# Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities

By

Iqra Zaheer

(Registration No: 00000362411)

Department of Biomedical Sciences

School of Mechanical and Manufacturing Engineering

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

# Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities

By

Iqra Zaheer

(Registration No: 00000362411)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in

Biomedical Sciences

Supervisor: Dr. Asim Waris

School of Mechanical and Manufacturing Engineering

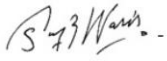
National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

## THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Regn No. 00000362411 Iqra Zaheer** of **School of Mechanical & Manufacturing Engineering (SMME)** has been vetted by undersigned, found complete in all respects as per NUST Statues/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis titled. **Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities**

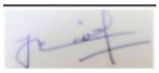
**Signature:** 

**Name (Supervisor):** Muhammad Asim Waris

**Date:** 26 - Jul - 2024

**Signature (HOD):** 

**Date:** 26 - Jul - 2024

**Signature (DEAN):** 

**Date:** 26 - Jul - 2024

**Form TH-4**

# National University of Sciences & Technology (NUST)

## MASTER'S THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: Iqra Zaheer (00000362411)
Titled: Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities be accepted in partial fulfillment of the requirements for the award of MS in Biomedical Sciences degree.

## Examination Committee Members

1. Name: Aneeqa Noor Signature:

2. Name: Adeeb Shehzad Signature:

**Supervisor:** Muhammad Asim Waris Signature:

Date: 26 - Jul - 2024

Head of Department 26 - Jul - 2024
Date

## COUNTERSINGED

26 - Jul - 2024
Date

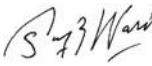
Dean/Principal

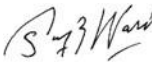
# CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis, entitled "Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities'' was conducted by Ms. Iqra Zaheer under the supervision of Dr. Asim Waris.

No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the Department of Biomedical Sciences (NUST) in partial fulfillment of the requirements for the degree of Master of Science in Field of Neuroscience, Department of Biomedical Sciences, National University of Sciences and Technology, Islamabad.

Student Name: Iqra Zaheer Signature: 

Supervisor Name: Dr. Asim Waris Signature: 

Name of Dean/HOD: Dr. Asim Waris Signature: 

# AUTHOR'S DECLARATION

I Iqra Zaheer hereby state that my MS thesis titled "Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities" is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: IQRA ZAHEER

Date: 26-07-2024

# PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "Comparative Analysis of Machine Learning Techniques for Phonocardiogram-Based Classification of Cardiac Abnormalities'' is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST), Islamabad towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and NUST, Islamabad has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student Signature: 

Name: Iqra Zaheer

*“Dedicated to my beloved parents Zaheer Ahmed and Nighat Asia, whose unwavering faith, enduring belief, and ceaseless encouragement have been my guiding light and anchor. To my dearest siblings, your consistent support has been my strength and inspiration.”*

# ACKNOWLEDGEMENTS

I am profoundly grateful to Almighty ALLAH (SWT) for bestowing the willpower and divine direction that guided me through the intricacies of this research project. My heartfelt appreciation goes to my beloved parents (Zaheer Ahmed and Nighat Asia), and siblings (Waleed, Sara, Inaaya, Talha, and Hammad) who have been unending pillars of support. Their moral and emotional encouragement helped me surmount disappointments, and their constant presence provided solace through thick and thin. This journey was made lighter by their belief in me and their enduring companionship.

I am deeply indebted to my supervisor, Dr. Asim Waris, for his continuous support, motivation, and extensive knowledge. His demonstrative supervision, invaluable guidance, and genuine encouragement have been instrumental in shaping the essence of this research work. I am grateful for the privilege of working under his mentorship.

Acknowledgment is also due to my close friend Rida, whose unwavering support and belief in me have been a constant source of strength. Whether it was celebrating small victories or offering a listening ear during challenges, her encouragement and friendship have meant the world to me.

Thanks to Urwah and other special friends, whose steadfast support and encouragement have been invaluable to me throughout this journey.

