Rapid ventricular ejection:** The third stage is called the rapid ventricular ejection. During this stage the aortic valve opens and blood rapidly ejects to the aorta. The pressure transmitted to the aorta due to the change of pressure in the ventricles. In this phase left ventricle and aorta behave as a single chamber.

- **Slow ventricular ejection:** The fourth stage is known as the slow ventricular ejection. In this stage mitral valve becomes closed. Pressure in the aorta also starts falling and the aortic valve remains open which leads to the slow ejection of blood to aorta.

- **Isovolumetric relaxation:** The fifth stage is known as the isovolumetric relaxation. In this stage ventricles start relaxing with closed valves and the ventricular pressure decreases rapidly. Still, the pressure in ventricles is high enough compared to the

pressure in the atrium.

- **Rapid passive ventricular filling:** The sixth stage is known as the rapid passive ventricular filling. In this stage pressure in the ventricle becomes less than the pressure in atrium which leads to the opening of the mitral valve. Blood rushes to ventricles from the atrium.

- **Slow passive ventricular filling:** The seventh and the final stage is known as the slow passive ventricular filling. In this stage atrioventricular valves open and blood directly moves to the ventricle.

2.2 PCG signal

PCG is an automatic computer-aided diagnosis tool that is the graphical depiction of heart sounds and murmurs. It helps to monitor various components of the heart sounds through the heart cycle. Heart sounds are produced due to the flow of blood across the heart valves, the opening and the closure of the heart valves, and from the mechanical actions of heart muscles. These heart sounds are primary monitoring technique for diagnosing different cardiac diseases. Doctors and cardiologists usually use stethoscope to hear the heart sounds before any clinical diagnosis. Heart sounds are identical in all healthy hearts. Abnormal heart sounds are related to cardiovascular diseases.

2.2.1 Shape of a PCG signal

A normal PCG signal consists of two fundamental heart sounds called the first heart sound (S1) and the second heart sound (S2), which are generated due to the closure of the atri-

oventricular valves and semilunar valves respectively. The interval from the starting point of S1 to the starting point of S2 is called the systole interval (S1-S2 interval) and the interval from the starting point of S2 to the starting point of S1 is called the diastole interval (S2-S1 interval) [2, 41]. Diastole interval is usually longer than the systole interval [42]. Beside S1 and S2, two extra heart sounds known as third and fourth heart sound (S3 and S4) can appear in both normal and pathological conditions. S3 appears just after S2, and S4 appears just before S1 [14, 41]. Table 1 shows different heart sounds and their properties.

Table 1: Different heart sounds and their properties

Sound	S1	S2	S3	S4
Frequency	30-100 Hz	Above 100 Hz	20-25 Hz	Bellow 30 Hz
Time	50-100 ms	25-50 ms	120-150 ms	90 ms before S1
Occurrence	Closure of the mitral and tricuspid valve.	Closure of the aortic and pulmonary valve.	Caused by the rapid ventricular filling in the early diastole.	Caused by the ventricular filling due to the atrial contraction.

Besides these heart sounds different kinds of heart murmurs may also present in the signal which are produced because of the turbulent flow of blood across the valves and related to the cardiac diseases. Murmurs may present in systole or diastole or in both intervals [41]. Murmurs usually have higher frequency compared to the heart sounds [43]. When the blood circulates through the heart valves and chamber, sometimes it produces innocent murmurs which is not related to any cardiac diseases. There are mainly three kinds of heart murmurs:

- **Systolic murmurs:** Start after S1 and ends before S2.
- **Diastolic murmurs:** Start after S2 and end before S1.
- **Continuous murmurs:** Usually occur throughout or some parts of the cardiac cycle.

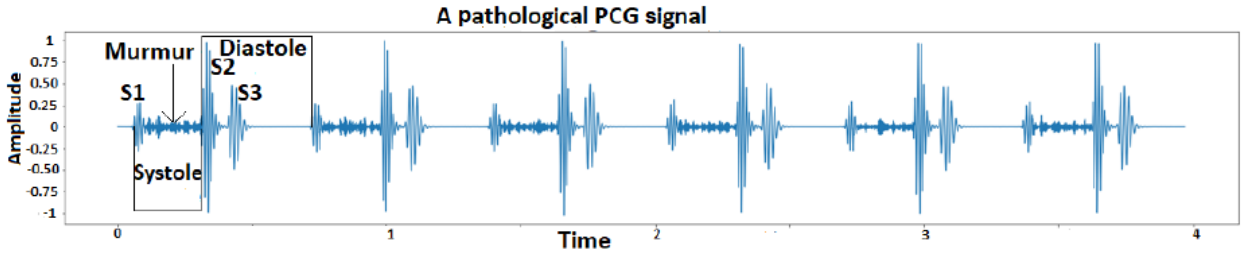

Figure 3: A pathological PCG signal

Fig. 3 shows a PCG signal with murmurs and S3 with reference to S1 and S2.

Th recording of a PCG signal is very straight forward. Digital stethoscope, different sensors or microphones can be placed on the skin of the chest to record heart sounds. The main difficulty while recording a PCG signal is the existence of different kinds noise that makes it difficult to segment and classify PCG signals.

2.3 ECG signal

ECG is the graphical record of changes in the magnitude and direction of the electrical activity of the heart. More specifically, the electric current that is generated by the depolarization and repolarization of the atria and ventricles can be monitored through the ECG signal. The ECG signal is captured through an array of electrode sensors known as leads. Leads are attached to the skin to detect the electrical activity of the heart. This information is recorded on a graph. As the electrical signal traverses through the heart, the graph shows each phase of the signal. Under normal condition, the ECG signal has a very predictable direction, duration, and amplitude. Any change in the ECG signal is related to cardiac cardiac abnormalities [8, 13]. Therefore, by analyzing ECG signal continuously, it is possible to detect any abnormal heart function in the primary stage which helps cardiologists for the proper clinical diagnosis. However, reliable and efficient clinical applications are highly dependent

on the accuracy of information extracted from the ECG recording. Usually ECG signals are subjected to contamination by various noises. The sources of noise may be either cardiac or extra-cardiac. Reduction or disappearance of the isoelectric interval, prolonged repolarization, and atrial flutter are responsible for cardiac noise. Whereas, respiration, changes of electrode position, muscle contraction, and power line interference cause extra-cardiac noise. Moreover continuous recording of ECG signals requires large amount of storage. So, compression and denoising of ECG signals without loosing any pathological information is prerequisite for the continuous monitoring of heart through ECG signals.

2.3.1 Shape of an ECG signal

ECG signal is composited from 5 waves - P, Q, R, S and T. This signal could be measured by electrodes from human body in typical engagement. Signals from these electrodes are brought to simple electrical circuits with amplifiers and analogue to digital converters. The muscle mass of the atria is small compared with the ventricles, and the electrical change of the atria is very small. Contraction of atria associated with the ECG wave is called 'P'. For the large mass of ventricular, it has large deflection which is called 'QRS' complex. The 'T' wave of the ECG is associated with the return of the ventricular mass to its resting electrical state [8, 13]. Fig. 4 shows the basic shape of a normal ECG signal.

Different ECG waves and their properties are given bellow:

- **P-wave:** It occurs due to the depolarization of atrial muscle. The amplitude of P wave is around 0.25 mV.

- **QRS-wave:** It occurs due to the repolarization of atria and depolarization of ventri-

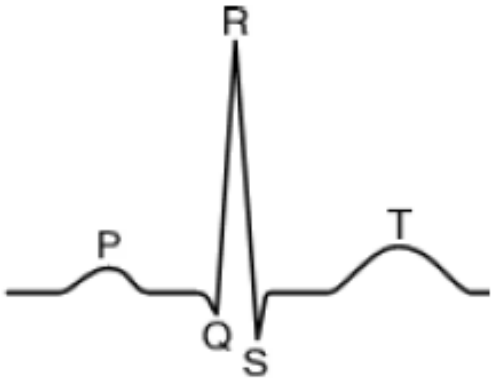

Figure 4: Basic shape of a normal ECG signal [8]

cles. The amplitude of R wave is around 1.60 mV. The amplitude of Q wave is around 25% of R wave. The time duration QRS interval is around 0.09 second.

- **T-wave:** It happens due to the ventricular repolarization. The amplitude of T wave is around 0.1 to 0.5 mV.

- **U-wave:** If present, it comes after potential in the ventricular muscle and represents repolarization of the purkinje fibers.

In a normal cardiac cycle, the p wave occurs first, followed by the QRS complex and the T wave. The section of the ECG between the waves and complexes are called segments and interval such as the PR segment, the ST segment, the TP segment, the PR interval, the QT interval, and the R-R interval. When the electrical activity of the heart is not being detected the ECG is straight, flat line [8,13]. Different intervals and their properties are given bellow:

- **P-Q interval:** Delay of excitation in the fiber near the AV node.

- **P-R interval:** Start from P wave to start of QRS complex. The time duration of the P-R interval is around 0.12 to 0.20 second.

- **Q-T interval:** Start of QRS complex to end of T wave. The time duration of the Q-T interval is around 0.35 to 0.44 second.

- **S-T interval:** End of QRS complex to start of T wave. The time duration of the S-T interval is around 0.05 to 0.15 second.

2.3.2 Recording of an ECG signal

The electrical signal from the heart is detected at the surface of the body through array of electrodes known as leads, which are joined to the ECG recorder by wires. Leads collect the electrical activity of the heart from a particular angle across the body, obtained by using different combination of wires. The most common technique to record an ECG signal is the 12-lead ECG method. The standard ECG has 12 leads. Six of the leads are considered as limb leads, as they are placed on the arms and/or legs of the individual. The other six leads are considered as chest leads because they are placed on the chest. The six limb leads are called lead I, II, III, aVL, aVR and aVF. The letter "a" stands for "augmented," as these leads are calculated as a combination of leads I, II and III. The six chest leads are called leads V1, V2, V3, V4, V5 and V6. Leads I, II and III are each making use of a pair of electrodes (bipolar), with one electrode measuring between itself and the other. Leads aVR, aVL, and aVF make use of all the connections to the patient. Each of the six pericardial or chest electrodes (V1-V6) represent six different views and unique information that can't be derived from other leads [8].

2.4 Relationship between ECG and PCG signals

PCG defines the mechanical activity of the heart and ECG defines electrical activity of the heart. Mechanical activity of the heart is known as the opening and closure of heart valves and the sound they produce during the cardiac cycle. This mechanical function relies on the electrical operation of the heart. So, if there is any defect in the electrical action of the heart, the mechanical function of the heart will also be interrupted. So, ECG and PCG signals are correlated with each other [10]. Fig. 5 shows the time domain relationship between ECG and PCG signals. From Fig. 5 we can see that the 1st heart sound (S1) appears 0.04 second to 0.06 second after the beginning of the QRS complex. The second heart sound (S2) starts at the end of the T wave. The third heart sound occurs after the T wave and before the P wave. The fourth heart sound (S4) occurs after the P wave and before the QRS complex. S3 and S4 both occur during the diastolic period. Based on this calculation it's possible to analyze the mechanical and electrical activities of the heart.

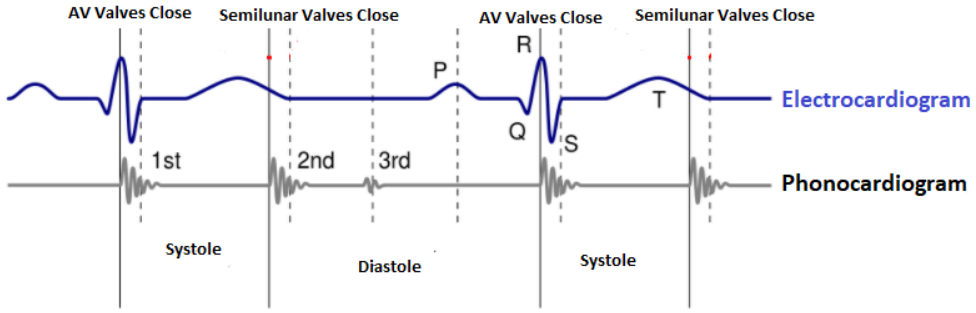

Figure 5: Time domain relationship between ECG and PCG signals [10]

CHAPTER 3

COMPRESSION AND DENOISING OF PCG AND ECG SIGNALS USING DWT

3.1 Signal compression

Telemedicine can play a vital role to the 24 hours cardiac monitoring of a patient suffering from cardiac abnormalities. But 24 hours online monitoring requires huge amount of storage to store the data and higher bandwidth to transfer the data from homes to physicians in clinics. So, the main aim of an efficient compression process is to remove all the redundant information without loosing any data containing pathological information. It can be achieved by reducing the number of significant coefficients thus it will save the memory and less bandwidth will be required [5, 7]. Computerized medical signal processing systems such as PCG and ECG acquires a large amount of data that is difficult to store and transmit. Thus, it is very important to use a compression technique to reduce the quantity of data without loss of any important cardiac information. All data compression algorithms aim to minimize the data storage by discarding the redundant information and noise. A high compression ratio is wanted with very less distortion. An efficient data compression algorithm must have the ability to compress the signal with acceptable fidelity. In biomedical data compression, the clinical acceptability of the reconstructed signal depends on its fidelity which can be measured by calculating the difference between the original signal and the reconstructed signal. In biomedical signal processing, the main goal of an optimized compression is to minimize the number of samples without losing the remarkable pathological information of

the original signal in order to achieve a correct clinical diagnosis.

3.1.1 Performance evaluation of a compression algorithm

Several Performance analysis techniques are available to evaluate the quality of a compression algorithm and also the fidelity of the reconstructed signal with respect to original signal. Following techniques were used in our research:

Compression Ratio (CR) and Percentage of Compression (PC):

The Compression Ratio (CR) and percentage of compression (PC) both are used to measure the compression efficiency of the compressed algorithm. The CR and PC are calculated using these formulas:

$$CR = \frac{a}{b} \tag{3.1.1}$$

$$PC = \frac{a-b}{a} \times 100 \tag{3.1.2}$$

where a is the size of the uncompressed original signal and b is the size of the compressed signal.

Percent Root-mean-square Difference (PRD):

Percent Root-mean-square Difference (PRD) is used to measure the fidelity of the reconstructed signal with respect to original signal. Mathematically PRD can be expressed as:

$$PRD = \sqrt{\frac{\sum_{n=1}^{N}(x_1(n) - x_2(n))^2}{\sum_{n=1}^{N}(x_1(n))^2}} \times 100 \tag{3.1.3}$$

where $x_1(n)$ is the original signal and $x_2(n)$ is the reconstructed signal. N is the length of the signal and n is an integer. The best result is achieved when the compression method has

high CR and PC with a very low PRD.

Quality Score (QS):

Quality score (QS) is used to compare the performance of a compression technique with other compression techniques. It is the ratio of the CR and PRD. QS can be expressed as:

$$QS = \frac{CR}{PRD} \tag{3.1.4}$$

A compression method is considered to be more efficient if it has higher QS compared to the other compression methods.

3.2 Signal denoising

Noise is an undesirable signal that usually does not contain any important information. Noise usually overlaps with any desired signals and makes it difficult to extract the original meaningful information from the original signal. Noise is the major factor that causes limitation in biomedical data transfer and effects the accuracy level of the result. Thus, noise elimination is necessary in biomedical signal processing. While recording, PCG and ECG signal are usually interrupted by different kinds of noise and unnecessary information that may cause inaccurate clinical diagnosis. So, before segmentation, it is required to remove these redundant pieces of information from ECG and PCG signals to evaluate the proper functionality of the heart [5, 7].

PCG signals usually suffer from different kinds of noise like the sounds by the mechanical actions of lung, breathing, movement of the patients, talking, inaccurate connection of the microphone to the human body, different environmental noises mixed with the signal and so on, thus make it very difficult to extract the correct diagnostic information and features

from PCG signals [5, 7, 25]. The common sources of noise in PCG signals are given below:

- **Electronic noise:** Noise come from electronic devices such as electrodes or sensors, inaccurate connection of the microphone to the human body.
- **Acoustic noise:** Noise due to vibrations such as mechanical actions of lung, breathing, movement and talking of patients.
- **Electromagnetic noise:** Due to the coupling of signals with radio-frequency spectral components while transmission.
- **Murmurs:** Due to the turbulent flow of blood across the heart valves. Murmurs are related to heart diseases.

The common sources of noise in ECG signals are given below [30]:

- **Baseline Shift noise:** It is a slow-moving and non-deterministic wave. It happens due to the respiration.
- **Muscle Noise:** It is caused by the random firing of muscle fibers.
- **Power line noise:** it is picked up by the electrode leads from neighboring equipment.
- **Electrode contact noise:** It is caused by loss of contact between the electrode and skin.
- **Motion artifacts:** Due to the change of the position of the electrodes in the skin and patient movements.
- **Noise generated by electronic devices:** It happens due to the sound coming from electronic devices.

3.2.1 Performance evaluation of a denoising algorithm

The performance of any denoising algorithm can be evaluated by using Signal to Noise Ratio (SNR). SNR is used to measure the amount of noise in the signal. SNR is defined as the ratio of signal power to noise power and expressed in decibels [7]. mathematically the formula of SNR can be written as:

$$SNR = 10 * log_{10}(\frac{E_{signal}}{E_{noise}})^2 \; db \tag{3.2.1}$$

where E_{signal} is the Root-mean-square amplitude of the signal and E_{noise} is the Root-mean-square amplitude of the noise.

3.3 DWT

DWT is a time-frequency signal analysis method based on the Fourier Transform. It has good localization in both frequency and time domains. DWT has been widely used in biomedical signal process. Although, FT and STFT are commonly used tools for examining stationary signals, but they show limited performance in examining non-stationary signals, as they are unable to provide simultaneous time and frequency localization [20]. Non-stationary signals like PCG and ECG can be analyzed properly by using DWT, as DWT provides very good time-frequency localization [20]. By using a multi-resolution analysis technique based on DWT, PCG and ECG signals can be decomposed into different sub-bands having different frequency ranges. The required sub-bands containing valuable clinical information can be picked up for further analysis. The other sub-bands, which contain murmurs, noise, or other unnecessary information, can be discarded [7, 44]. Multi-resolution analysis based on DWT

can analyze the signal with different resolution at different frequencies. At high frequencies, a good time resolution and poor frequency resolution can be achieved. Similarly at low frequencies, a good frequency resolution and poor time resolution can be achieved by using DWT. This feature is achieved by scaling and shifting of the mother wavelet. A mother wavelet $\psi(t)$ is defined as a function of zero mean:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{3.3.1}$$

When this mother wavelet is scaled (or dilated) by factor of a and translated (or shifted) by factor of b then the mother wavelet is denoted by $\psi_{a,b}(t)$ and is defined as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \tag{3.3.2}$$

In DWT, scaling and translation parameters are chosen such as the resulting wavelet set forms an orthogonal set. The scaling factors are chosen as power of two. So the values of a and b : $a = n * 2^m$ and $b = 2^m$, where $n, m \, \epsilon \, Z$. So the mother wavelet in DWT can be represented as:

$$\psi_{m,n}(t) = 2^{-\frac{m}{2}}\psi(2^{-m}t - n) \tag{3.3.3}$$

And the DWT of the signal x(t) is given by:

$$X_{DWT} = \int_{-\infty}^{\infty} x(t)2^{-\frac{m}{2}}\psi(2^{-m}t - n)dt \tag{3.3.4}$$

where m is the scaling factor and n is the translation factor. Scaling factor m depends on the width of the mother wavelet and different wavelets have different width. Narrow wavelets with high frequency are shifted by small steps and work like small data windows

to analyze high frequency components and wider wavelets with low frequency are shifted by large steps work like large data windows to analyze the low frequency components of the signal. The time-frequency analysis of a signal x by using DWT can be calculated by passing the signal through a series of high pass and low pass filter. First the samples of the signal x are decomposed simultaneously using a low pass filter of impulse response g and a high pass filter of impulse response h, which results the convolution of the signal x with the filters presented bellow:

$$Y_1[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k] \tag{3.3.5}$$

$$Y_2[n] = (x * h)[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] \tag{3.3.6}$$

where n is an integer. The block diagram of the filter analysis is shown in Fig. 6.

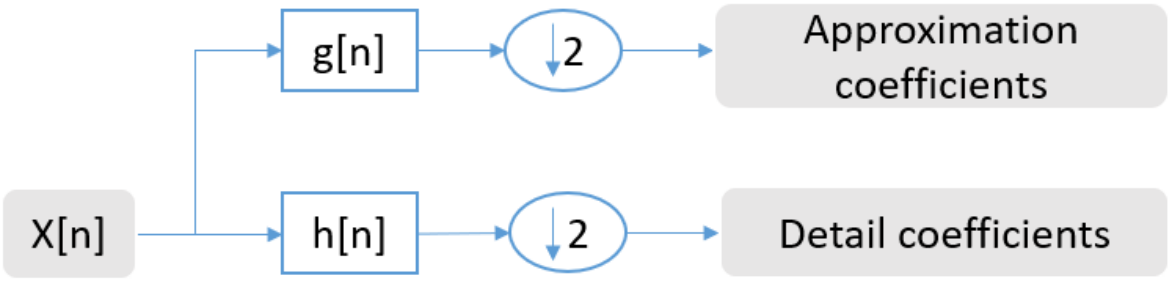

Figure 6: Block diagram of the filter analysis

After the process of filtering, as half of the frequencies have been removed, half of the total samples can be rejected according to the Nyquist's rule. So the signal is sub-sampled by 2 to discard half of the total samples from the signal. This process is called one level

signal decomposition and mathematically can be expressed as:

$$Y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \tag{3.3.7}$$

$$Y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \tag{3.3.8}$$

where $Y_{low}[n]$ and $Y_{high}[n]$ are the outputs of the low pass and high pass filters respectively after down-sampled by 2. As now only the half number of samples of the total samples are representing the entire signal so the time resolution is halved. However since total frequency spectrum now covers only half of the previous frequency spectrum, it doubles the frequency resolution. So multi-resolution analysis is achieved. The output from high pass and low pass filters are called the detailed and approximate coefficients respectively. These two filter are dependent to each other and called quadrature mirror filters. Choosing the best mother wavelet is also an important task of the filtering process as the detailed coefficients of the filters correlate to the coefficients of the mother wavelet. Decomposition technique can be repeated if further decomposition is required to get desired frequency spectrum from the input signal. At each level of decomposition, the signal will split into low and high frequency band, the time resolution will be halved and the frequency resolution will be doubled of the previous one by the filtering and sub-sampling process [7, 20, 44].

