Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

(Bioinstrumentation Engineering)

A Master's Thesis Report

On

Analysis of Heart Sound for Biometric Identification system using Bayesian Optimized Support vector machine (SVM)

A Thesis Report Submitted to School of Graduate Studies of Jimma University in Partial Fulfilment for the Requirements of the Degree of Master of Science in Biomedical Engineering

(Bioinstrumentation Engineering)

By: Kalkidan Gazahegn

May 31, 2024

Jimma, Ethiopia.

Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

(Bioinstrumentation Engineering)

Analysis of Heart Sound for Biometric Identification system using Bayesian Optimized Support vector machine (SVM)

A Final Thesis Report Submitted to School of Graduate Studies of Jimma University in Partial Fulfilment for the Requirements of the Degree of Master of Science in Biomedical Engineering

(Bioinstrumentation Engineering)

By: Kalkidan Gazahegn

Advisor: Dr. Janarthanan K .(Ph.D.)

Co-Advisor: Mrs. Werqnesh Hailu (M.Sc.)

May 31, 2024

Jimma, Ethiopia.

Declaration

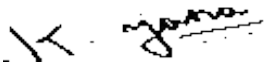
I declare this research with the title of "**Analysis of Heart Sound for Biometric Identification system using Bayesian Optimized Support vector machine (SVM)**" as my original work and I assure it with my signature

Candidate: Kalkidan Gazahegn

Signature  Date May 31,2024

On behalf of the School of Biomedical Engineering at Jimma Institute of Technology, we as university advisors confirmed below that this research is approved as M.Sc. Theis for the student.

Advisors

Name: Dr. Janarthanan K .(Ph.D.) Signature-  Date June 17, 2024

Name: Mrs. Werqnesh Hailu (M.Sc.) Signature  Date June 17, 2024

Examiners

External Examiner Name	Signature	Date
Melkamu Hunegnaw, PhD		June 17, 2024
Internal Examiner Name	Signature	Date
Ahmed Mohammed (M.sc)		June 18, 2024

Chair Person

Name	Signature	Date
Abel Worku (M.sc)		

Abstract

Researchers have found that the distinctiveness of the phonocardiogram (PCG) signal can be used for biometric identification, unlike advances in heart sound analysis for screening and diagnosing different heart problems. Biometrics, functioning as a measure and statistical test, proves invaluable for representing the distinctive behavioral and intellectual characteristics of individuals. Auscultation through a stethoscope is a non-invasive, and cost-effective technique for heart sound analysis in both diagnostic and identification applications. However, the non-stationary nature of heart sounds renders exclusive spectral or time domain analysis ineffective. That's why it's important to make the transition to time-frequency analysis.

This study introduces a method for heart sound analysis aimed at identifying subjects based on the uniqueness of their heart sounds. An electronic stethoscope was developed for signal acquisition, 560 heart sound recordings were obtained from 20 participants, At a uniform sampling frequency of 44.1 KHz and 16-bit depth. For machine learning based classification, data augmentation was performed using amplitude scaling techniques within a scaling range of [0.8, 1.2]. Discrete Wavelet Transform (DWT) based denoising of the raw PCG signal employed a Db10 wavelet family at 5^{th} level of decomposition and soft thresholding method. Feature extraction utilized DWT-based and Mel-Frequency Cepstral Coefficients (MFCC) techniques, followed by the application of a feature reduction algorithm to eliminate irrelevant features. The selected features underwent training using a cubic Support Vector Machine (SVM), chosen for its superior classification accuracy among various machine learning classifiers. Bayesian optimization was applied during the training phase.

System performance was evaluated on an independent dataset during the testing phase, resulting in a notable accuracy, precision, recall, and F1 score of 96.6%, 96.7%, 96.7%, and 96.6%, respectively. The proposed method demonstrates significant enhancement in representing distinct features crucial for identity recognition. As a future extension, the method will be refined using a larger dataset on heart sound to enhance its further capabilities on classification.

Key words : Identification, DWT, MFCC, Optimization, SVM, PCG

Acknowledgment

I express my deepest gratitude to the Almighty God, for providing me with the strength, wisdom, and perseverance throughout the journey of completing my MSc final thesis. My heartfelt appreciation goes to Ambo University College of Health science, and Medicine for the sponsorship that made this academic pursuit possible.

I am sincerely thankful to my main advisor Dr. Janarthanan K .(Ph.D.) for his invaluable guidance, mentorship, and unwavering support. His expertise and encouragement have been instrumental in shaping the trajectory of my research. I also extend my gratitude to my co-advisor Mrs. Werqnesh Hailu (M.Sc.) for her valuable insights and contributions, which significantly enriched the quality of my thesis. A special acknowledgment is extended to Hika Barki, a Ph.D. scholar, for his valuable contributions and support throughout this academic endeavor. I would also like to extend my sincere thanks to my Husband, and my family.

I am indebted to all those who played a role, whether big or small, in this academic endeavor. Your support has been a source of inspiration, and I am truly grateful for the collaborative efforts that have contributed to the successful completion of my MSc thesis.

Table of Contents

Declaration ..... **i**

Abstract ..... **ii**

Acknowledgment ..... **iii**

List of Figures ..... **vii**

List of Tables ..... **ix**

Acronyms and Abbreviations ..... **x**

CHAPTER ONE **1**

INTRODUCTON **1**

1.1. Background of the Study 1

1.2. Motivation of the Study 4

1.3. Statement of the problem 4

1.4. Objectives of the study 5

1.4.1. General objective 5

1.4.2.Specific Objective 5

1.5.Significance of the study 6

1.6.Scope of the Thesis 6

1.7.Document Organization of the study 6

CHAPTER TWO **7**

THE HEART 7

2.1.Hearts 7

2.1.1.Anatomy and physiology of Heart 8

2.2.Heart sound 9

2.2.1.Mechanism of heart sound production 10

2.2.2.AUSCULTATION PROCESS OF HEART SOUND 11

2.3.Signal Processing Techniques 12

2.3.1. Time – domain Analysis 12

2.3.2. Frequency - domain analysis 13

2.3.3. Time Frequency domain Analysis 14

CHAPTER THREE **19**

LITERATURE REVIEW 19

3.1. Related Works 19
3.2. Gap Analysis 24
CHAPTER FOUR **25**
METHODS AND MATERIALS **25**
4.1 Overview 25
4.2.Development of Electronic Stethoscope 26
4.3. Signal Acquisition 27
4.4. Data augmentation 28
4.5. Preprocessing of Heart Sound Signals using Discrete wavelet transform 29
4.6.Feature extraction 32
4.6.1 DWT based feature extraction 32
4.6.2 MFCC feature extraction 36
4.7. Feature Selection 39
4.8.Feature Normalization 40
4.9. Model Training 40
4.9.1. Support vector machine 41
4.10. Model optimization 42
4.11. Performance Evaluation 43
4.12. Materials 43
CHAPTER FIVE **44**
RESULTS and DISCUSSION **44**
5.1. Results 44
5.1.1. Building an Electronic Stethoscope 44
5.1.2. Acquisition of heart sound 48
5.1.3. Data Augmentation 49
5.1.4 Preprocessing 50
5.1.5. Feature Extraction 53
5.1.6 Feature reduction and normalization 56
5.1.7. Model training and Evaluation 56
5.1.8. Model Optimization 64
5.2. Discussion 67

CHAPTER SIX 68
CONCIUSION AND RECOMMENDATION 68
6.1. Conclusion 68
6.2. Recommendation 69
REFERENCES 70
Annex 76
Appendix A: Letter for Ethical Approval of Research Protocol 76

List of Figures

Figure 2.1 : The circulation of blood in the body [22]..... 7

Figure 2.2 : Chambers of the heart and valves [26]..... 8

Figure 2.3 : PCG signal illustration with heart sounds (s1, s2, s3, s4) [29] 9

Figure 2.4 Systole and diastole of heart..... 11

Figure 2.5 : Time- domain approach [32]..... 12

Figure 2.6 : Frequency – domain approach [33]..... 13

Figure 2.7 : A wave and wavelet [36]..... 16

Figure 4.1: A general procedure for subjects' identification 25

Figure 4.2: Image of stethoscope[55]..... 26

Figure 4.3 : Typical auscultation sites to place microphones [56]..... 28

Figure 4.4 : A method to denoise heart sound signals using wavelet-based techniques. 30

Figure 4.5: An overview sketch of svm algorithm [76]..... 41

Figure 5.1 : Materials utilized in crafting the electronic stethoscope..... 44

Figure 5.2 : An image showing head of an acoustic stethoscope with diaphragm and bell..... 45

Figure 5.3 : Stethoscope tubing connected to the head of stethoscope..... 45

Figure 5.4 : Image of a condenser microphone, and audio cable..... 46

Figure 5.5: Illustration of the condenser microphone connected to an audio cable. 46

Figure 5.6: Illustration of the completed construction of the electronic stethoscope. 47

Figure 5.7: Recording heart sound signal from subjects 48

Figure 5.8: Comparison between the original signal and its augmented counterpart 49

Figure 5.9 Wavelet decomposition of heart sound signals using db10 at 5th level 50

Figure 5.10 : Spectrograms of heart sound signals 52

Figure 5.11: Accuracy attained by various machine learning classifiers 57

Figure 5.12: Confusion matrix for classifiers (a) medium gaussian svm, (b) medium knn, and (c) subspace discriminant 63

List of Tables

Table 3.1 : A summary of recent related researches. 22

Table 5.1: SNR results for denoising heart sound signals using various wavelet functions at different decomposition levels with soft and hard thresholding methods (applying rigrsure threshold selection rule) 51

Table 5.2: A sample of extracted features by dwt based feature extraction 53

Table 5.3 : A sample of extracted features using mfcc feature extraction procedure 54

Table 5.4: A sample of features extracted by dwt and mfcc 55

Table 5.5: Classification report for machine learning models 56

Table 5.6: Number of correctly and misclassified observations for each subject 58

Table 5.7: Performance of cubic svm classifier per true class 59

Table 5.8: Performance of cubic svm classifier based on predicted classes 61

Table 5.9: Model summary based on precision, recall, accuracy, and f1 score 62

Table 5.10 : A number of correctly and misclassified observations for each class after optimization 64

Table 5.11: Performance of an optimized svm performed per true class 65

Table 5.12: Performance of an optimized svm per predicted class 66

Acronyms and Abbreviations

AR Autoregressive Modeling

AV Atrioventricular

BO Bayesian Optimization

CRO Cathode Ray Oscilloscope

CWT Continuous Wavelet Transform

DA Discriminant Analysis

DCT Discrete Cosine Transform

DFT Discrete Fourier Transform

DNA Deoxyribonucleic Acid

DNN Deep Neural Network

DR Dynamic Range

DWT Discrete Wavelet Transform

EB-Trees Ensemble of Bagged Decision Trees

ECG Electrocardiography

ED Euclidean Distance

EMD Empirical Mode Decomposition

EMG Electromyogram

FFT Fast Fourier Transform

FNR False Negative Rate

FR Fisher Ratio

FT Fourier Transform

GMM Gaussian Mixture Model

HHT Hilbert Huang Transform

HPF High Pass Filter

ICEEMDAN	Enhanced complete ensemble empirical mode decomposition
IDWT	Inverse Discrete Wavelet Transform
IMF	Intrinsic Mode Function
ISO	International Organization for Standardization
ISP	Independent Sub-band Function
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LPF	Low Pass Filter
LR	Logistic Regression
MFCC	Mel Frequency Cepstral Coefficient
MHS	Marginal Hilbert Spectrum
M-MFCC	Modified Mel Frequency Cepstral Coefficient
OVR	One versus rest
PCG	Phonocardiogram
PPG	Photoplethysmography
RCMDE	Refined composite multiscale dispersion entropy
RMS	Root Mean Square
SNR	Signal to noise ratio
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
THD	Total Harmonic Distortion
TNR	True Negative Rate
TPR	True Positive Rate
VQ	Vector Quantization
WPCC	Wavelet Packet cepstral coefficients
WT	Wavelet Transform

CHAPTER ONE

INTRODUCTON

1.1. Background of the Study

According to the International Organization for Standardization (ISO), a biometric system is "the automatic recognition of individuals based on their behavioral and biological traits."[1]. Various frequently utilized biometric characteristics include voice, gait, iris, palm print, DNA, and fingerprints. Biometric technologies have garnered substantial attention, particularly in fields where security and authentication have become increasingly crucial. This technology ensures the measurement of both physical and behavioral characteristics of a living being [2].

In our daily lives, human recognition is becoming increasingly vital, whether we must utilize our own equipment or provide proof of identity to third parties in order to access services or physical locations. Traditional authentication methods are divided into two categories: proving your knowledge (i.e., password-based authentication) and showing your ownership (i.e., biometric authentication) (i.e., token-based authentication) [3]. When a person is identified by traditional methods, their identity is linked to a different, less detailed representation, such as a password that can be misplaced, lost, or shared. These gaps were filled by physiological and behavioral biometrics, which use data gathered from precise measurements of human body parts.

Among these physiological biometrics are as follows; DNA is a molecule that contains the genetic code used by all known living things, including most viruses, to develop and function [4]. Ear identification is based on a person's unique ear shape and the structure of the mostly cartilaginous protruding region of the outer ear [5]. Infrared thermo grams of the face, hands, and hand veins are biometric measurements based on the heat radiation pattern of the human body that can be recorded by an infrared camera [6]. As a form of identification, a fingerprint is an imprint made on a surface by the traces or curves on a fingertip [7]. A palm print is an image of the hand's palm that has been captured. Similar to a fingerprint, it also has traces, textures, and marks that can be used to compare one palm to another [8].

The patterns of blood vessels on the retina are seen and analyzed by retina recognition technology [9]. The iris is the colorful area of the eye that lies between the pupil and the white part of the eye, the process of identifying a person based on the apparent pattern of his or her iris is known as iris recognition [9]. Face recognition, on the other hand, is the process of identifying a person based on some facial traits including pupils, the tip of the nose, the corners of the mouth, and the ends of the eyebrows [10]. On the other hand, the behavioral characteristics of a person are used as a biometric identification tool like Signature recognition: The manner a person signs his or her name is recognized as a trait of that person [11].

In order to identify a person, speech recognition systems employ the voice's features. A person's behavioral speech characteristics, such as voice pitch and speaking style, change with time as a result of factors including age, certain medical problems (such the common cold), emotional state, because the acoustic characteristics of speech that are used to distinguish between persons depend on the size and shape of the throat and mouth, voice recognition also falls under the category of (physiological biometrics) [12]. Convectional biometric systems, which use a mix of behavioral and physiological aspects, are vulnerable to fraud and can occasionally mistakenly detect copycats who are using human body parts as identification. Artificial masks, for instance, are extremely sensitive to facial characteristics, On the other hand, materials like silicone, nylon, polyurethane, and adhesives can be used to replicate fingerprints. To mimic the geometry and pattern of an iris, contact lenses with printed copies of its features can be utilized [13].

Biometrics based on Bio signals have piqued the interest of academics all over the world in recent years due to their inherent condition of aliveness detection, robustness against spoofing, and universality [14]. Researchers discovered a possible approach is to authenticate and identify users using biological features. PCG (phonocardiograph), ECG (electrocardiography), PPG (photo plethysmography), and EMG (electromyogram) security systems have the advantage of ensuring user movement with a higher assurance of the individual's aliveness [15].

Heart sounds are a reliable biometric because of their universality, distinctiveness, and fragility [16]. First and foremost, a defining characteristic of humans is that they all have beating hearts. Due to behavioral characteristics and hereditary reasons, each person has an own heart sound. Because heart sounds are vulnerable, it is more difficult to mimic them, which reduces fabrication. Comparing the heart sound signal to convectional biometrics, there are several benefits. It cannot be simply copied or changed because it is inherent in nature. To accurately reproduce someone else's heart sound, one must also take into account the structure of the heart and its surroundings. Health, age, and the physical makeup of the heart all have an impact on how loud the heart beats. Even people with the same heart disease condition have different heart sounds from one another[17], [18]. The greatest advantage of heart sound as a biometric trait is its universality, stability, uniqueness, and collectability [19].

The use of heart sound signals biometric systems to prevent identity fraud and forgery is one of the most active research areas at the moment, due to the advantages of biometric trait of heart sound. In this work, an effort has been made to increase the effectiveness of human identification using cardiac auscultation. Hence, multiresolution analysis of heart sound has been performed in this study in order to enhance human identification. For more accurate analysis of the non-stationary heart sound signals, a DWT based multiresolution decomposition was carried out. Important features were chosen and used as input to machine learning classifiers after signal acquisition, pre-processing, and the extraction of DWT-based and Cepstral features from the heart sound signals.

1.2. Motivation of the Study

There is surprise and amazement regarding the application of heart sounds for personal identification. An intriguing aspect is that phonocardiogram signals, as a biometric feature, possess the unique advantage of being unfeasible to artificially recreate. This study seeks to capitalize on the benefits of stethoscope integration with computers, aiming to swiftly acquire and analyze heart sounds for personnel identification.

The capacity to identify individuals through heart sounds proves particularly advantageous, especially when considering the challenges associated with other distinctive behavioral and physiological traits, which can be artificially reproduced, leading to potential misidentification.

The exploration of biometric identification based on phonocardiograms remains an active and evolving research area. Motivated by the prospect of enhancing identification capabilities, our research is inspired by the integration of a wavelet-based unique feature set with a Cepstral feature set. This dual approach compelled us to delve into human identification through heart sound analysis, envisioning an adaptation of stethoscope-based auscultation diagnosis for heart problems to more effectively leverage these sounds for identification purposes.

1.3. Statement of the problem

Nowadays, Biological signals such as heart sound, Pulse waves or ECG provide a greater robustness than image data because disguises are impossible, due to the reason that convectional biometrics are largely practiced for falsifiability, and can be easily forged. Studies showed that heart sound as biometric serve as a trait due to its universality, and uniqueness [14,16,19]. The most important merit is that the heart sound can be captured only from a living human body; therefore, it is difficult to forge or steal someone's PCG[3]. Despite significant advancements in biometric identification systems, the accurate and reliable identification of individuals based on heart sound analysis remains a challenging problem. Previous works in this domain have made valuable contributions for the implementation of Real-time processing capabilities, but several critical issues persist that hinder the practical implementation of heart sound analysis for biometric identification. These issues include limited accuracy, vulnerability to external noise and interference, lack of robustness across diverse populations which is a natural physiological variation among individuals from different demographic backgrounds can lead to inconsistencies in the method's performance, and inadequate scalability for real-world applications. Another

drawback in prior research involves lack of exploring various types of distinctive feature sets. The current methods frequently lack the efficiency and computational speed required for practical deployment, impeding their adoption in high-throughput identification applications.

Hence, it is imperative to explore diverse feature types to identify the most distinctive ones. This exploration aims to evaluate the effectiveness of heart sound analysis methods in biometric systems, ultimately enhancing efficiency and computational speed for practical deployment investigations.

1.4. Objectives of the study

1.4.1. General objective

The general objective of this research is to construct a model for a heart sound based human identification using machine learning through cardiac auscultation.

1.4.2.Specific Objective

The following specific objectives, which are geared toward resolving the research questions, were addressed in an effort to further the study's overall goal.

- To construct an electronic stethoscope by modifying the existing acoustic stethoscope
- To perform data collection of heart sound signals and get the information ready for training, and testing implementations
- To calculate the heart sound features in different domains

1.5.Significance of the study

This research holds numerous benefits for researchers in the field of physiological-based biometric system development. The unique characteristics of Phonocardiogram (PCG) signals open avenues for the creation of intriguing applications where continuous identification is a crucial element. Many studies have concentrated on auscultation and the attributes of heart sounds for identification, involving the adaptation and modification of diagnostic systems. Furthermore, this study has the potential to greatly assist researchers dedicated to exploring human identification through heart sound analysis, paving the way for the implementation of a biometric system with enhanced recognition performance.

1.6.Scope of the Thesis

The scope of this research was to identify and categorize subjects by analyzing their heart sounds. The heart sound records were obtained consistently from the pulmonary chest site for each of the subjects.

1.7.Document Organization of the study

The structure of this thesis comprises the following sections. **Chapter Two** introduces the heart sound production, signal processing techniques, and existing works on human identification systems using heart sounds. This chapter focuses on signal processing techniques and classification methods applied for personnel identification through heart sound processing. **Chapter Three** introduces literature review. While **Chapter Four** comprehensively covers the methods and materials employed in the research. It details the various techniques employed to achieve the research goals and introduces the tools crucial for the successful execution of the study. The results and discussion section are housed in **Chapter Five**, offering a comprehensive overview of the study's findings. Additionally, this chapter includes an in-depth discussion and interpretation of the results. Finally, **Chapter Six** concludes the thesis and provides recommendations. It encapsulates the overall conclusion of the study and proposes avenues for future research.

CHAPTER TWO

THE HEART

2.1.Hearts

Heart is a remarkable organ nestled within the chest, serves as the life giving power house of the human body [21]. Figure 2.1 shows the circulation of blood in the body. While the left side of the heart receives oxygenated blood from the lungs and pumps it to the body and brain, the right side of the heart receives blood from the body and brain's venous circulation and pumps it to the lung for oxygenation and carbon dioxide release [22].

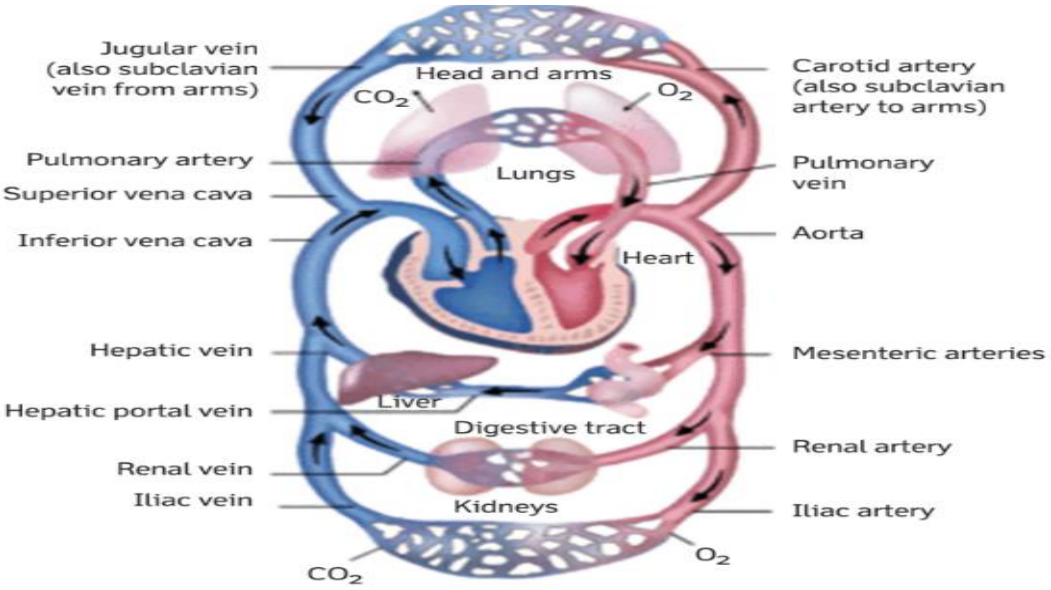


Figure 2.1 : The circulation of blood in the body [22].

The major diagnostic tool most frequently used to evaluate the condition of the heart is the stethoscope-based auscultation approach. since Laennec created the first stethoscope in 1816, it has been the accepted method of diagnosis [23]. Due to the fact that the heart's sounds encompass information about an individual's physiology, such signals possess the potential to offer a distinct identity for each person[24] .