# TABLE OF CONTENTS


**ACKNOWLEDGEMENTS** **VIII**

**TABLE OF CONTENTS** **IX**

**LIST OF TABLES** **XII**

**LIST OF FIGURES** **XIII**

**LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS** **XIV**

**ABSTRACT** **XV**

**CHAPTER 1. INTRODUCTION** **1**
**1.1 Cardiovascular Disease: A Global Challenge** **1**
1.1.1 Prevalence and Impact of CVDs 1
**1.2 Limitations of Traditional Diagnostic Methods** **1**
**1.3 PCG Signals: A Promising Alternative** **2**
1.3.1 Challenges in PCG Signal Analysis 2
**1.4 Advanced Signal Processing Techniques** **2**
1.4.1 The Power of Combining Techniques 3
**1.5 Problem Statement** **3**
**1.6 Objectives of the Study** **4**
**1.7 Research Questions and Hypothesis** **4**
1.7.1 Research Questions 4
1.7.2 Hypothesis 5
**1.8 Significance of the Study** **5**
1.8.1 Addressing Diagnostic Limitations: 5
1.8.2 Comprehensive Model Evaluation: 5
1.8.3 Practical Clinical Application: 5
1.8.4 Contribution to Biomedical Engineering: 6
**1.9 Literature Review** **6**
1.9.1 The Intricate Symphony of Your Heart: Unveiling Health Through Heart Sounds6
1.9.2 The Rhythmic Melody: Normal Heart Sounds 6
1.9.3 Beyond the Basics: A Look at Splitting and Murmurs 8
1.9.4 Classifying Heart Sounds: Unveiling Abnormalities 8
1.9.5 The Power of Listening: The Importance of Early Detection 9
1.9.6 Overview of Cardiac Pathologies 9
1.9.7 The Symphony of Diagnosis 13
1.9.8 Looking Forward: Technology's Role in Listening PCG Signals 13
1.9.9 Signal Processing Techniques in Cardiac Diagnostics 14
1.9.10 The Power of Deep Learning: Automating Heart Signal Analysis 15
1.9.11 Filling the Gap: A Comparative Approach 15

1.9.12 The Proposed Framework: A Symphony of Techniques 16
1.9.13 The Future of Deep Listening 16

**CHAPTER 2. METHODOLOGY 17**
**2.1 Database Description and Categorization 17**
2.1.1 Data Classification and Structure: 17
2.1.2 Data Collection Process and Data Sources: 18
2.1.3 Data Preprocessing: 19
**2.2 Audio Signal Processing 19**
2.2.1 Discrete Wavelet Transform (DWT) for Signal Denoising and Decomposition 20
**2.3 Feature Extraction for PCG Signals 22**
2.3.1 Mel Scale Power Spectrogram 22
2.3.2 Mel-Frequency Cepstral Coefficients (MFCC) 25
**2.4 Classification Models 28**
2.4.1 Deep Neural Networks (DNNs) for Cardiac Diagnosis: A Powerful Tool for Heart Sound Classification 28
2.4.2 The Heartbeat of Innovation: How AI Listens to Your Heart 30
**2.5 Performance of Models: Unveiling the Efficacy of DNN and CNN in PCG Classification 33**
2.5.1 Confusion Matrix: 33
2.5.2 Essential Metrics for Classification Performance 34

**CHAPTER 3. RESULTS 37**
**3.1 Denoising with Discrete Wavelet Transform (DWT) 37**
3.1.1 Wavelet Selection: 37
3.1.2 Coefficient Selection: 37
3.1.3 Signal-to-Noise Ratio (SNR) Improvement: 38
**3.2 Feature Extraction: Unveiling the Spectrogram and the Power of MFCCs 40**
3.2.1 Mel Scale Power Spectrogram: 40
42
3.2.2 Mel Frequency Cepstral Coefficients (MFCCs): 44
**3.3 Classification Results 48**
3.3.1 DNN and CNN in Action 48
50
50
3.3.2 Key Findings 51

**CHAPTER 4. DISCUSSION 52**
**4.1 Advancements in the Current Study 52**
4.11 Detailed Classification of Cardiac Abnormalities: 52
4.1.2 Comparative Model Evaluation: 53
**4.2 Enhanced Preprocessing and Feature Extraction: 53**
4.2.1 Denoising with DWT: 53
4.2.2 Feature Extraction Techniques: 53
**4.3 Drawbacks and Upcoming Projects: 54**
**4.4 Opportunities for Future Exploration 54**