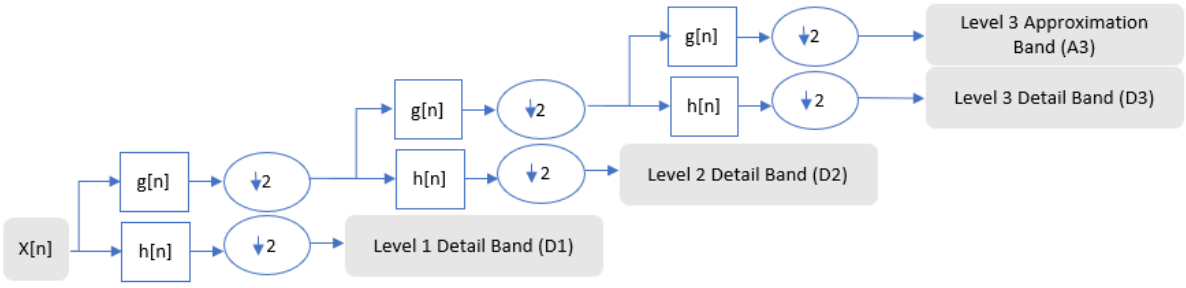

Figure 7: A DWT model of 3 levels decomposition

Fig. 7 depicts this process as three level filter bank. The detail and approximate coefficients of the decomposed signal enables to reconstruct a new signal. This reconstruction process is called Inverse Discrete Wavelet Transform (IDWT) which is just opposite to the DWT. The signals at each level are up-sampled by 2 and then passed through the high-pass and low-pass filter respectively before adding with each other to reconstruct a new signal.

CHAPTER 4

SEGMENTATION OF PCG SIGNALS USING SHANNON ENERGY ENVELOPE

4.1 Segmentation of PCG signals

A cardiac cycle or a complete heartbeat usually consists of the first heart sound (S1), the systolic interval, the second heart sound (S2), and the diastolic intervals. These parameters are very important to determine the cardiac condition of a person. Segmentation of the PCG signal facilitates to get the exact position and duration of the heart sounds, systole and diastole intervals. But the presence of murmurs which produced by the turbulent blood flow is the major reason for the false detection of the heart sounds. So the elimination of murmur is necessary before segmentation. After discarding the murmurs, different segmentation algorithms can be used to segment PCG signals.There are many PCG segmentation algorithms available in the scientific literature. Most of the segmentation algorithms use ECG signal as reference to identify the components of heart sound. But it is not convenient because of the additional hardware requirements and also it is very expensive. In our research we used a segmentation algorithm based on the normalized average Shannon energy. This algorithm does not need any reference signal. Furthermore, this technique is more efficient and computationally less expensive compared to other segmentation algorithms. After discarding the murmurs, Shannon energy envelope is used to detect the boundaries of each heart sound and zero crossing algorithm is used to detect the point where the signal crosses the zero, thus the location and duration of the heart sound can be attained. Then we locate the first heart

sound, the systolic interval, the second heart sound and diastolic interval. Then the peaks of the heart sounds are calculated to get their amplitude [5].

4.1.1 Shannon Energy Envelope

Different methods are available in the scientific literature to detect the envelope of the signal like absolute value of the signal, the energy (square), Shannon entropy and Shannon energy. The absolute value gives same weight to all the components thus it's difficult to separate the high amplitude signal from the low amplitude signals using this procedure. The energy (square) gives weight to high amplitude rather than low amplitude signals. Shannon entropy gives more weight to low intensity rather than high intensity signals. But Shannon energy gives better result compared to all other methods by giving emphasize on the signal having medium intensity and reduces the impact of low amplitude signal more than the signals having high amplitude. Thus it's possible to detect the difference of the envelope intensity of the high and low amplitude sounds. Because of these beneficial effect Shannon energy is used in our research to accurately identify boundaries of all the heart sounds (S1, S2, S3, S4) and to discard the low frequency noise from the signal [32, 33]. A threshold is set to discard the effect of the noise and the low amplitude signal. The mathematical expression of this procedure is given below:

$$Shannon\ Energy : E = -x^2 log x^2 \tag{4.1.1}$$

where is x is the denoised signal.The average Shannon energy is calculated in 0.02 second continuous segments of the whole PCG signal with segment overlapping of 0.01 seconds. The

average Shannon energy can be represented as:

$$E_s = -\frac{1}{N}\sum_{i=1}^{N} x^2(i) * logx^2(i) \tag{4.1.2}$$

where x(i) is the denoised PCG signal. N is the signal length or the number of coefficients and i is an integer. At last the normalized average Shannon energy is calculated to get the Shannon energy envelop of the signal. Mathematically normalized average Shannon energy can be written as:

$$E_n = \sum_{i=1}^{N}(\frac{E_s(i) - M(E_s(i))}{max(|E_s(i)|)}) \tag{4.1.3}$$

where N is the signal length or the number of coefficients and i is an integer. E_n is the normalized average Shannon energy. $E_s(i)$ is the average Shannon energy. $M(E_s(i))$ is the mean value of $E_s(i)$ and $max(|E_s(i)|)$ is the maximum absolute value among all the coefficients of $E_s(i)$. After calculating the boundary of the heart sounds accurately, a zero crossing algorithm is implemented to know the starting and stopping point of the signal by calculating the points where the sign of the boundaries changes from positive to negative or vice versa.

4.1.2 Peak detection

By detecting the peak of each heart sound the amplitude of the heart sounds can be computed. For peak detection, the package peakUtils in python is used which provides different utilities to detect the peak of an 1D signal [34]. This function detects the local maxima within a fixed distance and by using a threshold. A local maxima can be considered as a peak if it's distance to the nearest peak and it's amplitude is greater than the predefined dis-

tance and threshold respectively. Though this technique is used to detect the peak of all the heart sounds in the database, distinguishing S1 from S2 is not possible with this technique. S1 and S2 is differentiated based on the fact that distance from S1 to S2 (Systole interval) is smaller than the distance from S2 to S1 (Diastole interval). So a peak is considered to be S1 if it's distance to it's next peak is smaller than the distance to it's previous peak. Similarly a peak is considered to be S2 if it's distance to it's next peak is larger than the distance to it's previous peak. In this way if we can detect the location of S1 and S2, based on their location we can get the systole and diastole interval.

CHAPTER 5

AN OVERVIEW OF ML AND DL APPROACHES

Artificial intelligence (AI) refers to the intelligence of machines to perform any difficult task without the help of any human. AI enables machines to think. Without any human intervention the machine will be able to take his own decision using AI. With the help of AI, machines are capable of applying intelligence to a large amount of data and can derive meaningful results. ML is a subset of AI which provides us statistical tools to explore the data. ML algorithms are usually used to discover the similarity, patterns, and difference in the data. Manual analysis of these huge amounts of data is very difficult and cumbersome task, which ML algorithms can easily perform using it's automatic learning process. In recent days, DL has gained huge interest due to it's supremacy in terms of accuracy when trained with large amount of data. ML algorithms underperfom in presence huge amount of data. DL can solve this problem through it's hidden layer architecture. DL algorithms mimic the function of human brain. In traditional ML architectures, important features need to be identified by an domain expert in order to reduce the complexity of the data and increase the accuracy. The main advantage of DL architectures is the capability to learn high-level features from data when it is in large amounts and to eliminate the need of domain expertise to extract important features [35, 36]. Fig. 8 and 9 show classification process of ML and DL architectures.

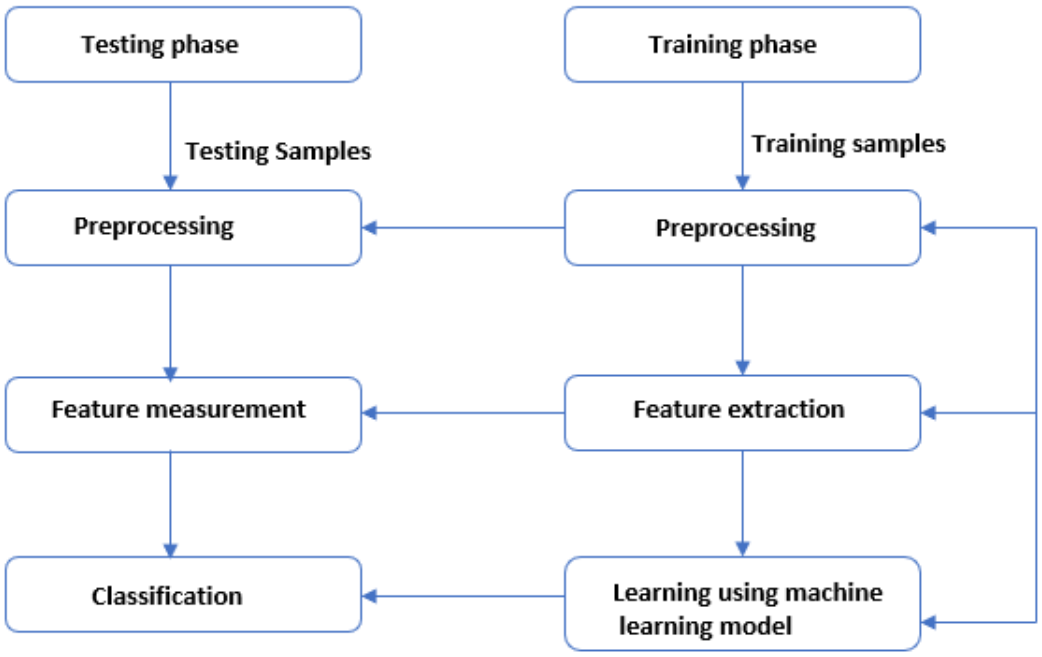

Figure 8: The classification process of ML algorithms

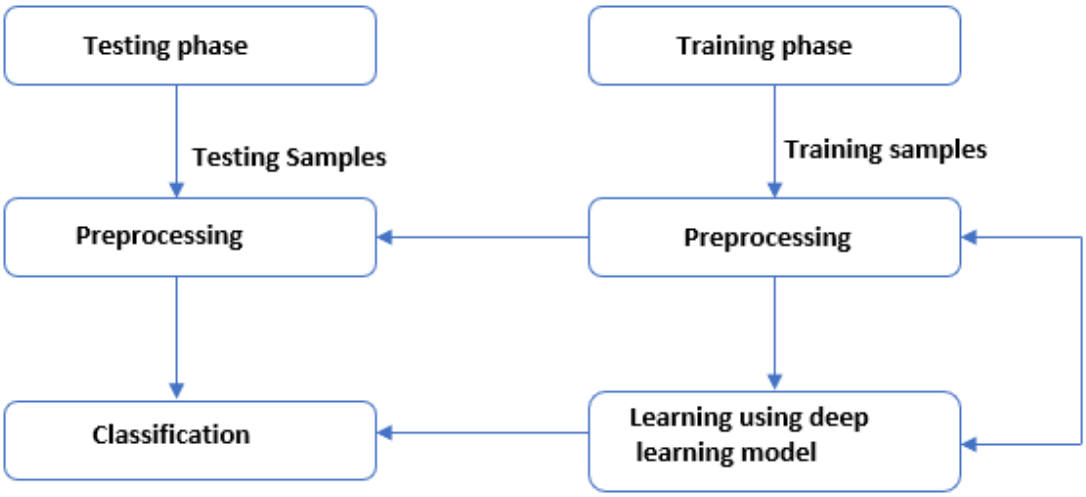

Figure 9: The classification process of DL algorithms

5.1 ML algorithms

ML algorithms are mainly grouped into three categories. They are supervised, semi-supervised, and reinforcement learning. Supervised learning algorithms are the most common ML algorithms which depends on the labels of the sample input. Based on the input data and labels it provides output. In biomedical signal processing supervised ML algorithms are widely used for different classification. Most commonly used classical ML algorithms are NB, KNN, LR, DT, AB, RF, and SVM [35, 36]. The concepts behind the most commonly used ML algorithms are discussed briefly.

- **LR:** LR is a ML algorithm which usually uses a logistic function to model a binary dependent variable and to overcome the limitation of the linear regression. In linear regression misclassifications happen due to the independence between the features and the outcome (dependent variable) and it can be any number. In LR, the log-odds of the probability of an event is a linear combination of independent variables and the outcome (dependent variable) has only a limited number of possible values basically either '0' or '1'. The outcomes of LR are obtained by utilizing the non-linear function which is called the sigmoid function.

- **NB:** NB is a powerful ML algorithm which is based on the principle of Bayes theorem with an assumption of independence among predictors. NB Classifier works on the assumption that attributes are conditionally independent given the class variable. Based on the prior information of conditions related with the event, it is possible to find the probability of occurrence of an event using NB classifier. One of the main

advantages of this approach is that it can classify efficiently using only a small amount of training data. The conditional probability A given B calculated by Bayes theorem is given below:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{5.1.1}$$

where, $P(A|B)$ is the posterior probability, $P(B|A)$ is the Likelihood, $P(A)$ is the Prior probability, $P(B)$ is the marginal probability. Thus, if the class variable y is dependent on features $x_1, x_2, x_3, ...x_n$, based on the conditional probability that the probability of y occurring given the condition $x_1, x_2, x_3, ...x_n$ can be represented as:

$$P(y|x_1, x_2, x_3, ...x_n) = \frac{P(x_1, x_2, x_3, ...x_n|y) * P(y)}{P(x_1, x_2, x_3, ...x_n)} \tag{5.1.2}$$

Using the above function, we can obtain the class, given the predictors.

- **KNN:** KNN is a ML algorithm which calculates the distance between two data points and based on the distance it classifies them in different groups. There are different methods to calculate the distance between two data points. Euclidean distance, Manhattan distance, Hamming distance, Minkowski distance, cosine similarity are different distances measurement techniques which can be used for calculating the distance between the points. Among all of these methods, the Euclidean distance is the most common and popular distance measurement method. This method calculates the straight-line distance between two points. According to the Euclidean distance formula, the distance between two points (x, y) and (a, b) is given by:

$$dist((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2} \tag{5.1.3}$$

With this very simple implementation and low dimension data, KNN performs well to classify data. The performance of this algorithm becomes limited when the data dimension is increased.

- **DT:** DT is another supervised ML algorithm where a training model can predict the class or the target variable by learning decision rules from prior data. In this algorithm a DT is constructed by continuously splitting the data. DT is comprised of leaves and nodes where leaves are the results of each decision and nodes are the decision processes. In DT, the calculation usually starts from the root of tree where the values of root attribute are compared with the record's attribute. The best outcome which classifies the training data is chosen as the leaf. This process is continuous and recursive until tree fits best for the training data. Each leaf is decided by calculating the distance in entropies of the parent node and the children nodes. This process is know as information gain which tries to estimate the entropy of each node. Entropy measures the uncertainty or randomness in the data. Entropy can be written mathematically by:

$$H(S) = -\sum_{i \epsilon X} (P(i) * log_2 P(i)) \tag{5.1.4}$$

where $H(s)$ is the Shannon entropy over a finite set S. $P(i)$ is the probability of an outcome i. The information gain can be calculated using the formula given below:

$$IG(S,U) = H(S) - \sum_{i=0}^{n} (p(i) * H(i)) \tag{5.1.5}$$

where $H(S)$ is the entropy of the entire set. P(i) is the probability of event i to occur. The value of the IG differs from 0 to 1. If the feature doesn't have any significance

than the IG should be 0 and if the feature has very high significance then IG will be 1.

- **AB:** AB is an ensemble ML classifier which helps to combine multiple weak classifiers into a single strong classifier. This is a boosting process where the performance of ML classifier can be boosted by iteratively to retrain the algorithm based on accuracy of previous training. At every iteration, a weight is given to every trained classifier based on the accuracy achieved by the classifier. AB works by putting more weight on the items which were not correctly classified and less weight on the items which were classified accurately. The items which have higher weight usually have higher probability in next classifier. A positive or zero weight is assigned to any classifier which has accuracy of 50% or more. The more accurate the classifier, the larger the weight. Similarly, a negative weight is assigned to any classifier which has accuracy of less than 50%. The accuracy of the classifier can be increased by increasing the number of iterations.

- **RF:** RF algorithm is considered as the best ML algorithm for the classification. It is a supervised learning algorithm where it builds the forest using ensemble of decision trees (DT). DTs are usually trained with the ensemble learning method like bagging where the overall result can be increased by the combination of learning models. This bagging method also reduces the overfitting of data while making these DTs. Initially, multiple DTs are created which are then merged to get the correct classification. Accuracy of the algorithm is directly proportional to the number of DTs. In other words, higher the number of DTs, higher the classification accuracy. RF usually gives very good result without using optimization technique or hyper parameter tuning. RF can be

used for both regression and classification tasks. The main limitation of RF is that it becomes slow and ineffective in real time when it produces a large number of DTs in the practical scenario.

- **SVM:** SVM is a ML algorithm which is used for classification and regression problems. The objective of the SVM is to find a hyper-plane with the support vectors that classifies the data points. Hyper-planes are decision boundaries to classify the data points. To separate the two classes of data points, it is possible to choose different hyper-planes. In SVM, the main objective is to find a plane that has the maximum distance between data points of both classes. Data points can be assigned to different classes if they fall on either side of the hyper-plane. The number of input features controls the dimension of the hyper-plane. If the number of input features is 2, then the hyper-plane is just a line. If the number of input features is 3, then the hyper-plane becomes a two-dimensional plane. When the number of input features exceeds 3, it becomes difficult to create the hyper-plane. SVM works very well when there is a clear margin of separation between classes. The performance of the SVM becomes limited where the number of features for each data point exceeds the number of training data samples and also if the data set is too large.

5.2 DL algorithms

DL algorithms use layers of neural-network to convert raw input data to higher-level information that increase the classification accuracy. There are different kinds of DL algorithms such as CNN, Deep Neural Network (DNN) Recurrent Neural Network (RNN), Long Short

Term Memory (LSTM) and so on. In our research we used DNN and CNN to classify PCG and ECG signals and got very good accuracy compared to other ML algorithms [35,36]. The concepts behind the DNN and CNN are discussed below:

- **DNN:** A DNN is the network of artificial neurons with multiple hidden layers between input and output layers. These neurons usually create a complex network of different layers. Neurons from one layer pass signals to other neurons in the next layer. Fig. 10 represents a DNN of N hidden layers.

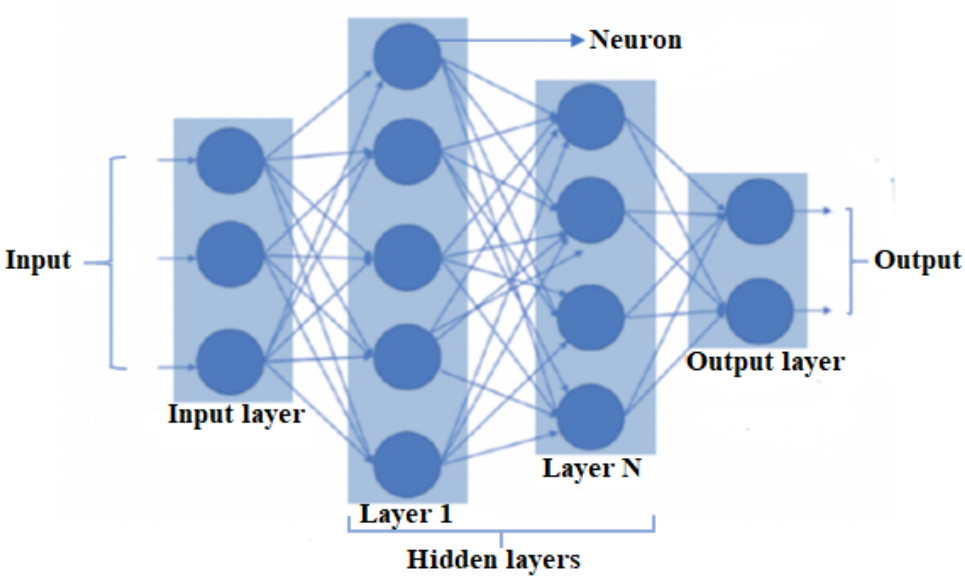

Figure 10: A DNN architecture with N hidden layers

From Fig. 10, we can see that the input data is fed into the neurons of the input layer. The output of the input layer works as input to the first hidden layer. This process will continue until the final layer. The output of the final layer will give the final prediction. Each layer can have one or more neurons and each neuron uses a threshold value in the form of an activation function to pass the signal to the next connected neuron.

Two neurons of consecutive layers are connected with a parameter called weight. The function of the weight is to transform the input data within the hidden layers. While training the model, DNN uses a backpropagation algorithm to provide feedback to the network based on the output. The goal of the backpropagation algorithm is to update each of the weights several times step-by-step, thereby minimizing the error and gradually increase the overall accuracy [39]. After nth iteration the error at the output of neuron p can be expressed as:

$$e_p(n) = d_p(n) - a_p(n) \tag{5.2.1}$$

where $d_p(n)$ and $a_p(n)$ are the desired and actual output of neuron p, respectively. The instantaneous error energy at the output layer is defined as:

$$E(n) = \frac{1}{N} \sum_{p=1}^{N} e_p^2(n) \tag{5.2.2}$$

The above error can be reduced by using gradient descent method. The gradient descent is the widely used optimization method to update the weights by calculating the derivative of the error with respect to the weights of the network. This process can be expressed as:

$$\Delta w_{p,j}(n) = -\eta \frac{\partial E(n)}{\partial w_{p,j}(n)} \tag{5.2.3}$$

where η is known as the learning rate. Learning rate is a hyper-parameter, which determines the adjustment of the weights with respect to the loss gradient. The range of the learning rate is between 0 to 1. This process of updating the weights will continue until the loss function is minimum. The final updated value of $\Delta w_{p,j}(n)$ can

be expressed as:

$$\Delta w_{p,j}(n+1) = w_{p,j}(n) + \Delta w_{p,j}(n) \tag{5.2.4}$$

Thus, by minimizing the error we can find the optimal values for the weights of each neuron that will give the best model performance.