2.1.1.Anatomy and physiology of Heart

The pericardium, myocardium, and endocardium are the three layers that make up the heart wall [25,26]. The heart is divided into right and left side and is composed of four chambers: the right atrium, the left atrium, the right ventricle and the left ventricle. The two lower chambers are known as the right and left ventricles, while the two top chambers are called the right and left atrium and left atrium (plural: atria) [26].

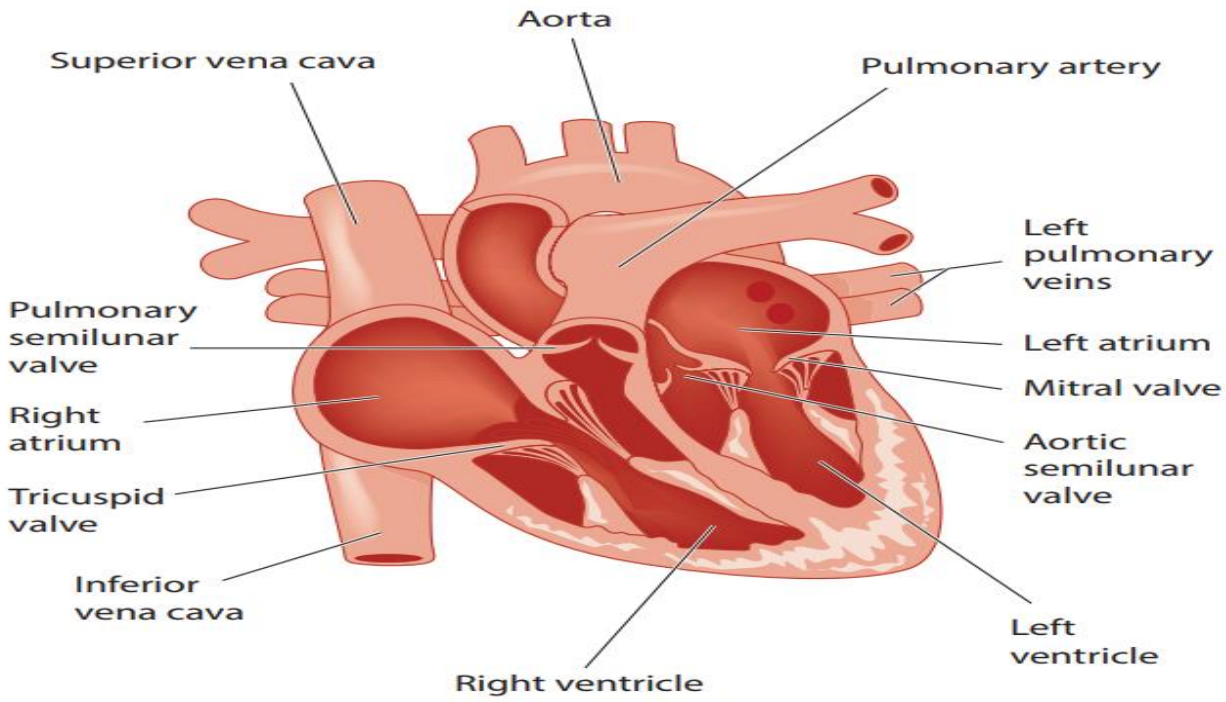


Figure 2.2 : Chambers of the heart and valves [26].

The vessels that carry blood to and from the heart are mostly of three sorts. These comprise the capillaries, veins, and arteries. Blood is transported to and from the heart by the veins and arteries, respectively.

The conduction system-coordinated alternating contraction and relaxation of the myocardium in the heart chamber walls during a single heartbeat is referred to as the cardiac cycle. The cardiac cycle's contraction phase is called diastole, while its relaxation phase is called systole. One cardiac cycle takes place every 0.8 seconds at a normal heart rate [27]. As the main organs of the human cardiovascular system, hearts are susceptible to a variety of heart diseases and heart sounds are crucial markers of both cardiac health and disease. Meaning that heart sound analysis can be used to easily diagnose conditions related to the heart valves in particular. The next section covers the mechanics of heart sounds, how they are produced, an overview of signal processing methods, and heart sounds.

2.2.Heart sound

Phonocardiography (PCG) signals are digital signals created from the heart sounds captured by an electronic stethoscope [28]. The acoustic vibrations of the valvular, muscle, vascular, and blood circulation provide the audible components of heart sound. The first heart sound (S1), the systolic pause segment (also known as the systolic pause period) following the sound S1, the second heart sound (S2), and the diastolic pause segment (also known as the diastolic pause period) following the sound S2 are the four main segments of a typical cardiac PCG cycle. Systolic and diastolic interval segments may contain the other additional heart sounds, which include the heart murmurs, the fourth heart sound (S4), and the third heart sound (S3) [29]. The heart sound components (s1, s2, s3, s4) of the PCG signal are displayed in Figure 2.3.

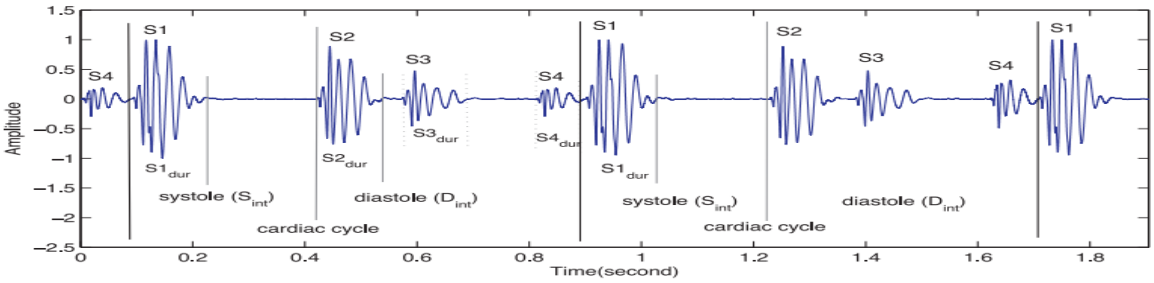


Figure 2.3 : PCG signal illustration with heart sounds (S1, S2, S3, S4) [29] .

2.2.1.Mechanism of heart sound production

The human heart operates as a four-chambered pump: two atria receiving blood from veins and two ventricles propelling blood out to arteries. Its mechanical orchestration relies on an electrical signal originating from specialized pacemaker cells in the right atrium (known as the Sino-atrial node). This signal navigates through the atria to the AV-node (a delay junction) before reaching the ventricles. The heart's rhythmic activity emerges from a complex interplay among pressure gradients, blood flow dynamics, and the compliance of cardiac chambers and blood vessels. Within this mechanism, distinct sounds and murmurs which is a pathology play a role, notably S1 and S2, both easily detectable and lasting approximately 150 ms and 120 ms, respectively. Falling within a frequency range of 20 to 150 hertz, these sounds are pivotal. S1 aligns with the isovolumetric contraction of the ventricles, correlated with the closure of the mitral-tricuspid valve [28,29]**.** during the ventricular phase of isovolumetric relaxation, the heart sound S2 correlates with the closure of the aortic and pulmonary valves. Contrarily, the sounds S3 and S4, faint and often imperceptible, carry very low frequencies. Specifically, S3, occurring early in diastole, doesn't originate from a valve, earning it the name "Proto diastolic sound." Blood influx into the left ventricle’s triggers valve vibrations, manifesting as S3, and S4, the heart sound occurring in late diastole, corresponding with atrial contraction.

S3 is typically found in children and young individuals, it typically diminishes by middle age. Yet, its emergence later in life signals a risk of Proto-diastolic gallop, indicating potential left ventricular failure. On the other hand, S4, prevalent among healthy youngsters, becomes rare in adults. When present, it's termed presystolic gallop. During diastole, the ventricles undergo relaxation, causing a decrease in pressure within these chambers. This drop in pressure prompts the opening of the Mitral and Tricuspid valves when the atrial pressure falls below that of the ventricles, enabling blood to move from the atria into the ventricles. Subsequently, during systole, the ventricles contract. The surge in pressure prompts the closure of the Tricuspid and Mitral valves, while the abrupt rise in pressure leads to the opening of the Aortic and Pulmonary valves, facilitating the ejection of blood out of the heart [28,29,30].

A murmur is one of the more common abnormal phenomena of the heart activities. The substantial difference between heart sounds and murmurs is that murmurs are noisy and have a longer duration. Heart murmurs can be classified as a systolic murmur occurring in the systolic interval, and diastolic murmur or continuous murmurs occurring in the diastolic interval [30]. Figure 2.4 shows the difference between systolic and diastolic.

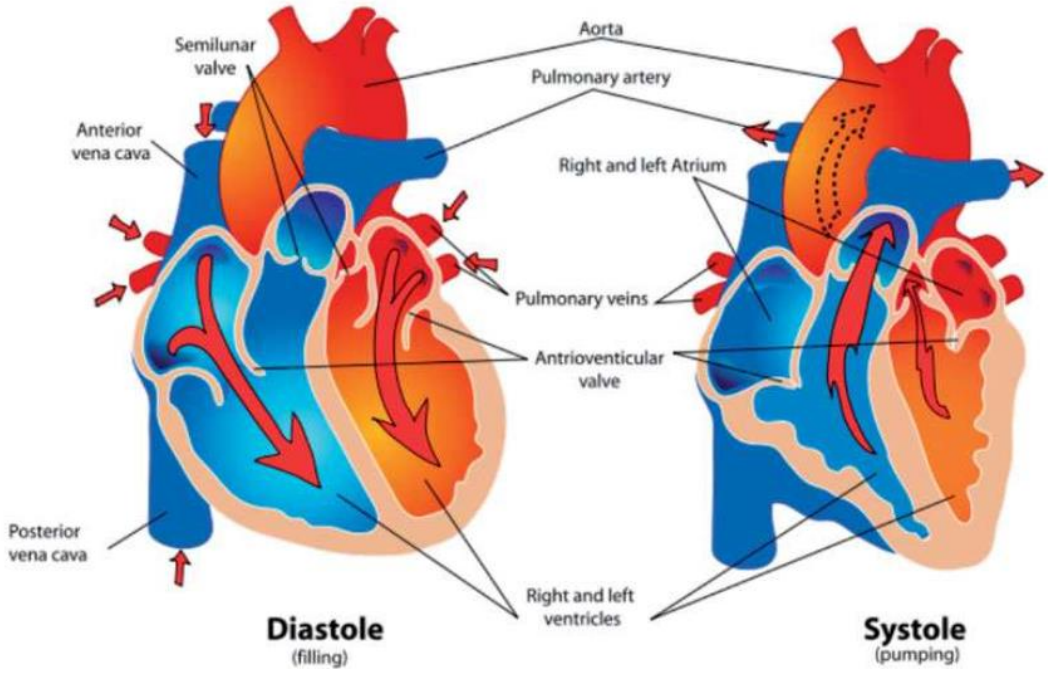


Figure 2.4 Systole and diastole of heart

2.2.2.AUSCULTATION PROCESS OF HEART SOUND

Auscultation involves listening to internal body sounds, typically using a digital stethoscope. Within the chest region, there are four key auscultation sites: the Aortic, Pulmonary, Lower Sternal Border, and Mitral areas. The primary heart sounds, S1 (lub) and S2 (dub), each have distinct origins. S1 arises from the abrupt halt of blood flow caused by the closure of the Atrioventricular valves (comprising the Tricuspid and Mitral valves) at the beginning of ventricular contraction (Systole). Conversely, S2 stems from the blockage of blood flow due to the closure of the

Semilunar valves at the conclusion of ventricular systole, marking the onset of ventricular diastole [31].

2.3.Signal Processing Techniques

Signal processing techniques encompass a broad range of methods used to manipulate, analyze, and interpret signals. techniques for signal processing include a wide range of approaches designed to extract relevant information from signals in different fields. These methods allow for the understanding, interpretation, and manipulation of signals. They include frequency-domain transformations and time-domain analysis. Analyzing a signal's amplitude and temporal properties across time is known as time-domain analysis. On the other hand, frequency-domain methods like Fourier analysis examine signal components at various frequencies in order to reveal spectral information and hidden patterns in the signal. For localized signal analysis, other techniques such as wavelet analysis provide a time-frequency resolution balance.

2.3.1. Time – domain Analysis

Time domain analysis involves the examination of data across a specified time period. This method is applied to study various functions like electronic signals, market behaviors, and biological systems [32]. Time domain analysis of an electronic signal primarily involves examining either the voltage-time plot or the current-time plot. This analysis always correlates a variable against time. Various devices are employed for time domain analysis, with the cathode ray oscilloscope (CRO) being the most prevalent tool for scrutinizing electrical signals in this domain [32].

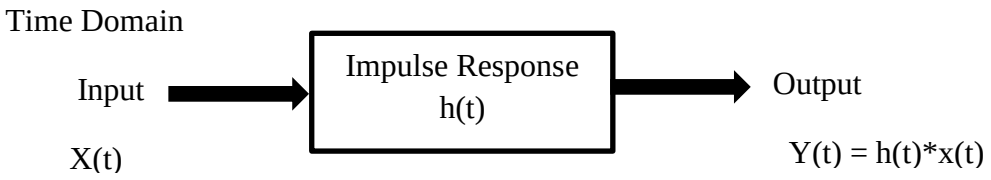


Figure 2.5 : Time- domain approach [32]

When examining an audio signal using the time domain approach, the X-axis represents time, while the value on the Y-axis reflects the signal's variation over time.

2.3.2. Frequency - domain analysis

Frequency domain analysis primarily examines signals or functions that exhibit periodicity over time, although it's not limited to analyzing solely periodic signals. Transformations form a crucial aspect of frequency domain analysis, facilitating the conversion between time and frequency domain functions. Among these transformations, the Fourier transformation stands out as the most prevalent in frequency domain analysis. It excels in converting signals of varying shapes into a composite of an infinite number of sinusoidal waves [33].

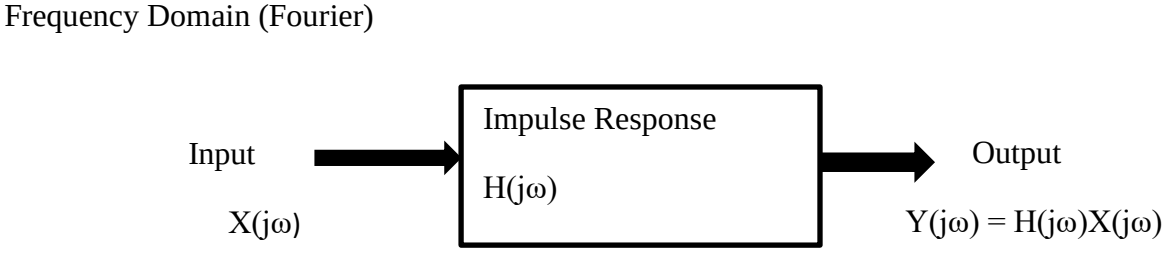


Figure 2.6 : Frequency – Domain approach [33].

Equation (2.1) presents the Fourier analysis representation of an energy function x(t) with a finite duration as the sum of sinusoids, denoted as $e^{j\omega}$ [34].

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(\omega)e^{j\omega t}\, d\omega \quad \text{.....Equation (2.1)}$$

The amplitude x(ω) for every sinusoidal wave function $e^{j\omega t}$ equals the correlation between $e^{j\omega t}$ and x, referred to as the Fourier transform, as detailed in Equation (2.2) [34].

$$x(j\omega) = \frac{2}{2\pi}\int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt \quad \text{.....Equation (2.2)}$$

The smoother the function x(t), the quicker the decay of the amplitude |x(ω)| as the frequency ω increases. This pair of equations is termed a Fourier transform pair. While the Fourier transform serves as a robust frequency domain analysis tool for examining time-invariant signals, it falls short when dealing with non-stationary signals. In cases like heart sounds, which exhibit non-stationarity, relying solely on Fourier analysis proves challenging. Hence, to extract crucial

diagnostic insights from heart sounds, a comprehensive analysis combining both time and frequency is imperative. Alternative techniques like Short-Time Fourier Transform (STFT) and wavelet analysis step in to facilitate this crucial time-frequency analysis of signals.

2.3.3. Time Frequency domain Analysis

Time-frequency analysis is a powerful method used in signal processing to examine how the frequency content of a signal evolves over time. Unlike traditional Fourier analysis that provides frequency information but loses temporal details, time-frequency analysis captures both time and frequency characteristics simultaneously, offering a more comprehensive understanding of signals, especially those that change their frequency components over time. Techniques like Short-Time Fourier Transform (STFT), wavelet transforms, and spectrograms are employed to dissect signals into their time-varying frequency components, enabling the identification of transient events, non-stationary behavior, and complex signal structures that might be missed by conventional frequency analysis alone.

This approach finds wide application in various fields such as biomedical signal processing, audio processing, communication systems, and vibration analysis, where understanding the time-varying nature of signals is crucial for accurate interpretation and decision-making.

2.3.3.1. Short time Fourier transforms (STFT)

When dealing with continuous nonstationary signals, the Fourier Transform (FT) proves unsuitable due to its inability to offer simultaneous time and frequency information. Consequently, the Short-Time Fourier Transform (STFT) emerged as a solution to this limitation. The concept involves segmenting the continuous signal into smaller, equal-length sections using a window function w(t- τ), where τ represents the segment duration. Subsequently, the Fourier Transform is computed for each segment, utilizing methods such as the Discrete Fourier Transform (DFT) or the Fast Fourier Transform (FFT). This approach enables the analysis of localized frequency components within specific time intervals, facilitating a more comprehensive understanding of time-varying signals [35]. Equation 2.3 represents the mathematical formulation of the continuous-time Short-Time Fourier Transform (STFT) [35].

$$STFT\{x(t)\} = x(\tau, w) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-i\omega t}dt \quad \text{.....Equation (2.3)}$$

In this context, X(τ, w) denotes the Fourier Transform (FT) of x(t) convolved with the window function w(t - τ). The function w(τ) is centered at zero, while τ represents the time index. The spectrogram visually presents the magnitude of the Short-Time Fourier Transform (STFT) over time and is expressed as Equation 2.4 [35].

$$spectrogram\,\{x(t)\} = |x(\tau, w)|^2 \quad \text{Equation (2.4)}$$

When dealing with discrete domain data, it involves breaking the data into smaller segments, computing the Discrete Fourier Transform (DFT) for each segment, and gathering the outcomes in matrix form. Equation 2.5 illustrates this process [35].

$$STFT\{x[n]\} = x(m, w) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-i\omega t} \quad \text{Equation (2.5)}$$

The Short-Time Fourier Transform (STFT) addresses the shortcomings of the Fourier Transform (FT), yet its effectiveness hinges on the chosen window size, determining time and frequency resolutions. Opting for a smaller window yields better time-scale resolution but sacrifices frequency precision. Conversely, a larger window enhances frequency accuracy at the cost of time resolution. This presents a trade-off between time and frequency resolutions, where selecting an appropriate window size stands as the primary challenge in STFT. Consequently, STFT encounters limitations when analyzing nonstationary signals like Phonocardiogram (PCG), which exhibit frequent variations in frequency. Wavelet Transform (WT) steps in to mitigate the drawbacks of STFT, an aspect elaborated upon in the subsequent subsection [35].

2.3.3.2. Wavelet Transform

A wavelet, distinct from a sinusoid, embodies a swiftly decaying wave-like oscillation characterized by a zero mean and infinite extension across an unbounded duration. Wavelets exhibit diverse sizes and shapes, such as morlet, daubechies, coiflets, biorthogonal, Mexican hat, and symlets. Presently, wavelets find extensive application in analyzing transient, time-varying, or nonstationary phenomena. The fundamental concept of Wavelet Transform (WT) involves identifying a basis function that offers an effective computational framework. Unlike the Short-Time Fourier Transform (STFT), where fixed-size windows limit the extraction of very small frequency components crucial in signals like Phonocardiogram (PCG), WT introduces variable-sized regions. This approach enhances frequency resolution in the low-frequency range while preserving commendable time resolution at higher frequencies [35].

The wavelet is manipulated in two ways: translation and scaling. Translation involves shifting the central portion of the wavelet along the time axis, facilitating the extraction of time-related information from a signal. On the other hand, scaling alters the amplitude and time duration of the wavelet functions to derive frequency-related details. Consequently, through these operations of translation and scaling, the wavelet becomes localized simultaneously in both the time and frequency domains [36].

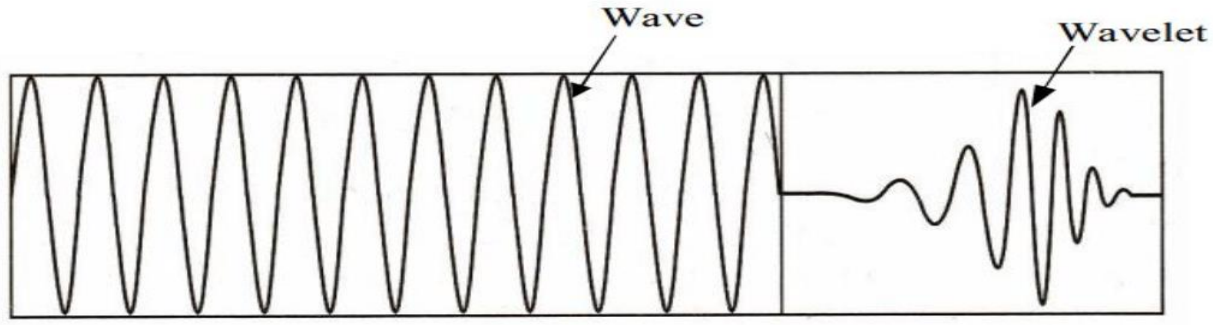


Figure 2.7 : A wave and wavelet [36]

Therefore, Wavelet Transform (WT) proves more suitable for heart sound analysis. It is categorized into two types: Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT)[35].

2.3.3.2.1 Continuous wavelet transform

The Continuous Wavelet Transform (CWT) of a time-domain signal x(t) is described as the integral transform of x(t) using a family of wavelet functions $\psi_{a,b}$(t) [35].

$$x(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\, \psi * \left(\frac{t-b}{a}\right) dt \quad \text{.....Equation (2.6)}$$

In this context, the function ψ(t) signifies the mother wavelet, while $\psi_{a,b}$ (t) is referred to as the daughter wavelet.

Additionally, ψ* denotes the complex conjugate of the wavelet function ψ. Daughter wavelets are generated through operations of scaling and shifting applied to the mother wavelet. Parameters a and b represent the scaling factor and shifting factor, respectively. In the discrete domain, the Continuous Wavelet Transform (CWT) can be formulated as equation 2.7 [35].

$$x[n,s] = \frac{1}{\sqrt{|s|}} \sum_{i=\, n-\frac{M}{2}}^{n+\frac{M}{2}} x[i]\psi\left[\frac{i-n}{s}\right] \text{ .....Equation (2.7)}$$

In this context, x signifies the signal undergoing transformation, ψ represents the mother wavelet function, n denotes the dilation parameter, and s stands for the translation parameter. M signifies the duration of the wavelet when s = 1. The Continuous Wavelet Transform (CWT) diverges from other transforms by presenting its graphical representation in terms of time versus scale rather than time versus frequency. The Wavelet Scale Spectrum (WSS) finds expression in a discrete domain as equation 2.8 [35].

$$WSS(x[n,s]) = \sum_{n=1}^{N} |x[n,s]|^2 \text{ . . . . . Equation (2.8)}$$

N represents the number of samples in the signal. Utilizing the Continuous Wavelet Transform (CWT) often leads to highly repetitive wavelet coefficients, demanding substantial processing time. To mitigate redundancy and expedite computations, the Discrete Wavelet Transform (DWT) is employed.