4.4.1 Data Expansion and Diversity: 54
4.4.2 Computational Efficiency for Real-time Applications: 55
4.4.3 Feature Engineering and Exploration: 55

**CHAPTER 5. CONCLUSIONS AND FUTURE RECOMMENDATION 56**

**SUMMARY OF RESEARCH WORK 57**

**REFERENCES 59**

# LIST OF TABLES

Table 1.1: Characteristics of Heart Sounds ..... 7
Table 1.2: Overview of some Common Cardiac Diseases Diagnosed using PCG ..... 12
Table 2.1: Dataset Details ..... 18

# LIST OF FIGURES

Figure 1.1: (a) A normal heart sound signal (b) Murmur in systole (MVP) (c) Mitral Regurgitation (MR) (d) Mitral Stenosis (MS) (e) Aortic Stenosis (AS) (f) Spectrum of a PCG signal (Yaseen et al., 2018) ..... 9
Figure 1.2: A healthy heart and Aortic Valve Stenosis condition ..... 10
Figure 1.3: A healthy heart mechanism compared with Mitral Valve Stenosis and Mitral Valve Regurgitation ..... 11
Figure 1.4: A healthy heart anatomy given in contrast with a Mitral Valve Prolapse Condition ..... 13
Figure 2.1: Phonocardiogram Signal Processing Workflow ..... 17
Figure 2.2: Feature Extraction Pipeline from Mel Scale Power Spectrogram ..... 22
Figure 2.3 Flowchart for MFCC Feature Extraction ..... 26
Figure 3.1: DWT for Normal PCG Signal ..... 38
Figure 3.2: DWT for a Aortic Stenosis PCG Signal ..... 38
Figure 3.3: DWT for Mitral Regurgitation PCG Signal ..... 39
Figure 3.4: DWT for Mitral Valve Prolapse PCG Signal ..... 39
Figure 3.5 DWT for Mitral Stenosis PCG Signal ..... 40
Figure 3.6: Mel Power Spectrogram for Normal PCG Signal ..... 41
Figure 3.7: Mel Power Spectrogram for MS PCG Signal ..... 42
Figure 3.8 Mel Power Spectrogram for MR PCG Signal ..... 42
Figure 3.9 Mel Power Spectrogram for AS ..... 43
Figure 3.10 Mel Power Spectrogram Signal for MVP ..... 43
Figure 3.11: MFCC for Normal PCG Signal ..... 45
Figure 3.12: MFCC for MS PCG Signal ..... 45
Figure 3.13 MFCC for AS PCG Signal ..... 46
Figure 3.14 MFCC for MVP PCG Signal ..... 47
Figure 3.15 MFCC for MR PCG Signal ..... 47
Figure 3.16 Classification Report for DNN Model for Five Classes ..... 49
Figure 3.17 Confusion Matrix representing the predicted 5 Classes for DNN Model ..... 49
Figure 3.18 Classification Report for Five Classes for CNN Model ..... 50
Figure 3.19 Confusion Matrix representing the predicted 5 Classes for CNN Model ..... 50

# LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| CVD | Cardiovascular Diseases |
| PCG | Phonocardiogram |
| AS | Aortic Stenosis |
| MR | Mitral Regurgitation |
| MS | Mitral Stenosis |
| MVP | Mitral Valve Prolapse |
| DWT | Discrete Wavelet Transform |
| MFCC | Mel Frequency Cepstral Coefficient |
| DNN | Deep Neural Network |
| CNN | Convolutional Neural Network |