- **CNN:** A CNN is a DL algorithm which uses a series of convolutions with different filters to automatically learn important features directly from the raw data. A CNN architecture usually composed of 3 layers. These layers are convolution layer, pooling layer, and fully connected layer. Convolution can be 1D convolution, 2D convolution, or 3D convolution. In our research we used 1D convolution. Convolution layer contains filters that passes over the data to capture the optimal features. If the the dimension of the feature is high it can be reduced by using pooling function. Pooling can be min, max, or average pooling. Finally, the pooling features are passed into a fully connected layer for the final classification. For a 1-D signal $x(n) = [x(1), x(2), x(3)....., x(N)]$, if it has K number of classes to classify and N is the signal length then initially a 1D Convolution method is used to extract the optimal features from the raw input data by applying a series of 1-D convolutions with different 1-D kernels. This process is achieved by sliding a kernel h(n) with length of W samples along the input data. In this way, the ith output $y_i(n)$ from the Conv1D layer can be expressed by:

$$y_i(n) = \sum_{k=0}^{W-1} h_i(k)x(n-k) \tag{5.2.5}$$

where $i = 1, ..., L$ and L is the number of filters. By adding a pooling layer to the

output of the convolutional layer, it is possible to extract only the effective features from the feature map. Different kinds pooling layers can be used based on the application. These significant features are then fed into the fully connected layer which concatenates its inputs data into a long single features vector. For example, if we use L kernels in the CNN model, then we have L different outputs from the Conv1D and Pooling layers. Consequently, the output vector V of the FC layer can be represented as follows:

$$V = [V_1(n), V_2(n), V_3(n), ..., V_L(n)] \tag{5.2.6}$$

A non linear activation function called Rectified Linear Unit (ReLu) is applied to the output of the convolutional and Pooling layers to introduce non-linearity. To obtain the estimated probability for each class, Sigmoid or Softmax activation function is used at the final output of the fully connected layer. If the classification is binary then Sigmoid activation function is used. If it is a multi-class classification then Softmax activation should be used. Fig. 11 represents a simple CNN architecture.

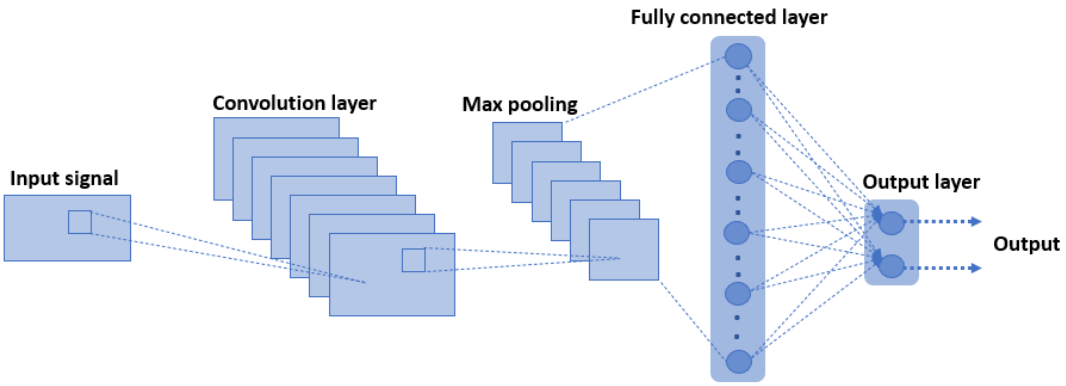

Figure 11: A CNN architecture

5.3 DHL algorithms

DHL combines the advantages of traditional ML and DL techniques. Though DL models usually outperform ML models in classifying PCG and ECG signals, optimal classification can be achieved if we combine these two architectures to build a single classification model. A CNN is used along with seven traditional ML methods including LR, RF, KNN, DT, NB, SVM, and AB to build hybrid PCG and ECG classification models. DHL models are less computationally expensive and have less time complexity compared to traditional DL algorithms. The final classification layer of a DL model usually results in overfitting when the model is fed with unstructured or less data. This overfitting problem increases the time and computational complexity of traditional DL models, which is not present in traditional ML algorithms. In DHL models, fully connected neural networks in the DL model are followed by the ML models. Thus, DHL models are faster and do not require additional time for processing compared to traditional standalone DL models. In addition, DHL models remove the need for feature engineering techniques on which all traditional ML algorithms are dependent. Fig. 12 represents a DHL architecture of N hidden layers. Alabandi [36] implemented DHL on four different data sets such as small human activity dataset, large human activity dataset, small emotion dataset, and the dataset for emotion analysis using eeg, physiological and video signals (DEAP). Their experimental results showed that their hybrid approach outperformed DL and traditional ML algorithms when those are used separately. Bhattacharya [37] also proposed a fusion approach by combining ML with DL to diagnose various diseases from large chest x-ray image dataset. He also showed, DHL networks performed better than standalone DNN and ML models with respect to precision, recall, and

accuracy. Sengupta et al. [38] also proposed DHL pipelines for accurate diagnosis of ovarian cancer based on nuclear morphology image dataset and achieved 100% testing accuracy.

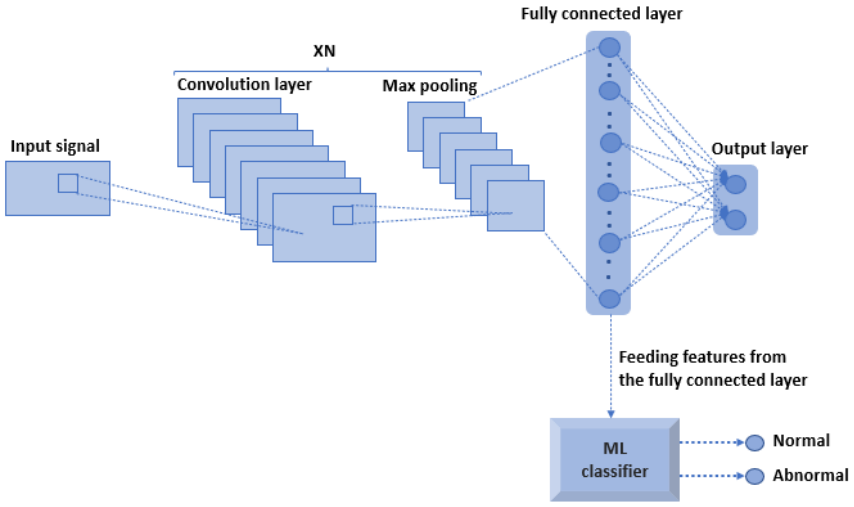

Figure 12: A DHL architecture

5.4 Performance evaluation

Classification accuracy of any learning model can be evaluated by investigating the confusion matrix. Several parameters such as sensitivity/recall, specificity, and precision are used to analyze the prediction accuracy of any classification models. The sensitivity/recall indicates the true positive rate and measures the proportion of the correctly identified actual positives. The specificity indicates the true negative rate and measures the proportion of the correctly identified actual negatives. Out of predictive positive, how many of them are actual positive is defined by the precision metric. All of these parameters can be calculated by using the confusion matrix. Another important metric used to evaluate the classification model is known as accuracy, which is the number of correctly predicted data points out of all the data points [5, 6, 12, 35, 36]. Sensitivity, specificity, and accuracy can be calculated by using

these formulas:

$$Sensitivity/Recall = \frac{TP}{TP + FN} \tag{5.4.1}$$

$$Specificity = \frac{TN}{FP + TN} \tag{5.4.2}$$

$$Precision = \frac{TP}{TP + FP} \tag{5.4.3}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.4.4}$$

where TP (True Positive) is the number of sick people correctly identified as sick, TN (True Negative) is the number of healthy people correctly identified as healthy, FP (False Positive) is the number of healthy people incorrectly identified as sick, and FN (False Negative) is the number of sick people incorrectly identified as healthy.

5.5 Feature extraction

Feature extraction is a process of deriving a compact and useful representation of the information from any signal. Feature extraction considered to be one of the main steps of PCG and ECG classification systems. The performance of classifiers highly depend on the extracting features and the performance will decrease if the features are not selected properly. PCG and ECG signals are redundant in nature and highly non-stationary signals. So, we need to extract the required and meaningful representation from the signal to train the model. In the feature extraction process the dimension of the signal should be reduced to a lower dimension that contains useful information to differentiate different signals. For

the classification of ECG signals, we used the ECG beat as the input to the classification model. We did not need to extract any feature from each ECG beat manually, as our deep intelligent model used different layers of CNN and more computation to identify important features from each heart beat to classify different heartbeats. The heart sound signals in the database are redundant in nature and contain lots of noise and unnecessary information. Therefore, we need to extract the necessary and meaningful features from PCG signals to train the model. Mel-scaled power spectrogram and MFCC are used in our research to extract important and meaningful features from PCG signals. The features of the Mel-scaled power spectrogram and the MFCC are biologically inspired and resemble the resolution of the human auditory system, which (features) are proven to be more efficient to discriminate between two different sound signals [72].

5.5.1 Mel-scaled Power Spectrogram

Time vs. Frequency representation of a signal is called the spectrogram of the signal. A spectrogram visually represents the change of the frequency of a signal with respect to time, which helps the model to recognize the sound accurately [72]. The Mel-scale aims to mimic the non-linear human ear perception of sound, by being more distinctive at lower frequencies and less distinctive at higher frequencies. The Mel-scaled filters are non-uniformly placed in the frequency axis to simulate human ear properties. Thus, there are more filters in the low-frequency region and fewer filters in the high-frequency region. A Mel-scaled power spectrogram of a signal can be found by applying Mel-scaled filters to the power spectrum of a signal and the neural network works much better if the Mel-scaled power spectrogram is used instead of the spectrogram [72]. First, the signal is divided equally into small sections

of short duration (20 to 30 ms) known as frames. Then, each frame is multiplied by the Hamming window. The Hamming window can be expressed as:

$$w[n] = 0.54 - 0.46cos(\frac{2\pi n}{N-1}) \tag{5.5.1}$$

where $0 \leq n \leq N-1$, and N is the window length. Then, the Discrete Fourier transform (DFT) is applied to convert the signal from the time domain to the frequency domain. The Mel-scale filter-banks are computed as follows:

$$m = 2595\ ln(\frac{f}{700} + 1) \tag{5.5.2}$$

where f is the frequency in the linear scale, and m is the resulting frequency in Mel-scale. Now, the Mel-scaled power spectrogram of the signal is obtained by applying Mel-scale filter-banks to the power spectrum of the signal and the log of the energy output of each filter. This can be expressed as:

$$S[m] = log(\sum_{k=0}^{N-1} |x[k]|^2 H_m[k]) \tag{5.5.3}$$

where $H_m[k]$ is the filter-banks, and m is the number of the filter-bank. The process of obtaining the Mel-scaled power spectrogram of a signal is shown in Fig. 13.

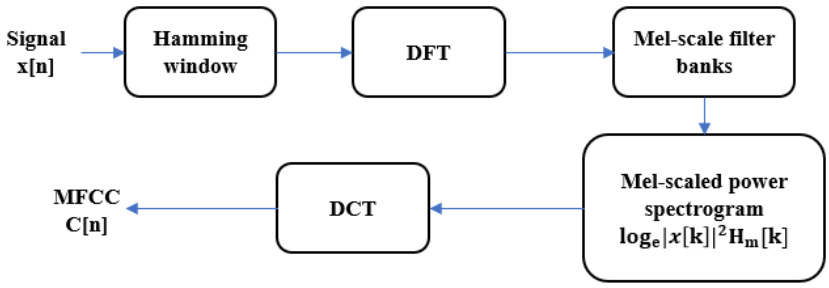

Figure 13: Feature extraction process by using Mel-scaled power spectrogram and MFCC

Fig. 14 shows the Mel-scaled power spectrogram of a normal signal (without murmurs) and an abnormal signal (with murmurs), respectively.

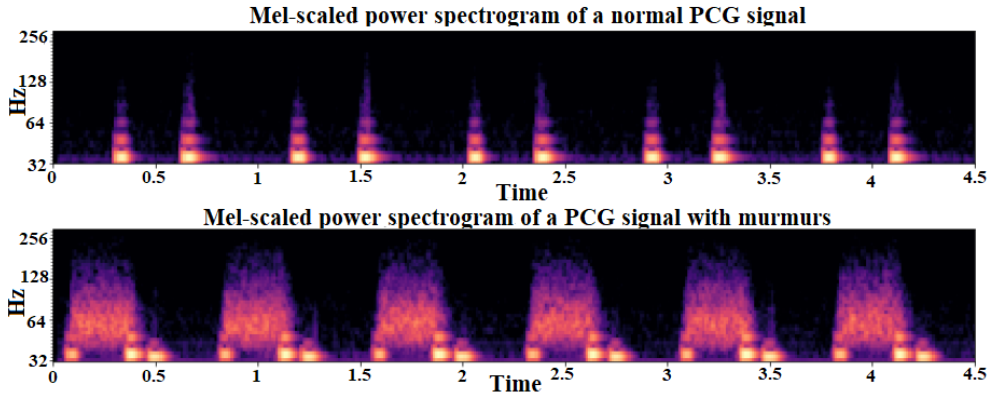

Figure 14: a) Mel-scaled power spectrogram of a normal PCG signal b) Mel-scaled power spectrogram of a PCG signal with murmurs

5.5.2 MFCC

MFCC is the compressed representation of the Mel-scaled power spectrogram, which can be found by taking the Discrete Cosine Transform (DCT) of a log power spectrum on a nonlinear Mel-scale of frequency [72]. The process of obtaining the Mel-scaled power spectrogram of a signal is shown in Fig. 15. The DCT of the spectrum to obtain the MFCC can be represented as:

$$c[n] = \sum_{m=0}^{M-1} S[m]\ cos(\frac{\pi n}{M}(m - \frac{1}{2})), n = 0, 1, 2, ..., M \tag{5.5.4}$$

where M is the total number of filter banks. Fig. 15 shows the MFCC of a normal signal (without murmurs) and an abnormal signal (with murmurs), respectively.

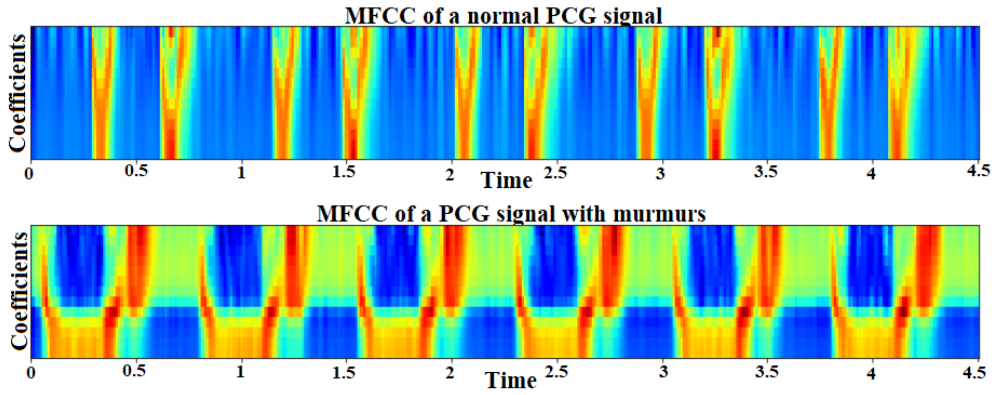

Figure 15: a) MFCC of a normal PCG signal b) MFCC of a PCG signal with murmurs

5.6 Hyper parameters tuning and resampling techniques

Hyper parameters play a very pivotal role to design ML and DL models. They are the variables which define the basic structure of the model. The training processes of the ML and DL models also depend on hyper parameters. In our research, we used hyper parameters optimization technique to tune different parameters for better performance. We used grid search method to tune the following parameters for the optimized final result:

- **Hidden Layers:** Hidden layers stay between input and output layers. We tuned the number of hidden layers with regularization technique to increase the accuracy by reducing underfitting.

- **Dropout:** It is a regularization technique which we tuned to increase validation accuracy by avoiding overfitting.

- **Weight Initialization and Activation Function:** Weight initialization and activation function help to prevent activation outputs from exploding or vanishing gradients during forward propagation and back propagation.Thus it will take long time to converge. We tuned different weight initialization techniques and activation functions to

avoid these problems.

- **Learning Rate:** The learning rate defines the step size by which a network update its parameter to minimize the loss function. A higher learning rate may increase the speed but network won't converge. On the other hand, a lower learning rate may cause under fitting. So, choosing the optimal learning rate is critical to design a model. We tuned the learning rate to increase the accuracy.

- **Number of Neurons and Epochs:** There is no technique yet to determine the number of neurons and epochs which will give the best result. The preferred way is to repeat the process by changing the number of neurons and epochs. We tuned the number of neurons and number of epochs to get our desired result.

- **Batch Size:** Batch size defines the number of training samples sent to the model after which parameters get updated. We also checked with different batch sizes to get the optimized result.

The data set we used in our research is not balanced. The imbalance ratio of normal signal to abnormal signal is 1:4. We needed equal number of samples for each class to increase the classification accuracy as ML and DL algorithms fail if the data set is not balanced. So, beside hyper parameters tuning, we also used over sampling and under sampling techniques to make the data set balanced.

CHAPTER 6

RESULTS AND ANALYSIS

In this section, we verified the performance of the proposed compression, denoising, segmentation, and classification algorithms, respectively. An extensive simulation was carried out using the Python programming language to implement these algorithms.

6.1 PCG dataset

In this paper, the well-known University of Michigan Heart Sound and Murmur Library [73] and the 2016 PhysioNet Computing in Cardiology Challenge database [59] were used for evaluating the performance of different algorithms. There are 23 PCG signals in the University of Michigan Heart Sound and Murmur Library database including 5 normal and 18 pathological. The 2016 PhysioNet Computing in Cardiology Challenge database consists of 6 datasets (A through F) containing a total of 3,240 unique heart sound recordings. The recordings from these 2 databases were collected from both healthy people and patients with confirmed cardiac diseases. A total of 123 unique PCG signals were used from these 2 databases to validate our segmentation algorithm. For the classification, we used all the 3,240 PCG signals available in the 2016 PhysioNet Computing in Cardiology Challenge database. These 2 databases are not balanced. The imbalance ratio of normal heart sounds to abnormal heart sounds is 1:4.

6.2 ECG dataset

we used the labeled MIT- BIH Arrhythmia dataset [74]. This dataset has been widely used to classify different cardiac arrhythmias. This dataset consists of 48 ECG recordings collected from 47 different patients of Beth Israel Hospital (BIH). Each recording is a minimum of 30 minutes long and sampled at 360 Hz. Each heartbeat was annotated by at least two cardiovascular specialists. Annotations in this dataset were used to create 1,09,449 heart beats. This database is not balanced. The imbalance ratio of normal beat to other beat categories is 1:4. Table 2 shows the number of beats in each class. Fig. 16 shows 1 beat ECG of every category.

Table 2: Breakdown of the heartbeats of five different classes

Type	Number of beats
Normal	90592
Atrial premature	2781
Premature ventricular contraction	7235
Fusion of ventricular	802
Fusion of paced/unclassifiable	8039
Total	109449

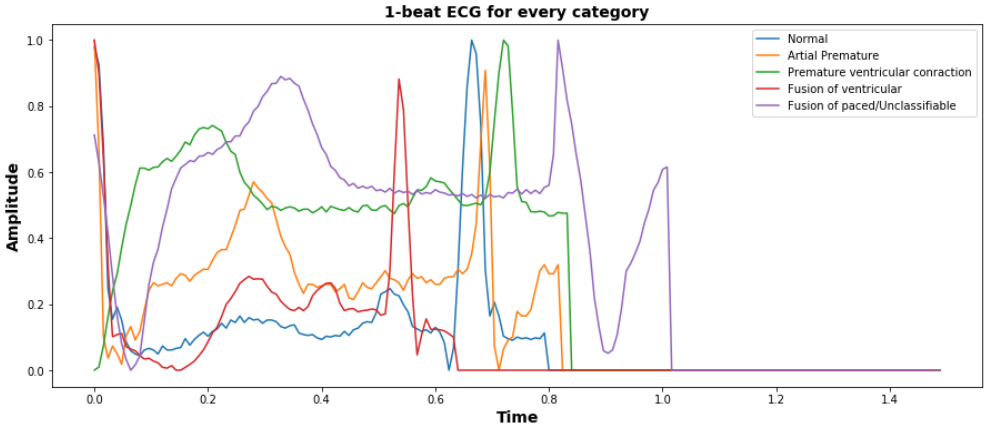

Figure 16: 1 beat ECG of total five categories

6.3 Results of the compression and denoising of PCG signals

The main aim of an efficient compression process is to remove all the redundant information from the signal without losing any data containing pathological information [45]. It can be accomplished by decreasing the significant coefficients, thus it will save the memory, less bandwidth will be required, and the transmission will be faster [44,46,47]. Furthermore, the presence of different kinds of noise and murmurs make it very difficult to extract the correct diagnostic information and features from PCG signals. In some cases it is almost impossible to segment PCG signals because of the noise. Therefore, it is necessary to eliminate noise and isolate murmurs from PCG signals before segmentation. These isolated murmurs can be picked up for further processing. Our proposed compression algorithm consists of decomposing the acquired signal first using DWT, then replacing the small wavelet coefficients by using an adaptive multilevel thresholding method based on the energy compaction property of the wavelet coefficients. Run-length encoding (RLE) is used to encode and to reduce the physical size of the repeating characters in the sender end. This compressed, denoised and encoded signal will be transferred to the physician through the wireless network. At the receiver end this compressed signal will be decoded by Run-length decoding (RLD) and then the signal will be reconstructed by using IDWT for proper cardiac diagnosis [47, 48]. Thus maximal redundant data retrenchment and denoising is possible by preserving all the pathological information [5, 7, 44].

6.3.1 DWT of PCG signals

PCG signals are decomposed with 6 layers multi-scale wavelet transform. The optimized compression of a signal depends on the decomposition level which is related to the sampling frequency of the signal [29]. According to the sampling theorem, the highest frequency of a signal is half of its sampling frequency. So, the highest frequency of all the PCG signals in the database is 22050 Hz. Now the signals are decomposed in such a way that their approximation band (lowest resolution band) contain most of the information as well as the energy. Fig. 17 shows the frequency spectrum of a normal PCG signal , which shows most of the information of the signal are within $0-250$ Hz range. So, the signal is decomposed up to 6th level to cover $0-250$ Hz by approximation band. Each sub-band and their frequency spectrum are shown in the Table 3.