2.3.3.2.1. Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) combines both a low-pass filter (LPF) and a high-pass filter (HPF). This process breaks down the input signal into two key coefficients: the details coefficient and the approximate coefficient. High-frequency components reside within the detail coefficient, while the low-frequency components are found in the approximation coefficients. By iterating through multiple levels of the DWT, the transformation relies on the approximation coefficients from one level to the next, offering varying levels of signal decomposition The Discrete Wavelet Transform (DWT) differs from the Continuous Wavelet Transform (CWT) by employing scale and position values that adhere to powers of 2 and the values are a=2^j, and $b = k^* 2^j, (j,k) \in z^2$ given as [35].

$\Psi_{a,b}(t) = \frac{1}{\sqrt{2^j}} \Psi\left(\frac{t-k*2^j}{2^j}\right)$ ……Equation (2.9)

The Discrete Wavelet Transform (DWT) serves to break down signals into high- and low-frequency components, enhancing signal analysis. To reconstruct the original signal, an inverse DWT operation is necessary. The degree of wavelet decomposition hinges on the chosen hierarchical levels. Selection of the decomposition level is influenced by the desired cut-off frequency. Moreover, specifying the mother wavelet and DWT level is critical for extracting specific features from Phonocardiogram (PCG) signals [35].

In this study, we utilized the wavelet multiresolution analysis technique to examine heart sound signals through discrete wavelet transforms (DWT). DWT stands out as the optimal method in wavelet signal processing for conducting multiresolution analysis of signals. Its capability to decompose signal components, offering fine frequency resolution for low-frequency elements and fine time resolution for high-frequency components, renders it highly suitable. This DWT-based multiresolution analysis aids in a comprehensive understanding of signals and proves valuable in applications involving feature extraction. Compared to Fourier transform and Short-Time Fourier Transform (STFT), employing Wavelet signal processing via DWT enables a more thorough multiresolution analysis of non-stationary heart sound signals.

CHAPTER THREE

LITERATURE REVIEW

3.1. Related Works

This section provides a review of earlier studies focused on analyzing heart sounds using distinct techniques aimed at identifying individual uniqueness.

Zhidong, et al. [37], utilized a feature set derived from new marginal spectrum coefficients, resulting a remarkable recognition rate of 94.4%. This marked a substantial improvement compared to the traditional Fourier spectrum's recognition rate of 84.32%. Their study, conducted on a database comprising 280 heart sounds from 40 participants, utilized VQ (vector quantization) as the classification model. Likewise, Abo-zahhad, et al. [38], introduced wavelet packet cepstral coefficients as novel features for heart sound signals in biometric applications. Their system incorporated tests with Vector Quantization (VQ), Gaussian Mixture Models (GMM), Discriminant Analysis with linear boundaries, and Support Vector Machine (SVM). Their findings highlighted the system's peak classification accuracy at 91.05%, achieved specifically using the LDA and Bayes decision rule classifier.

Swati Verma, and Tanuja Kashyap. [39], conducted human identity identification by employing Mel-frequency cepstral coefficients (MFCC) for feature extraction. They utilized a linear support vector machine (SVM) as the classifier, and TPR and TNR obtained for 30 individuals are 0.96 and 0.99 respectively in their study focused on identifying human identities through heart sound signals. Similarly, S. Zahra Fatemian, et al. [40], assessed the distinctive properties of ECG and PCG signals through a bimodal biometric system. Their findings indicated an impressive overall accuracy of 97% for the fused configuration, surpassing the individual performance of each biometric trait. El-Bendary, et al. [41], preprocessed 400 heart sounds utilizing DWT and reported precision scores of 93% with the KNN classifier and 80.7% with the MSE classifier. Julian, et al. [42], conducted an analysis of heart sound signals in the time-frequency domain employing a wavelet transform. They extracted the feature vector using a Shannon energy envelogram.

Likewise, Muhammed, et al. [43], utilized empirical mode decomposition (EMD) to analyze a self-acquired signals of heart sounds for robust biometric human identification. They attained a remarkable accuracy of 95.4% by employing a quadratic kernel for SVM. Cheng, et al. [44], developed a double-header, two-way voice auscultation detection device for signal acquisition. They introduced a unique human feature extraction method based on an enhanced circle convolution slicing algorithm combined with independent sub-band function (ISP).

Their approach claimed an identification accuracy of 85.7%. S. Bindu, et al. [45], concentrated on analyzing PCG signals for biometric identification. They utilized the Shannon entropy method to extract the S1 and S2 features of heart sounds. Sherif, et al. [46], utilized wavelet denoising for preprocessing and incorporated two novel cepstral feature sets, namely Modified Mel frequency cepstral coefficient (MFCC) features and wavelet packet cepstral coefficients (WPCC), into the recognition process. They tested vector quantization (VQ), Gaussian mixture models (GMM), and discriminant analysis (DA), concluding that the most optimal recognition rate was achieved through discriminant analysis. Cheng, et al. [47], implemented the transformation of one-dimensional heart sounds into a 2D Phonocardiogram (PCG) through wavelet noise reduction, amplitude normalization, and image processing technology. Their approach, incorporating the application of inflection point sequence code, resulted in a notable recognition accuracy of 95.67%. Imran et al. [48], introduced the AR Burg method and proposed essential features, namely Marginal Hilbert Spectrum (MHS), Mel-Frequency Cepstral Coefficients (MFCC), and Modified MFCC (M-MFCC). The authentication process was optimized using the EB-Trees classifier, establishing it as an effective method for human authentication. In the preprocessing stage, PCG signals underwent Envelope Extraction through Hilbert-Huang Transform (HHT), and wavelet denoising was employed to detect peaks. The user identification achieved an accuracy of 86.7%. Xiefeng Cheng, et al. [49]. proposed an innovative approach for heart sound biometric recognition using multimodal multiscale dispersion entropy.

They initiated the process by employing an enhanced complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) to decompose heart sounds into intrinsic mode functions (IMFs). Subsequently, refined composite multiscale dispersion entropy (RCMDE) was calculated following the segmentation of IMFs into frames. Feature selection involved the use of the Fisher ratio (FR). Ultimately, an impressive accuracy rate of 96.08% was achieved through Euclidean distance (ED) and the closure principle.

Additionally, Mizuho Tagashira, and Takafumi Nakagawa.[50],conducted biometric identification using heart sounds on ten healthy participants and 19 datasets with heart murmurs. They attained a recognition rate of 90% for the ten healthy participants and a perfect 100% recognition rate for the 19 cases involving heart murmurs. Tien-En Chen, et al. [51], presented a recognition system based on acoustic features. Their primary emphasis was on two sets of data: a training set comprising 16 participants with 626 heart sounds and a test set with 6 subjects and 120 heart sounds. To enhance the features, they employed MFCC and k-means for extraction. These characteristics were then fed into five distinct classifiers: LR, GMM, KNN, SVM, and DNN. The trial results indicated that DNN achieved the highest accuracy, reaching 91.12%.

In the field of heart sound analysis, researchers employ various signal analysis techniques to analyze heart sounds. These techniques include spectral analysis methods such as Fourier Transform (FT), AR-Burg, Mel-frequency cepstral coefficients (MFCC), and among others. These methods are commonly used to analyze recorded heart sounds in the context of biometric systems [37,39,46]. Xiefeng Cheng, et al. have introduced the concept of multimodal multiscale dispersion entropy, utilizing enhanced complete ensemble empirical mode decomposition with adaptive noise. This technique has gained popularity as a heart sound analysis method for exploring the identity of subjects. Table 2.1 provides a concise overview of recent studies conducted by diverse scholars. It summarizes the methods employed by researchers, including preprocessing, feature extraction, and classification techniques, along with the corresponding accuracy achieved in each study.

Table 3.1 : A summary of recent related researches.

Authors	Applied methods			Database	Results
	Preprocessing	**Feature Extraction**	**Classification**		
Sherif, et al., 2016.[46]	DWT	M-MFCC, and WPCC	VQ, GMM,DA,SVM	Bio sec PCG database of 21 subjects	CRR= 99.3%
Xiefeng Cheng, et al., 2016.[52]	GRAY background, normalization, Binarization refinement	Vertical, and horizontal ratio of coordinate sequence code of inflection point	Euclidean Distance	Heart sound recording of 16 subjects	Acc= 95.67%
Imran Farad, et al. , 2019.[48]	DWT Envelope Etraction: Hilbert- Hung Transform (HHT) AR BURG modeling	MFCC M-MFCC	Q-SVM F-KNN EB-Trees	HSCT-11	Acc= 86.7%
Xiefeng Cheng, et al., 2020.[49]	Down sampling Denoising, and	RCMDE	SVM KNN	Michigan, Washington, and Littman	Acc=96.08%

	cycle segmentation. Wavelet packet multi threshold denoising	ICEEMDAN-RCMDE- FR	ED	heart sound database	
Muhammed, et al., 2020.[43]	EMD	Time- Domain, Frequency-Domain, and cepstral features	Q-SVM	Heart sound recording of 30 subjects	Acc=95.4%

Certain researchers employed time-domain analysis techniques, including statistical methods and empirical mode decomposition (EMD), to analyze heart sounds [43,50]. Certain investigators incorporated modified Mel frequency cepstral coefficients (M-MFCC) and wavelet packet cepstral coefficients (WPCC) into the recognition process [38,46]. however, relying solely on frequency-domain or time-domain analysis techniques falls short in extracting sufficient information, particularly when applied to nonstationary signals like heart sounds. Therefore, to extract crucial distinctive information from heart sounds, it becomes imperative to incorporate time-frequency domain analysis of heart sounds. Some researchers have employed time-frequency analysis techniques, including Short Time Fourier Transform (STFT) and Wavelet Transform (WT), to effectively analyze lung sounds [41,42,46]

The optimal processing method for analyzing non-stationary signals such as heart sounds is the Discrete Wavelet Transform (DWT) [38,53]. DWT stands out because it can break down high-frequency components of a signal with precise time resolution and broad frequency resolution, as well as decompose low-frequency components with detailed frequency resolution and less precise time resolution. This capability allows DWT to facilitate multiresolution analysis of non-stationary heart sound signals. The process involves decomposing the signal into detail (high frequency) coefficients and approximation (low frequency) coefficients through successive high-pass and

low-pass filtering operations. Hence, this study proposes a method involving DWT-based analysis of heart sounds to further identify the distinctive characteristics of heart sound uniqueness.

3.2. Gap Analysis

While numerous studies have focused on PCG signal processing to analyze the uniqueness of heart sounds, it has been observed in some literature that many of the previous works lacked proper preprocessing or denoising steps [37]. Another identified gap in some literature is the use of a limited number of features [39,43,45]. Some of the studies conducted in this area [44,53], attained limited accuracy and utilized a small number of datasets. This research draws upon the insights from the reviewed and other relevant recent literature to inform its investigation.

CHAPTER FOUR

METHODS AND MATERIALS

4.1 Overview

Figure 4.1 illustrates the general procedure used for personal identification based on heart sound analysis. It comprises the following steps: building an electronic stethoscope; acquiring signals; preprocessing; feature extraction; selecting features; visualizing and normalizing features; splitting data; training and evaluating models; optimizing models; and classifying subjects.

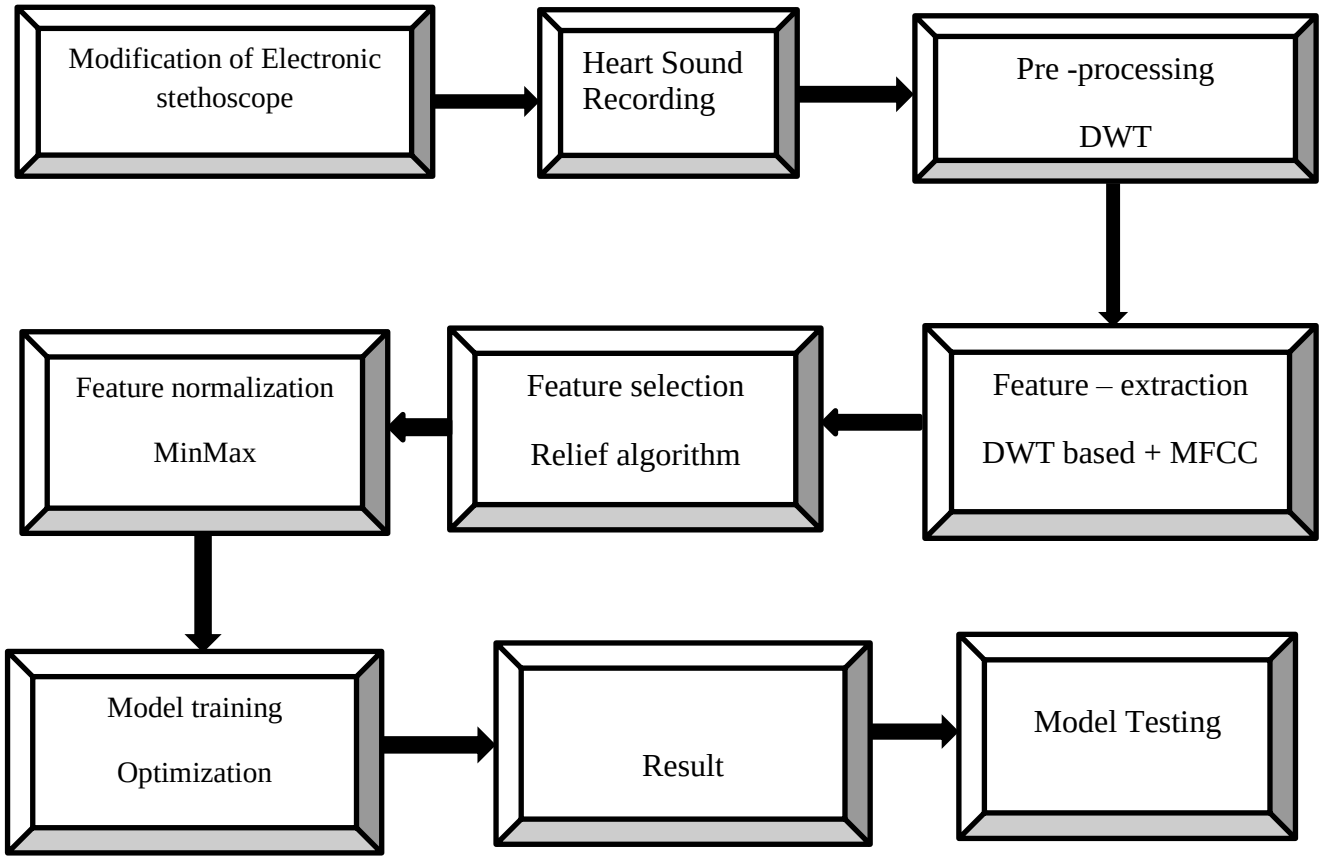


Figure 4.1: A general procedure for subjects' identification

4.2.Development of Electronic Stethoscope

With the use of a stethoscope, a medical instrument, one can listen to various internal body sounds, including heart and lung sounds as well as Korotkoff sounds [54]. The ear tips, ear tube, tubing, headset, stem, chest piece, diaphragm, and bell are the main parts of the stethoscope. When these components are assembled correctly, a working stethoscope is produced, enabling medical personnel such as doctors and nurses to listen to a variety of sounds.

Figure 4.2: Image of Stethoscope[55].

Nevertheless, the capabilities of an acoustic stethoscope—which are widely accessible in developing countries' healthcare settings—are restricted to the analog filtering and amplification of bio-sounds for subsequent interpretation by qualified medical personnel[55]. It is crucial to build an electronic stethoscope that differs from the conventional acoustic stethoscope in a few key ways in order to circumvent these limits and subjectivity problems. An electronic stethoscope's primary purpose is to transform analog sound signals into digital ones.

In order to transform the acoustic stethoscope into an electronic one, a transduction device that converts acoustic waves into an equivalent electrical signal is needed. We used a condenser microphone because of its inexpensive cost, superior sensitivity, high SNR, and flat frequency response [55]. An electret condenser microphone, the predominant transducer for sound signal detection, converts sound into an electrical signal. Similar to a capacitor, it contains two plates whose distance varies with environmental sound intensity. Operating within a frequency range of 20 Hz to 16,000 Hz, it captures heart sounds detectable by stethoscopes, transforming them into electrical signals. Hence, in this study, we employed an electret microphone with a frequency range of 20 Hz to 16,000 Hz, high sensitivity and a Signal-to-Noise Ratio (SNR) exceeding 62 db.

4.3. Signal Acquisition

In this study, Like the majority of earlier researchers, we gathered our own dataset since inherent physiological differences between people from various demographic backgrounds can cause inconsistencies in the method's effectiveness. We employed our electronic stethoscope to capture local heart sounds during a collection session held at Sant Luke Catholic Hospital College of Nursing, and Midwifery. 560 heart sounds in all, 28 recordings of each heart sound from a group of 20 people, consisting of 13 males and 7 females aged between 19 and 35. considering the inclusion criteria of the study's focus on conducting it on healthy subjects Participants underwent preliminary screening conducted by the clinical collaborator using Stethoscope to evaluate their cardiac condition. The duration of each recording varied from 30 to 120 seconds. There are four different typical auscultation sites to place microphones these sites are Aortic, pulmonary, LLSB, and mitral [56]. For this study all heart sounds acquired from the same location: the pulmonary of the heart. We chose this location because, in some cases, particularly in those with high body mass indices, the pulmonary area near the left upper sternal border can provide improved access to heart sounds, also it may be more convenient or comfortable for subjects to provide data from the pulmonary area than from the chest wall. . It's noteworthy that every recording was sourced from individuals with verified good health.

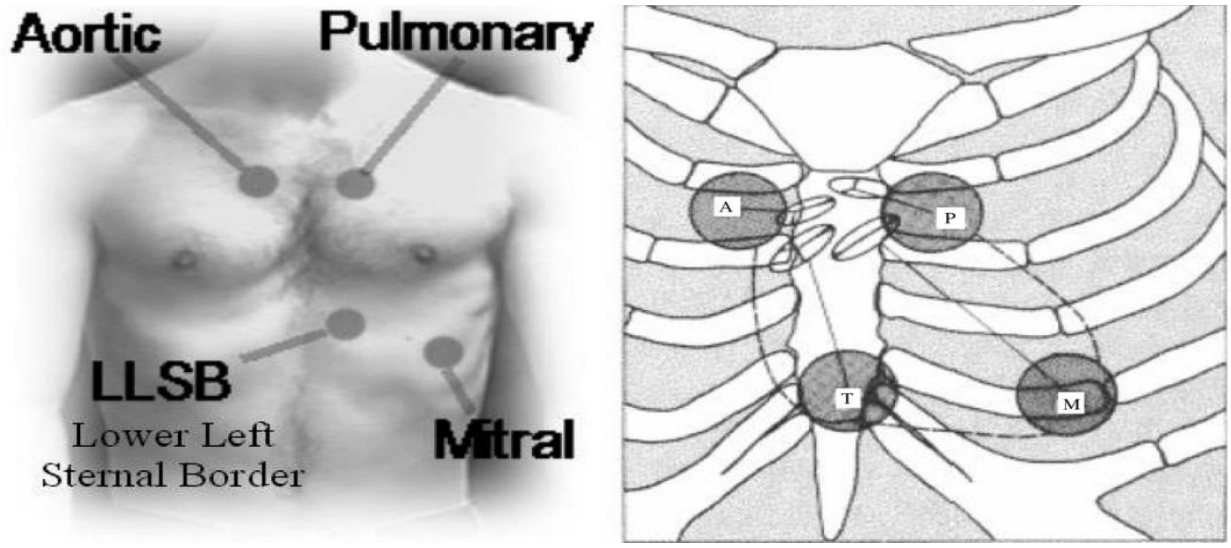


Figure 4.3 : Typical auscultation sites to place microphones [56].

4.4. Data augmentation

Data augmentation involves creating additional samples by applying various transformations to existing data. The purpose is to augment the dataset, thereby enhancing the accuracy and resilience of classifiers. Audio signal augmentation techniques play a crucial role in enhancing the diversity and quantity of audio data for training machine learning models. These methods involve applying various transformations to original audio signals, such as time stretching, pitch shifting, noise injection, amplitude scaling, and spectrogram modifications [56]. Since amplitude scaling provides a simple and efficient means of controlling the signal's level without appreciably changing its fundamental properties. Amplitude scaling merely modifies the signal's intensity, as opposed to time stretching or pitch shifting, which have the ability to alter the signal's temporal or frequency content. For this reason, the chosen data augmentation technique for this study is amplitude scaling with scale range of [0.8,1.2]. This makes it possible to precisely adjust the amplitude of the signal, which is helpful for highlighting particular aspects or reducing noise without adding extraneous artifacts. Amplitude scaling is a feasible option for audio signal augmentation in this situation because it is also simple to apply and computationally efficient.

4.5. Preprocessing of Heart Sound Signals using Discrete wavelet transform

Signal pre-processing, especially denoising, stands as a crucial phase in signal processing, pivotal for eliminating disruptive artifacts that might compromise the integrity of acquired signals. Stethoscope recordings, in particular, are susceptible to various forms of noise in the process to collect heart sound signals, it is Vulnerable to external acoustic signals and electrical noise interference, in particular, the friction caused by subjects breathing or body movement. Since The wavelet transform is a time-frequency signal analysis method based on Fourier transforms. It has good localization in both frequency and time domains. This research employes a wavelet-based denoising technique. This method proves instrumental in eliminating disruptive signals that might otherwise distort the time and frequency characteristics of heart sounds. Heart sound denoising via Discrete Wavelet Transform (DWT) involves a sequence of steps: decomposition, detail coefficient thresholding, and reconstruction. Initially, a pertinent mother wavelet is selected for the decomposition phase. We applied soft thresholding to the detail (high frequency) coefficients of the signal *S* through comparing the signal to noise ratio. Subsequently, the process concludes with reconstruction, which involves modifying the detail coefficients of levels 1 through N and restoring the original approximation coefficients of level N. This method ensures the restoration of the heart sound signal while effectively managing the noise across its different levels. We executed the DWT-based denoising of Heart Sound signals by adhering to the procedure depicted in Figure 4.4.

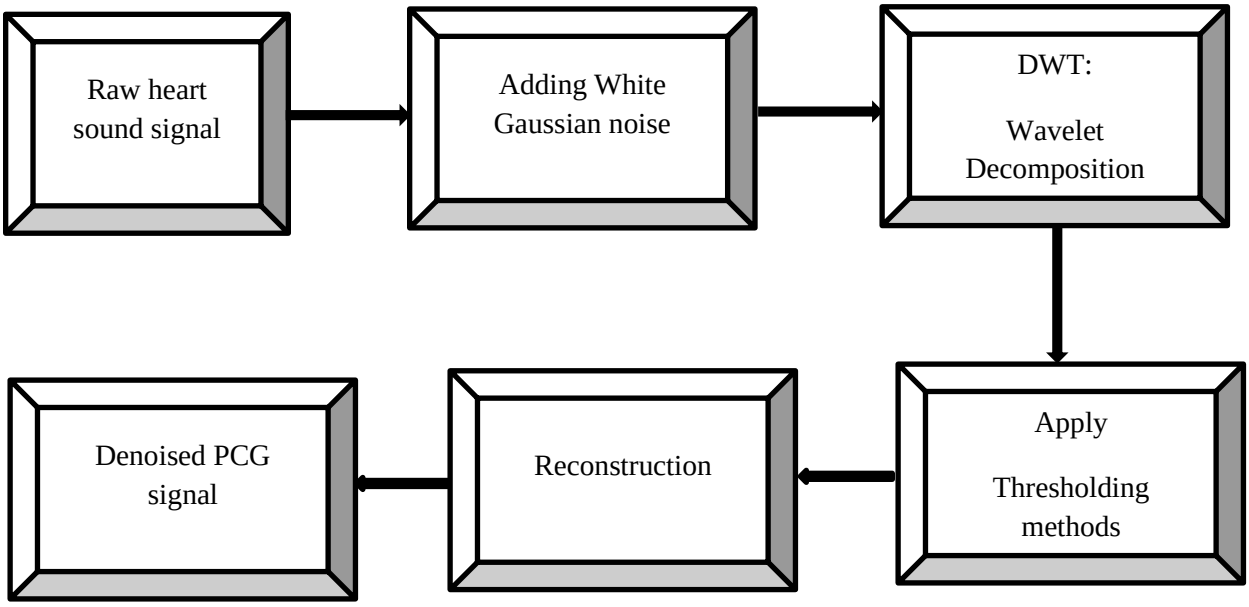


Figure 4.4 : A method to denoise heart sound signals using wavelet-based techniques.