# ABSTRACT

Cardiovascular Diseases (CVDs) remain a leading cause of morbidity and mortality worldwide, necessitating early and accurate detection for effective disease management. This work employs advanced signal processing techniques in conjunction with machine learning methodologies to classify normal and particular cardiac conditions—Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), and Mitral Valve Prolapse (MVP)—using phonocardiogram (PCG) signals. Preprocessing involved denoising using the Discrete Wavelet Transform (DWT) technique with the db8 wavelet and cA2 component, optimizing noise reduction while retaining valuable features for further analysis. Feature extraction was performed using Mel-Frequency Cepstral Coefficients (MFCC) and Mel Power Spectrogram, providing a robust and efficient representation of heart sounds. Two machine learning models—Deep Neural Network (DNN) and Convolutional Neural Network (CNN)—were used to assess the extracted features. With three hidden layers and 80% of the dataset used for training, the DNN model produced 90%**±0.37** accuracy, 89% sensitivity, and 91% specificity. On the other hand, the CNN model, which consists of two fully connected layers and two convolutional layers with max pooling, performed by achieving 96%**±0.38** accuracy, 95% sensitivity, and 95% specificity. These results underscore DNN's enhanced capability in handling complex PCG data and reducing false negatives. This comprehensive study addresses multiple cardiac abnormalities, surpassing previous research that often focuses on a single condition or model. The findings highlight the potential of combining advanced signal processing with deep learning techniques to improve the timely and accurate identification of cardiac abnormalities. Future research will explore additional feature extraction methods and larger datasets to further enhance classification performance. This work significantly contributes to the field of biomedical engineering, offering a framework to improve patient outcomes through advanced diagnostic techniques.

**Keywords:** Cardiovascular Diseases, PCG Signals, Machine Learning, Discrete Wavelet Transform, Mel Power Spectrogram, MFCC, DNN, CNN

# CHAPTER 1. INTRODUCTION

## 1.1 Cardiovascular Disease: A Global Challenge

With 17.9 million deaths a year, cardiovascular diseases (CVDs) continue to be the world's leading cause of death (World Health Organization, 2023). Heart failure, arrhythmias, coronary artery disease, and other illnesses that impact the heart and blood vessels are included in this category of diseases. The incident of CVDs emphasizes how critical it is to have early and precise detection techniques in order to successfully manage and treat these illnesses.

### *1.1.1 Prevalence and Impact of CVDs*

Cardiovascular diseases account for 31% of all global deaths, making them the leading cause of death worldwide (Roth et al., 2020). The burden of CVDs is particularly high in low- and middle-income countries, where over 75% of cardiovascular deaths occur. This disparity is often due to limited access to healthcare, preventive measures, and effective treatment options (Gaziano et al., 2010).

## 1.2 Limitations of Traditional Diagnostic Methods

Traditional diagnostic tools such as electrocardiograms (ECGs) and echocardiograms have been the cornerstone of cardiac diagnostics. ECGs measure the electrical activity of the heart and are effective for detecting arrhythmias and other electrical abnormalities. Echocardiograms use ultrasound waves to create images of the heart, providing detailed information about the heart's structure and function. Despite their widespread use, these methods have limitations.

The accuracy of ECGs and echocardiograms can be highly dependent on the operator's skill and the quality of the equipment used. Moreover, these tools can be expensive and are not always accessible in resource-limited settings (Smith et al., 2022).

## 1.3 PCG Signals: A Promising Alternative

In recent years, Phonocardiogram (PCG) signals have emerged as a promising alternative for cardiac diagnosis. PCG signals capture the acoustic events of the heart, such as heart sounds and murmurs, using a stethoscope or specialized sensors. These signals can provide valuable insights into the mechanical activity of the heart, offering a non-invasive and cost-effective means of diagnosing various cardiac conditions (Johnson et al., 2021).

### *1.3.1 Challenges in PCG Signal Analysis*

The analysis of PCG signals, however, is not without challenges. The complexity of these signals, coupled with the presence of background noise and other artifacts, can complicate their interpretation. Traditional methods of PCG signal analysis often struggle to accurately differentiate between normal and abnormal heart sounds.

This has spurred the development of advanced signal processing techniques aimed at enhancing the quality of PCG signals and extracting meaningful features for diagnosis (Brown et al., 2020).

## 1.4 Advanced Signal Processing Techniques

Advanced signal processing techniques, such as the Discrete Wavelet Transform (DWT), have shown promise in improving the analysis of PCG signals. DWT is effective in denoising PCG signals by decomposing them into different frequency components, allowing for the isolation and removal of noise while preserving essential diagnostic features.