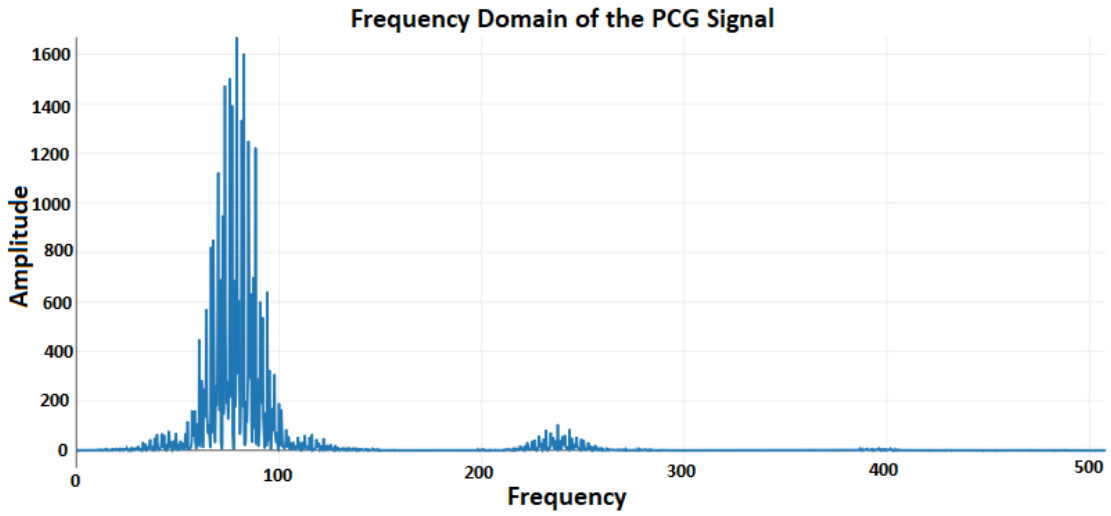

Figure 17: Frequency spectrum of a normal PCG signal

The selection of the best mother wavelet is also very crucial for the perfect reconstruction of the compressed signal [49, 50]. To evaluate the performance of the best mother wavelet an extensive simulation is carried out among 20 wavelets from Daubechies family, 5 wavelets

Table 3: Different levels of PCG signals and their corresponding coefficients and frequency spectrum

Levels	Frequency Range (Hz)	Coefficients	Sub-Bands
6	0 to 344.531	3135	App. Band (A_6)
6	344.531 to 689.025	3135	Detail Band (D_6)
5	689.0625 to 1378.125	6235	Detail Band (D_5)
4	1378.125 to 2756.25	12435	Detail Band (D_4)
3	2756.25 to 5512.5	24836	Detail Band (D_3)
2	5512.5 to 11025	49638	Detail Band (D_2)
1	11025 to 22050	99242	Detail Band (D_1)

from Coiflets family and 15 wavelets from both Biorthogonal and Reverse-biorthogonal families (total 55 orthogonal wavelets). Among all of these wavelets db18 wavelet was chosen from the Daubechies family as it outperformed all other wavelets by giving the smallest value of PRD with maximum energy in the approximation band. The performances of 14 different mother wavelets are shown in Table 4, which shows that best result was achieved with db18 as it gave very low PRD and highest energy in the approximation band compared to the other wavelets.

Table 4: Performance Analysis of 14 different wavelets to reconstruct PCG signals

Wavelets	PRD (%)	App. Band Energy (%)	Wavelets	PRD (%)	App. Band Energy (%)
db1	10.5011	95.6588	bior2.2	0.5949	99.8650
db3	0.4303	99.9243	bior3.7	0.1003	99.9880
db11	0.0676	99.9935	bior4.4	0.1458	99.9813
db18	**0.0613**	**99.9964**	rbio2.2	5.0183	98.8315
db20	0.0651	99.9944	rbio3.7	0.2036	99.9423
coif1	1.5591	99.6369	rbio4.4	0.3055	99.9389
coif3	0.0959	99.9845	rbio6.8	0.0924	99.9829

6.3.2 Threshold of the DWT coefficients

Among all the DWT coefficients of different sub-bands, only a small number of coefficients hold significant diagnostic information while the others hold negligible details [51]. So, the goal of the thresholding is to extract those all important coefficients by ignoring others to

discard the redundant information from the signal. Though a high data shrinkage is possible by using a large threshold value, it will produce a low quality reconstructed signal. Similarly, a small threshold will give a high quality reconstructed signal but it will give very low data compression [53]. So, choosing an optimal threshold value is the prerequisite for an efficient data compression. In our research, an adaptive multilevel thresholding is used based on the energy of the coefficients [23, 52, 54, 55]. From Table 3, we can see that approximation sub-band has the lowest frequency spectrum but it contains most of the information as well as the energy of the signal. On the other hand detail sub-bands contain less information as well as less energy with low amplitude.

The energy packing efficiency (EPE) is a percentage quantity which is known as the ratio of the energy of the coefficients of a fixed sub-band and the energy of all the coefficients of the signal [45, 54]. The mathematical expression of EPE is given below:

$$EPE = \frac{\sum_{n=1}^{k_i}[c(n)]^2}{\sum_{n=1}^{k}[c(n)]^2} \times 100\% \tag{6.3.1}$$

where n is an integer, k_i is the number of DWT coefficients in ith sub-band and k is the number of all the DWT coefficients. The contribution of the energy of each sub-band with the number of coefficients is shown in Table 5. Different signals in different sub-bands are shown in Fig. 18.

Table 5: EPE of different sub-bands

Sub-bands	Energy	Value of EPE (%)	Coefficients
App. Band (A_6)	6259.3340	99.9964	3135
Detail Band (D_6)	0.2257	3.605×10^{-05}	3135
Detail Band (D_5)	0.0024	3.689×10^{-07}	6235
Detail Band (D_4)	1.7669×10^{-05}	2.823×10^{-09}	12435
Detail Band (D_3)	5.9396×10^{-06}	9.489×10^{-10}	24836
Detail Band (D_2)	9.6516×10^{-06}	1.542×10^{-09}	49638
Detail Band (D_1)	1.5877×10^{-05}	2.5365×10^{-09}	99242

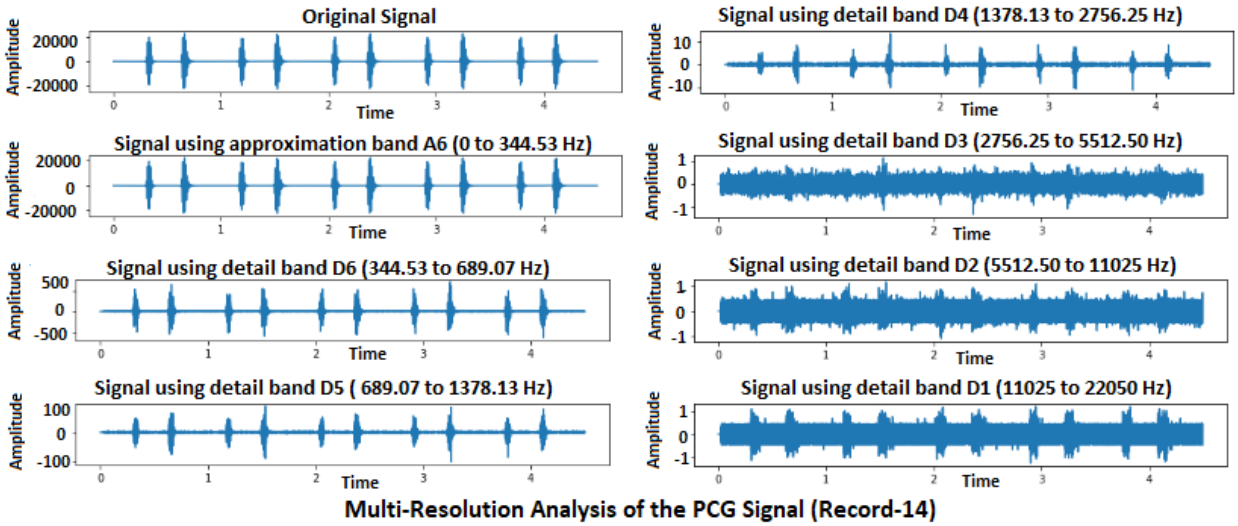

Figure 18: Different levels and their corresponding signals

The total energy of the signal is 6259.5620 with 198656 coefficients. Table 5 shows that about 99.9964% of the energy is stored in the 3135 coefficients of the approximation band while the other 195521 detail coefficients contain only the 0.0032% of the total energy of the signal. But we can't discard all the detail coefficients as it will cause signal distortion. So, sub-bands are separated into two groups for thresholding. Group A includes the approximation and detail band coefficients (A6 and D6) of level 6 and group B consists of the detail coefficients from level 1 to 5 (D1-D5). The sub-bands under group A are kept unchanged and the value of the threshold for all the sub-bands under group B is calculated in such a way that after thresholding, the conserved energy in all the detail sub-bands of group B is α% of their previous energy level before thresholding. The value of α can be adjusted further to change the threshold. Then the value of the coefficients below the threshold level are converted to zero in each sub-band belongs to group B. Thus the number of zero increased after the thresholding. The number of total coefficients, significant coefficients, and zeros

before and after the thresholding in the decomposed signal are shown in Table 6.

Table 6: Significant coefficients and zero coefficients before and after thresholding

	Total Coefficients	Significant Coefficients	Zero Coefficients
Before Threshold	198656	197401	1285
After Threshold	198656	6236	192420

From Table 6, it can be observed that after thresholding, the number of significant coefficients decreased about 96.84%, that means only 3.14% coefficients of the total number of coefficients are now representing the whole signal without any distortion.

6.3.3 RLE technique to encode signals

In order to take advantage of DWT for signal compression, encoding is necessary. RLE is used in our research as it is the simplest encoding method and very easy to implement. RLE method compresses any signal by reducing the size of any type of repeating data sequence [48, 56]. It represents any consecutive runs of the identical value in the data as the count followed by the value, thus compresses the size. Repeating data values are typically encoded into two bytes. If a data value is represented by d and it runs n consecutive times then by using RLE it can be represented as <n d>. Here both count and values are 1 byte respectively [47, 48]. An example of RLE is shown in Fig. 19.

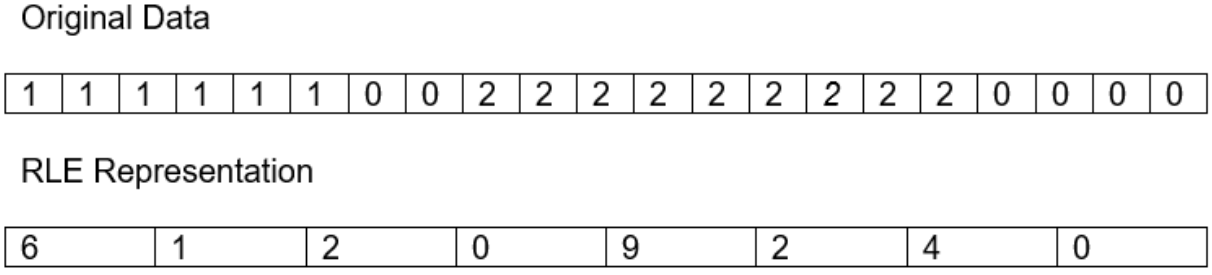

Figure 19: RLE of a signal with repeating characters

After RLE the total number of coefficients decreased from 198656 to only 6236 that means the number of coefficients reduced about 96.85%. If we convert this signal to an audio signal, we will find that the size of the compressed audio signal is only 25048 bytes whereas the size of the signal before the compression was 396944 bytes. That means the physical size of the signal decreased by 93.70%. So, it will require less storage to store the signal and less bandwidth to transfer the signal from sender end to receiver end. Another advantage of this encoding process is that the signal can be transferred securely by keeping all the information hidden. Only the receiver can decode the signal. So, secure signal transmission is also possible with our compression algorithm.

6.3.4 Signal reconstruction

After receiver received the encoded signal it's information is extracted by using RLD and IDWT in the receiver end. RLD is just opposite of RLE. RLD decompress the signal (restore all the zeros) and then the signal is reconstructed using IDWT [7].

6.3.5 Adding noise

Our proposed compression algorithm will not only compress and encrypt PCG signals but will also denoise PCG signals by removing high frequency noises and murmurs. We also checked the effectiveness of DWT to denoise PCG signals by adding some noises in PCG signals. The best way for evaluating the performance of a denoising algorithm is to add noise with the signal [7, 14]. Fig. 20 represents a normal original PCG signal and the signal after adding noise.

From the Fig. 20 we can see that white Gaussian noise is added to the signal to evaluate

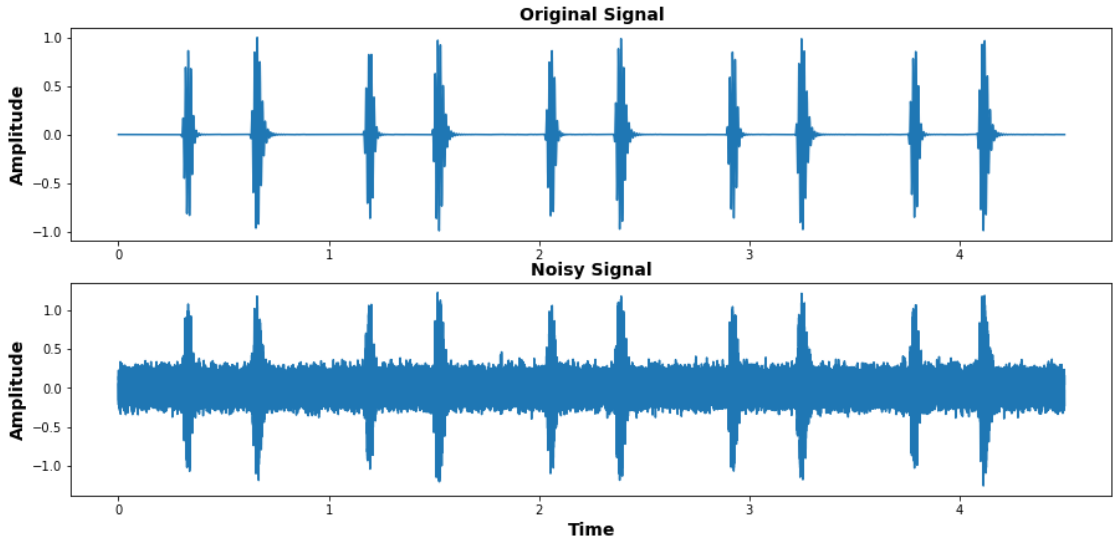

Figure 20: a) Original PCG signal b) Noisy PCG signal

the performance of the denoising algorithm. It can be defined as:

$$X_W(n) = X(n) + W(n) \tag{6.3.2}$$

where X(n) is the original PCG signal, W(n) is the White Gaussian noise and $X_W(n)$ is the noisy PCG signal.

6.3.6 DWT and filtering

The noisy signal is again decomposed with 6 layers multi-scale filter bank for getting the approximation and the detail sub-bands shown in Table 3. Then this reconstructed signal is filtered using digital lfilter available in SciPy library in python to discard the low frequency noise below 344.531 Hz [57]. The choice of mother wavelet is also very important for de-noising, as a better signal to noise ratio (SNR) can be achieved by using the wavelet which resembles the signal most and has high oscillation [29, 58]. A total of 55 orthogonal wavelets

were tested and bior4.4 wavelet from Bi-orthogonal family was selected as it outperformed all other wavelets. The performance of 14 different mother wavelets for denoising the PCG signal is shown in Table 7.

Table 7: Performance analysis of 14 different wavelets to denoise PCG signals

Wavelet	SNR (db)	Wavelet	SNR (db)
db1	15.79	bior2.2	21.12
db3	22.22	bior3.7	21.90
db11	20.10	**bior4.4**	**22.87**
db18	15.68	rbio2.2	7.85
db20	14.66	rbio2.2	7.85
coif1	17.30	rbio3.7	21.62
coif3	21.40	rbio6.8	21.37

6.3.7 Simulation result

The compression algorithm is applied to all of the signals available in the database and both high compression and SNR are achieved by maintaining the signal fidelity. Fig. 21 shows the original and compressed form of a PCG signal. Fig. 22 represents a noisy PCG signal and that PCG signal after removing noise.

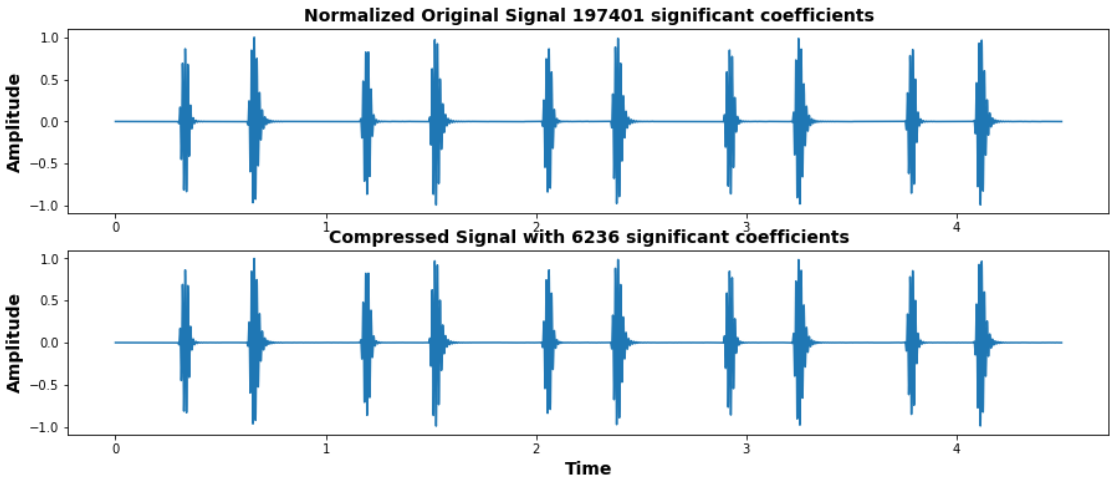

Figure 21: a) Original PCG signal b) Compressed PCG signal

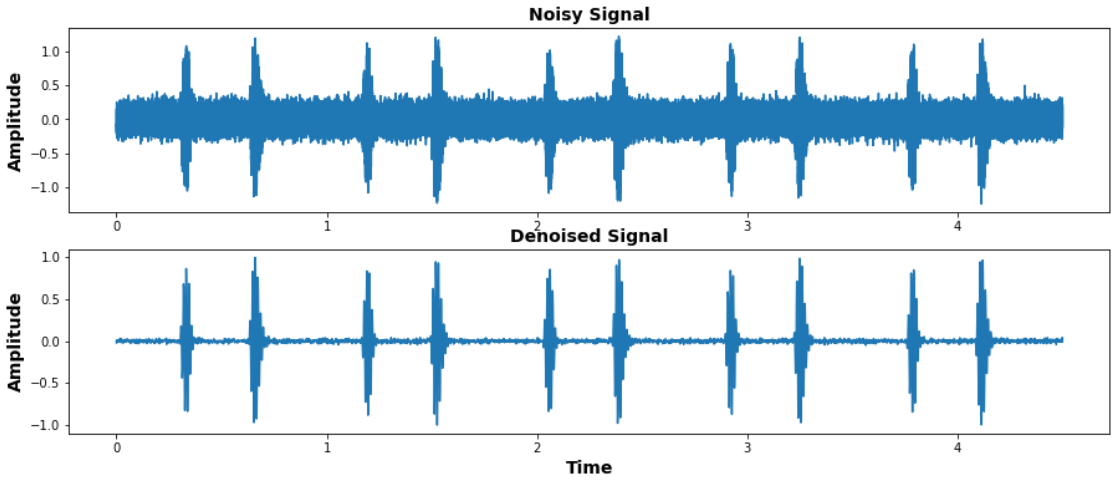

Figure 22: a) Noisy PCG signal b) Denoised PCG signal

The values of CR, PC, PRD, SNR before denoising and SNR after denoising of different recordings are presented in Table 8.

Table 8: The values of CR, PC, PRD and SNR of different PCG records

Records	CR	PC (%)	PRD (%)	SNR Before Denoising	SNR After Denoising
1	15.80	93.66	0.038	1.95	19.40
2	15.86	93.68	0.075	3.94	21.17
3	15.86	93.68	0.084	4.25	21.88
4	15.86	93.68	0.077	3.00	20.37
5	16.02	93.76	0.040	2.05	19.73
6	15.86	93.68	0.091	3.53	14.84
7	15.86	93.68	1.115	3.56	15.01
8	15.86	93.68	0.495	4.42	20.01
9	15.86	93.68	0.747	4.93	16.87
10	15.86	93.68	0.310	3.00	20.42
11	15.86	93.68	0.677	3.17	18.55
12	15.86	93.68	0.825	3.41	16.00
13	15.79	93.66	0.434	0.92	18.77
14	15.87	93.69	0.061	4.99	22.87
15	15.85	93.68	0.154	1.73	18.5
16	15.85	93.68	0.161	1.27	19.03
17	15.85	93.68	0.602	1.82	17.50
18	15.85	93.68	0.061	4.99	12.98
19	15.85	93.68	0.071	4.29	21.87
20	15.85	93.68	0.066	3.97	21.90
21	15.80	93.66	0.605	3.48	20.90
22	15.85	93.68	0.133	4.65	20.83
23	15.80	93.66	0.665	3.38	20.53

From the above figures (Fig. 21 and 22) and Table 8, we can see that the signals are compressed about 93.68% with an average small PRD of $< 0.50\%$ and also all the pathological information is preserved in the signal. Moreover, a high SNR can also be achieved by using

our proposed compression algorithm.

6.3.8 Comparison with other PCG compression methods

The performance of this compression technique is compared with other compression techniques presented in the paper [9]. By comparing the QS of different algorithms we can evaluate their performance. Table 9 shows the QS of our proposed algorithm along with the QS of those algorithms reported in paper [9].