The DWT process decomposes Heart sound signals into high and low frequency components through consecutive high pass and low pass filtering steps. This signal decomposition involves subjecting the signal to a sequence of high pass filters for assessing its high frequency elements and low pass filters for evaluating its low frequency elements. After evaluating various wavelet functions, we identified a 5th level decomposition of heart sound signals using the db10 wavelet function for the best outcome. The evaluation of different wavelet family through comparing SNR is presented under chapter five table 5.1. The chosen wavelet function is assessed by comparing the signal-to-noise ratio, and the decision to select the wavelet family is based on its performance across various studies, demonstrating superior results as observed in this research. To conduct thresholding on the signals, both soft and hard thresholding methods, the two most prevalent approaches, were utilized. Additionally, four distinct threshold selection rules were employed

within this study to assess their efficacy in signal denoising [57]. Furthermore, this study delved into assessing the efficacy of these thresholding methods in denoising heart sounds. This evaluation involved employing four established standard threshold selection rules: heursure, rigrsure, minimaxi, and sqtwolog.

- Rigrsure: This method determines the threshold by employing Stein's unbiased risk estimate (SURE) quadrature loss function.
- Sqtwolog: The threshold here is established at a level that yields minimax performance, multiplied by a small factor proportional to the logarithm of the signal's length (s). usually, $\sqrt{2log(length(s))}$.
- Heursure: This technique selects the threshold using a combination of the first two methods.
- Minimax: With this approach, a fixed threshold is chosen to achieve minimax performance concerning mean square error compared to an ideal procedure. These methods are all available within the MATLAB software toolbox.

In our research, we implemented the soft thresholding method utilizing the rigrsure threshold selection rule for denoising heart sound signals. Following decomposition and thresholding, the ultimate phase involves reconstructing the original signal. This step yields the noise-free heart sound signal or the denoised heart sound signal by executing the inverse discrete wavelet transform (IDWT). The most effective method to gauge the impact of added noise on a heart sound signal involves the addition of white Gaussian noise. When evaluating signal denoising, the standard approach to showcasing the denoising algorithm's efficacy is by computing

Signal-to-Noise Ratio (SNR) values. These values were calculated using the mathematical equation outlined in Equation (4.1) [58].

$$SNR_{db} = 10log_{10}\left(\frac{Xsignal}{Xnoise}\right)^2 \quad \text{.....Equation 4.1}$$

In this context, $Xsignal$ represents the root mean square amplitude of the signal, while $Xnoise$ represents the root mean square amplitude of the noise.

4.6.Feature extraction

Feature extraction stands as a cornerstone in machine learning, aiming to identify discriminative parameters, amalgamated into what's known as a feature vector [59]. In this research, we employed a feature fusion approach, features extracted from time domain, frequency domain, statistical domain, and MFCC (Mel Frequency Cepstral Coefficients) analysis methods to extract various features. The discrete wavelet transform enables the capture of both time and frequency characteristics, while MFCC, a robust method in speech and audio processing, it captures perceptual information of the sound. By combining these distinct feature extraction techniques, we aimed to enhance the discriminative power of the feature vectors, crucial in classifying different subjects heart sound classes or within machine learning frameworks.

4.6.1 DWT based feature extraction

The DWT feature extraction technique proves exceptionally beneficial in extracting wavelet features from non-stationary heart sound signals. Consequently, when dealing with these non-stationary heart sound signals, it becomes crucial to employ a suitable method for analysis and feature extraction. Hence, the selection of the DWT feature extraction technique serves the purpose effectively in this context. Primarily, the DWT-based feature extraction process comprises two fundamental steps. Initially, the given signal is decomposed into N levels using filtering and decimation, yielding both approximation and detailed coefficients. This decomposition involves a sequence of high pass and low pass filtering operations. Subsequently, the signal undergoes reconstruction, achieved through the inverse discrete wavelet transform (IDWT). The features derived from the IDWT of signals are esteemed as valuable inputs for classifiers due to their effective representation of non-stationary signals in the time-frequency domain.

The DWT feature extraction algorithm involves a sequence of steps, outlined as follows:

Step 1: Decomposition: Heart sound signals are decomposed into six detail sub-bands using DWT at level $N = 5$, resulting in high-frequency (detail) and low-frequency (approximation) coefficients.

Step 2: Sub-band Analysis: The approximation coefficients undergo further decomposition to extract localized information using a fifth-level db10 wavelet decomposition in this study.

Step 3 : Coefficient Processing: All fifth-level detail band coefficients are collected for subsequent analysis.

Step 4 : Frequency Vector Extraction: The frequency vectors (in radians/sample) for the fifth detail sub-bands are extracted using the MATLAB periodogram function.

Step 5 : Signal Reconstruction: Signals are reconstructed using the Inverse Discrete Wavelet Transform (IDWT) after decomposition.

Step 6 : Feature Computation: Features are computed from the reconstructed signals either by utilizing specific syntax or implementing formulated procedures.

Step 7: Feature Tabulation: Extracted features from all heart sound classes are tabulated into a feature table for classification purposes.

A variety of features were retrieved from the heart sound signals to be fused with features extracted by MFCC, following the steps shown above. Below is a detailed explanation of these retrieved traits, which are specifically meant to be used for subject identification through heart sound analysis.

Entropy : Entropy is a metric for signal complexity, irregularity, and randomness. The entropy employed in this investigation is derived from Shannon and spectral entropies [60], which may be computed analytically as follows:

$Entropy = -\sum_{i=1}^{N}(pi \log_2 pi)$ …… Equation (4.2)

where, P = The probability distribution of the signal & N = Total number of distinct classes present in that distribution

Kurtosis : Kurtosis represents a statistical time domain property of a signal that is of higher order. It can also be understood as the duration of a signal distribution [61], which indicates how prone the signal is to outliers. Mathematically, kurtosis can be determined using Equation (4.3)

$\sum_{i\equiv N}^{N} = \left(\frac{(x(i)-m)}{s}\right)^4 . \frac{1}{N}$…….. Equation(4.3)

X(*i*) is an individual score; "*m*" is the population mean; "*s*" is the population standard deviation; *N* is the population size.

Skewness: represents a significant higher-order statistical feature in the time domain of a signal, denoting the asymmetry within its distribution. Essentially, it quantifies the degree of asymmetry present in the data. Mathematically, one can compute skewness using Equation 4.4.

$$x_{skew} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_{mean})^3}{\left(\frac{1}{N}\sum_{i=1}^{N}(x_i - x_{mean})^2\right)^{\frac{3}{2}}} \quad \text{……Equation (4.4)}$$

Mean : The mean stands as a fundamental statistical time domain feature within a signal. It signifies the average value encapsulated by the signal's data. Computationally, it is derived by summing all samples and dividing by the total number of samples (*N*). Mathematically, the arithmetic mean of a signal can be obtained using Equation (4.5).

$$x_{mean} = \frac{\sum_{i=1}^{N} x_i}{N} \quad \text{……Equation(4.5)}$$

Root mean square(RMS) : The RMS value of a signal can be determined by squaring each individual value, computing the arithmetic mean of these squared values, and finally, taking the square root of the outcome. Essentially, the RMS signifies the signal's average power, expressed mathematically through Equation (4.6).

$$X_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \quad \text{.....Equation(4.6)}$$

Standard deviation: It is defined as the amount of variation in the set of data values

Peak amplitude : refers to the highest positive or negative deviation of a sinusoidal waveform from its zero-reference level. In symmetrical sinusoidal waves, the positive peak aligns with the negative peak. Typically, peak amplitude serves as a spectral feature derived from these peak amplitudes, signifying the maximum value within the signal's amplitudes.

Variance: a large variance denotes that the data points are widely dispersed about the mean and away from one another, whereas a small variance suggests that the data points tend to be relatively close to the mean and thus to each other.

Mid-frequency : refers to the frequency value that is reached at the greatest value of the power spectral density [36].

Dynamic range : The dynamic range of a waveform is the ratio of its maximum value to its minimum value. Max is the signal's maximum value.

THD: Harmonic distortion measurement assesses the presence of distortion in a signal. It is quantified as the ratio of the combined powers of harmonic components to the power of the fundamental frequency.

Maximum frequency: It represents the highest frequency value within the spectrum's energy distribution.

Average Frequency: $$f_{avg} = \frac{\sum_{i=1}^{n} f_i x p_i}{\sum_{i=1}^{n} p_i} \quad \text{. . . . . Equation(4.7)}$$

Here, p denotes power spectral density, and f represents the frequency vector.

4.6.2 MFCC feature extraction

For this study we used the standard mfcc computation. MFCCs, short for Mel-Frequency Cepstrum Coefficients, stand out as one of the most prevalent parametric representations of audio signals, as highlighted by (Davis and Mermelstein).Heart sound is an acoustic signal, and many methods used in human recognition tasks today are derived from speech recognition techniques. Mel Frequency Cepstral Coefficients (MFCC), which map the signal onto a Mel-Scale that is non-linear and resembles human hearing, are the most effective and widely used option for feature extraction of acoustic signals [62].

The concept of employing Mel Frequency Cepstral Coefficients (MFCC) as the feature set for a PCG biometric system stems from the triumph of MFCC in speaker identification tasks [63]. and due to the fact that both speech and PCG are auditory signals. the MFCC is founded on the fact that human hearing is limited to frequencies below 1 kHz. Stated differently, the MFCC relies on the established fluctuation of the critical bandwidth of the human ear with frequency. The two filter types used in MFCC are logarithmic in spacing above 1000 Hz and linear in spacing at low frequencies below 1000 Hz [13]. The MFCC feature extraction algorithm involves a sequence of steps, outlined as follows:

Pre _ emphasis : To enhance the high-frequency components of an audio signal, a pre-emphasis operation employs a Finite Impulse Response (FIR) filter. The pre-emphasis operation applied to the audio signal s[n], in the time domain can be represented as in equation(4.8):

$S_0(n) = S(n) - \alpha S(n-1)$, $0.9 \leq \alpha \leq 1$ ……...Equation(4.8)

Where α represents the pre_emphasis filter constant

Framing : framing is a fundamental step involving the segmentation of the audio signal into short, overlapping frames. This technique ensures that the signal is analyzed in smaller, more manageable portions, typically around 20-40 milliseconds each**.** By overlapping these frames, usually by 50% or less, it helps capture transient information that might get lost when considering isolated segments[66]. In this research the signals underwent resampling to 2000 Hz and were segmented into 25ms intervals. Consequently, the frame size for MFCC extraction, in samples, equates to 0.025 × 2000 = 50. The frame step remained constant at 10ms (0.010 × 2000 = 20 samples), resulting in a frame rate of 2000/20 = 100 frames per second.

Windowing : In the MFCC feature extraction process, windowing is a critical step applied to each framed segment of the signal. This technique involves multiplying the audio samples within these segments by a window function, such as the Hamming, Hanning, Triangular or Rectangular window, to taper the signal at both ends. Windowing minimizes spectral leakage that occurs when abruptly cutting segments of the signal, ensuring a smoother transition between frames. By attenuating signal discontinuities, windowing reduces artifacts in the Fourier Transform analysis, allowing for a more accurate representation of the signal's frequency content within each frame. This procedure enhances the reliability of MFCC computation by mitigating the impact of spectral distortion caused by abrupt segment boundaries, contributing to a more precise characterization of the signal's spectral features. Typically, the Hamming window is used. In mathematical terms, the Hamming window is expressed as [67].

$w(i) = 0.54 - 0.46\cos\left(\frac{2\pi i}{n}\right), \quad 0 \leq i \leq n$ .....Equation (4.9)

FFT : Within MFCC, after pre-emphasis to balance the spectrum and framing the signal into smaller segments, FFT is applied to each frame to obtain the power spectrum. The Fast Fourier Transform (FFT) plays a pivotal role in converting a signal from the time domain into the frequency domain. The FFT algorithm efficiently computes the Discrete Fourier Transform (DFT), breaking down the signal into its constituent frequencies.

The Fast Fourier Transform (FFT) is an optimized algorithm employed for computing the Discrete Fourier Transform (DFT). It excels in performing calculations with reduced steps while maintaining a commendable signal representation comparable to that achieved through the conventional DFT transformation [67].

Mel Filter Bank : is designed to mimic human auditory perception by dividing the audio spectrum into specific frequency bands [68]. It consists of a set of triangular filters spaced evenly in the Mel-scale, which is a perceptually relevant frequency scale. These filters overlap and cover the entire frequency range of interest, typically spanning from the lower to higher frequencies. Each filter's triangular shape allows it to capture varying levels of energy within specific frequency ranges, providing a robust representation of the spectral characteristics present in the audio signal. The output of the Mel filter bank serves as the basis for further feature extraction steps in MFCC computation, facilitating the extraction of essential information while reducing the dimensionality of the signal for efficient processing in various audio analysis tasks. The correlation between the tangible frequency, f_h in Hertz, and the perceived frequency, f_{mel} , measured in mel-scale, is expressed as in equation (3.10).

$Fmel = 2595\log_{10}\left(1 + \frac{Fh}{700}\right)$ . . . . . Equation(4.10)

DCT : This method utilizes the Discrete Cosine Transform (DCT) to shift from the frequency domain to the time domain. The resulting outcome comprises Mel Frequency Cepstrum Coefficients (MFCC). With cepstral analysis, the local spectral characteristics of the signal are effectively captured alongside frame analysis**.** [65]. The Discrete Cosine Transform (DCT) is used in this operation to translate the log Mel spectrum into the time domain. The Mel spectrum coefficients are real integers, as are their logarithms [64]. The formula for the MFCC parameters is expressed as in equation (4.11)

$$C_{J} = \sum_{i=1}^{m} X_i \cos\left(j\left(i-\frac{1}{2}\right).\frac{\pi}{M}\right) \quad \text{Equation (4.11)}$$

In this context, M represents the quantity of filters in the filter bank, J denotes the count of computed cepstral coefficients, and xi is defined as the log-energy output of the i^{th} filter. During the training phase, these feature vectors can be utilized to create a reference model, and during the testing phase, they can be used to determine the subject's identity.

Delta and Energy Spectrum

Changes occur in both in the signal and the frames, such as alterations in the formant's slope during transitions. Hence, it is essential to incorporate features associated with the temporal evolution of cepstral features.

4.7. Feature Selection

The thesis employs the Relief function for feature selection, a filter-method approach within the feature selection algorithm pioneered by Kira and Rendell back in 1992 [69]. The initial design targeted its application towards binary classification challenges encompassing discrete or numerical features. Relief computes individual feature scores, enabling the selection of top-ranking features. Its assessment hinges on detecting variations in feature values among nearest-neighbor instance pairs. Ultimately, the Relief function furnishes both the ranking and weights of predictors within the input data matrix [69]. The choice of the Relief algorithm for this study stems from its implement ability, effectiveness, and widely acknowledged approach to feature ranking. Notably, the Relief algorithm's versatility extends beyond class-related problems [71,72]

4.8.Feature Normalization

This step aims to standardize the range or distribution of feature values . It is particularly important when dealing with features that have different scales or units [72]. Normalization techniques, such as min-max scaling or z-score normalization, transform the values of features to a common scale, often between 0 and 1 or with a mean of 0 and a standard deviation of one [72]This process ensures that each feature contributes proportionally to the learning process, preventing features with larger magnitudes from dominating the model's training. Normalization facilitates efficient convergence during training and improves the stability and generalization of the model. Before training the classifier, we standardized the features to ensure a uniform scale. Following that, we proceeded with data splitting.

4.9. Model Training

The prepared dataset underwent fitting with, K-Nearest Neighbor (KNN), Ensemble methods, and Support Vector Machines (SVM). Every model underwent training and testing with the identical dataset. During the model training phase, specific parameters like kernel type, number of neighbors, and criterion for splitting nodes were adjusted and implemented based on the model type. Additionally, a one-versus-rest (OVR) or One-vs-All classifier was employed, meaning the method splits a multi-class dataset into multiple binary classification problems. Then a binary classifier is trained on each binary classification problem so that predictions are made using the most confident model. [73]. For training the models data set were split into 70% training set and 30% of testing set. Moreover, to guarantee dependable and impartial assessments, each model utilized a 10-fold cross-validation technique. This approach entailed dividing the data into ten subsets, training the models on nine of these subsets, and validating their performance on the tenth subset. This process was iterated ten times, ensuring a robust performance estimate. The models, K-nearest Neighbor(KNN), It categorizes the object by evaluating the distance between the new object and the predefined objects.

The assignment of the object to a specific class, denoted as class k, is determined by identifying the nearest neighbor, i.e., the class k with the shortest distance to the new object [74]. Ensemble classifiers, Ensemble classifiers amalgamate outcomes from numerous weak learners to create a unified, high-quality ensemble model. In the initial run of results, it was observed that the Cubic SVM, particularly, emerged as the most effective model, showcasing superior classification accuracy.

4.9.1. Support vector machine

Support Vector Machine (SVM) stands as a widely adopted supervised machine learning technique proficient in handling both classification and regression tasks, leveraging the kernel method [75]. As illustrated in Figure 4.5 (a), numerous separating hyperplanes might exist. SVM addresses this by prioritizing the smallest distance between any data sample and the decision boundary [76].

The optimal hyperplane for an SVM refers to the one that possesses the widest margin between the two classes. This margin represents the maximal width of the space parallel to the hyperplane, devoid of any internal support vectors. Support vectors, the data points closest to the separating hyperplane, reside on the boundary of this space, as demonstrated in Figure 4.5. This figure provides an illustrative overview of the SVM algorithm.

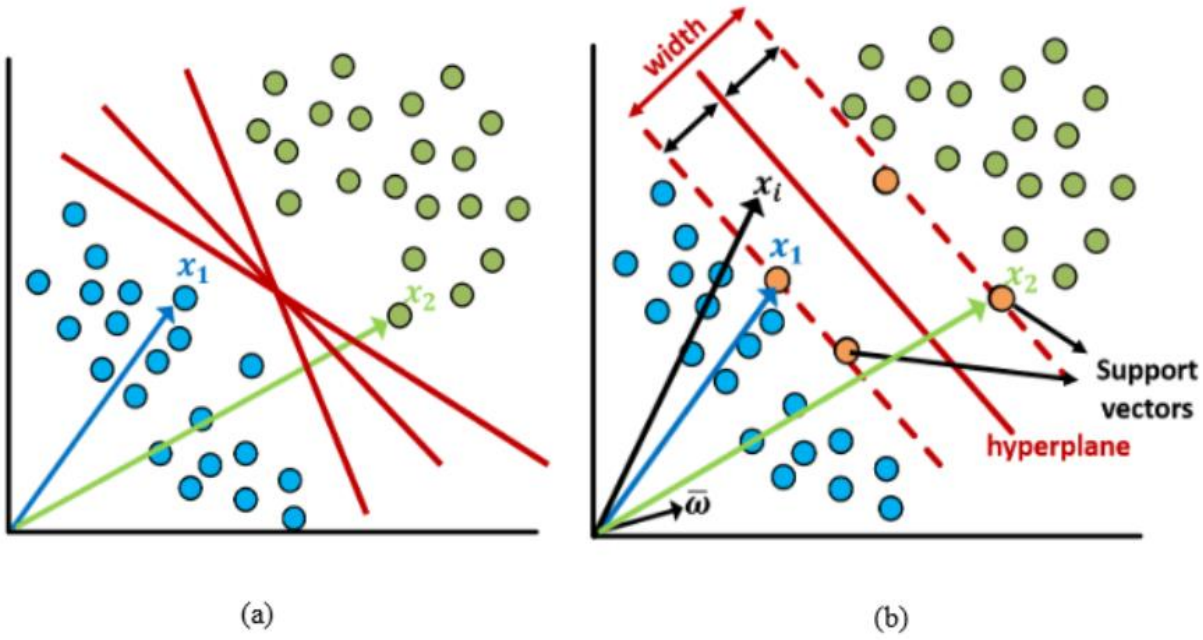


Figure 4.5 : An overview sketch of SVM algorithm [76].

4.10. Model optimization

Optimization involves identifying a set of inputs for an objective function that leads to either maximum or minimum function evaluations. This process specifically revolves around adjusting hyperparameters to minimize the cost function, employing various optimization techniques to achieve this objective Several common strategies for hyperparameter tuning in machine learning encompass grid search, random search, and Bayesian optimization [78].

Bayesian Optimization (BO) refers to a sequential method employed for optimizing the parameters of a surrogate function, f(x), which is a representative of an objective function. Notably, the surrogate function is easy to calculate than the objective function. In our case, say, we would like to optimize the hyper parameters based on 'accuracy' of the model (objective function). So, we use, $\mathrm{P}(\mathrm{A} = '\mathrm{hyperparameters}'|\mathrm{B} =' \mathrm{accuracy}')$, as the surrogate function or posterior beliefs. The probability of prior beliefs, $\mathrm{P}(\mathrm{B} =' \mathrm{accuracy}'|\mathrm{A} = '\mathrm{hyperparameters}')$, and the marginal probability, $\mathrm{P}(\mathrm{A} = '\mathrm{hyperparameters}')$ are assumed to come from different probability distributions (E.g Multivariate Gaussian distribution). By sampling from the assumed distributions of $\mathrm{P}(\mathrm{B}|\mathrm{A})$ and $\mathrm{P}(\mathrm{A})$, and optimizing or maximizing the posterior beliefs, iteratively, we arrive at an optimized combination of 'hyperparameter'. This process is iterated until a stopping criterion is achieved for the objective function, i.e., 'accuracy' [79]. The foundation for the optimization process, reliant on Bayes' Theorem, is encapsulated in Equation (4.12), which relates to a model A and observation B [28].

$$P(A/B) = \frac{\left(P(B|A)P(A)\right)}{P(B)} \quad \text{Equation 4.12}$$

In this context, P(A|B) signifies the likelihood of A given B, P(B|A) represents the likelihood of B given A, P(A) denotes the prior probability of A, and P(B) signifies the marginal probability of B. Bayesian Optimization serves the purpose of identifying the minimum or maximum of a function within a constrained set [80]. This study utilized the classification learner app to automatically construct an optimized model, employing the Bayesian optimization scheme. Consequently, the Bayesian optimization technique was employed in this research to enhance the accuracy of the selected model by optimizing the hyper parameters of SVM .

4.11. Performance Evaluation

The selected model's performance was assessed using a confusion matrix, evaluating metrics like accuracy, precision, recall, and F1 score via equations (4.11, 4.12, 4.13, 4.14) [[81]]. Accuracy assesses the overall model performance by determining the percentage of correctly classified samples. Precision gauges the percentage of predicted positive class instances that actually belong to that class. Conversely, recall identifies the percentage of positive samples that the classifier correctly identifies as positive. The f1-score amalgamates precision and recall into a single metric.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad \text{..... Equation (4.11)}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{.....Equation (4.12)}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{..... Equation (4.13)}$$

$$F1\ score = \frac{2*(precision*recall)}{precision+recall} \quad \text{..... Equation (4.14)}$$

4.12. Materials

Various software and hardware resources were instrumental in the successful execution of this research. Software tools played a vital role in solving mathematical computations, code and file management, data processing, and signal analysis. MATLAB (short for "matrix laboratory") and Audacity were two pivotal software utilized in this study. Specifically, MATLAB R2021a facilitated the tasks mentioned above. Furthermore, the most recent version of Audacity software (Audacity 3.2.5) was utilized. Moreover, to assemble the electronic stethoscope, hardware components including the head of an acoustic stethoscope, a condenser microphone, an audio cable, and a computer equipped with Audacity software were employed.