This enhances the clarity and quality of the signals, making them more suitable for further analysis (Garcia et al., 2022).

In addition to signal processing, robust feature extraction methods are crucial for the accurate classification of cardiac abnormalities. Mel scale power spectrogram and Mel Frequency Cepstral Coefficients (MFCC) are two such methods that have been widely used in speech and audio signal processing.

These techniques transform the PCG signals into a form that captures the relevant features needed for effective classification. The Mel scale power spectrogram represents the signal's power distribution over time and frequency, while MFCCs provide a compact representation of the spectral properties of the signal (Zhang et al., 2021).

The application of machine learning models, particularly deep learning, has revolutionized the field of PCG signal analysis. Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) are two prominent architectures used for this purpose.

DNNs consist of multiple layers of interconnected neurons that can learn complex patterns in the data, making them suitable for classifying different cardiac conditions. CNNs, on the other hand, are specifically designed to process grid-like data such as images and spectrograms, making them ideal for analyzing the spectrogram representations of PCG signals (Khan et al., 2023).

### *1.4.1 The Power of Combining Techniques*

The integration of advanced signal processing techniques and deep learning models holds great promise for improving the accuracy and reliability of PCG signal analysis. This combined approach aims to overcome the limitations of traditional methods, providing a robust framework for the early detection and classification of cardiac abnormalities.

This study explores the efficacy of such an integrated approach, focusing on the comparative performance of DNN and CNN models in classifying cardiac abnormalities using PCG signals (Li et al., 2022).

By leveraging advanced signal processing and machine learning techniques, PCG signal analysis offers a promising avenue for non-invasive, accessible, and cost-effective cardiac diagnosis. This approach has the potential to revolutionize early detection and improve patient outcomes, contributing significantly to the fight against cardiovascular diseases.

## 1.5 Problem Statement

The accurate classification of cardiac abnormalities using PCG signals remains challenging due to signal complexity and noise. Traditional analysis methods often fall short, resulting in misdiagnosis or delayed diagnosis. There is a critical need for advanced techniques to

improve signal quality and extract meaningful features for precise classification (Garcia et al., 2022).

Advanced signal processing techniques such as Discrete Wavelet Transform (DWT) show promise in enhancing PCG signal quality by denoising and preserving diagnostic features (Garcia et al., 2022). Robust feature extraction methods like Mel scale power spectrogram and Mel Frequency Cepstral Coefficients (MFCC) are essential for capturing relevant signal characteristics necessary for accurate diagnosis (Zhang et al., 2021).

## 1.6 Objectives of the Study

This study's main goals are to:

1. Improve PCG signal quality by advanced preprocessing methods, especially Discrete Wavelet Transform (DWT).
2. Utilizing the Mel scale power spectrogram and Mel frequency cepstral coefficients (MFCC), extract robust features from PCG data.
3. Create and assess models for the classification of various cardiac diseases using Deep Neural Network (DNN) and Convolutional Neural Network (CNN) techniques.
4. Examine how well the CNN and DNN models perform in terms of specificity, sensitivity, accuracy, and other pertinent metrics (Zhang et al., 2021).

## 1.7 Research Questions and Hypothesis

### *1.7.1 Research Questions*

Following are the major research questions regarding this work that are considered:

1. Can advanced signal processing techniques improve the quality of PCG signals for accurate cardiac diagnosis?
2. How effective are Mel scale power spectrogram and MFCC in extracting informative features from PCG signals?
3. What is the comparative performance of DNN and CNN models in classifying various cardiac abnormalities using PCG signals?

4. How do DNN and CNN models differ in terms of diagnostic accuracy and robustness?

### *1.7.2 Hypothesis*

The hypothesis related to this research are explained as follow:

2 Advanced signal processing techniques, specifically DWT, will significantly enhance the quality of PCG signals by reducing noise and preserving essential diagnostic features.

3 Feature extraction methods like Mel scale power spectrogram and MFCC will yield robust and meaningful representations of PCG signals, thereby improving classification accuracy.