Table 9: Performance comparison of different compression algorithms

Compression Technique	CR	PRD (%)	QS
Amplitude Zone Time Epoch Coding	10.00	28.00	0.36
Improved Amplitude Zone Time Epoch Coding	9.91	7.99	1.24
Coordinate Reduction Time Encoding System	4.80	7.00	0.69
Turning Point	2.00	5.10	0.40
Wavelet packet compression	8.00	2.60	3.08
Set Partitioning in Hierarchical Trees Algorithm	8.00	1.18	6.78
Linear prediction of the wavelet coefficients	11.60	5.30	2.19
Perceptual masks & Discrete Cosine Transform	3.50	1.24	2.82
Neural Network	12.74	0.61	20.89
Huffman coding	11.06	2.73	4.05
Wavelet Transform	12.00	0.98	12.25
Discrete Cosine Transform (Min CR)	6.20	1.50	4.13
Discrete Cosine Transform (Max CR)	10.90	3.00	3.63
ASCII character encoding	15.72	7.89	1.99
JPEG2000	20.00	3.26	6.134
DC equalization and complexity sorting	8.00	0.86	9.30
Mother wavelet modification	23.10	1.60	14.43
Fourier descriptors	7.40	7.00	1.05
Fast Fourier Transform	6.28	0.75	8.37
Fourier Transform (fixed strategy)	14.67	1.06	13.84
Fourier Transform (adaptive strategy)	16.58	1.07	15.49
Proposed model	15.84	0.50	31.68

From the Table 9, we can see that the average QS of the proposed method is around 31.68 while the QS of the other algorithms range from 0.357 to 20.885 which proved that the proposed compression algorithm outperformed all the other compression algorithms. Reddy et al. [75] used Fourier descriptors to compress the signal and achieved a QS of 1.05. They converted each signal as closed contour. Coordinates of each contour were represented as complex sequences. Then they did the FT of those complex sequences and calculated the

significant coefficients which are also known as Fourier descriptors. Shinde et al. [76] have used Fast Fourier Transform (FFT) to compress signals and got a QS of 8.37. Sadhukhan et al. [77] have increased the QS by applying fixed strategy and adaptive strategy of FT. The fixed strategy was based on the selection of a fixed band-limiting frequency, and the adaptive strategy was dependent on the spectral energy distribution of signals. They encoded significant coefficients to optimize the usage of bit and compress signals. By applying the significant coefficients they achieved a QS of 13.84 which was further improved to 15.49 by applying the adaptive strategy method. The AZTEC algorithm converts raw signals into plateaus. The amplitude and length of each plateau are stored for reconstruction. A poor QS of 0.36 was achieved through the AZTEC technique which was then improved to 1.24 through the improved AZTEC technique. Improved AZTEC algorithm optimizes the trade off between CR and PRD. The TP data reduction algorithm reduces the sampling frequency of signals to compress the signal which has very low QS of 0.40. The CORTES algorithm is a hybrid approach of AZTEC and TP to achieve high CR of the AZTEC and the low reconstruction error of the TP technique. The DCT technique compress signals by restoring the signal information in a fixed number of DCT coefficients. Other techniques such as Huffman coding, ASCII character encoding, JPEG2000, Set Partitioning in Hierarchical Trees Algorithm (SPIHT), DC equalization and complexity sorting, and Neural Network (NN) are also applied to compress 1-D signals [9]. DWT outperformed these previous techniques to compress 1-D non-stationary signals by giving high CR with low PRD.

6.4 Results of the segmentation of PCG signals

Segmentation of the PCG signals facilitates to get the exact position and duration of the heart sounds, systole and diastole intervals. But the presence of murmurs which produced by turbulent blood flow is the major reason for the false detection of heart sounds. So the elimination of murmur is necessary before segmentation. After discarding the murmurs, Shannon energy envelope is used to detect the boundaries of each heart sound and zero crossing algorithm is used to detect the point where the signal crosses the zero and thus the location and duration of the heart sound can be attained. Then the peak of the heart sounds is calculated to get their amplitude. Fig. 23 shows the block diagram of the segmentation algorithm.

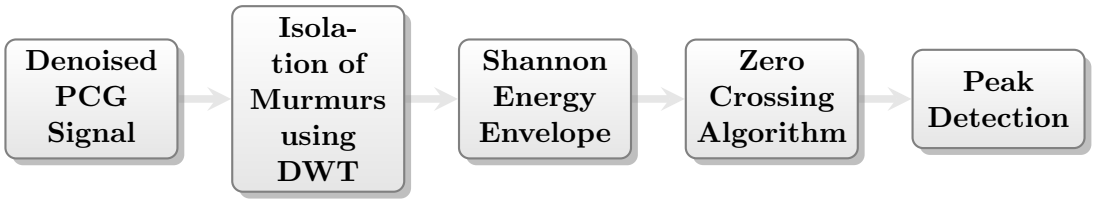

Figure 23: Function Diagram of the segmentation algorithm

6.4.1 DWT to separate murmurs

Murmurs usually exhibit higher frequency than the heart sounds. So to isolate murmurs we need to discard the frequencies higher than the frequency range of heart sounds. Fig. 24 represents the frequency spectrum of a normal and an abnormal PCG signal. Fig. 24 shows that heart sounds stay within the frequency range of 10 to 150 Hz and murmurs exhibit higher frequency than heart sounds. So by using DWT at the 7th level of decomposition the frequency range of the approximation band (0 Hz to 172.28 Hz) covers the frequency spec-

trum of the heart sounds. Murmurs with the higher frequency are eliminated by discarding all other sub-bands except the approximation band. Fig. 25 represents a pathological PCG signal with murmurs and after the isolation of the murmurs from that PCG signal.

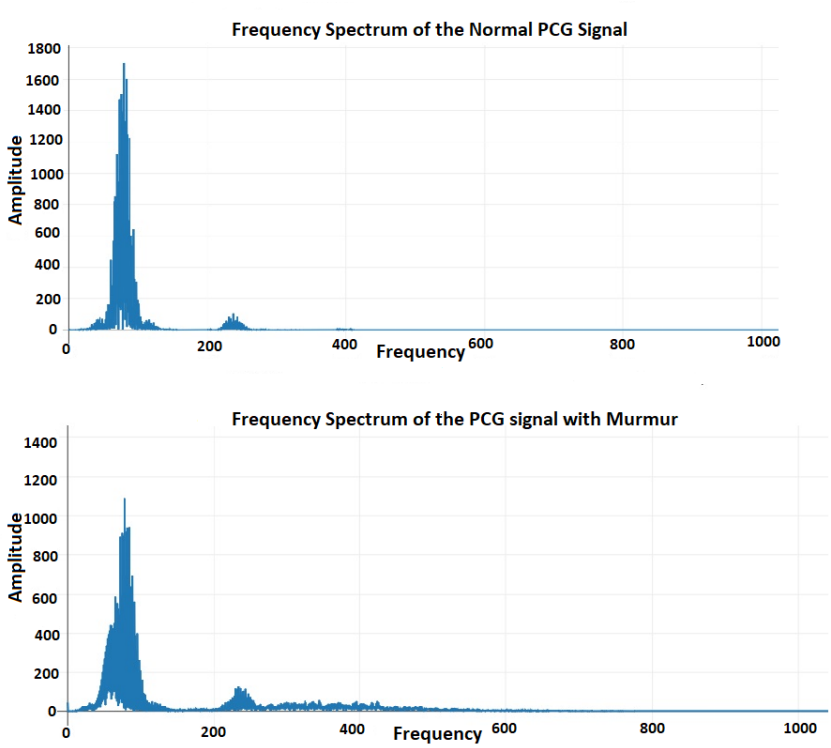

Figure 24: Frequency Spectrum of PCG signals: a) Without murmurs b) With murmurs

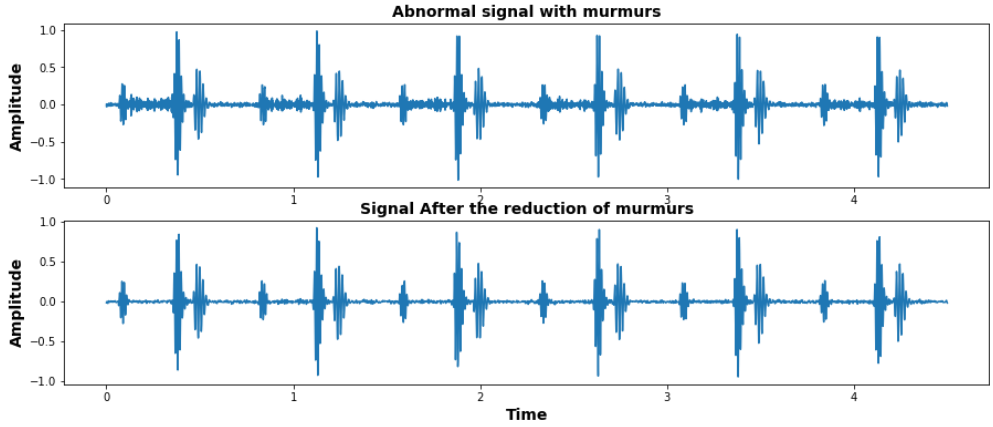

Figure 25: a) Pathological PCG signal with murmurs b) PCG signal after the isolation of the murmurs

6.4.2 Simulation results

The segmentation technique, based on the Shannon energy envelope and the zero-crossing algorithm, effectively extracted all the important characteristics of PCG signals. A threshold was set to discard the effect of the noise and the low amplitude signal. Fig. 26 and 27 represent the original signal, the signal after removing noise and murmurs, Shannon energy envelope, zero-crossing, and peak detection of a normal and an abnormal PCG signal, respectively. The time duration and amplitude of the four basic heart sounds, systole interval, diastole interval, one cardiac cycle, and heart rate information extracted from a normal and an abnormal PCG signal are shown in Table 10.

Table 10: Extraction of the cardiac information from PCG signals

Cardiac parameter	Value in a normal signal	Value in an abnormal signal
Duration of S1	0.06 second	0.04 second
Duration of S2	0.08 second	0.07 second
Duration of S3	–	0.07 second
Amplitude of S1	0.94	0.30
Amplitude of S2	1.10	0.98
Amplitude of S3	–	0.52
Systole interval	0.32 second	0.30 second
Diastole interval	0.54 second	0.47 second
One cardiac cycle	0.86 second	0.77 second
Heart rate	70 beats per minute	80 beats per minute

6.5 Results of the classification of PCG signals

In this section, we verified the performance of the proposed PCG classification models. We have used the Python programming language to implement these architectures. To evaluate the effectiveness of the proposed DNN and CNN models, we have compared the performance of these DL models with other traditional ML algorithms in terms of sensitivity/recall, specificity, and accuracy.

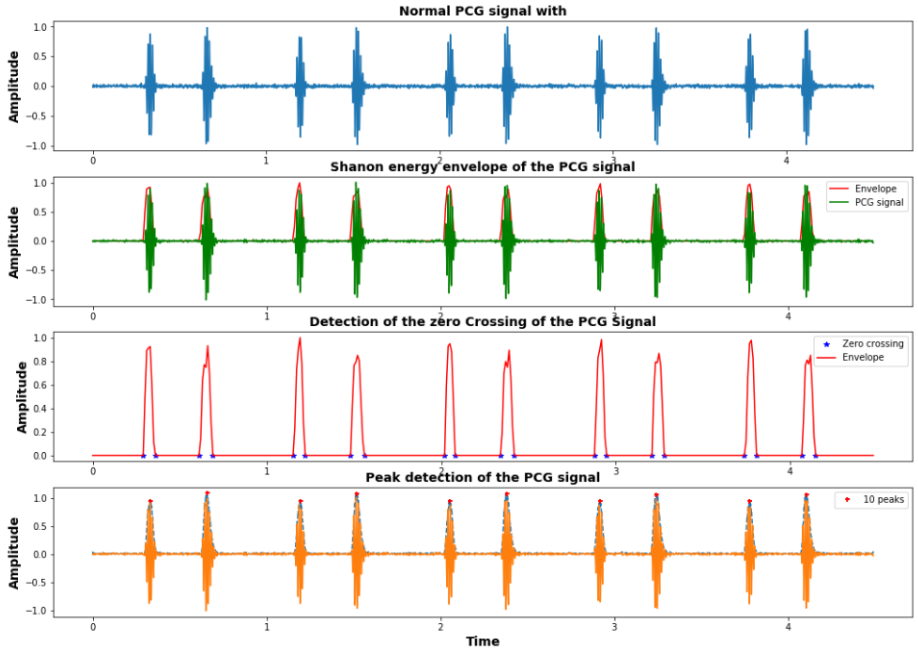

Figure 26: a) Original normal PCG signal b) Shannon energy envelope of the PCG signal c) Zero-crossing of the PCG signal d) Peak detection of the heart sounds

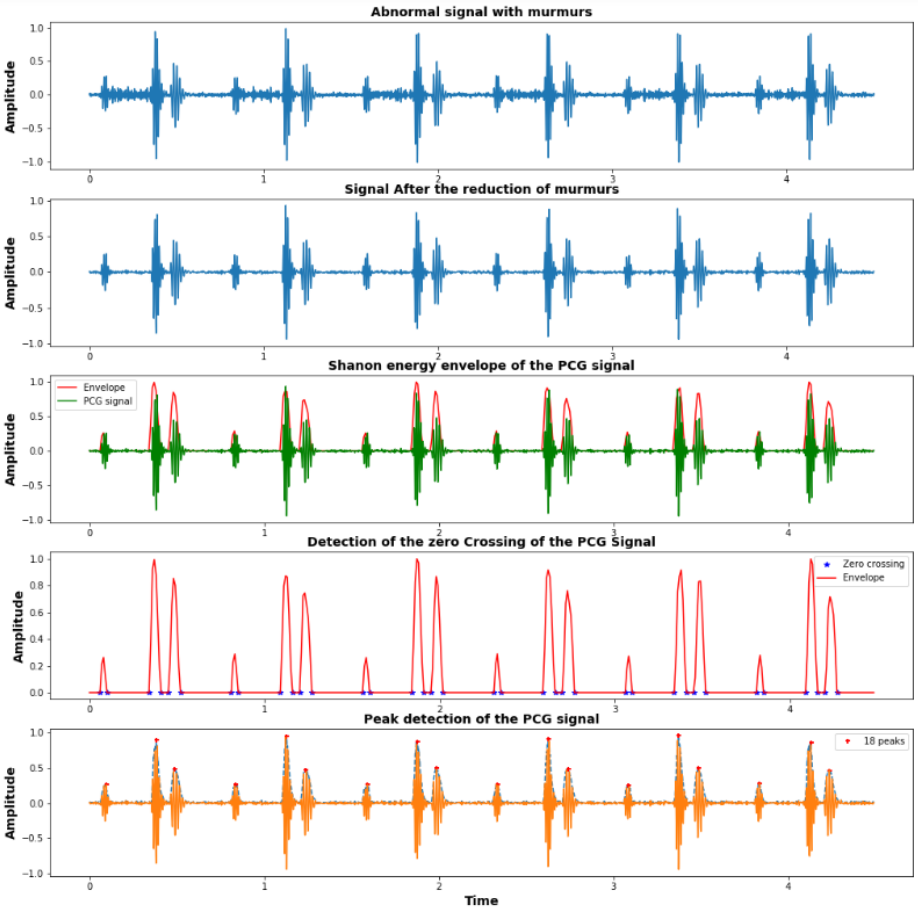

Figure 27: a) Original abnormal PCG signal. b) Denoising and isolation of the murmurs from the PCG signal after reconstruction. c) Shannon energy envelope of the PCG Signal d) Zero-crossing of the PCG signal e) Peak detection of the heart sounds

6.5.1 Proposed DNN model to classify PCG signals

A 5-layer sequential feed-forward DNN model trained by Keras was used in our research to classify the PCG signal into two categories, either normal or abnormal. Keras is the high-level API of TensorFlow, which we used to train our classifying model with great speed. Mel-scaled power spectrogram and MFCC were used to extract meaningful features from each heart sound of the database. A total of 40 features were achieved from each of the PCG signals using these two methods, which were then fed into the DNN to train the model. Afterwards, 5 hidden layers with 256, 512, 768, 1,024, and 1,280 filters were implemented with the ReLU activation function for non-linearity. In the output layer, the Sigmoid activation function was used to get the probability distribution, which we applied on the cross-entropy cost function. The cross-entropy cost function was used to measure how far apart the output of the model was from that of the desired or target output. The Adam optimizer was used to minimize the cost function. The training started with a learning rate of 0.0001 and continued until it reached the maximum number of epochs. The dropout technique was used in the model to reduce independent learning among the neurons and to handle overfitting. Hyper parameter optimization technique was used to get optimal hyper parameters for the DNN model. After training the model its prediction capability was tested on the testing set. Fig. 28 shows the proposed DNN model of five layers to classify PCG signals.

6.5.2 Proposed CNN model to classify PCG signals

A 4-layer sequential CNN model trained by Keras was used in our research to classify the PCG signal into two categories, either normal or abnormal. Keras is we used to train our

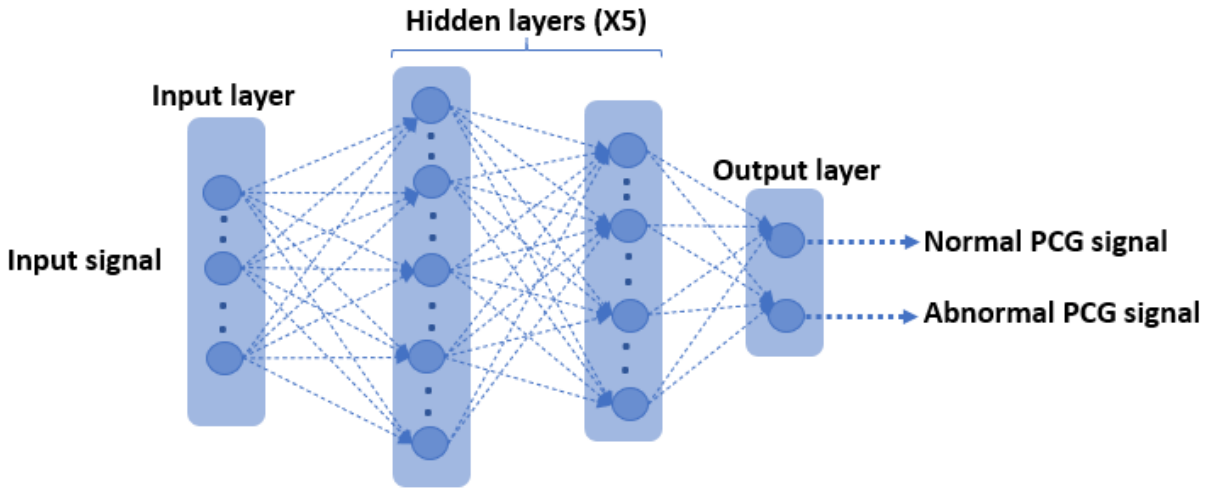

Figure 28: Proposed DNN model to classify PCG signals

classifying model with great speed. Mel-scaled power spectrogram and MFCC were used to extract meaningful features from each heart sound of the database. A total of 40 features were achieved from each of the PCG signals using these two methods, which were then fed into the CNN to train the model. We proposed a 1D CNN model of 4 hidden layers to classify each PCG signal in the database. Four hidden layers with 256, 512, 1024, and 2048 filters were implemented with the ReLU activation function. Each layer had the same kernel size of 4 to combine the number of input features with the number of new output features. A max-pooling layer of kernel size 2 was also used in each layer to downsample the input and to reduce the number of dimensions. In the output layer, the Sigmoid activation function was used to predict the final classification output. A regularization technique called dropout was used in the model to reduce independent learning among the neurons and to handle overfitting. While training the model, the dropout technique usually deletes random samples of the activation by making them zero and helps the network to learn robust features that are useful to increase classification accuracy. The training started with a learning

rate of 0.0001 and continued until it reached the maximum number of epochs. Network weights had been updated iteratively in each epoch based on training data using the Adam optimization algorithm. Hyper parameter optimization technique was used to get optimal hyper parameters for the CNN model. After training, the model's prediction capability was tested on the testing set. Fig. 29 shows the proposed CNN model of four layers to classify PCG signals.

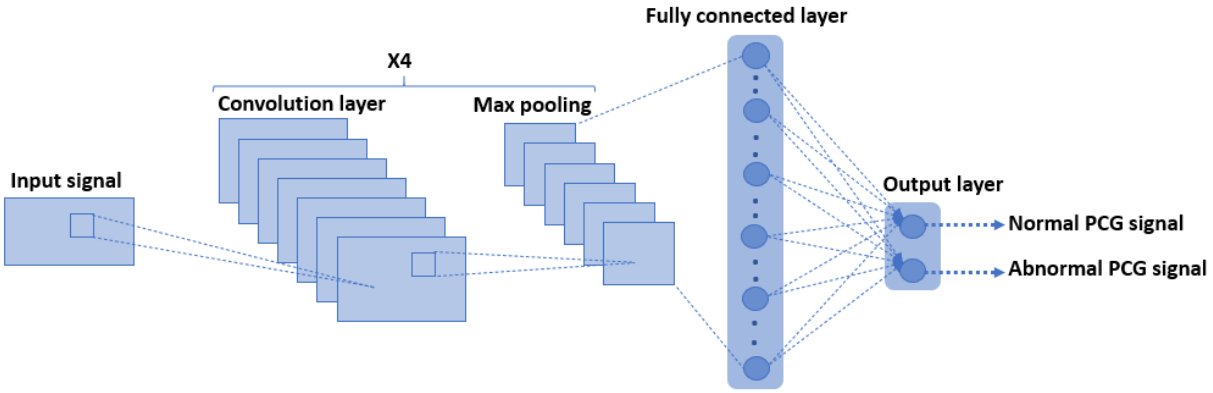

Figure 29: Proposed CNN model to to classify PCG signals

6.5.3 Proposed DHL models to classify PCG signals

Fig. 30 illustrates the design of the proposed deep hybrid architecture. It has four 1D convolution layers with 256, 512, 1,024, and 2048 filters, respectively. All the layers have the same kernel size of 2. Each convolution layer uses a Rectified Linear Unit (ReLU) activation function, and each max pooling layer is of size 2. A dropout layer with 10% dropping rate is connected with each convolution layer to handle overfitting. After the convolution and max pooling, the learned features are flattened to one long vector and are passed to a fully connected layer with 512 filters. The fully connected layer also used the ReLU activation function and 20% dropout rate to reduce overfitting problems. The fully connected layer

works as a buffer between the learned features and the output. The cost function is minimized by using the Adam optimizer. The number of training epochs, batch size, and the learning rate are set to 100, 64, and 0.0001, respectively. After optimizing and training the CNN, the learned features from the fully connected layer are passed on to the ML classifiers for the final prediction task.