CHAPTER FIVE

RESULTS and DISCUSSION

5.1. Results

5.1.1. Building an Electronic Stethoscope

In this research, we developed an electronic stethoscope for heart sound acquisition through the modification of a standard acoustic stethoscope. The construction of this stethoscope has been done by utilizing accessible materials such as the head of an acoustic stethoscope, a condenser microphone, an audio cable, and a personal computer equipped with Audacity software. A modified device was developed. the integration of these components resulted in an electronic stethoscope poised to revolutionize cardiac data acquisition. Figure 5.1 shows the materials used for the development of an electronic stethoscope.

Figure 5.1 : Materials utilized in crafting the electronic stethoscope

The image depicted in Figure 5.2 displays the head of an acoustic stethoscope, designed for precise placement on the selected auscultation site, and it contains the diaphragm and bell.

Figure 5.2 : an image showing head of an acoustic stethoscope with diaphragm and bell

To establish the connection between the head of an acoustic stethoscope and the condenser microphone, we utilized the readily available acoustic stethoscope tubing. Specifically, we tailored the tubing by cutting it into a smaller size suitable for the connection purpose. In Figure 5.3, the illustration displays the connection of one end of the stethoscope tubing to the stethoscope's head.

Figure 5.3 : Stethoscope tubing connected to the head of stethoscope

The heart of this setup lies in the condenser electret microphone, designed explicitly to translate acoustic heart sounds into electrical signals. This pivotal component serves as the bridge, converting the physiological sound into an electrical counterpart. Strategically connected at one end to the head of the acoustic stethoscope and at the other to an audio cable, its purpose extends to facilitating signal transfer to a computer. This crucial connection is carefully shielded by tubing, ensuring the fidelity and integrity of the electrical signal during transmission. Such a configuration exemplifies a meticulous integration of components, enabling the conversion and preservation of

heart sounds for comprehensive analysis and interpretation on a digital platform. Figure 5.4 shows a condenser microphone, and an audio cable separately, while In Figure 5.5, there is an illustration depicting the connection between a condenser microphone sensor and an audio cable.

Figure 5.4 : Image of a condenser microphone, and audio cable

Figure 5.5: Illustration of the condenser microphone connected to an audio cable.

Following the completion of the electronic stethoscope assembly, the acquired heart sound signals are processed using Audacity software. Audacity, a versatile and user-friendly audio editing program, plays a pivotal role in visualizing and analyzing the captured cardiac signals. Leveraging its diverse functionalities, this software allows for the precise examination and manipulation of the recorded heart sounds. With its interface tailored for waveform visualization and spectral analysis,

Audacity enables researchers to delve into the intricacies of the obtained signals. This digital platform proves invaluable, providing tools for enhanced signal interpretation and aiding in the extraction of meaningful insights from the acquired cardiac data. For this research we have used an audacity software version 3.2. Figure 5.6, shows the completed electronic stethoscope depicted, fully prepared for recording purposes.

Figure 5.6: Illustration of the completed construction of the electronic stethoscope.

The constructed electronic stethoscope represents a significant leap in cardiac signal acquisition, offering advantages comparable to commercially available counterparts. Its tailored design, crafted by integrating an acoustic stethoscope head with a condenser microphone and audio cable connected to a computer with Audacity software, provides a cost-effective alternative. This innovation allows for efficient recording and analysis of heart sounds, akin to high-end electronic stethoscopes. While maintaining similar functionality, the advantage lies in its adaptability and affordability, making advanced cardiac signal acquisition accessible in research settings with constrained resources. Its performance aligns closely with commercial electronic stethoscopes, showcasing potential for widespread adoption in clinical and research environments seeking cost-effective yet reliable tools for cardiac assessment and analysis.

5.1.2. Acquisition of heart sound

In this research study, data from a pool of 20 healthy subjects was gathered internally. Sequential recordings of heart sounds were obtained from each participant while they were in a seated position as shown in figure 5.7. recording in a seated position minimizes potential confounding factors such as body movement or posture changes that may occur when recording in other positions, thereby enhancing the reliability and consistency of the recorded data. The data collection process was conducted on various days to ensure diversity and consistency.

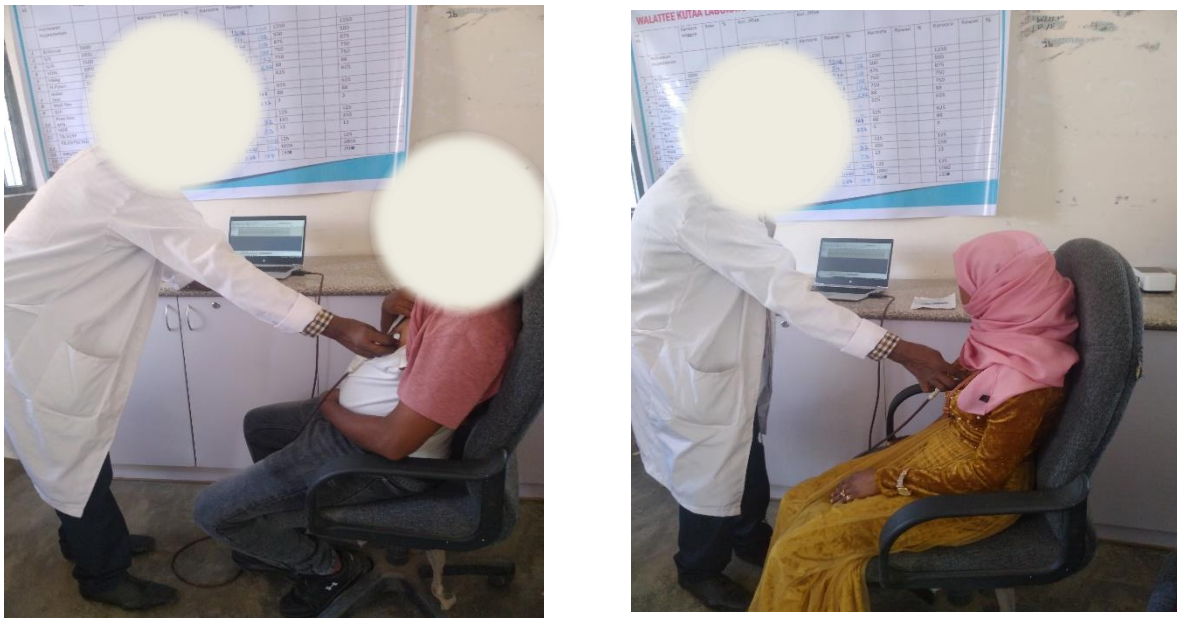


Figure 5.7: Recording heart sound signal from subjects

5.1.3. Data Augmentation

In this study, a dataset comprising 560 in-house heart sound recordings was utilized, employing augmentation techniques to expand the dataset for model training purposes. The augmentation process involved selecting a subset of the original signals, and applying amplitude scaling. This scaling was achieved using an amplitude factor range of [0.8, 1.2]. Consequently, although not all 560 signals were augmented, the total number of signals after augmentation reached 1360. The augmented signals are used only for training the model, These augmented signals are pivotal for enhancing the robustness and diversity of the dataset, contributing significantly to the training of the model. The amplitude scaling technique employed within the specified range ensured variability in signal amplitudes, potentially enriching the dataset's representation of real-world scenarios and aiding the model in learning diverse features. A sample of augmented signals are shown in figure 5.8 below.

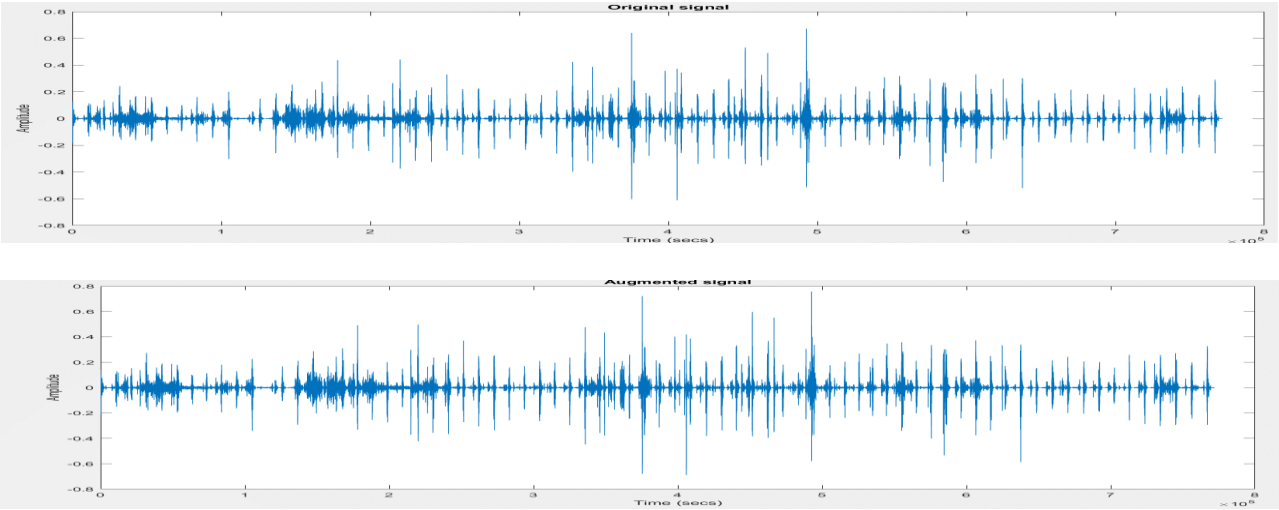


Figure 5.8: Comparison between the original signal and its augmented counterpart

5.1.4 Preprocessing

Using a discrete wavelet transform, the recorded Phonocardiogram (PCG) signals were first denoised before being subjected to feature extraction. Three basic steps made up this method: decomposition, thresholding of the detail coefficient, and reconstruction afterward. The first step in the DWT-based denoising process, signal decomposition, was crucial to this process. Figure 5.9 provides an illustrative representation of the wavelet decomposition process applied to heart sound signals.

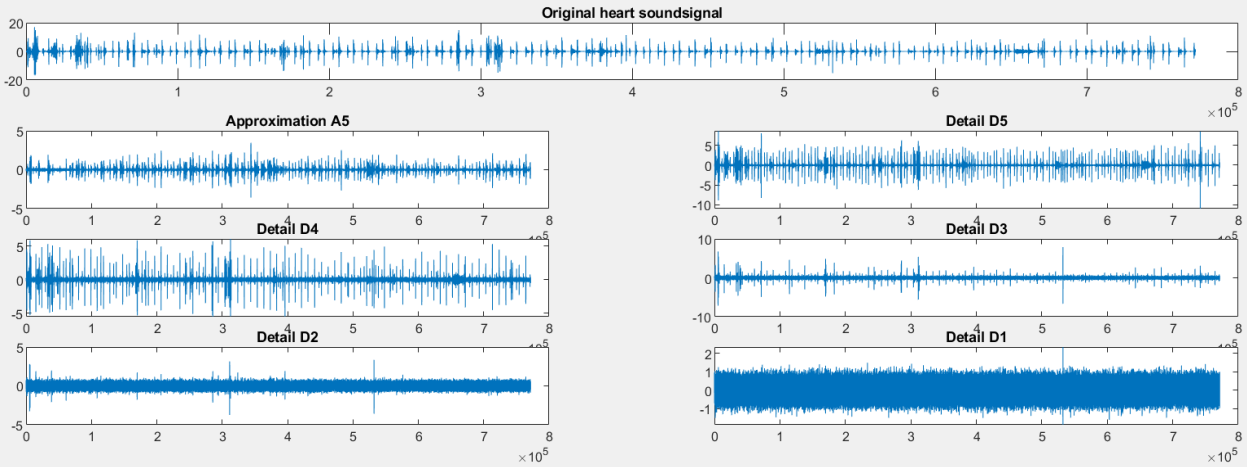


Figure 5.9 wavelet decomposition of heart sound signals using db10 at 5th level

After the decomposition process, denoising was executed using the db10 wavelet function, employing the soft thresholding method alongside the rigrsure threshold selection rule.

Additionally, the impact of employing two thresholding techniques (soft and hard thresholding) was investigated by assessing the Signal to Noise Ratio (SNR) outcomes while denoising various heart sound signals with different wavelet functions. The SNR results achieved during the denoising of heart sound signals using sym6, Db4, Db10, and sym5 wavelet functions at various decomposition levels with both thresholding methods are presented in Table 5.1.

Table 5.1: SNR Results for Denoising heart Sound Signals Using Various Wavelet Functions at Different Decomposition Levels with Soft and Hard Thresholding Methods (Applying Rigrsure Threshold Selection Rule)

SNR(dB)								
Wavelet function	**Level4**		**Level5**		**Level6**		**Level7**	
	Soft	**Hard**	**Soft**	**Hard**	**Soft**	**Hard**	**Soft**	**Hard**
Db4	14.0805	13.988	15.0977	13.0532	10.8104	11.0801	9.8765	9.7663
Sym6	13.0704	13.766	14.9879	14.0608	11.0405	11.0405	9.0675	9.0043
Db10	14.0898	14.0278	**15.9874**	**15.4304**	12.0043	12.0043	10.5343	10.0089
Sym5	13.0757	13.0459	14.0075	13.9885	11.9898	11.9898	9.8753	8.9877

Table 5.1 presents the Signal-to-Noise Ratio (SNR) outcomes for various wavelet functions across different decomposition levels ranging from the 4^{th} to the 7^{th} level. Each SNR value was computed using the Rigrsure threshold selection rule and both thresholding methods (soft and hard). Notably, the most favorable SNR results were observed specifically at the 5th level of decomposition using the Db10 wavelet function with soft thresholding. Two crucial factors impacting the SNR value are the level of decomposition and the type of thresholding used.

Consequently, for the preprocessing of heart sound signals in this research, the Db10 wavelet function at the 5th level with soft thresholding was chosen. In conclusion, the impact of four threshold selection rule. rigrsure, sqtwolog, Heursure, and min max was examined while denoising heart sound signals with the chosen Db10 wavelet function. Among these rules, the Rigrsure threshold selection displayed the highest performance in denoising the heart sound signals from various healthy subjects in this research. To illustrate the clarity achieved after implementing the proposed denoising algorithm, Figure 5.10 displays spectrograms comparing the noisy heart sound signals with their denoised counterparts. This visual representation showcases the enhanced quality of the heart sound segments. The figure presents spectrograms for the noisy and denoised PCG signals to show the clarity of the heart sound components obtained after applying the proposed denoising algorithm. In the denoised PCG signal spectrogram, the heart sounds are clear.

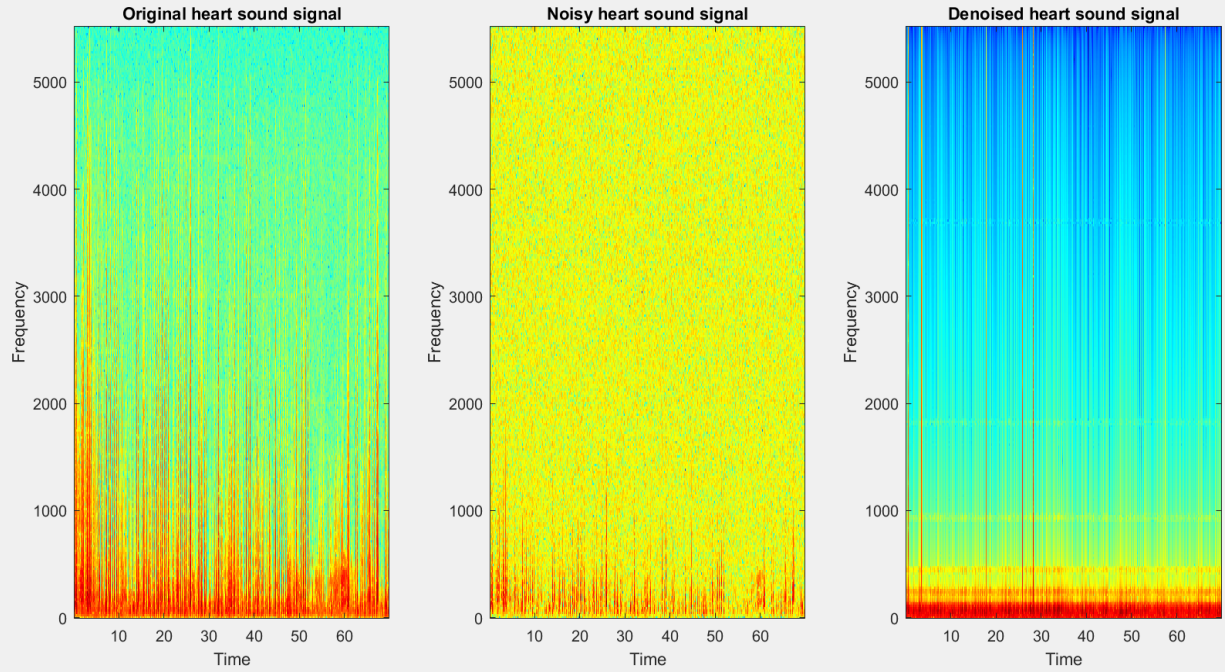


Figure 5.10 : Spectrograms of Heart Sound Signals

Upon implementing the Db10 wavelet function at the 5th level alongside the soft thresholding method using the Rigrsure threshold selection rule, the spectrogram of the denoised heart sound signals reveals a notable clarity in the heart sound segments.

5.1.5. Feature Extraction

5.1.5.1 Features extracted by DWT based feature extraction procedure

Various features were extracted across time, frequency, and statistical domains, employing multiple MATLAB built-in functions and formulas for computation. These extracted features encompass a range including entropy, root mean square, mean, standard deviation, kurtosis, skewness, variance, among others. Table 5.2 illustrates a sample of these features extracted in the study, displaying numerical values representing the extracted features.

Table 5.2: A sample of extracted features by DWT based feature extraction

	A	B	C	D	E	F	G	H	I	J	K	L	M
	Entropy	Kurtosis	Skewness	Mean	RMS	STD	Peak.Amp	Variance	Mid.Freq	DR	THD	Max.Freq	Avg.Freq
	1650.531	0.460702	1.002625	9.131024	-4.11493	219.8667	0.501343	0.05233	0.228757	0.000341	0.228758	0.05233	0.000566
	1443.802	0.403074	0.877045	7.987328	-4.11491	219.8666	0.438568	0.040042	0.200105	0.000283	0.200106	0.040042	0.00057
	1643.756	0.458608	0.998505	9.093577	-4.11494	219.8675	0.499298	0.051901	0.227818	0.000324	0.227819	0.051902	0.000565
	1614.609	0.450664	0.980804	8.932286	-4.11493	219.866	0.490448	0.050077	0.223778	0.000318	0.22378	0.050077	0.000565
	1522.083	0.286815	1.004273	9.111896	-12.5401	234.7642	0.502777	0.050124	0.223884	-0.00118	0.223883	0.050123	0.018165
	1616.898	0.304809	1.066834	9.679545	-12.5401	234.7658	0.534088	0.056563	0.237831	-0.00127	0.237829	0.056563	0.018166
	1239.411	0.233443	0.81778	7.419695	-12.5401	234.7641	0.409393	0.033235	0.182306	-0.00098	0.182305	0.033235	0.018163
	1734.648	0.326786	1.144501	10.38441	-12.5401	234.7646	0.572968	0.065102	0.25515	-0.00136	0.255148	0.065101	0.01816
0	1575.871	0.26646	1.003845	8.999383	-12.6839	231.6838	0.502411	0.051068	0.225981	-0.00121	0.22598	0.051067	0.007454
1	1435.263	0.242823	0.914246	8.196381	-12.684	231.6846	0.457581	0.042361	0.205818	-0.00112	0.205816	0.04236	0.007464
2	1289.793	0.218229	0.821625	7.36573	-12.684	231.6858	0.411224	0.034209	0.184958	-0.001	0.184957	0.034209	0.007453
3	1321.917	0.223735	0.842102	7.549187	-12.6839	231.6855	0.421478	0.035935	0.189565	-0.00103	0.189563	0.035934	0.007455
4	1422.139	0.284089	1.004364	8.713401	-11.257	229.5174	0.502533	0.045315	0.212874	-0.00116	0.212872	0.045314	-0.02069
5	1354.75	0.270666	0.956757	8.300584	-11.257	229.5186	0.478699	0.041123	0.202787	-0.00112	0.202785	0.041122	-0.02069
6	1573.182	0.314519	1.111023	9.638854	-11.257	229.5175	0.555878	0.055452	0.235482	-0.00129	0.23548	0.055451	-0.0207
7	1590.058	0.317634	1.122955	9.74229	-11.257	229.5179	0.561859	0.056648	0.238009	-0.00131	0.238007	0.056648	-0.02069
8	1371.547	0.522662	1.003326	10.19135	-12.5809	271.7777	0.502014	0.05105	0.225942	-0.00024	0.225944	0.051051	0.007583
9	1191.156	0.453593	0.871368	8.850926	-12.5808	271.7768	0.436005	0.038504	0.196226	-0.00023	0.196227	0.038505	0.007584
0	1484.593	0.565426	1.085999	11.03133	-12.5809	271.7764	0.543396	0.059812	0.244565	-0.00028	0.244567	0.059813	0.007582
1	1114.703	0.424618	0.81543	8.282853	-12.5809	271.7775	0.40802	0.03372	0.183631	-0.00021	0.183632	0.033721	0.007582
2	1633.238	0.590994	1.003845	9.423335	-11.4469	243.2757	0.502411	0.044249	0.210354	0.000439	0.210355	0.044249	0.027512

5.1.5.2 Features extracted by MFCC feature extraction procedure

In this research, a comprehensive feature extraction process was undertaken, specifically focusing on Mel-frequency cepstral coefficients (MFCCs). Employing the established MFCC feature extraction procedure, a set of distinct MFCC features was successfully derived. These features were obtained through a series of well-defined steps intrinsic to the MFCC extraction process. In Table 5.3, a representative sample of the extracted MFCC coefficients has been presented. This meticulous approach to feature extraction contributes significantly to the thorough analysis and understanding of the underlying characteristics within the heart sound signals, forming a vital component of this thesis report.

Table 5.3 : A sample of extracted features using MFCC feature extraction procedure

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6	MFCC7	MFCC8	MFCC9	MFCC10	MFCC11	MFCC12	MFCC13
3.086061	-5.83625	3.884995	3.604505	1.145625	1.753193	0.358776	0.673511	-0.3786	0.294273	-0.61684	-0.06413	-0.42336
3.086062	-6.27342	3.642049	3.380297	1.081147	1.649638	0.346016	0.640314	-0.34323	0.285425	-0.56588	-0.04929	-0.38502
3.085993	-5.85093	3.877377	3.597429	1.143604	1.750259	0.358224	0.672624	-0.37739	0.29373	-0.61519	-0.0636	-0.42219
3.086076	-5.91432	3.843923	3.566809	1.135321	1.735637	0.356784	0.667684	-0.3722	0.29275	-0.60842	-0.06139	-0.41719
3.225768	-5.72642	3.808561	3.572755	1.126473	1.729378	0.21471	0.678026	-0.35924	0.305676	-0.55271	-0.02432	-0.44573
3.225749	-5.48611	3.920176	3.677038	1.155451	1.777181	0.215764	0.693409	-0.3756	0.309925	-0.57558	-0.03013	-0.46493
3.225788	-6.39687	3.447486	3.235803	1.031422	1.574316	0.209743	0.626983	-0.30776	0.291612	-0.48167	-0.0065	-0.38539
3.225776	-5.17783	4.051133	3.800264	1.190019	1.832854	0.21721	0.71131	-0.39554	0.313954	-0.60153	-0.03825	-0.48775
3.175706	-5.88543	3.719929	3.59003	1.087282	1.812607	0.224156	0.733939	-0.36361	0.262981	-0.55462	-0.05031	-0.41124
3.175731	-6.19782	3.553654	3.430047	1.044196	1.735482	0.221463	0.707117	-0.3391	0.258607	-0.52092	-0.04093	-0.38409
3.175677	-6.50418	3.372899	3.256302	0.996868	1.651247	0.218125	0.678135	-0.31312	0.253187	-0.48517	-0.03014	-0.35543
3.175681	-6.43801	3.413356	3.295476	1.007714	1.670152	0.218729	0.684725	-0.31874	0.254102	-0.49307	-0.03257	-0.36196
3.547797	-5.92729	3.644016	3.415169	1.163459	1.831314	0.272557	0.670795	-0.34363	0.237389	-0.65062	-0.0821	-0.52606
3.547745	-6.0935	3.558208	3.334402	1.13916	1.789802	0.26982	0.658572	-0.33136	0.235871	-0.6306	-0.07577	-0.50926
3.547802	-5.53902	3.827958	3.587437	1.216165	1.919146	0.278204	0.697171	-0.37111	0.240998	-0.69417	-0.09566	-0.56317
3.547781	-5.49443	3.847573	3.606264	1.221979	1.927985	0.279198	0.699699	-0.37377	0.241265	-0.69943	-0.09663	-0.5675
3.245142	-4.71625	4.238493	3.562919	1.088943	1.527261	0.045588	0.742262	-0.2887	0.218801	-0.63402	-0.07263	-0.50513
3.245152	-5.36615	3.984238	3.350877	1.033291	1.443361	0.054861	0.70794	-0.25805	0.217	-0.58167	-0.05573	-0.46098
3.245165	-4.29342	4.378375	3.679486	1.119219	1.572906	0.039741	0.760264	-0.30633	0.218956	-0.66389	-0.08269	-0.53034
3.245151	-5.62992	3.865255	3.252508	1.005829	1.404261	0.058973	0.69106	-0.24424	0.215786	-0.55772	-0.04845	-0.44138
3.489903	-5.44573	3.607296	3.727772	1.163612	1.736833	0.099961	0.687602	-0.27443	0.287823	-0.51956	-0.08006	-0.59801

5.1.5.3. Features extracted by DWT and MFCC

In this study, a comprehensive feature extraction strategy was employed to gather diverse insights from heart sound signals. Leveraging the Discrete Wavelet Transform (DWT), features spanning time, frequency, and statistical domains were meticulously derived. These features were then fused with the extensive Mel-frequency cepstral coefficients (MFCCs) obtained through the established MFCC feature extraction procedure. The integration of these distinct feature sets was depicted in table 5.4, showcasing the combination of DWT-derived features with the rich MFCC feature set.

Table 5.4: A sample of features extracted by DWT and MFCC

C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Skewness	Mean	RMS	STD	Peak.Amp	Variance	Mid.Freq	DR	THD	Max.Freq	Avg.Freq	MFCC1	MFCC2	MFCC3	MFCC4	MFCC5	MFCC6
1.002625	9.131024	-4.11493	219.8667	0.501343	0.05233	0.228757	0.000341	0.228758	0.05233	0.000566	3.086061	-5.83625	3.884995	3.604505	1.145625	1.753193
0.877045	7.987328	-4.11491	219.8666	0.438568	0.040042	0.200105	0.000283	0.200106	0.040042	0.00057	3.086062	-6.27342	3.642049	3.380297	1.081147	1.649638
0.998505	9.093577	-4.11494	219.8675	0.499298	0.051901	0.227818	0.000324	0.227819	0.051902	0.000565	3.085993	-5.85093	3.877377	3.597429	1.143604	1.750259
0.980804	8.932286	-4.11493	219.866	0.490448	0.050077	0.223778	0.000318	0.22378	0.050077	0.000565	3.086076	-5.91432	3.843923	3.566809	1.135321	1.735637
1.004273	9.111896	-12.5401	234.7642	0.502777	0.050124	0.223884	-0.00118	0.223883	0.050123	0.018165	3.225768	-5.72642	3.808561	3.572755	1.126473	1.729378
1.066834	9.679545	-12.5401	234.7658	0.534088	0.056563	0.237831	-0.00127	0.237829	0.056563	0.018166	3.225749	-5.48611	3.920176	3.677038	1.155451	1.777181
0.81778	7.419695	-12.5401	234.7641	0.409393	0.033235	0.182306	-0.00098	0.182305	0.033235	0.018163	3.225788	-6.39687	3.447486	3.235803	1.031422	1.574316
1.144501	10.38441	-12.5401	234.7646	0.572968	0.065102	0.25515	-0.00136	0.255148	0.065101	0.01816	3.225776	-5.17783	4.051133	3.800264	1.190019	1.832854
1.003845	8.999383	-12.6839	231.6838	0.502411	0.051068	0.225981	-0.00121	0.22598	0.051067	0.007454	3.175706	-5.88543	3.719929	3.59003	1.087282	1.812607
0.914246	8.196381	-12.684	231.6846	0.457581	0.042361	0.205818	-0.00112	0.205816	0.04236	0.007464	3.175731	-6.19782	3.553654	3.430047	1.044196	1.735482
0.821625	7.36573	-12.684	231.6858	0.411224	0.034209	0.184958	-0.001	0.184957	0.034209	0.007453	3.175677	-6.50418	3.372899	3.256302	0.996868	1.651247
0.842102	7.549187	-12.6839	231.6855	0.421478	0.035935	0.189565	-0.00103	0.189563	0.035934	0.007455	3.175681	-6.43801	3.413356	3.295476	1.007714	1.670152
1.004364	8.713401	-11.257	229.5174	0.502533	0.045315	0.212874	-0.00116	0.212872	0.045314	-0.02069	3.547797	-5.92729	3.644016	3.415169	1.163459	1.831314
0.956757	8.300584	-11.257	229.5186	0.478699	0.041123	0.202787	-0.00112	0.202785	0.041122	-0.02069	3.547745	-6.0935	3.558208	3.334402	1.13916	1.789802
1.111023	9.638854	-11.257	229.5175	0.555878	0.055452	0.235482	-0.00129	0.23548	0.055451	-0.0207	3.547802	-5.53902	3.827958	3.587437	1.216165	1.919146
1.122955	9.74229	-11.257	229.5179	0.561859	0.056648	0.238009	-0.00131	0.238007	0.056648	-0.02069	3.547781	-5.49443	3.847573	3.606264	1.221979	1.927985
1.003326	10.19135	-12.5809	271.7777	0.502014	0.05105	0.225942	-0.00024	0.225944	0.051051	0.007583	3.245142	-4.71625	4.238493	3.562919	1.088943	1.527261
0.871368	8.850926	-12.5808	271.7768	0.436005	0.038504	0.196226	-0.00023	0.196227	0.038505	0.007584	3.245152	-5.36615	3.984238	3.350877	1.033291	1.443361
1.085999	11.03133	-12.5809	271.7764	0.543396	0.059812	0.244565	-0.00028	0.244567	0.059813	0.007582	3.245165	-4.29342	4.378375	3.679486	1.119219	1.572906
0.81543	8.282853	-12.5809	271.7775	0.40802	0.03372	0.183631	-0.00021	0.183632	0.033721	0.007582	3.245151	-5.62992	3.865255	3.252508	1.005829	1.404261
1.003845	9.423335	-11.4469	243.2757	0.502411	0.044249	0.210354	0.000439	0.210355	0.044249	0.027512	3.489903	-5.44573	3.607296	3.727772	1.163612	1.736833

5.1.6 Feature reduction and normalization

Within this study, the Relief algorithm was implemented to select the most discriminative features. This strategic selection aimed to curtail dimensionality, eliminate redundancy, and alleviate computational load. Skewness, mid-frequency, variance, and total harmonic distortion, which held the lowest ranks with nearly zero or negative weights, are excluded from consideration as they are the least prioritized features. The selected features include: maximum frequency, average frequency, dynamic range, peak amplitude, standard deviation, RMS, mean, Entropy, and the 13 MFCCs are selected. Post feature reduction, the chosen attributes must align in scale. Hence, to achieve this, feature normalization using the min-max scaling method was employed.

5.1.7. Model training and Evaluation

For training the models data set were split into 70% training set and 30% of testing set. Following the division of the dataset, training and testing were conducted using K-nearest neighbor (KNN), Ensemble, and Support Vector Machine (SVM). Furthermore, to ensure reliable and unbiased evaluations, a 10-fold cross-validation technique was employed for each model. This method involved splitting the data into ten subsets, training the models on nine subsets, and validating them on the remaining subset, repeating the process ten times to obtain robust performance estimate. The outcomes of employing these machine learning models for the ultimate classification of data are presented in Table 5.5 and Figure 5.11.

Table 5.5: Classification Report for Machine Learning Models

Classifiers	**Algorithms**	**Accuracy**
Ensemble	Boosted Trees	87.0%
	Bagged Trees	89.5%
	Subspace Discriminant	94.5%
	Cubic KNN	89.0%

KNN	Medium KNN	89.2%
SVM	**Cubic SVM**	**95.8%**
	Medium Gaussian SVM	95.0%

Table 5.5 illustrates the accuracy comparison among various machine learning classifiers employed for subject identification based on their heart sounds. The SVM model, specifically utilizing the Cubic SVM algorithm, achieved the highest classification accuracy. Additionally, Figure 5.11 depicts the ascending order of accuracy for each classifier.

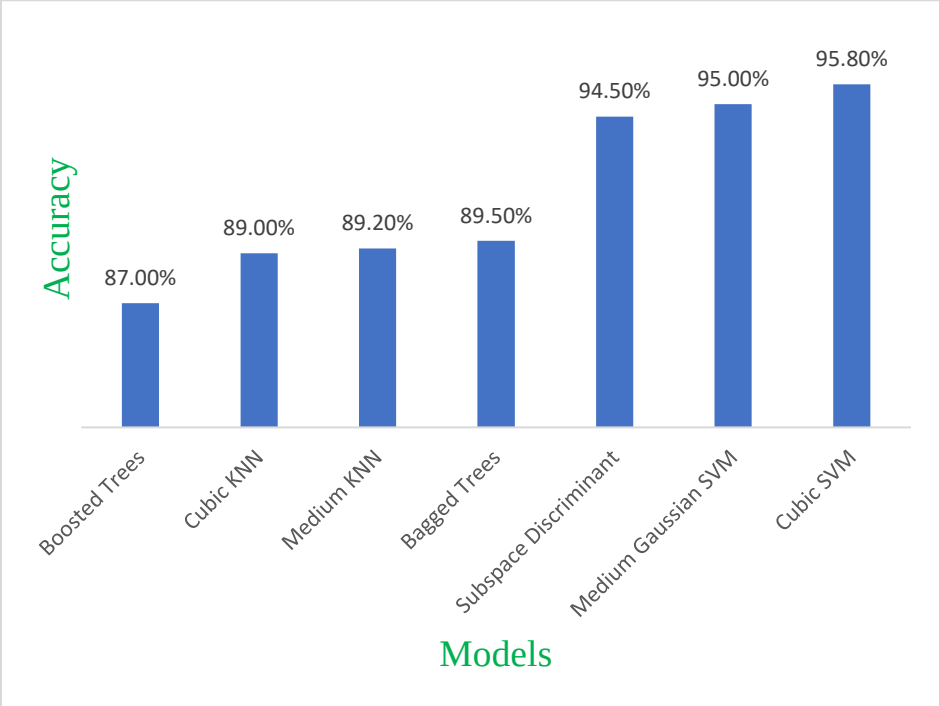


Figure 5.11: Accuracy Attained by Various Machine Learning Classifiers

The column chart, as depicted in Figure 5.11, illustrates a comparison of the accuracy attained by the three selected models, along with their respective optimal algorithms for the specific target. Notably, the highest accuracy in identification was attained through the utilization of the Cubic SVM. Specifically, the Cubic SVM stands out as the optimal model, boasting a superior classification accuracy of 95.8%. Subsequently, we evaluated the classification performance of the chosen classifier, Cubic SVM, utilizing confusion matrices.

Table 5.6: Number of Correctly and Misclassified Observations for Each Subject

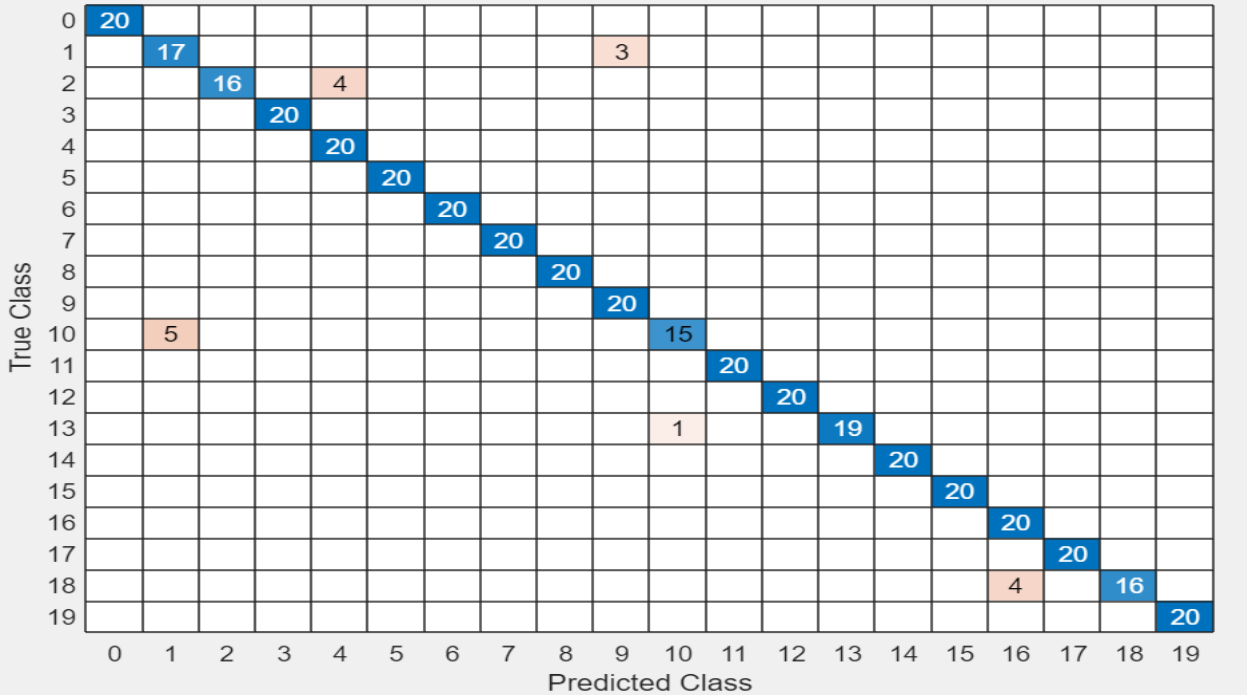
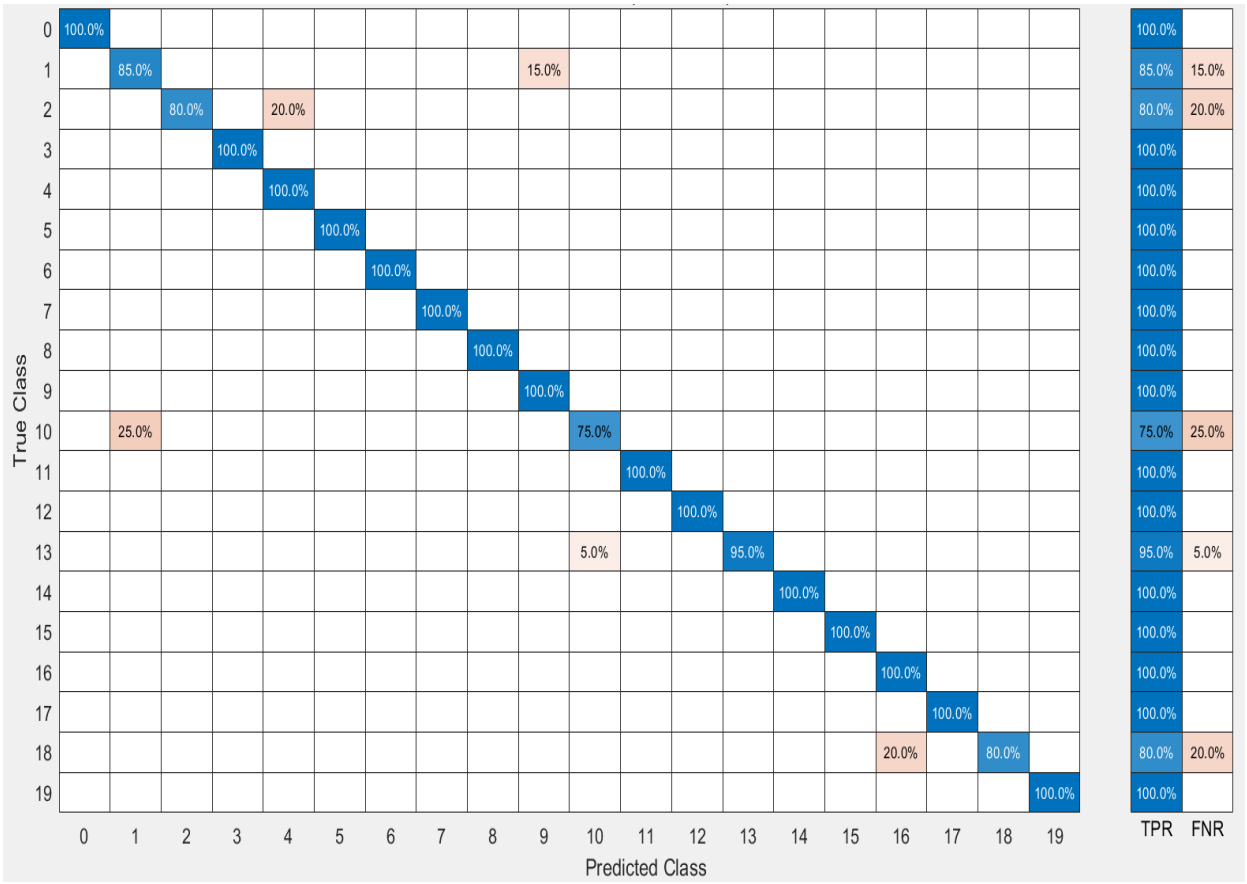


Table 5.6 displays the count of accurately and inaccurately classified observations for each subject during the testing of Cubic SVM. Each subject was designated a class ID, ranging from subject one with class ID 0, subject two with class ID 1, and so forth, extending up to subject 20 assigned to class ID 19. In the evaluation presented in the above table, 20 heart sound records obtained from different subjects underwent testing, revealing varied classification outcomes. For Subject 2, 17 records were accurately classified, while 3 records were misclassified during testing. Similarly, Subject 3's 20 heart sound records resulted in 16 correct classifications and 4 misclassifications. Subject 11's testing involved 20 heart sound records, with 15 correctly classified and 5 misclassified. In the case of Subject 14, 19 heart sound records were accurately classified, but 1 record was misclassified. Subject 19 had 16 correctly classified records, with 4 misclassified out of the 20 heart sound records tested. On the other hand, Subjects 1, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 17, 18, and 20 all demonstrated perfect classification, with all 20 heart sound recordings for each subject being correctly classified. This detailed breakdown provides a comprehensive overview of the testing outcomes for each individual subject.

Additionally, we visually examined the classifier's performance concerning true classes through the analysis of true positive rates (TPR) and false negative rates (FNR). TPR represents the ratio of correctly classified observations per true class, while FNR represents the proportion of incorrectly classified observations per true class. The performance of the classifier per true class is detailed in the confusion matrix presented in Table 5.7 Specifically, the rightmost two columns of Table 5.7 highlight the TPR and FNR for each true class. The first column corresponds to TPR, while the second column corresponds to FNR. This structured representation offers a clear insight into the TPR and FNR metrics per true class.

Table 5.7: Performance of Cubic SVM Classifier per True Class



From Table 5.7, consider the second row from the top, which encompasses all heart sound records attributed to the true class of subject 2. The column reveals predicted classes, with 85% representing the True Positive Rate (TPR) for accurately classified points in Class ID 1, denoted in the green cell within the TPR column. The remaining 15% of heart sound records in the Class ID 1 row are misclassified, contributing to a False Negative Rate (FNR) of 15%, depicted in the pink cell within the FNR column. Similarly, the third row from the top corresponds to all heart sound records designated with true class ID 2. Here, 80% signifies the TPR for correctly classified points in the class, highlighted in the green cell within the TPR column. The remaining 20% of heart sound records in the Class ID 2 row are inaccurately classified, resulting in an FNR of 20%, illustrated in the purple cell within the FNR column. This interpretation applies uniformly to each row in Table 5.7.

We conducted an evaluation of the classifier's performance based on predicted classes rather than true classes, specifically examining false discovery rates (FDR). In addition to FDR, we delved into positive predictive values (PPV). FDR represents the proportion of incorrectly classified observations per predicted class, while PPV signifies the proportion of correctly classified observations per predicted class. The performance analysis based on predicted classes is depicted in the confusion matrix presented in Table 5.8. Notably, the bottom two rows of Table 5.8 display the PPV and FDR per predicted class. The first row corresponds to PPV, while the second row corresponds to FDR. This organized presentation provides a detailed overview of the PPV and FDR metrics for each predicted class.