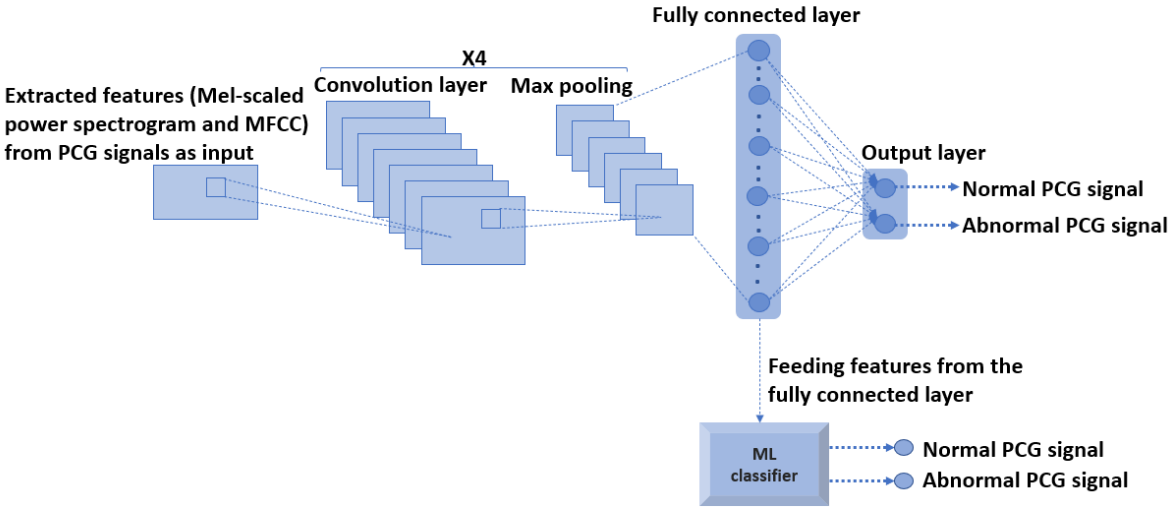

Figure 30: The proposed DHL model architecture to classify PCG signals

6.5.4 Simulation results

Our proposed DL models can discriminate between normal and abnormal PCG signals with great accuracy compared to other traditional ML algorithms. Our proposed DL models outperformed ML algorithms in terms of sensitivity, specificity and accuracy. We have used the 10-fold cross validation technique to test the performance of ML, DL, and DHL models. We got almost similar accuracy with our proposed DNN and CNN models. Our proposed DNN model detected normal and abnormal PCG signals with a very good testing accuracy of 91.70%. The achieved sensitivity/recall and the specificity of the proposed DNN model

were 85.45% and 92.93%, respectively. The CNN model gave an accuracy of 92.00%. The achieved sensitivity/recall and the specificity of the CNN model were 87.69% and 93.05%, respectively.

Table 11: Comparison of the proposed DL models with ML models to classify PCG signals

Classification Model	Sensitivity (%)	Specificity (%)	Accuracy (%)
DT	70.30	88.00	84.60
RF	78.13	90.70	88.30
LR	71.85	63.47	65.10
SVM	76.50	58.85	62.30
KNN	75.00	82.70	81.20
AB	76.50	84.23	82.70
NB	34.38	78.07	69.40
CNN	**87.69**	**93.05**	**92.00**
DNN	85.45	92.93	91.70

Table 11 shows the performance of the proposed DNN and CNN models with the other traditional ML models. The performance of the proposed hybrid models and other traditional ML and DL models with separate implementations are shown in Table 12.

Table 12: Comparison of the proposed DHL models with ML and DL classification models implemented separately

Classification Model	Sensitivity (%)	Specificity (%)	Accuracy (%)
CNN	87.69	93.44	92.00
LR	72.33	66.25	67.50
CNN-LR	94.58	92.27	92.70
RF	75.48	85.35	83.33
CNN-RF	**92.03**	**94.83**	**94.30**
KNN	75.19	80.12	79.10
CNN-KNN	95.33	90.79	91.72
DT	78.19	76.42	76.79
CNN-DT	92.03	92.47	92.37
NB	70.22	67.30	75.33
CNN-NB	97.00	90.17	91.60
SVM	77.29	80.15	79.56
CNN-SVM	93.30	91.80	92.10
AB	74.14	81.13	79.70
CNN-AB	94.14	92.23	92.62

As shown in Table 12, among all the ML classifiers, LR classifier performed worst with 67.50% accuracy, and best result is achieved by the RF classifier with 83.33% accuracy. The classifier learned with a single CNN model performed better than all other ML models with

a classification accuracy of 92.00%. Our proposed deep hybrid CNN-ML methods were able to improve the accuracy obtained from a single CNN model and seven ML models which were implemented separately. The best classification accuracy of 94.30% is achieved by CNN-RF models followed by CNN-LR (92.70%), CNN-AB (92.62%), CNN-DT (92.37%), and CNN-SVM (92.10%) models. Though CNN-KNN and CNN-NB hybrid models showed around 92.00% accuracy, which is the same as the single CNN model, they showed significant improvement in the sensitivity of the model. The sensitivity of CNN-KNN and CNN-NB models are 95.33% and 97.00%, respectively. The sensitivity of both models is higher than that of the single CNN model (87.69%). The specificity of the single CNN model is 93.44%. For the specificity, there is a 3.00% reduction from the single CNN to the CNN-KNN and CNN-NB models, respectively. Fig. 31 shows the accuracy of the proposed DHL model and Fig. 32 shows the reduction of the cost with respect to epochs of the proposed DHL model.

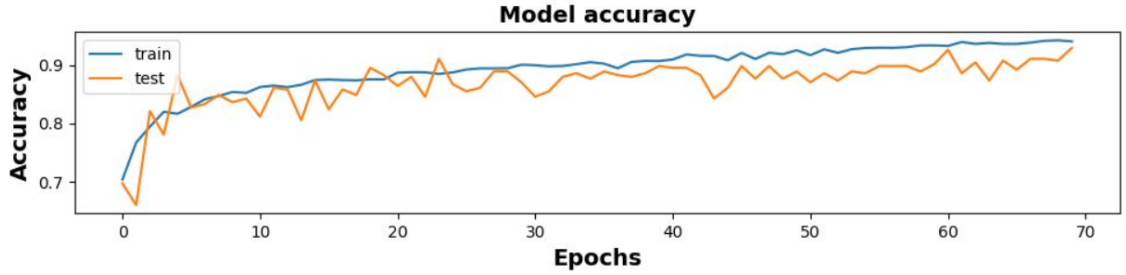

Figure 31: Training and testing accuracy of the proposed DHL model with respect to epochs

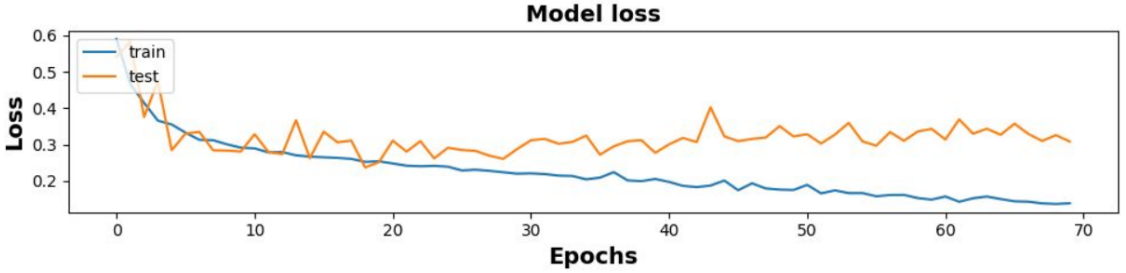

Figure 32: Training and testing loss of the proposed DHL model with respect to epochs

Based on the result, our proposed DHL models can be a promising solution to detect early-stage heart diseases by picking up potential abnormal PCG signals from a series of normal PCG signals.

6.5.5 Comparison with other PCG classification models

The main goal of the 2016 PhysioNet Computing in Cardiology Challenge [59] was to build a robust intelligent system that can detect anomaly in the PCG signal and can classify a PCG signal as normal or abnormal based on its features. The best overall accuracy achieved in the official phase of the 2016 PhysioNet Computing in Cardiology Challenge was 86.02% with sensitivity and specificity of 94.24% and 77.81%, respectively. Table 13 shows the comparison of our proposed PCG classification model with 12 other state-of-the-art PCG classification models. All these models used the same dataset published by the 2016 PhysioNet Computing in Cardiology Challenge. Presently, this is the largest database of PCG signals in the world.

Table 13: Comparison of the proposed DHL model with other PCG classification models

Author	Approach	Sen. (%)	Spec. (%)	Acc. (%)
Potes, (2016)	Time-Freq. features & AB-CNN	94.24	77.81	86.02
Nassralla, (2017)	Time-Freq. features & RF	78.00	98.00	92.00
Whitaker, (2017)	Sparse coding & SVM	90.00	88.45	89.26
Langley, (2017)	FFT-WT & DT	77.00	80.00	79.00
Han, (2018)	Segmentation & CNN	98.33	84.67	91.50
Tang, (2018)	Time-Freq. features & SVM	88.00	87.00	88.00
Dominguez, (2018)	Images & AlexNet	93.20	95.12	97.00
Sotaquirá, (2018)	DNN & Probability comparisons	91.30	93.80	92.60
Singh, (2019)	Time-Freq. features & KNN	93.00	90.00	90.00
Nogueira, (2019)	WT & SVM	90.45	85.25	87.85
Sing, (2020)	Time-Freq. features & AB	94.08	91.95	92.47
Krishnan, (2020)	Segmentation & DNN	86.73	84.75	85.65
Proposed Model	Time-Freq. features & CNN-RF	**92.03**	**94.83**	**94.30**

As shown in Table 13, the classification accuracy achieved from the previous models varied between 79.00% to 97.00%, whereas the range of the sensitivity and specificity varied

between 77.00% to 98.33% and 77.81% to 98.00%, respectively. It is noteworthy to mention that the AB-CNN model proposed by Potes et al. [60] was ranked 1st in the 2016 PhysioNet Computing in Cardiology Challenge. Nassaralla et al. [61] extracted time and frequency features of PCG signals to build a learning model using RF and DNN. Nassaralla et al obtained a very good accuracy and specificity of 92.00% and 98.00%, respectively, but low sensitivity of 78.00% using the RF classifier. On the other hand, Han et al. [64] reached an overall good accuracy and sensitivity of 91.50% and 98.33%, respectively, but with less specificity of 84.67%. They used complex segmentation of heart sounds and CNN to identify PCG signals. Krishnan et al. [71] also implemented segmentation of the cardiac cycle and achieved 85.65% accuracy. Sotaquirá et al. [67] used DNN and weighted probability comparison of each card cycle and got high accuracy of 92.60%. Langley et al. [63] obtained 79.00% accuracy without using complex segmentation technique. They used threshold-based classification tree for PCG classification. Singh et al. [68] initially applied KNN on unsegmented heart sounds recording and got 90.00% accuracy. Later, Singh et al. [70] improved the accuracy to 92.47% by applying AB classifiers. Whitaker et al. [62], Tang et al. [65], and Nogueira et al. [69] employed SVM with different structures to build their models and achieved 89.26%, 88.00%, and 87.85% accuracy, respectively. Dominguez et. al. [66] attained a great accuracy of 97.00% by employing the modified version of the AlexNet model but this model has high computational complexity.

6.6 Results of the compression and denoising of ECG signals

6.6.1 DWT of ECG signals

ECG signals are decomposed with 2 layers multi-scale DWT. The optimized compression of a signal depends on the decomposition level which is related to the sampling frequency of the signal [29]. According to the sampling theorem, the highest frequency of a signal is half of its sampling frequency. So, the highest frequency of all the ECG signals in the database is 180 Hz. The signal was decomposed in such a way that their approximation band (lowest resolution band) contain most of the information as well as the energy. So, signals were decomposed up to 2nd level to cover $0 - 45$ Hz by the approximation band. Each sub-band and their frequency spectrum are shown in the Fig. 33 and Table 14.

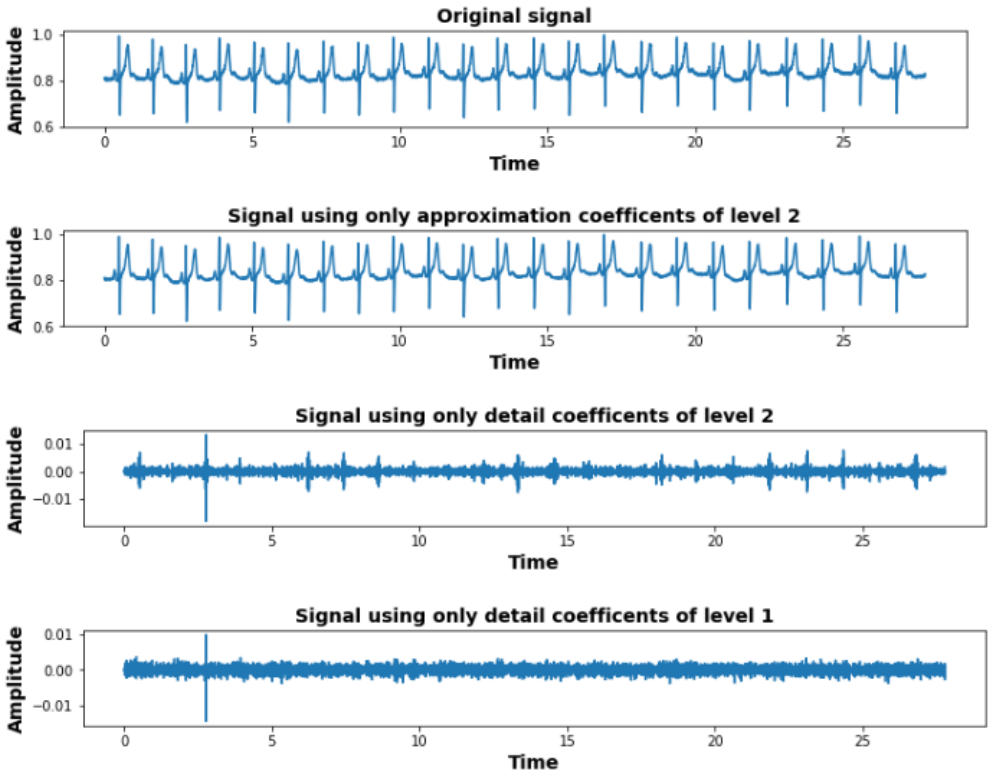

Figure 33: Different levels of ECG signals and their corresponding graphs

To evaluate the performance of the best mother wavelet an extensive simulation was carried out among 33 wavelets from Daubechies family, 10 wavelets from Coiflets family and 7 wavelets from both Biorthogonal (total 50 orthogonal wavelets). Among all of these wavelets coif6 wavelet is chosen from the coiflet family as it outperformed all other wavelets by giving the smallest value of PRD with maximum energy in the approximation band. The performances of 14 different mother wavelets are shown in Table 15, which shows that best result was achieved with coif6 as it gave very low PRD and highest QS compared to the other wavelets.

Table 14: Different levels of ECG signals and their corresponding coefficients and frequency spectrum

Levels	Frequency range (Hz)	Coefficients	Sub-bands
2	0 to 45	2526	App. Band (A_2)
2	45 to 90	2526	Detail Band (D_2)
1	90 to 180	5017	Detail Band (D_1)

Table 15: Performance analysis of 14 different wavelets to reconstruct ECG signals

Wavelets	PRD (%)	CR (%)	Wavelets	PRD (%)	CR (%)
db1	0.91	3.98	bior2.2	0.34	3.98
db3	0.35	3.97	bior1.5	0.96	3.95
db11	0.20	3.91	bior1.3	0.94	3.97
db18	0.19	3.87	coif3	0.23	3.93
db20	0.19	3.85	coif10	0.18	3.78
coif1	0.50	3.97	bio4.4	0.24	3.95
coif6	**0.17**	**3.95**	bio6.8	0.21	3.93

6.6.2 Threshold of the DWT coefficients

From Table 14, we can see that approximation sub-band has the lowest frequency spectrum but it contains most of the information as well as the energy of the signal. On the other hand detail sub-bands contain less information as well as less energy with low amplitude.

The contribution of the energy of each sub-band with the number of coefficients are shown in Table 16.

Table 16: EPE of different sub-bands

Sub-bands	Energy	Value of EPE (%)	Coefficients
App. Band (A_2)	7066.6520	99.9996	2526
Detail Band (D_2)	0.01339	1.89×10^{-06}	2526
Detail Band (D_1)	0.0088	1.24×10^{-06}	5017

The total energy of the signal is 7066.6743 with 10069 coefficients. Table 16 shows that about 99.9964% of the energy is stored in the 2526 coefficients of the approximation band while the other 7543 detail coefficients contain only the 0.0032% of the total energy of the signal. But we can't discard all the detail coefficients as it will cause signal distortion. So sub-bands are separated into two groups for thresholding. Group A includes the approximation band coefficients (A2) of level 2 and group B consists of the detail coefficients from level 1 and 2 (D1 & D2). The sub-bands under group A are kept unchanged and the value of the threshold for all the sub-bands under group B is calculated in such a way that after thresholding, the conserved energy in all the detail sub-bands of group B is α% of their previous energy level before thresholding. The value of α can be adjusted further to change the threshold. Then the value of the coefficients below the threshold level are converted to zero in each sub-band belongs to group B. Thus the number of zero increased after the thresholding. The number of total coefficients, significant coefficients, and zeros before and after the thresholding in the decomposed signal are shown in Table 17.

Table 17: Significant coefficients and zero coefficients before and after thresholding

	Total Coefficients	Significant Coefficients	Zero Coefficients
Before Threshold	10069	10069	0
After Threshold	10069	2531	7538

From Table 17, it can be observed that after thresholding, the number of significant coefficients decreased about 74.87%, that means only 25.16% coefficients of the total number of coefficients are now representing the whole signal without any distortion. This compressed signal is then encode with RLE method. After RLE the total number of coefficients decreased from 10069 to only 2531 that means the number of coefficients reduced about 74.57%. The advantage of this encoding process is that the signal can be transferred securely by keeping all the information hidden. Only the receiver can decode the signal. So, secure signal transmission is also possible with our compression algorithm.

6.6.3 Signal reconstruction

After receiver received the encoded signal, it's information can be extracted by using RLD and IDWT in the receiver end.

6.6.4 Simulation result

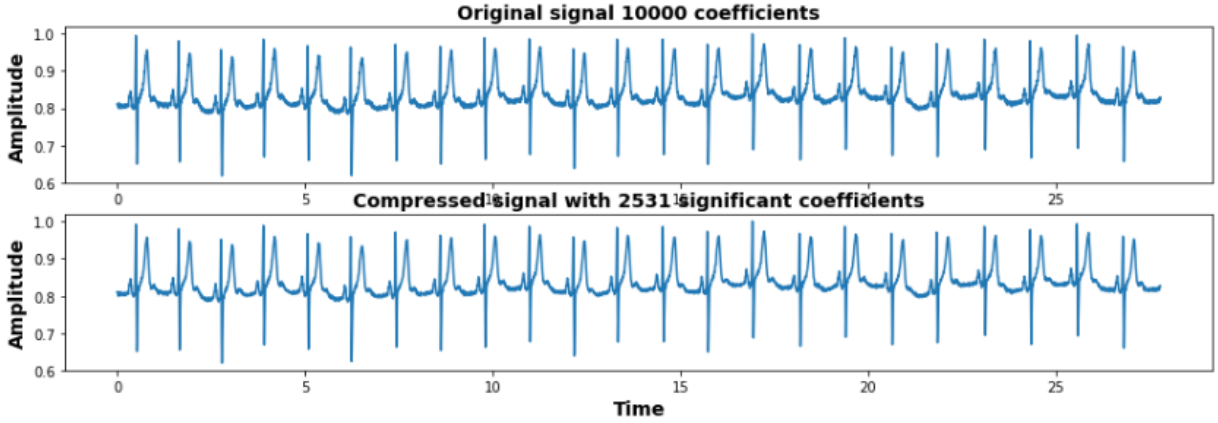

Figure 34: a) Original ECG signal b) Compressed ECG signal

The compression algorithm is applied to several ECG signals available in the database and a high compression is achieved by maintaining the signal fidelity. Fig. 34 shows the

original and compressed form of record 117. The values of CR, PC and PRD of several recordings of the database are presented in Table 17. From the Fig. 34 and Table 18, we can see that the signal is compressed about 74.80% with an average small PRD of 0.20%.

Table 18: The values of CR, PC, and PRD after compressing different ECG records

Records	CR	PC (%)	PRD(%)
105	3.97	74.78	0.35
109	3.96	74.76	0.25
111	3.93	74.76	0.23
112	3.97	74.54	0.16
117	3.97	74.80	0.16
121	3.92	74.50	0.12
124	3.97	74.78	0.18
201	3.95	74.70	0.18
202	3.93	74.52	0.22
205	3.97	74.78	0.21
207	3.96	74.76	0.22
210	3.96	74.61	0.22
223	3.97	74.80	0.24
230	3.97	74.80	0.20
232	3.97	74.78	0.21
233	3.97	74.67	0.17
234	3.97	74.70	0.15

This ECG compression method will preserve all the pathological information and other sub-bands with noise and unnecessary information will be discarded from the original signal [6]-[8]. This technique will not only compress and denoise the signal but also the signal will be encrypted while transferring from home to clinic. These compressed and denoised ECG signals are then used to classify heartbeats into five different arrhythmias by using AI.