Table 5.8: Performance of Cubic SVM Classifier Based on Predicted Classes

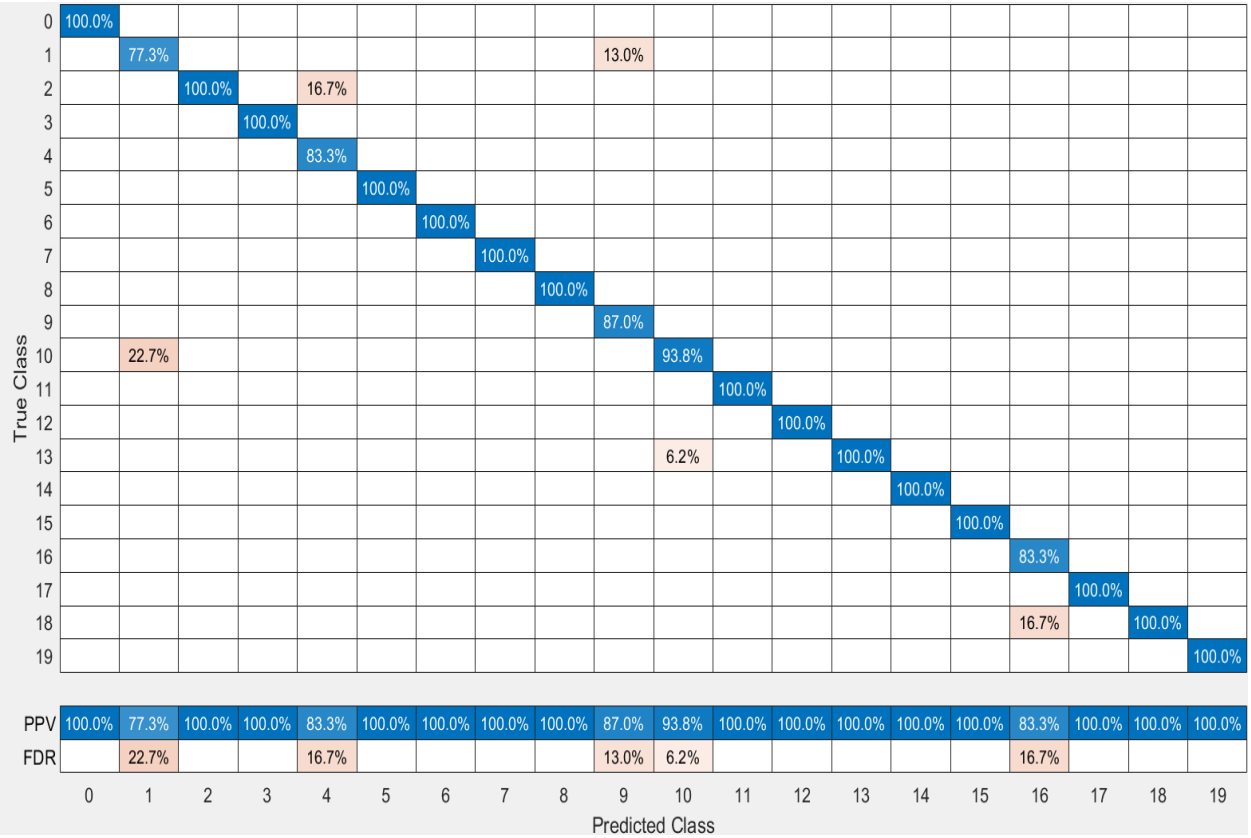
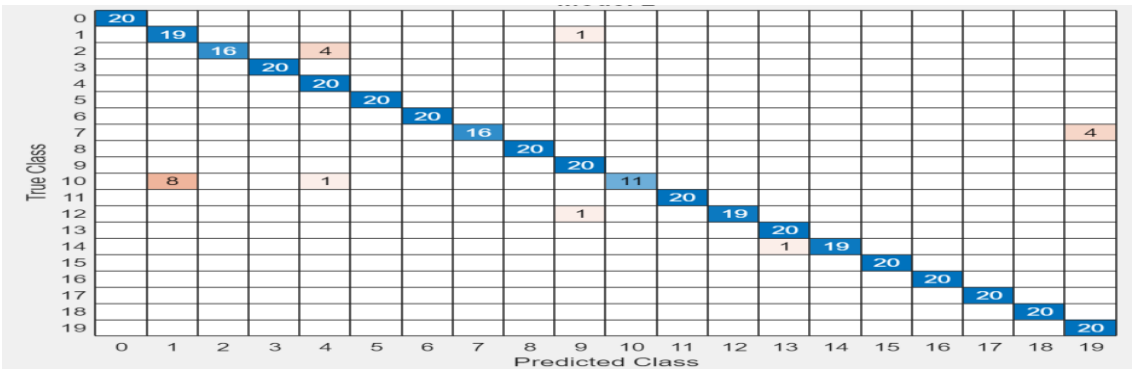


Table 5.8 provides an assessment of classifier performance based on predicted classes, specifically focusing on Positive Predictive Values (PPV) and False Discovery Rates (FDR). In this representation, PPV, denoting correctly predicted points in each class, is highlighted in green, while FDR, representing incorrectly predicted points in each class, is emphasized in pink. The evaluation of the Cubic SVM and the tested model's performance has been conducted using a confusion matrix. Table 5.9 presents a comprehensive overview of the classification models' performance, considering metrics such as Accuracy, Precision, Recall, and F1 Score. The Cubic SVM stands out as the model with the highest accuracy at 95.8%, accompanied by a precision of 96.2%, recall of 95.7%, and an F1 score of 95.7%. Notably, the Boosted Trees model exhibits the lowest performance among the models.

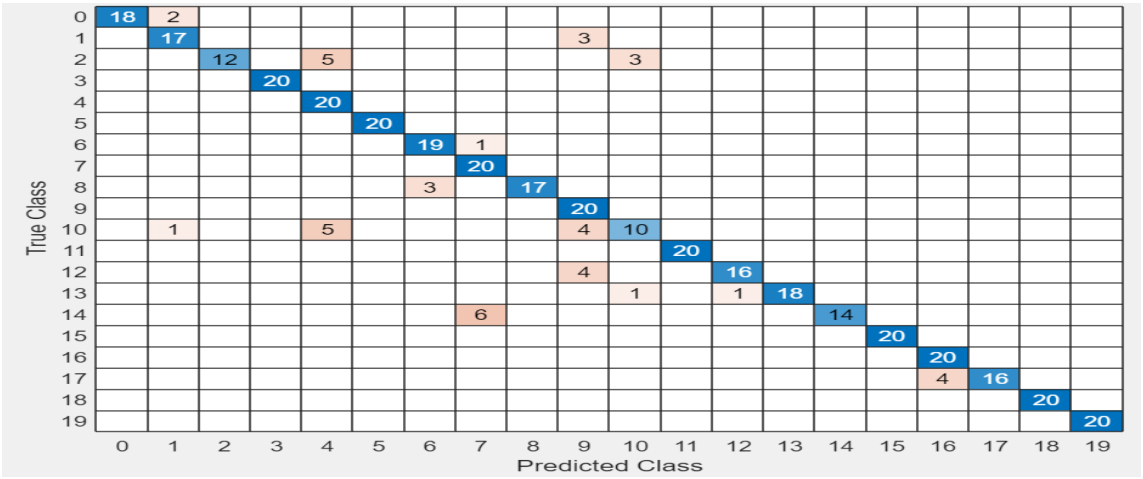
Table 5.9: Model Summary Based on Precision, Recall, Accuracy, and F1 Score

No	Model	Precision	Recall	Accuracy	F1 Score
1	Cubic SVM	96.2%	95.7%	95.8%	95.7%
2	Medium Gaussian SVM	96%	95%	95%	94.8%
3	Cubic KNN	90.5%	89%	89%	88.9%
4	Medium KNN	91%	89.2%	89.2%	89.0%
5	Boosted Trees	88.5%	87%	87%	86.9%
6	Bagged Trees	91%	89.5%	89.5%	89.5%
7	Subspace Discriminant	96%	94.5%	94.5%	94%

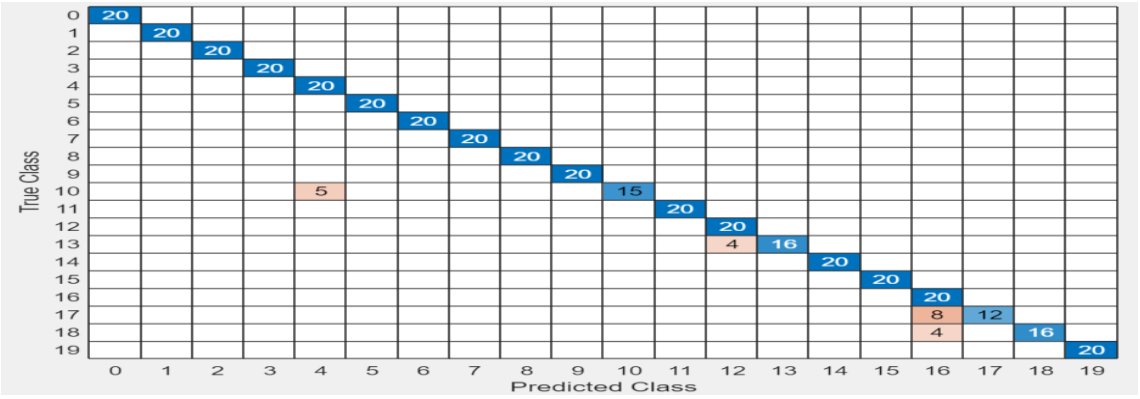
The study results underwent additional evaluation through the use of confusion matrices, as depicted in Figure 5.12. These matrices played a crucial role in enhancing the assessment of classifier performance. The findings revealed that the Cubic SVM exhibited superior performance, surpassing the other classifiers in the study.



(a) Confusion matrix for medium gaussian SVM



(b) confusion matrix for medium KNN



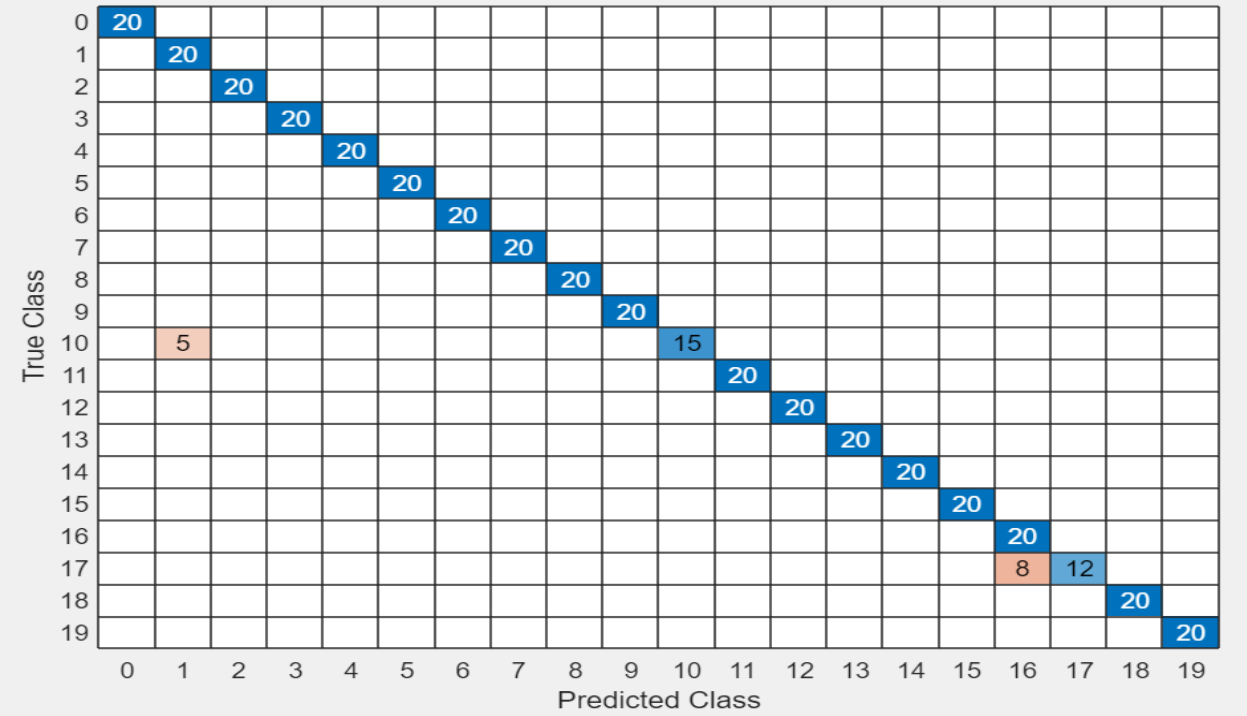
(c) Confusion matrix for Subspace Discriminant

Figure 5.12: Confusion matrix for classifiers (a) medium Gaussian SVM, (b) medium KNN, and (c) Subspace Discriminant

5.1.8. Model Optimization

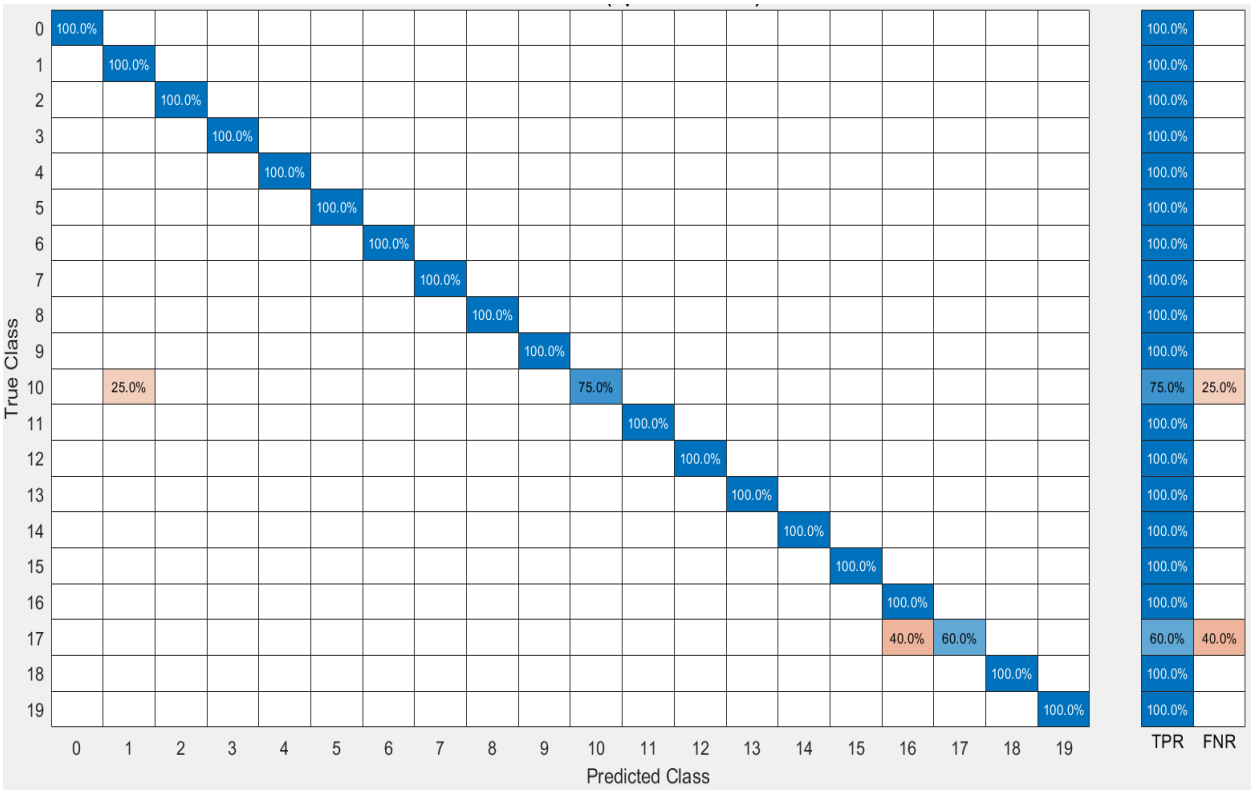
The evaluation of the initial selected and tested model's performance involved the use of a confusion matrix. Subsequently, efforts were made to enhance its accuracy through optimization. In this study, Bayesian optimization was employed to optimize the hyperparameters which are the kernel type for Support Vector Machine (SVM). Following the optimization of the model, we proceeded to export the final optimized model for predictions on new data. The export process involved transferring the optimized model from the classification learner to the MATLAB workspace, resulting in an optimized trained structure. This structure can now be utilized to make predictions effectively with previously unseen data. Subsequent experimentation with parameters and hyperparameters using Bayesian optimization techniques, the accuracy of our model increased to 96.8%. The performance of the optimized BO-SVM model was then assessed using four distinct metrics: accuracy, precision, recall, and F1 score, mirroring the approach taken in prior evaluations.

Table 5.10 : A number of correctly and misclassified observations for each class after optimization



Observations extracted from Table 5.10 reveal that out of the 20 heart sound records obtained from class ID10, 15 were accurately classified, and 5 were misclassified. Likewise, for class ID 17, 12 records were correctly classified, but 8 records were misclassified. Conversely, for the remaining subjects categorized under various class IDs, all 20 heart sound records were successfully classified during testing. The optimized classifier's performance was further assessed on a class-specific basis by visualizing true positive rates (TPR) and false negative rates (FNR). Table 5.11 displays the corresponding confusion matrix, providing a comprehensive overview of the classifier's performance across true classes. The last two columns on the right side of Table 5.11 specifically present the TPR and FNR for each true class. The TPR column occupies the first position, followed by the FNR column.

Table 5.11: Performance of an optimized SVM performed per true class



For instance, referencing Table 5.11, the second row from the top corresponds to heart sound records classified under the true class, class ID 1. Notably, 100% of heart sound records from class ID 1 were accurately classified as belonging to class ID 1, resulting in a TPR of 100%, as indicated by the green cell in the TPR column. In contrast, the unoptimized model exhibited TPR and FNR values of 85% and 15%, respectively, for class ID 1. This improvement signifies the enhanced capability of the model in accurately classifying records under the true class ID 1. The interpretation of Table 5.11 remains consistent across each row. The evaluation of the optimized classifier extended to an examination of its performance based on predicted classes, shifting the focus towards false discovery rates (FDR) and positive predictive values (PPV). Table 5.12 exhibits the corresponding confusion matrix, providing insights into the classifier's performance per predicted class. The last two rows of the table specifically denote the PPV and FDR for each predicted class, with the initial row dedicated to PPV and the subsequent row to FDR.

Table 5.12: Performance of an optimized SVM per predicted class

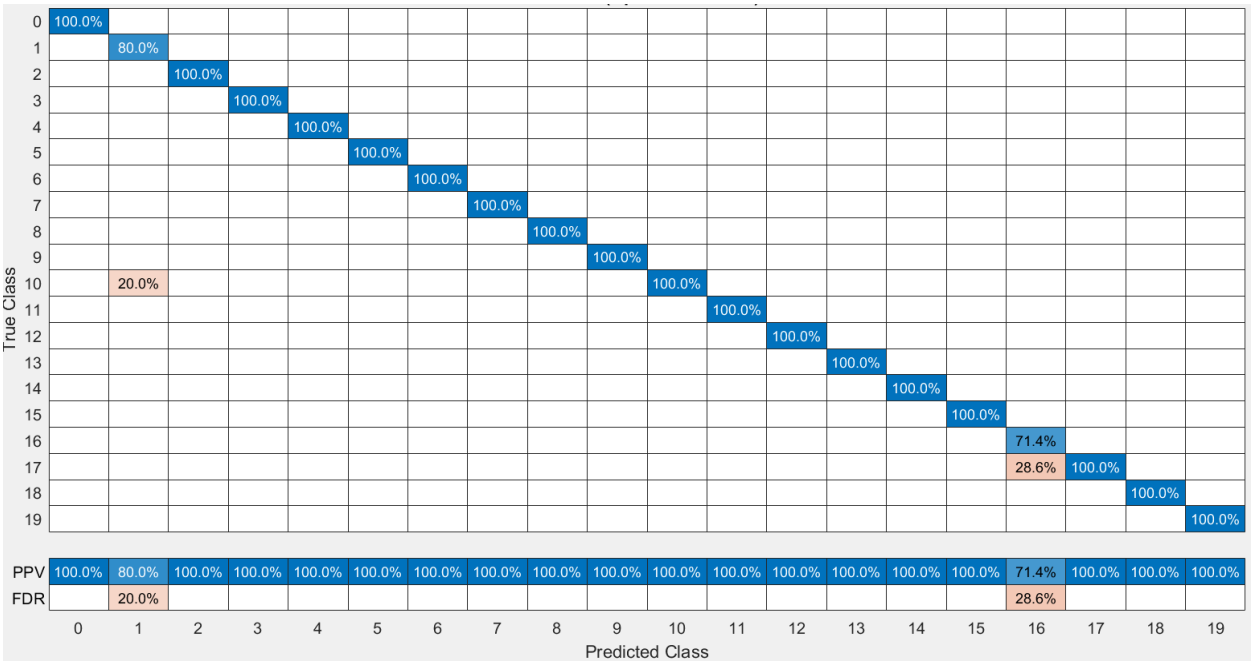


Table 5.12 provides a comprehensive overview of the optimized classifier's performance with a focus on predicted classes, assessing Positive Predictive Values (PPV) and False Discovery Rates (FDR). Correctly predicted points within each class are highlighted in green to represent PPV, while incorrectly predicted points are displayed in pink to indicate FDR. A closer examination of Table 5.12 reveals notable enhancements in the optimized model's ability to accurately predict various classes, including but not limited to class ID1 and class ID2. In conclusion, the precision, recall, accuracy, and F1 score of the final optimized model on the unseen data were 97.5%, 96.7%, 96.8%, and 96.6%, respectively.

5.2. Discussion

The study aimed to develop a model for heart sound based human identification using machine learning through cardiac auscultation. The modified electronic stethoscope was connected to a personal computer equipped with Audacity software version 5.2 for data recording.

To develop the identification model, seven types of models were trained and tested. According to table 5.9 shows the performance of the models were tested using precision, recall, accuracy, and F1 score. The models were trained with one-versus-rest classifier (OVR), additionally a 10-fold cross validation was applied to ensure the performance of each model.

Table 5.9 on the result section shows the identification report without applying Bayesian optimization technique on the models. High model accuracy was obtained from cubic SVM, medium gaussian SVM, and subspace discriminant classifiers. The average accuracy , precision, recall, and F1 score for cubic SVM was 95.8% ,96.2%, 95.7%, and 95.7% respectively.

Bayesian optimization was applied to optimize the cubic SVM to enhance its accuracy, and after optimization the accuracy of the model becomes 96.8%. according to table 5.10 on the result section shows a number of observations for each class ID after optimization, and misclassification is observed in class ID 10, and class ID17 .

The optimized classifiers performance was further assessed on specific basis by visualizing true positive rates (TPR), and false negative rate (FNR), and table 5.11 shows the performance of an optimized SVM performance per true class.

CHAPTER SIX

CONCIUSION AND RECOMMENDATION

6.1. Conclusion

This thesis analysis heart sounds for biometric identification, employing an optimizable Support Vector Machine (SVM). The study underscored the potential of heart sounds as a dependable biometric trait, citing their universality, distinctiveness, and fragility. The unique characteristics of heart sounds, influenced by cofounding factors, make them challenging to imitate or manipulate. Furthermore, the universality, stability, and ease of collection of heart sounds enhance their effectiveness as a biometric identifier.

The research employed a multiresolution analysis of heart sounds using the Discrete Wavelet Transform (DWT). The evaluation of the denoising algorithm's performance involved the computation of Signal-to-Noise Ratio (SNR) values at various levels of decomposition. Additionally, the study conducted feature extraction based on DWT and Mel Frequency Cepstral Coefficients (MFCC) to extract distinct features from the heart sound signals. Subsequently, the extracted features were fused together for a comprehensive analysis. With the extracted features in hand, the next steps involved feature selection and ranking using the relief algorithm. Following this, the selected features underwent normalization through the minmax normalization technique.

The normalized features served as the basis for training various machine learning models. In the classification process, the Cubic Support Vector Machine (SVM) algorithm demonstrated an initial accuracy of 95.8%. Subsequently, We optimized hyper parameters of the selected model using Bayesian optimization techniques, resulting in an enhanced accuracy of 96.8%. Finally, the performance of the optimized model was assessed using precision, recall, accuracy, and F1 score, achieving values of 96.7%, 96.7%, 96.8%, and 96.6%, respectively, in making predictions for new data.

The results of this study underscore the viability of employing heart sound signals within biometric systems as a preventive measure against identity fraud and forgery. Leveraging the distinctive attributes of heart sounds, coupled with advanced signal processing techniques and machine learning algorithms, presents a compelling prospect for improving the overall efficacy of biometric identification. The proposed method exhibits significant weaknesses, primarily stemming from the data acquisition system employed in this study. The heart sound signals were obtained using a single-channel data acquisition system, refers to a setup where only one sensor or transducer is used to capture the heart sound signal. Furthermore, it is noteworthy that the study exclusively focuses on healthy subjects, limiting the generalizability of its findings.

6.2. Recommendation

It is recommended to develop a user interface (UI) and establish a connection between the proposed model and the knowledge base, aiming to create an intelligent system for accurate subject identification.

Due to the limitations like resource constraints in this study, future research should focus on applying this methodology to extensive datasets and diverse population groups to validate its effectiveness and generalizability in biometric identification based on heart sounds. Additionally, the design of dedicated hardware, including portable and user-friendly devices for heart sound capture, is essential for the practical integration of the biometric identification system. This technological enhancement would facilitate real-time and non-intrusive identification using heart sounds.