6.6.5 Comparison with other ECG compression methods

The performance of this ECG compression technique is compared with 21 other compression techniques presented in the paper [9]. By comparing the QS of different algorithms we can evaluate their performance. Table 19 shows the QS of our proposed algorithm along with the QS of those algorithms reported in the paper [9]. From the Table 19 we can see that the average QS of the proposed method is around 19.40 while the QS of the other algorithms

range from 0.357 to 20.885. The best QS was 20.885 followed by 14.43, which proved that the proposed compression algorithm outperformed 20 other compression algorithms out of 21 algorithms. Reddy et al. [75] used Fourier descriptors to compress the signal and achieved a

Table 19: Performance comparison of different compression algorithms

Compression Technique	CR	PRD (%)	QS
Amplitude Zone Time Epoch Coding	10.00	28.00	0.36
Improved Amplitude Zone Time Epoch Coding	9.91	7.99	1.24
Coordinate Reduction Time Encoding System	4.80	7.00	0.69
Turning Point	2.00	5.10	0.40
Wavelet packet compression	8.00	2.60	3.08
Set Partitioning in Hierarchical Trees Algorithm	8.00	1.18	6.78
Linear prediction of the wavelet coefficients	11.60	5.30	2.19
Perceptual masks & Discrete Cosine Transform	3.50	1.24	2.82
Neural Network	12.74	0.61	20.89
Huffman coding	11.06	2.73	4.05
Wavelet Transform	12.00	0.98	12.25
Discrete Cosine Transform (Min CR)	6.20	1.50	4.13
Discrete Cosine Transform (Max CR)	10.90	3.00	3.63
ASCII character encoding	15.72	7.89	1.99
JPEG2000	20.00	3.26	6.134
DC equalization and complexity sorting	8.00	0.86	9.30
Mother wavelet modification	23.10	1.60	14.43
Fourier descriptors	7.40	7.00	1.05
Fast Fourier Transform	6.28	0.75	8.37
Fourier Transform (fixed strategy)	14.67	1.06	13.84
Fourier Transform (adaptive strategy)	16.58	1.07	15.49
Proposed model	3.96	0.20	19.40

QS of 1.05. They converted each signal as closed contour. Coordinates of each contour were represented as complex sequences. Then they did the FT of those complex sequences and calculated the significant coefficients which are also known as Fourier descriptors. Shinde et al. [76] have used FFT to compress signals and got a QS of 8.37. Sadhukhan et al. [77] have increased the QS by applying fixed strategy and adaptive strategy of FT. The fixed strategy was based on the selection of a fixed band-limiting frequency, and the adaptive strategy was dependent on the spectral energy distribution of signals. They encoded significant coefficients to optimize the usage of bit and compress signals. By applying the significant coefficients they achieved a QS of 13.84 which was further improved to 15.49 by applying the adaptive strategy method. The AZTEC algorithm converts raw signals into plateaus.

The amplitude and length of each plateau are stored for reconstruction. A poor QS of 0.36 was achieved through the AZTEC technique which was then improved to 1.24 through the improved AZTEC technique. Improved AZTEC algorithm optimizes the trade off between CR and PRD. The TP data reduction algorithm reduces the sampling frequency of signals to compress the signal which has very low QS of 0.40. CORTES algorithm is a hybrid approach of AZTEC and TP to achieve high CR of the AZTEC and the low reconstruction error of the TP technique. DCT compress signals by restoring the signal information in a fixed number of DCT coefficients. Other techniques such as Huffman coding, ASCII character encoding, JPEG2000, SPIHT, and DC equalization and complexity sorting are also applied to compress 1-D signals [9]. DWT outperformed these previous techniques to compress 1-D non-stationary signals by giving high CR with low PRD. The highest QS of 20.89 was achieved by using the NN method.

6.7 Results of the classification of ECG signals

In this section, we verified the performance of the proposed ECG classification models. We have used the Python programming language to implement these architectures. To evaluate the effectiveness of the proposed DNN and CNN models, we compared the performance of these DL models with other traditional ML algorithms in terms of sensitivity/recall, precision, and accuracy.

6.7.1 Proposed DNN model to classify ECG beats

A 5-layer sequential feed-forward DNN model trained by Keras was used in our research to classify ECG heartbeats into five categories. Keras is the high-level API of TensorFlow, which

we used to train our classifying model with great speed. In our proposed method we used each ECG beat as input to our deep learning model. A total of 168 features were achieved from each heartbeat, which were then fed into the DNN to train the model. Afterwards, 5 hidden layers with 32,64, 128, 256, 512 filters were implemented with the ReLU activation function for non-linearity. In the output layer, the softmax activation function was used to get the probability distribution, which we applied on the cross-entropy cost function. The cross-entropy cost function was used to measure how far apart the output of the model was from that of the desired or target output. The Adam optimizer was used to minimize the cost function. The training started with a learning rate of 0.0001 and continued until it reached the maximum number of epochs. The dropout technique was used in the model to reduce independent learning among the neurons and to handle overfitting. Hyper parameter optimization technique was used to get optimal hyper parameters for the DNN model. After training the model its prediction capability was tested on the testing set. Fig. 35 shows the proposed DNN model of five layers to classify heartbeats.

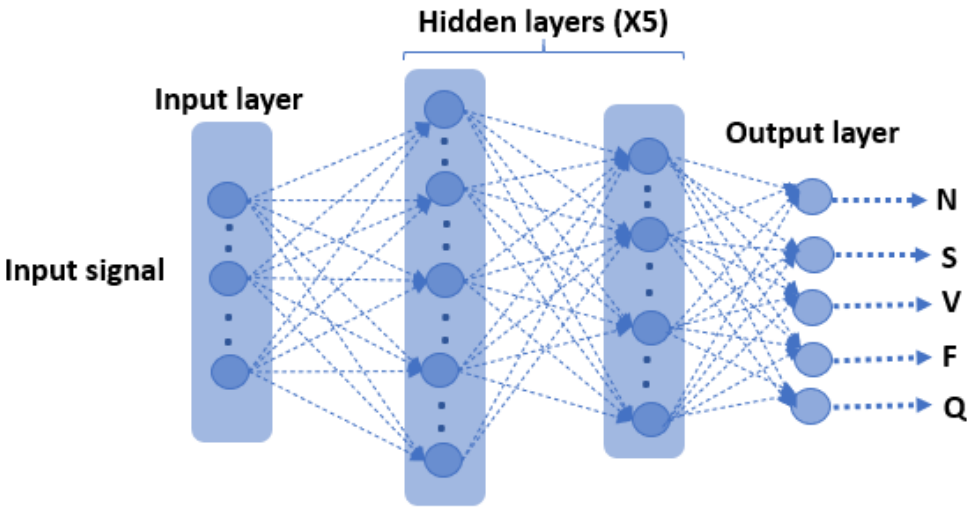

Figure 35: Proposed DNN model to classify heart beats

6.7.2 Proposed CNN model to classify ECG beats

We used a 1D CNN model of 4 hidden layers in our research to classify each heartbeat in the database. Four hidden layers (1st layer with 32 filters, 2nd layer with 64 filters, 3rd layer with 128 filters, 4th layer with 256 filters) were implemented with the ReLU activation function for non-linearity. Each layer had the same kernel size of 4 to combine the number of input features with the number of new output features. A max-pooling layer of kernel size 2 was also deployed in each layer to downsample the input and to reduce the number of dimensions. The last hidden layer of the CNN was connected with a final dense layer having 64 nodes. In the output layer, the SoftMax activation function was used to predict output class probabilities. The dropout technique was used in the model to reduce independent learning among the neurons and to handle overfitting. Adam optimization algorithm was used to update network weights iteratively based on training data. Hyper parameter optimization technique was used to get optimal hyper parameters for the CNN model. After training the model its prediction capability was tested on the testing set. Fig. 36 shows the proposed CNN model of four layers to classify heartbeats.

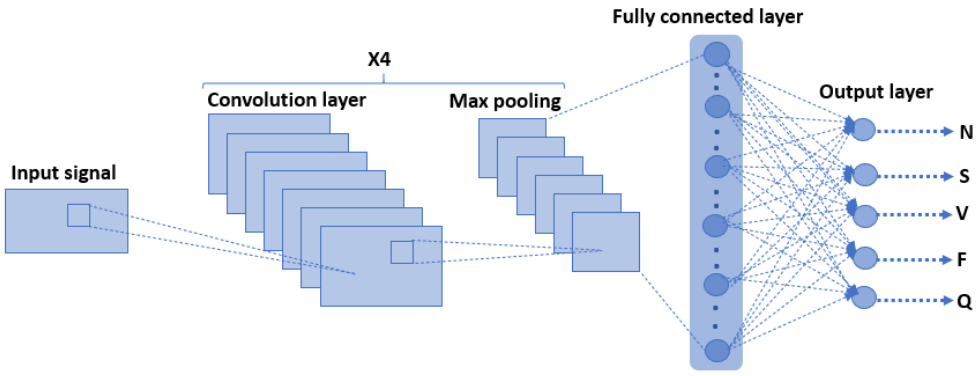

Figure 36: Proposed CNN model to classify heart beats

6.7.3 Proposed DHL models to classify ECG beats

Fig. 37 illustrates the design of the proposed deep hybrid architecture. It has four 1D convolution layers with 32, 64, 128, and 256 filters, respectively. All the layers have the same kernel size of 2. Each convolution layer uses a Rectified Linear Unit (ReLU) activation function, and each max pooling layer is of size 2. A dropout layer with 10% dropping rate is connected with each convolution layer to handle overfitting. After the convolution and max pooling, the learned features are flattened to one long vector and are passed to a fully connected layer with 512 filters. The fully connected layer also used the ReLU activation function and 20% dropout rate to reduce overfitting problems. The fully connected layer works as a buffer between the learned features and the output. The cost function is minimized by using the Adam optimizer. The number of training epochs, batch size, and the learning rate are set to 20, 64, and 0.0001, respectively. After optimizing and training the CNN, the learned features from the fully connected layer are passed on to the ML classifiers for the final prediction task.

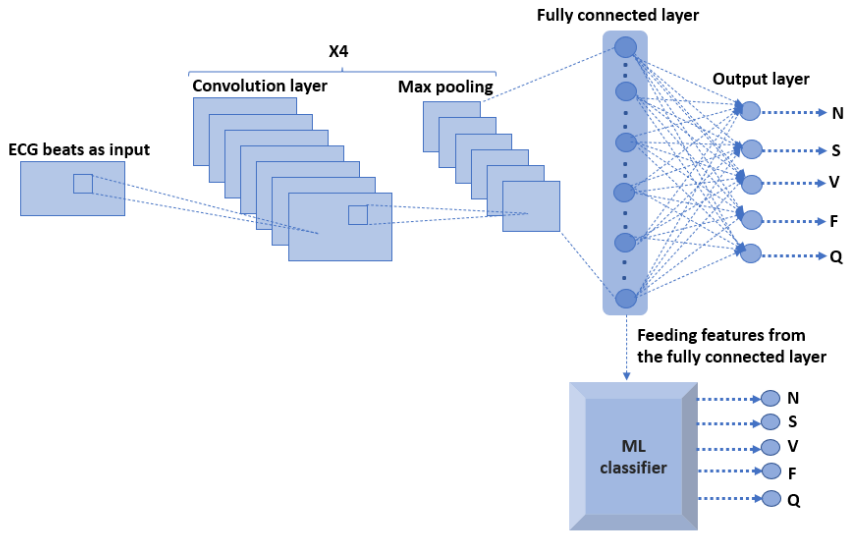

Figure 37: The proposed DHL models architecture to classify heart beats

6.7.4 Simulation results

Our proposed DL models can classify five different heartbeats with great accuracy compared to other traditional ML algorithms. Our proposed DL models outperformed ML algorithms in terms of sensitivity/recall, precision and accuracy. We used 80% of the data as the training set to develop the prediction ability of the model, and the remaining 20% of the data was used as the testing set to validate the model. Our proposed DNN model classified heart beats with a very good testing accuracy of 94.70%. The achieved precision and the recall of the proposed DNN model were 96.30% and 93.90%, respectively. We got better classification accuracy with our proposed CNN model compared to DNN model. The CNN model gave a testing accuracy of 97.90%. The achieved precision and the recall of the model were 98.10% and 97.80%, respectively.

Table 20: Comparison of the proposed DL models with ML models to classify ECG beats

Classification Model	Precision (%)	Recall (%)	Accuracy (%)
DT	94.50	93.70	93.70
RF	97.80	97.80	97.80
LR	88.10	66.00	66.00
SVM	95.70	91.00	91.00
KNN	95.80	94.30	94.30
AB	88.20	52.90	52.90
NB	79.90	19.00	19.00
DNN	96.30	93.90	94.70
CNN	**98.10**	**97.80**	**97.90**

Table 20 shows the performance of the proposed DNN, CNN, and the other traditional ML models. The performance of the proposed DHL models and other traditional ML and DL models with separate implementations are shown in Table 21. As shown in Table 21, among all the ML classifiers, NB classifier performed worst with 19.00% accuracy, and best result is achieved by the RF classifier with 97.80% accuracy. Our proposed deep hybrid CNN-ML methods were able to improve the accuracy obtained from a single CNN model and 7 ML

Table 21: Comparison of the proposed DHL models with ML and DL classification models implemented separately

Classification Model	Precision (%)	Recall (%)	Accuracy (%)
CNN	98.10	97.80	97.90
LR	88.10	66.00	66.00
CNN-LR	98.30	98.30	98.30
RF	97.80	97.80	97.80
CNN-RF	**98.50**	**98.60**	**98.60**
KNN	95.80	94.30	94.30
CNN-KNN	98.20	98.10	98.10
DT	94.50	93.70	93.70
CNN-DT	97.60	97.50	97.50
NB	79.90	19.00	19.00
CNN-NB	98.20	98.30	98.20
SVM	95.70	91.00	91.00
CNN-SVM	98.30	98.30	98.30
AB	88.20	52.90	52.90
CNN-AB	97.40	97.50	97.50

models which were implemented separately. The best classification accuracy of 98.60% is achieved by CNN-RF models followed by CNN-LR (98.30%), CNN-SVM (98.30%), CNN-NB (98.20%), and CNN-KNN (98.10%) models. CNN-DT (97.50%) and CNN-AB (97.50%) model didn't show any improvement over single ML and DL models. Fig. 38 shows the accuracy and Fig. 39 shows the reduction of the cost of the proposed DHL model.

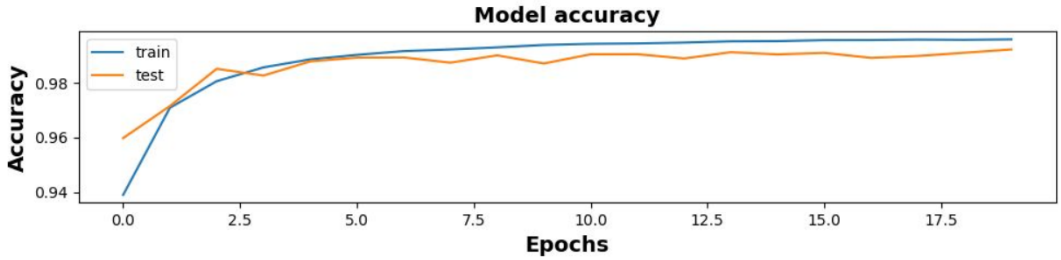

Figure 38: Training and testing accuracy of the proposed DHL model with respect to epochs

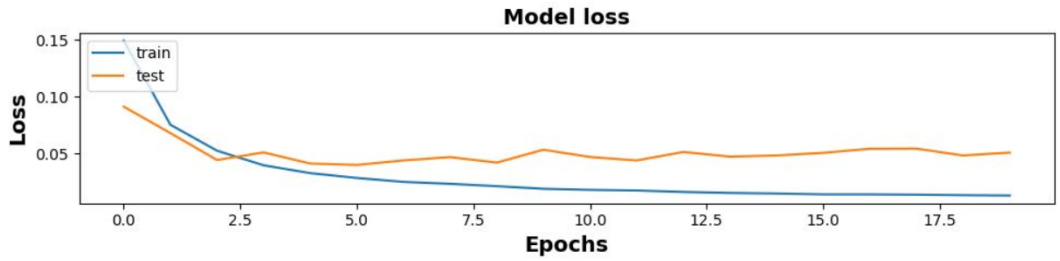

Figure 39: Training and testing loss of the proposed DHL model with respect to epochs

6.7.5 Comparison with other ECG heartbeat classification models

Table 22 shows the comparison of our proposed ECG heart beat classification models with 11 other state-of-the-art heart beat classification models. All these models used the same MIT-BIH Arrhythmia dataset. As shown in Table 22, the classification accuracy achieved from the previous models varied between 92.70% to 99.28%. Among all the models, DL models performed better to classify ECG heartbeats due to their potential to extract important and meaningful features automatically from raw data. Martis et al. [80, 87–89] used different approaches to classify ECG beats and achieved the accuracy of 94.52%, 98.11%, 93.48%, and 99.28% respectively. They used principal component analysis (PCA) and Independent Component Analysis (ICA) to reduce the dimensions of DWT coefficients, segmented ECG beats, and error signals of linear prediction model. These features were used to classify ECG beats using DNN and Least Square-Support Vector Machine (LS-SVM). They got 98.11% accuracy by applying PCA on segmented ECG beats and then sending those features to a DNN model [87]. Later, they applied PCA on bispectrum features of ECG beats. They sent those features to SVM and DNN models and got 93.48% and 94.52% accuracy, respectively [80, 88]. They achieved state of the art accuracy of 99.28% by applying ICA on the DWT coefficients of ECG beats. These reduced features were sent to a DNN model for the final classification [89]. Li et al. [81] achieved 94.61% accuracy by using DWT and RF classifier. Elhaj et al. [82] used ICA to reduce the dimension of non-linear features such as high order statistics and cumulants. Then PCA was used to reduce the dimension of DWT. They got 98.91% accuracy by sending these features to a DNN model. Zubair et al. [84] has achieved 92.70% accuracy by directly sending raw signals to a CNN model having 7 hidden

layers. Acharya et al. [79] developed a 9-layer deep CNN to classify heartbeats and got 93.47% testing accuracy. Kachuee et al. [78] used deep residual CNN consisting of 13 weight layers and achieved an accuracy of 93.40%. Yang et al. [85, 86] used ICA and PCA networks (ICA-PCANets) to extract features from raw ECG signals. After extracting features, they used different ML classifiers such as SVM, KNN, and RF to classify ECG beats. Among them, the combination of ICA-PCANet and linear SVM achieved the highest accuracy of 98.63% followed by the combination of PCA and SVM which is 97.80%.

Table 22: Comparison of the proposed DHL model with other ECG classification models

Author	Methods	Precision (%)	Recall (%)	Accuracy (%)
Martis, (2012)	PCA & DNN	–	99.90	98.11
Martis, (2013)	Bispectrum-PCA & SVM	–	99.27	93.48
Martis, (2013)	PCA & DNN	–	98.61	94.52
Martis, (2013)	DWT-ICA & DNN	–	99.97	99.28
Li, (2016)	DWT & RFC	–	–	94.61
Elhaj, (2016)	ICA-DWT-PCA & SVM	–	98.91	98.91
Zubair, (2016)	CNN	–	–	92.70
Acharya, (2017)	Augmentation & CNN	–	96.01	93.47
Kachuee, (2018)	Deep residual CNN	–	–	93.40
Yang, (2018)	PCA & SVM	–	–	97.80
Yang, (2020)	ICA-PCA & SVM	–	–	98.63
Proposed Model	**CNN-RF**	**98.50**	**98.60**	**98.60**

CHAPTER 7

CONCLUSION AND FUTURE WORKS

7.1 Conclusion

PCG and ECG signals have been used for decades to detect cardiac abnormalities. The continuous extraction of important cardiac information from these two major signals and the detection of the abnormality in the primary stage can play a vital role to decrease the death rate caused by cardiovascular diseases. The main goal of this research is to develop algorithms to accurately detect cardiac abnormalities in the primary stage while continuous monitoring of heart. We have combined the application of signal processing, ML, and DL approaches in PCG and ECG signals to detect abnormalities in the heart during the very primary stage without the need of any doctor or cardiologist. The proposed heart monitoring technique has the following three steps:

- **Compression, Denoising, and Encryption**: We have used DWT to compress and denoise both PCG and ECG signals without loosing any pathological information. Thus, it will not only save the storage but will also reduce the noise and unnecessary information that make the data collection process very difficult while continuous monitoring. We have compressed PCG and ECG signals to 93.67% and 74.57%, respectively. An encoding process named RLE process is also applied with DWT to encrypt the data and to ensure patient confidentiality while transferring data from home to clinic.

- **Segmentation:** A segmentation algorithm based on Shannon energy envelope is used to obtain important cardiac parameters of the heart from PCG signals. These cardiac

parameters are crucial to get the overall picture of the heart. With our proposed PCG segmentation algorithm we have successfully determined the durations and amplitudes of basic heart sounds as well as the duration of systole interval, diastole interval, and cardiac cycles properly. This method will also continuously calculate the heart beats of a patient and will notify the patient immediately if it finds any inconsistent cardiovascular parameters. One of the main advantage of this segmentation algorithm is that it doesn't require any ECG signal as reference to segment heart sounds.

- **Classification:** We have combined DL and ML models to build DHL models which can classify PCG and ECG signals with great accuracy. Our proposed DHL algorithm showed better testing accuracy compared to ML and DL models implemented separately. As PCG is an audio signal, Mel-scaled power spectrogram and MFCC are employed to extract informative features from the PCG signal, which are then fed into a classifier to classify each PCG signal into a normal or an abnormal signal. For the classification of heart beats we have directly used the raw ECG beats to classify heartbeats into five different beat categories. For the classification of PCG and ECG signals, we got around 94.30% and 98.60% accuracy, respectively, which is better than many other state-of-the-art PCG and ECG classification methods. This automatic classification process can certainly help doctors and cardiologists to detect different cardiac abnormalities and irregular heartbeats in the initial stage.

7.2 Summary of the research work

Summary of the research work is given below:

- Our proposed method is robust enough to compress PCG and ECG signals without loosing any pathological information. Denoising of PCG signals with high SNR is also achieved using our PCG compression algorithm.

- Using our proposed encryption technique it's possible to encrypt signals while transferring from home to clinic. Thus it can maintain patient confidentiality.

- Our proposed segmentation algorithm can automatically segment PCG signals and provide the values of different important cardiac parameters to detect early stage heart diseases.

- Our proposed PCG classification algorithm can detect the structural defects of the heart valves by classifying PCG signals into two categories: as normal and abnormal.

- Our proposed ECG classifier can accurately detect the irregular heartbeats in the primary stage by classifying ECG signals into five different arrhythmia.