REFERENCES

[1] A. Bonissi, R. D. Labati, L. Perico, R. Sassi, F. Scotti, and L. Sparagino, "A preliminary study on continuous authentication methods for photoplethysmographic biometrics," in *2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications, BioMS 2013 - Proceedings*, IEEE Computer Society, 2013, pp. 28–33. doi: 10.1109/BIOMS.2013.6656145.

[2] J. J. Winston and D. J. Hemanth, "A comprehensive review on iris image-based biometric system," *Soft comput*, vol. 23, no. 19, pp. 9361–9384, Oct. 2019, doi: 10.1007/s00500-018-3497-y.

[3] F. Beritelli and A. Spadaccini, "Human Identity Verification based on Heart Sounds: Recent Advances and Future Directions," May 2011, [Online]. Available: http://arxiv.org/abs/1105.4058

[4] R. Soram, "Biometric DNA and ECDLP-based Personal Authentication System: A Superior Posse of Security."

[5] A. Abaza, A. Ross, C. Hebert, M. A. F. Harrison, and M. S. Nixon, "A survey on ear biometrics," *ACM Computing Surveys*, vol. 45, no. 2. Feb. 2013. doi: 10.1145/2431211.2431221.

[6] C. Cabrera, M. R. Guillen, and M. Adjouadi, "Engineering for a Smart Planet, Innovation, Information Technology and Computational Tools for Sustainable Development."

[7] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, "Extract from," 2009.

[8] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognit*, vol. 42, no. 7, pp. 1408–1418, Jul. 2009, doi: 10.1016/j.patcog.2009.01.018.

[9] J. Daugman, "New methods in iris recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 5, pp. 1167–1175, Oct. 2007, doi: 10.1109/TSMCB.2007.903540.

[10] P. S. Prasad, R. Pathak, V. K. Gunjan, and H. V. Ramana Rao, "Deep Learning Based Representation for Face Recognition," in *Lecture Notes in Electrical Engineering*, Springer Verlag, 2020, pp. 419–424. doi: 10.1007/978-981-13-8715-9_50.

[11] D. Liu, Y. J. Ge, and X. Y. Zhang, "On-line signature verification utilizing 3-axis force features extraction based on DWT," in *Applied Mechanics and Materials*, 2013, pp. 2990–2995. doi: 10.4028/www.scientific.net/AMM.241-244.2990.

[12] Z. Saquib, N. Salam, R. Nair, and N. Pandey, "Voiceprint Recognition Systems for Remote Authentication-A Survey," 2011.

[13] N. Erdogmus and S. Marcel, "I SPOOFING IN 2D FACE RECOGNITION WITH 3D MASKS AND ANTI-SPOOFING WITH KINECT Spoofing in 2D Face Recognition with 3D Masks and Anti-spoofing with Kinect," 2013.

[14] D. E. Mancilla-Palestina, J. A. Jimenez-Duarte, J. M. Ramirez-Cortes, A. Hernandez, P. Gomez-Gil, and J. Rangel-Magdaleno, "Embedded system for bimodal biometrics with fiducial feature extraction on ECG and PPG signals," in *I2MTC 2020 - International Instrumentation and Measurement Technology Conference, Proceedings*, Institute of Electrical and Electronics Engineers Inc., May 2020. doi: 10.1109/I2MTC43012.2020.9128394.

[15] *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2019.

[16] A. S. Rathore, Z. Li, W. Zhu, Z. Jin, and W. Xu, "A Survey on Heart Biometrics," *ACM Computing Surveys*, vol. 53, no. 6. Association for Computing Machinery, Feb. 01, 2021. doi: 10.1145/3410158.

[17] M. U. Khan, S. Aziz, A. Zainab, H. Tanveer, K. Iqtidar, and A. Waseem, "Biometric System using PCG Signal Analysis: A New Method of Person Identification," in *2nd International Conference on Electrical, Communication and Computer Engineering, ICECCE 2020*, Institute of Electrical and Electronics Engineers Inc., Jun. 2020. doi: 10.1109/ICECCE49384.2020.9179257.

[18] Institute of Electrical and Electronics Engineers. Kerala Section, IEEE Region 10, and Institute of Electrical and Electronics Engineers, *Proceedings of the TENCON 2019 : Technology, Knowledge, and Society : 17-20 October 2019, Grand Hyatt Kochi Bolgatti, Kerala, India.*

[19] X. Cheng, P. Wang, and C. She, "Biometric identification method for heart sound based on multimodal multiscale dispersion entropy," *Entropy*, vol. 22, no. 2, Feb. 2020, doi: 10.3390/e22020238.

[20] O. Mitchell, "Experimental Research Design," in *The Encyclopedia of Crime and Punishment*, Wiley, 2015, pp. 1–6. doi: 10.1002/9781118519639.wbecpx113.

[21] "310118-Cardiac-system-1-anatomy-and-physiology (1)".

[22] J. Reimer-Kent, "Heart sounds: are you listening? Part 1.," *Can J Cardiovasc Nurs*, vol. 23, no. 2, 2013.

[23] C. D. Papadaniil and L. J. Hadjileontiadis, "Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features," *IEEE J Biomed Health Inform*, vol. 18, no. 4, pp. 1138–1152, 2014, doi: 10.1109/JBHI.2013.2294399.

[24] K. Phua, J. Chen, T. H. Dat, and L. Shue, "Heart sound as a biometric," *Pattern Recognit*, vol. 41, no. 3, pp. 906–919, Mar. 2008, doi: 10.1016/j.patcog.2007.07.018.

[25] Selina Jarvis, "Cardiac system : anatomy and physiology," *Nurs Times*, 2018.

[26] Christiana Wuche, "The cardiovascular system and associated disorders," *British Journal of Nursing* , vol. 31, no. 17, 2022.

[27] "Fundamentals_of_Anatomy_and_Physiology_2".

[28] M. Nabih-Ali, E. S. A. El-Dahshan, and A. S. Yahia, “A review of intelligent systems for heart sound signal analysis,” *Journal of Medical Engineering and Technology*, vol. 41, no. 7. Taylor and Francis Ltd, pp. 553–563, Oct. 03, 2017. doi: 10.1080/03091902.2017.1382584.

[29] V. N. Varghees and K. I. Ramachandran, “A novel heart sound activity detection framework for automated heart sound analysis,” *Biomed Signal Process Control*, vol. 13, no. 1, pp. 174–188, 2014, doi: 10.1016/j.bspc.2014.05.002.

[30] Chakrabarti Tamal, Sourav Saha, Sathi Roy, and Ishita Chel, “PHONOCARDIOGRAM SIGNAL ANALYSIS - PRACTICES, TRENDS AND CHALLENGES: A CRITICAL REVIEW,” *IEEE*, pp. 1–4, 2015.

[31] S. Singh Badghare, “Analysis of PHONOCARDIOGRAM SIGNAL FOR BIOMETRIC APPLICATION.”

[32] “Time Domain Versus Frequency Domain Analysis Time Domain vs Frequency Domain.”

[33] B. Osgood, “Lecture Notes for EE 261 The Fourier Transform and its Applications.”

[34] C. H. Chen, W. T. Huang, T. H. Tan, C. C. Chang, and Y. J. Chang, “Using K-nearest neighbor classification to diagnose abnormal lung sounds,” *Sensors (Switzerland)*, vol. 15, no. 6, pp. 13132–13158, Jun. 2015, doi: 10.3390/s150613132.

[35] S. Goel, *Microelectronics and Signal Processing*. CRC Press, 2021. doi: 10.1201/9781003168225.

[36] B. Ahmed Ali, T. Jemal Co-Advisor, and B. Lingaiah, “Classification of Heart Sounds Associated With Murmur for Automatic Diagnosis of Cardiac Valve Disorders,” 2018.

[37] Z. Zhao, Q. Shen, and F. Ren, “Heart sound biometric system based on marginal spectrum analysis,” *Sensors (Switzerland)*, vol. 13, no. 2, pp. 2530–2551, 2013, doi: 10.3390/s130202530.

[38] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, “Biometrics from heart sounds: Evaluation of a new approach based on wavelet packet cepstral features using HSCT-11 database,” *Computers and Electrical Engineering*, vol. 53, pp. 346–358, Jul. 2016, doi: 10.1016/j.compeleceng.2016.05.004.

[39] S. Verma and T. Kashyap, “ANALYSIS OF HEART SOUND AS BIOMETRIC USING MFCC & LINEAR SVM CLASSIFIER,” 2007. [Online]. Available: www.ijareeie.com

[40] M. IEEE Systems and IEEE Biometrics Council., *BTAS 2010 : IEEE Fourth International Conference on Biometrics : Theory, Applications, and Systems*.

[41] Hideyuki. Takagi, *Proceedings : 2010 Second World Congress on Nature and Biologically Inspired Computing : NaBIC 2010, Kitakyushu, Japan, December 15-17, 2010*.

[42] *2010 International Conference on Electronics and Information Engineering*. I E, 2010.

[43] M. U. Khan, S. Aziz, A. Zainab, H. Tanveer, K. Iqtidar, and A. Waseem, “Biometric System using PCG Signal Analysis: A New Method of Person Identification,” in *2nd International Conference on*

Electrical, Communication and Computer Engineering, ICECCE 2020, Institute of Electrical and Electronics Engineers Inc., Jun. 2020. doi: 10.1109/ICECCE49384.2020.9179257.

[44] X. F. Cheng, Y. W. Tao, and Z. J. Huang, "Heart sound recognition - A prospective candidate for biometric identification," in *Advanced Materials Research*, 2011, pp. 433–436. doi: 10.4028/www.scientific.net/AMR.225-226.433.

[45] Institute of Electrical and Electronics Engineers Pune Section, Sinhgad College of Engineering, Sinhgad College of Engineering Department of Electronics and Telecommunication Engineering, International Conference on Pervasive Computing 2015.01.08-10 Pune, and ICPC 2015.01.08-10 Pune, *International Conference on Pervasive Computing (ICPC), 2015 8-10 Jan. 2015, Pune, India*.

[46] S. N. Abbas, M. Abo-Zahhad, S. M. Ahmed, and M. Farrag, "Heart-ID: Human identity recognition using heart sounds based on modifying mel-frequency cepstral features," *IET Biom*, vol. 5, no. 4, pp. 284–296, Dec. 2016, doi: 10.1049/iet-bmt.2015.0033.

[47] IEEE Staff, *2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC)*. IEEE, 2016.

[48] Institute of Electrical and Electronics Engineers. Kerala Section, IEEE Region 10, and Institute of Electrical and Electronics Engineers, *Proceedings of the TENCON 2019 : Technology, Knowledge, and Society : 17-20 October 2019, Grand Hyatt Kochi Bolgatti, Kerala, India.*

[49] X. Cheng, P. Wang, and C. She, "Biometric identification method for heart sound based on multimodal multiscale dispersion entropy," *Entropy*, vol. 22, no. 2, Feb. 2020, doi: 10.3390/e22020238.

[50] Mizuho Tagashira and Takafumi Nakagawa, "Biometric Authentication Based on Auscultated Heart Sounds in Healthcare," *IAENG*, vol. 47, no. 3, 2020.

[51] T. E. Chen *et al.*, "S1 and S2 heart sound recognition using deep neural networks," *IEEE Trans Biomed Eng*, vol. 64, no. 2, pp. 372–380, Feb. 2017, doi: 10.1109/TBME.2016.2559800.

[52] IEEE Staff, *2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC)*. IEEE, 2016.

[53] Hideyuki. Takagi, *Proceedings : 2010 Second World Congress on Nature and Biologically Inspired Computing : NaBIC 2010, Kitakyushu, Japan, December 15-17, 2010*.

[54] K. Mangion, "Kenneth Mangion The stethoscope," 2007.

[55] A. Rao, E. Huynh, T. J. Royston, A. Kornblith, and S. Roy, "Acoustic Methods for Pulmonary Diagnosis," *IEEE Reviews in Biomedical Engineering*, vol. 12. Institute of Electrical and Electronics Engineers, pp. 221–239, 2019. doi: 10.1109/RBME.2018.2874353.

[56] Lucas Ferreira-Paiva, Elizabeth Alfaro-Espinoza, Vinicius M. Almeida, Leonardo B. Felix, and Rodolpho V. A. Neves, "A Survey of Data Augmentation for Audio Classification," in *Procedings*

do XXIV Congresso Brasileiro de Automática, SBA Sociedade Brasileira de Automática, 2022. doi: 10.20906/cba2022/3469.

[57] M. N. Ali, E. S. A. El-Dahshan, and A. H. Yahia, "Denoising of Heart Sound Signals Using Discrete Wavelet Transform," *Circuits Syst Signal Process*, vol. 36, no. 11, pp. 4482–4497, Nov. 2017, doi: 10.1007/s00034-017-0524-7.

[58] H. A. R. Akkar, W. A. H. Hadi, and I. H. M. Al-Dosari, "A novel wavelet thresholding method for implementation of signal compressor design," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Apr. 2019, pp. 12–17. doi: 10.1145/3321289.3321292.

[59] B. Ghojogh *et al.*, "Feature Selection and Feature Extraction in Pattern Analysis: A Literature Review," May 2019, [Online]. Available: http://arxiv.org/abs/1905.02845

[60] M. Borowska, "Entropy-based algorithms in the analysis of biomedical signals," *Studies in Logic, Grammar and Rhetoric*, vol. 43, no. 56, pp. 21–32, 2015, doi: 10.1515/slgr-2015-0039.

[61] M. Guven and F. Uysal, "A New Method for Heart Disease Detection: Long Short-Term Feature Extraction from Heart Sound Data," *Sensors*, vol. 23, no. 13, Jul. 2023, doi: 10.3390/s23135835.

[62] A. Babiker, A. Hassan, and H. Mustafa, "Heart Sounds Biometric System," *Journal of Biomedical Engineering and Medical Devices*, vol. 02, no. 02, 2017, doi: 10.4172/2475-7586.1000129.

[63] J. Martinez, H. Perez, E. Escamilla, and M. M. Suzuki, "Speaker recognition using Mel Frequency Cepstral Coefficients (MFCC) and Vector Quantization (VQ) Techniques."

[64] S. Verma and T. Kashyap, "ANALYSIS OF HEART SOUND AS BIOMETRIC USING MFCC & LINEAR SVM CLASSIFIER," 2007. [Online]. Available: www.ijareeie.com

[65] Adhiparasakthi Engineering College. Department of Electronics and Communication Engineering, Institute of Electrical and Electronics Engineers. Madras Section, and Institute of Electrical and Electronics Engineers, *ICCSP-2016 : 4th-6th April, 2016*.

[66] F. A. Khan, A. Abid, A. Abid, M. S. Khan, and M. S. Khan, "Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features," *Physiol Meas*, vol. 41, no. 5, May 2020, doi: 10.1088/1361-6579/ab8770.

[67] M. C. Al Fajar, M. Fatmawati, P. Wulandari, and D. Astharini, "Analysis of DFT and FFT Signal Transformation with Hamming Window in LabVIEW," in *2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering, BCWSP 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020, pp. 79–83. doi: 10.1109/BCWSP50066.2020.9249466.

[68] T. Marutkumar and P. P. Patel, "A Review on Feature Extraction with Classification Methods for Cardiac and Pulmonary Diseases," 2018. [Online]. Available: www.ijcrt.org

[69] R. P. L. Durgabai, "Feature Selection using ReliefF Algorithm," 2014. [Online]. Available: www.ijarcce.com

[70] "RELIEF: Feature Selection Approach." [Online]. Available: www.ijird.com

[71] Institute of Electrical and Electronics Engineers and IEEE Geoscience and Remote Sensing Society, *2016 IEEE International Geoscience & Remote Sensing Symposium : proceedings : July 10-15, 2016, Beijing, China*.

[72] D. Bollegala, "Dynamic Feature Scaling for Online Learning of Binary Classifiers," Jul. 2014, [Online]. Available: http://arxiv.org/abs/1407.7584

[73] W. Tao, J. Yongjia, and R. Xiangsheng, "A Novel Two-Level One-vs-Rest Classifier," in *2019 2nd International Conference on Information Systems and Computer Aided Education, ICISCAE 2019*, Institute of Electrical and Electronics Engineers Inc., Sep. 2019, pp. 645–648. doi: 10.1109/ICISCAE48440.2019.221714.

[74] F. Kamal Alsheref and W. Hassan Gomaa, "Blood Diseases Detection using Classical Machine Learning Algorithms," 2019. [Online]. Available: www.ijacsa.thesai.org

[75] N. Najwa Mohd Rizal, G. Hayder, M. Mnzool, B. M. E. Elnaim, A. O. Y. Mohammed, and M. M. Khayyat, "Comparison between Regression Models, Support Vector Machine (SVM), and Artificial Neural Network (ANN) in River Water Quality Prediction," *Processes*, vol. 10, no. 8, Aug. 2022, doi: 10.3390/pr10081652.

[76] "Support Vector Machines for Classification SVM from a Geometric Perspective."

[77] X. Yang, Q. Yu, L. He, and T. Guo, "The one-against-all partition based binary tree support vector machine algorithms for multi-class classification," *Neurocomputing*, vol. 113, pp. 1–7, Aug. 2013, doi: 10.1016/j.neucom.2012.12.048.

[78] H. Cho, Y. Kim, E. Lee, D. Choi, Y. Lee, and W. Rhee, "Basic Enhancement Strategies When Using Bayesian Optimization for Hyperparameter Tuning of Deep Neural Networks," *IEEE Access*, vol. 8, pp. 52588–52608, 2020, doi: 10.1109/ACCESS.2020.2981072.

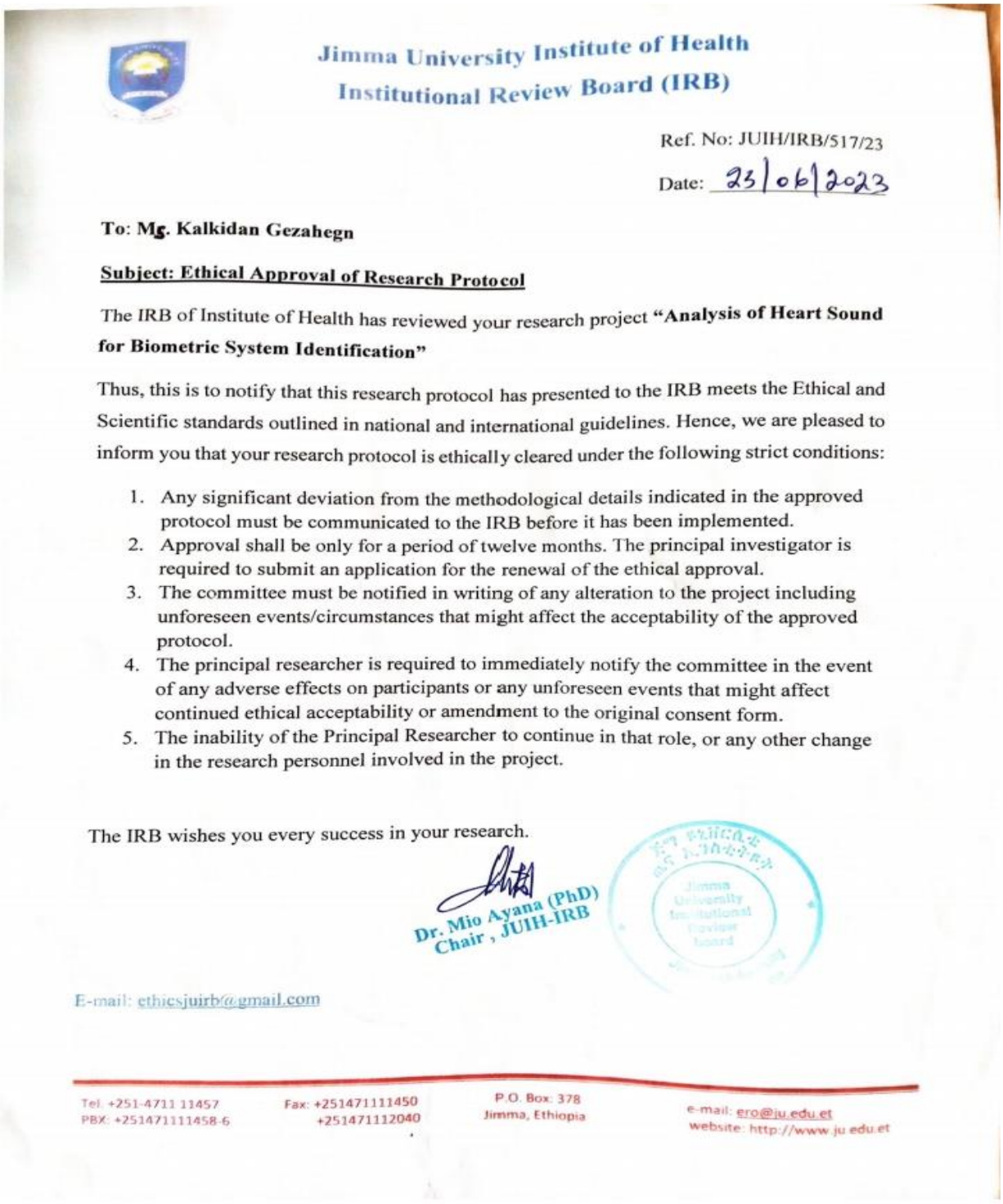
[79] C. Gambella, B. Ghaddar, and J. Naoum-Sawaya, "Optimization problems for machine learning: A survey," *European Journal of Operational Research*, vol. 290, no. 3. Elsevier B.V., pp. 807–828, May 01, 2021. doi: 10.1016/j.ejor.2020.08.045.

[80] A. H. Victoria and G. Maragatham, "Automatic tuning of hyperparameters using Bayesian optimization," *Evolving Systems*, vol. 12, no. 1, pp. 217–223, Mar. 2021, doi: 10.1007/s12530-020-09345-2.

[81] C. Camara, P. Peris-Lopez, M. Safkhani, and N. Bagheri, "ECGsound for human identification," *Biomed Signal Process Control*, vol. 72, Feb. 2022, doi: 10.1016/j.bspc.2021.103335.

Annex

Appendix A: Letter for Ethical Approval of Research Protocol



Jimma University Institute of Health
Institutional Review Board (IRB)

Ref. No: JUIH/IRB/517/23

Date: 23/06/2023

To: Mg. Kalkidan Gezahegn

Subject: Ethical Approval of Research Protocol

The IRB of Institute of Health has reviewed your research project **"Analysis of Heart Sound for Biometric System Identification"**

Thus, this is to notify that this research protocol has presented to the IRB meets the Ethical and Scientific standards outlined in national and international guidelines. Hence, we are pleased to inform you that your research protocol is ethically cleared under the following strict conditions:

1. Any significant deviation from the methodological details indicated in the approved protocol must be communicated to the IRB before it has been implemented.
2. Approval shall be only for a period of twelve months. The principal investigator is required to submit an application for the renewal of the ethical approval.
3. The committee must be notified in writing of any alteration to the project including unforeseen events/circumstances that might affect the acceptability of the approved protocol.
4. The principal researcher is required to immediately notify the committee in the event of any adverse effects on participants or any unforeseen events that might affect continued ethical acceptability or amendment to the original consent form.
5. The inability of the Principal Researcher to continue in that role, or any other change in the research personnel involved in the project.

The IRB wishes you every success in your research.

Dr. Mio Ayana (PhD)
Chair, JUIH-IRB

E-mail: ethicsjuirb@gmail.com

Tel. +251-4711 11457
PBX: +251471111458-6
Fax: +251471111450
+251471112040
P.O. Box: 378
Jimma, Ethiopia
e-mail: ero@ju.edu.et
website: http://www.ju.edu.et