- Our proposed DHL models are less computationally expensive and have less time complexity compared to traditional DL algorithms. The final classification layer of a DL model usually results in overfitting when the model is fed with unstructured or less data. This overfitting problem increases the time and computational complexity of traditional DL models, which is not present in traditional ML algorithms. In our proposed DHL models, fully connected neural networks in the DL model are followed by the ML models. Thus, our proposed DHL models are faster and do not require additional time for processing compared to traditional standalone DL models. In addition, our proposed DHL models remove the need for feature engineering techniques

on which all traditional ML algorithms are dependent. This automatic classification process can better help doctors and cardiologists to detect cardiac abnormalities and irregular heartbeats in the initial stage.

- This technique can also be applied to analyze brain signal and obstructive sleep apena.

7.3 My contributions

My contributions to this research are given below:

- To develop this unique technique which is a complete package of compression, denoising, encryption, and classification of PCG and ECG signals. Until now, very little research has been done on analyzing PCG and ECG signals which covers all four major techniques.

- To develop the proposed PCG segmentation algorithm that doesn't need any ECG signal as reference. This segmentation technique can not only segment PCG signals but also extracts cardiac information from raw PCG signals. These cardiac information can be analyzed to detect early stage cardiac diseases.

- To develop DHL algorithms that combine the advantages of ML and DL algorithms and overcome their limitations to classify PCG signals and ECG heartbeats with high accuracy.

7.4 Future works

Although, we have achieved a certain level of result, additional research is needed to further improve the overall result and accuracy. Our future works are given below:

- **Noise reduction:** The main limitation of our proposed algorithm is the existence of noise within the frequency range of the basic heart sounds. The efficiency of the PCG segmentation algorithm becomes limited in the presence of a large number of murmurs and noise overlapping with heart sounds. In this case, the Shannon energy is affected by noise and it is very difficult to accurately identify the boundaries of each heart sound. In the presence of high intensity noise, the Shannon energy envelope is too noisy to read and will provide incorrect output. The DWT reconstructs the PCG signal by separating high-frequency murmurs and noise from the low-frequency heart sounds. However, when the heart sounds, murmurs. and noise share the same frequency band, separating murmurs and noise from the PCG signal will eliminate some of the major details of heart sounds. This will cause potential loss of the cardiac information. Additional research is needed to solve this problem.

- **Require more PCG signals for the classification:** It should be noted that the proposed method requires a large amount of data to train the model. We got better accuracy in the ECG heartbeat classification compared to the PCG classification. The main reason behind this is that we didn't have enough PCG data to train the DL models. We had 1,09,449 heartbeats to classify ECG heartbeats whereas the number of PCG signals was only 3240 which was inadequate to train the model. Therefore, in future work, it is necessary to evaluate the performance of our proposed model by

using PCG signals from more subjects.

- **Feature extraction:** Clinically, it is essential to extract as many possible features from the PCG signal for the correct classification. However, not all of the features carry important information, and there can be some redundancy. We will focus on exploring other important features to improve the classification performance.

- **Find the best deep learning model for the classification:** In the future, other neural network models such as RNN, LSTM, and GRU will be combined with ML models to increase the sensitivity, specificity, and accuracy of the DHL models if possible, near 100%.

A PUBLICATION

In this appendix section, we have provided the list of the publications related to our research. We have published 6 papers listed as,

- Md. Chowdhury, K. Poudel and Y. Hu, "Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals," IEEE Access 2020.

- Md. Chowdhury, K. Poudel and Y. Hu, "Phonocardiography Data Compression using Discrete Wavelet Transform," IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, 2018.

- Md. Chowdhury, K. Poudel and Y. Hu, "Automatic Phonocardiography Analysis using Discrete Wavelet Transform," International Conference on Vision, Image, and Signal Processing (ICVISP), Vancouver, Canada, 2019.

- Md. Chowdhury, K. Poudel and Y. Hu, "Automatic ECG Signal Analysis Based on Discrete Wavelet Transform and Deep Convolutional Neural Network," IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA,2020.

- Md. Chowdhury, K. Poudel and Y. Hu, "Detecting Abnormal PCG Signals and Extracting Cardiac Information Employing Deep Learning and Shanon Energy Envelope," IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, 2020.

- Md. Chowdhury, C. Li, and K. Poudel, "Combining Deep Learning with Traditional Machine Learning to Improve Phonocardiography Classification Accuracy," IEEE SPMB Symposium., PA, 2021.

BIBLIOGRAPHY

[1] S. S. Virani, A. Alonso, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, et al, " Heart disease and stroke statistics—2020 update: a report from the American Heart Association," *Circulation*, pp. 139–596, 2020.

[2] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram signal analysis: a review," *Critical Reviews in Biomedical Engineering*, vol. 15, pp. 211–236, 1987.

[3] Steven McGee, "Auscultation of the Heart: General Principles," *Evidence-Based Physical Diagnosis (Fourth Edition), Elsevier,*, Ch. 39, pp. 327-332, 2018.

[4] L.G. Durand, and P. Pibarot, "Digital Signal Processing of the Phonocardiogram: Review of the Most Recent Advancements," *Critical Reviews in Biomedical Engineering.*, vol. 23, pp. 163–219, 1995.

[5] Md. Chowdhury, K. Poudel and Y. Hu, "Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals," *IEEE Access*, 2020.

[6] Md. Chowdhury, K. Poudel and Y. Hu, "Detecting Abnormal PCG Signals and Extracting Cardiac Information Employing Deep Learning and Shannon Energy Envelope," *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, 2020.

[7] Md. Chowdhury, K. Poudel and Y. Hu, "Automatic Phonocardiography Analysis using Discrete Wavelet Transform," *International Conference on Vision, Image, and Signal Processing (ICVISP)*, Vancouver, Canada, 2019.

[8] J. R. Hampton, "THE ECG MADE EASY,"Edinburgh : Churchill Livingstone, 2003.

[9] B. Singh, A. Kaur, and J. Singh, "A review of ECG data compression techniques," *International Journal of Computer Application*, vol. 116, no. 11, pp. 1-6, 2015.

[10] C. J. Wiggers, Wiggers diagram. [Online]. Available: https://en.wikipedia.org/wiki /Wiggers diagram

[11] G. Bebis, Short Time Fourier Transform. [Online]. Available: https://www.cse.unr.edu/ bebis/CS474/Lectures/ShortTimeFourierTrans

[12] Md. Chowdhury, K. Poudel and Y. Hu, "Automatic ECG Signal Analysis Based on Discrete Wavelet Transform and Deep Convolutional Neural Network," *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, 2020.

[13] M. S. Hossain, "ECG Compression Using Rule Based Thresholding of Wavelet Coefficients," M.Sc. dissertation, Department of Electrical and Electronic Engineering, Bangladesh University of Science and Technology, Dhaka, 2008.

[14] M. N. A. M. Nawar, "Time-Frequency analysis of Different types of signals," M.Sc. dissertation, Department of Physics, Ain Shams University, Egypt, 2016.

[15] Z. Tong, "An Integrated Framework for Cardiac Sounds Diagnosis," M.Sc. dissertation, Department of Electrical and Computer Engineering, Western Michigan University, Michigan, 2015.

[16] P. K. Kumari, " A Study on Novel Techniques for Heart Sound and Murmur Classification, Search and Retreival," Ph.D. dissertation, DEPARTMENT OF COMPUTER

SCIENCE AND ENGINEERING, Dr. M.G.R. Educational and Research Institute University, Chennai, 2012.

[17] F. Meziani, S. M. Debbal and A. Atbi, "Analysis of Phonocardiogram Signals using Wavelet Transform." *Journal of Medical Engineering and Technology.*, vol. 36, pp. 283–302, 2012.

[18] S. M. Debbal, and F. B. Reguig, "Computerized Sound Analysis," *Computers in Biology and Medicine.*, vol. 38, pp. 63–90, Mar. 2008.

[19] B. S. Emmanuel, "Discrete Wavelet Mathematical Transformation Method for Non-stationary Heart Sounds Signal Analysis," *ARPN Journal of Engineering and Applied Sciences.*, vol. 07, pp. 1021–1028, Aug. 2012.

[20] R. Polikar, "The wavelet tutorial," 1996. [Online]. Available: http://engineering.rowan.edu/polikar/WAVELETS/WTtutorial.html.

[21] Y. Meyer, "Wavelets: Algorithms and applications," *Society of Industrial and Applied Mathematics*, vol. 04, ch. 3 and 7, Feb. 1993.

[22] M. Unser, A. Aldroubi, and H. J. Shaw, "A Review of Wavelets in Biomedical Applications," *Proceedings of the IEEE.*, vol. 84, pp. 626–638, May. 1996.

[23] A. Djohan, T. X. Nguyen and W. J. Tompkins, "ECG Compression using Discrete Symmetric Wavelet Transform," in *17th International Conference of the Engineering in Medicine and Biology Society,* Montreal, Quebec, Canada, vol. 17, no. 09, 1995, pp. 167–168.

[24] C. Sawant, and H. T. Patii, "Wavelet Based ECG Signal De-noising," in *2014 First International Conference on Networks Soft Computing (ICNSC 2014),* Guntur, AP, India , pp. 20–24, Aug. 2014.

[25] G. Mishra, K. K. Biswal, and A. K. Mishra, "Denoising of Heart Sound Signal Using Wavelet Transform," *International Journal of Research in Engineering and Technology.*, vol. 02, no. 04, pp. 719–723, Jul. 2013.

[26] J. Zhong and F. Scalzo, "Automatic Heart Sound Signal Analysis with Reused Multi-Scale Wavelet Transform," *International Journal of Engineering and Science*, vol. 02, pp. 50–57, Jan. 2013.

[27] G. Peyre, Signal Denoising with Wavelets, 2010. [Online]. Available: http://www.numerical-tours.com/matlab/denoisingwav.

[28] M. Nabih, E. A. El-Dahshan and A. S. Yahia, "Denoising of Heart Sound Signals Using Discrete Wavelet Transform," *Circuits, Systems and Signal Processing*, vol. 36, pp. 4482–4497, 2017.

[29] B. Ergen, "Comparison of Wavelet Types and Thresholding Methods on Wavelet Based Denoising of Heart Sounds," *Journal of Signal and Information Processing.*, vol. 04, pp. 164–167, 2013.

[30] M. D. Aloia, A. Longo, M. Rizzi "Noisy ECG Signal Analysis for Automatic Peak Detection," *Information*, vol. 10, Jan. 2019.

[31] H. Nazeran, "Wavelet-based Segmentation and Feature Extraction of Heart Sounds for

Intelligent PDA-based Phonocardiography," *Methods of Information in Medicine*, vol. 46, pp. 135–141, Feb. 2007.

[32] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram," *Proceedings of The IEEE Computers in Cardiology*, vol. 24, pp. 105–108, 1987.

[33] N. Shankar, and M. S. Sangeetha, "Analysis of Phonocardiogram for Detection of Cardiac Murmurs using Wavelet transform," *International Journal of Advanced Scientific and Technical Research*, vol. 01, no. 03, pp. 350–357, 2013.

[34] PeakUtils 1.1.1 documentation. [Online]. Available: https://peakutils.readthedocs.io/en/latest/generated/scipy.signal.lfilter.htmlscipy.signal.lfilter.

[35] R. vinaykumar, " Deep Learning Approaches to Detect Advanced Cyber Attacks," Ph.D. dissertation, Center for Computational Engineering and Networking, Amrita Vishwa VidyaPeetham University, Tamil Nadu, June. 2019.

[36] G. A. H. Alabandi, " Combining Deep Learning with Traditional Machine Learning to Improve Classification Accuracy," M.Sc. dissertation, Department of Computer Science, Texas State University, June. 2019.

[37] A. Bhattacharya, " Deep Hybrid Learning — a fusion of conventional ML with state of the art DL," GIDS.AI/ML and Data Live 2020, July. 2020.

[38] D. Sengupta, S. N.Ali, A. Bhattacharya, J. Mustafi, A. Mukhopadhyay, K. Sengupta ," A deep hybrid learning pipeline for accurate diagnosis of ovarian cancer based on nuclear morphology," PLOS ONE, January. 2022.

[39] B. Benuwa, Y. Z. Zhan, B. Ghansah, D. K. wornyo, and F. K. Banaseka, "A Review of Deep Machine Learning," *International Journal of Engineering Research in Africa*, vol. 24, pp. 124–136, 2016.

[40] K. Courtemanche, V. Millette, and N. Baddour, "Heart Sound Segmentation Based on Mel-Scaled Wavelet Transform," in *31st Conference of the Canadian Medical and Biological Engineering Society* ,Montreal, Quebec, Canada 2008.

[41] J. Chebil and J. A. Nabulsi, "Classification of Heart Sound Signals Using Discrete Wavelet analysis," *International Journal of Soft Computing*, vol. 02, pp. 37–41, 2007.

[42] L. H. Cherif, S. M. DEBBAL, and F. B. Reguig, "Segmentation of Heart Sounds and Heart Murmurs," *Journal of Mechanics in Medicine and Biology*, vol. 08, pp. 549–559, Dec. 2008.

[43] X. Zhang, L. G. Durand, L. Senhadji, H. C. Lee, and J. L. Coatrieux, "Time-Frequency Scaling Transformation of the Phonocardiogram Based of the Matching Pursuit Method," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 08, pp. 972–979, Sep. 1998.

[44] M. T. H. Chowdhury, K. N. Poudel, and Y. Hu, "Phonocardiography Data Compression Using Discrete Wavelet Transform," in *2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB),* Philadephia, PA, USA, pp. 01–03, Dec. 2018.

[45] M. A. Zahad and B. A. Rajoub, "An Effective Coding Technique for the Compression of One-Dimensional Signals using Wavelet Transforms.," *Medical Engineering Physics.*, vol. 24, pp. 185–199, 2002.

[46] B. A. Rajoub, "An Efficient Coding Algorithm for the Compression of ECG signals using the Wavelet Transform," *Medical Engineering Physics.*, vol. 49, no. 04, pp. 355–362, 2002.

[47] H. M. Tun and W. K. Moe and Z. M. Naing, "Analysis on ECG Data Compression Using Wavelet Transform Technique," *International Journal of Psychological and Brain Sciences.*, vol. 02, no. 06, pp. 127–140, 2017.

[48] C. S. Ahmed, M. M. Benaiad and A. T. Ahmed, "Run length encoding and wavelet transform based ECG compression algorithm for transmission via IEEE802.11b WLAN channel," in *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies,* Barcelona, Spain, Oct. 26 - 29, 2011.

[49] T. Omari and F. B. Reguig, "An Automatic Wavelet Selection Scheme for Heart Sounds Denoising," in *International Work-Conference on Bioinformatics and Biomedical Engineering,* Granada, Spain, Apr. 2014, pp. 1450–1462.

[50] R. Khanam and S. N. Ahmad, "Selection of wavelets for evaluating SNR, PRD and CR of ECG signal," *International Journal of Engineering Science and Innovative Technology (IJESIT).*, vol. 02, pp. 112–119, Jan. 2013.

[51] A. A. Shrouf, M. A. Zahhad and S. M. Ahmed, "A Novel Compression Algorithm for Electrocardiogram Signals based on the Linear Prediction of the Wavelet Coefficients," *Digital Signal Processing.*, vol. 26, pp. 604–622, 2003.

[52] P Yip and K. Rao, "Energy Packing Efficiency for the Generalized Discrete Transforms," *IEEE Transactions on Communications.*, vol. 26, no. 08, pp. 1257–1262, 1978.

[53] M. S. Hossain and T. Aziz and M. Haque, "ECG compression using multilevel thresholding of wavelet coefficients," in *Intelligent Sensors, Sensor Networks and Information Processing,* Australia, 2009, pp. 321–326.

[54] M. S. Hossain, "ECG Signal Compression using Energy Compaction Based Thresholding of the Wavelet Coefficients," *DUET Journal.*, vol. 01, 2011.

[55] M. A. Zahhad and B. A. Rajoub, "ECG Compression Algorithm Based on Coding and Energy Compaction of the Wavelet Coefficients," in *The 8th IEEE International Conference on Electronics Circuits and Systems,* Malta, vol. 01, 2001, pp. 441–444.

[56] L. B. Portols, "Lossles Compression of ECG signals - Performance Analysis in a Wireless Network," Master Thesis, Dept. Biom. Eng., Linkopings Univ., Cambridge, Sweden, 2009.

[57] Scipy.signal.lfilter. [Online]. Available: https://docs.scipy.org/doc/scipy/reference generated/scipy.signal.lfilter.htmlscipy.signal.lfilter.

[58] S. R. Messer, J. Agzarian, and D. Abbott, "Optimal wavelet Denoising for Phonocardiograms," *International Journal of Biomedical and Biological Engineering.*, vol. 32, pp. 931–941, Oct. 2001.

[59] C. Liu , D. Springer, Q. Li, B. Moody, R. Juan, F. Chorro, F. Castells, J. Roig, I. Silva, A. Johnson, Z. Syed, S. Schmidt, C. Papadaniil, L. Hadjileontiadis, H. Naseri, A. Moukadem, A. Dieterlen, C. Brandt, H. Tang, M. Samieinasab, M. R. Samieinasab, R. Sameni, R. G. Mark, and G. D. Clifford, "An open access database for the evaluation

of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, pp. 2181—2213, Nov. 2016.

[60] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," *2016 Computing in Cardiology Conference (CinC)*, pp. 621–624, 2016.

[61] M. Nassralla, Z. El Zein, and H. Hajj, "Classification of normal and abnormal heart sounds," *in Advances in Biomedical Engineering (ICABME), 2017 Fourth International Conference on Advances in Biomedical Engineering*, IEEE, pp.1–4, 2017.

[62] B. M. Whitaker, P. B. Suresha, C. Liu, G. D. Clifford, and D. V. Anderson, "Combining sparse coding and time-domain features for heart sound classification," *Physiological Measurement*, vol. 38, 2017.

[63] P. Langley and A. Murray, "Heart sound classification from unsegmented phonocardiograms," *Physiological Measurement*, vol. 38, no. 08, Jul. 2017.

[64] W. Han, Z. Yang, J. Lu, and S. Xie, " Supervised threshold-based heart sound classification algorithm," *Physiological Measurement*, vol. 39, no. 11, Nov. 2018.

[65] H. Tang, Z. Dai, Y. Jiang, T. Li, and C. Liu, "PCG Classification Using Multidomain Features and SVM Classifier," *BioMed Research International*, pp. 1–14, 2018.

[66] J. P. Dominguez-Morales, A. F. Jimenez-Fernandez, J. M. Dominguez-Morales, and G. Jimenez-Moreno, " Deep Neural Networks for the Recognition and Classification of Heart Murmurs Using Neuromorphic Auditory Sensors," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, pp. 24–34, 2018.

[67] M. Sotaquirá, D. Alvear, and M. Mondragón, "Phonocardiogram classification using deep neural networks and weighted probability comparisons, " *Journal of Medical Engineering Technology*, vol. 42, no. 7, pp. 510–517, 2018.

[68] S. A. Singh and S. Majumder, "Classification of unsegmented heart sound recording using KNN classifier, "*Journal of Mechanics in Medicine and Biology*, vol. 19, no. 4, pp. 24–34, 2019.

[69] D. M. Nogueira, M. N. Zarmehri, C. A. Ferreira, A. M. Jorge, and L. Antunes, "Heart Sounds Classification Using Images from Wavelet Transformation," *Progress in Artificial Intelligence*, Aug. 2019.

[70] S. A. Sing and S. Majumder,"Short unsegmented PCG classification based on ensemble classifier," *Turkish Journal of Electrical Engineering and Computer Sciences*, pp. 875–889, 2020.

[71] P. T. Krishnan, P. Balasubramanian, and S. umapathy, "Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, Feb. 2020.

[72] J. L. C. Loong, K. S. Subari, M. K. Abdullah, N. N. Ahmad, and R. Besar, "Comparison of MFCC and Cepstral Coefficients as a Feature Set for PCG Biometric Systems," *International Journal of Biomedical and Biological Engineering*, vol. 04, no. 08, pp. 335–339, Jan. 2010.

[73] Heart Sound Murmur Library, Apr. 2014. [Online]. Available:

https://open.umich.edu/find/open-educational-resources/medical/heart-sound-murmur-library.

[74] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[75] B.R.S. Reddy and I.S.N. Murthy, "ECG data compression using Fourier descriptors," *IEEE Transactions on Biomedical Engineering*, vol. 33, no. 04, pp. 428-434, 1986.

[76] A. A. Shinde and P. M. Kanjalkar, "The comparison of different transform based methods for ECG data compression," *International Conference on Signal Processing, Communication, Computing and Networking Technologies*, pp. 332-335, 2011.

[77] D. Sadhukhan, M. Mitra, and S. Pal , "Electrocardiogram data compression using adaptive bit encoding of the discrete Fourier transforms coefficients," *IET Science, Measurement Technology*, vol. 09, 2015.

[78] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, New York, NY, pp. 443-444, 2018.

[79] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in biology and medicine*, vol. 89, pp. 389–396, 2017.

[80] R. J. Martis, U. R. Acharya, C. M. Lim, K. Mandana, A. K. Ray, and C. Chakraborty, "Application of higher order cumulant features for cardiac health diagnosis using ECG signals," *International journal of neural systems*, vol. 23, no. 04, 2013.

[81] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, 2016.

[82] F. A. Elhaj, N. Salim, A. R. Harris, T. T. Swee, and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 52–63, Apr. 2016.

[83] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016.

[84] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," *6th Int. Conf. IT Converg. Secur. (ICITCS)*, Prague, Czech Republic, pp. 1–6, Sep. 2016.

[85] W. Yang, Y. Si, D. Wang, and G. Zhang, "A novel method for identifying electrocardiograms using an independent component analysis and principal component analysis network," *Measurement*, vol. 152, Feb. 2020.

[86] W. Yang, Y. Si, D. Wang, and B. Guo, "Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine," *Computers in Biology and Medicine*, vol. 101, no. 1, pp. 22–32, Oct. 2018.

[87] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray and C. Chakraborty, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health," *Expert Systems with Applications*, vol. 39, no. 14, pp. 11792–11800, 2012.

[88] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, C. Chakraborty, "Cardiac decision making using higher order spectra," *Biomedical Signal Processing and Control*, vol. 08, pp. 193-203, Mar. 2013.

[89] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform," *Biomedical Signal Processing and Control*, vol. 08, no. 05, pp. 437-448, 2013.