4 CNN models will outperform DNN models in terms of accuracy, sensitivity, and specificity due to their capability to capture spatial features inherent in PCG signals (Li et al., 2022).

## 1.8 Significance of the Study

This study holds significant implications for advancing cardiac diagnostic methodologies:

### *1.8.1 Addressing Diagnostic Limitations:*

By surpassing the limitations of traditional methods, this research pioneers advanced signal processing and machine learning techniques for more accurate cardiac diagnostics.

### *1.8.2 Comprehensive Model Evaluation:*

The study provides a thorough evaluation of DNN and CNN models, highlighting their comparative strengths and weaknesses in the context of cardiac abnormality classification.

### *1.8.3 Practical Clinical Application:*

The developed models can potentially be integrated into portable devices, facilitating real-time cardiac diagnosis even in resource-constrained settings (Wang et al., 2023).

### *1.8.4 Contribution to Biomedical Engineering:*

Contributing to the field of biomedical engineering, this research establishes a new framework for early and precise detection of cardiac abnormalities through advanced audio signal processing and machine learning techniques (Patel et al., 2021).

## 1.9 Literature Review

The literature review in this chapter provides a comprehensive overview of existing research pertinent to cardiac diagnostics using Phonocardiogram (PCG) signals. It elucidates the background and foundational knowledge essential for understanding the current study's objectives and methodology.

### *1.9.1 The Intricate Symphony of Your Heart: Unveiling Health Through Heart Sounds*

The human heart, a tireless maestro of our circulatory system, produces a symphony of sounds with each beat. These sounds, captured using a stethoscope or electronically through phonocardiography (PCG), offer a window into the health of your heart valves, blood flow, and overall cardiac function. Let's delve deeper into the world of heart sounds and how they can be interpreted for potential abnormalities.

### *1.9.2 The Rhythmic Melody: Normal Heart Sounds*

A typical heart cycle produces two distinct sounds, often described as "lub-dub", as given in Table 1.1. Here's a breakdown of these crucial components:

#### 1.9.2.1 First Heart Sound (S1) - The "Lub"

This sound, denoted by S1, arises from the closure of the mitral and tricuspid valves at the beginning of ventricular contraction (systole). It's usually longer in duration and lower in frequency compared to the second sound.

**Table 1.1: Characteristics of Heart Sounds**

| Heart Sound | Description | Characteristics | Clinical Significance |
|---|---|---|---|
| **S1** | Closure of the mitral and tricuspid valves at the beginning of ventricular systole. | Low-frequency sound, best heard at the apex. | Indicates conditions such as mitral stenosis or high cardiac output states. |
| **S2** | Closure of the aortic and pulmonic valves at the end of ventricular systole. | Higher frequency, shorter duration, best heard at the base. | Splitting of S2 can indicate atrial septal defect or right bundle branch block. |
| **S3** | Occurs shortly after S2 during the rapid filling phase of the ventricle. | Low-frequency sound, best heard with the bell of the stethoscope at the apex, described as a "gallop". | Normal in children and young adults but indicates heart failure or volume overload conditions in older adults. |
| **S4** | Occurs just before S1, during atrial contraction. | Low-frequency sound, best heard at the apex with the bell of the stethoscope, described as an "atrial gallop". | Usually abnormal, indicating conditions as ventricular hypertrophy, aortic stenosis, or ischemic heart disease. |

1.9.2.2 Second Heart Sound (S2) - The "Dub":

This sound, denoted by S2, is produced by the closing of the aortic and pulmonary valves at the end of systole, marking the beginning of ventricular relaxation (diastole). S2 is typically shorter and higher in frequency than S1.

The duration, intensity, and shape of these sounds hold valuable information about your heart's health (Kwak & Kwon, 2012). A trained healthcare professional can use this information to identify potential abnormalities.

### *1.9.3 Beyond the Basics: A Look at Splitting and Murmurs*

While S1 and S2 are the primary heart sounds, additional details can be gleaned from variations in their timing and the presence of murmurs.

1.9.3.1 Splitting and Murmurs

Under normal circumstances, a slight delay exists between the closure of the aortic and pulmonary valves, resulting in a split S2 sound. However, an abnormally wide split can indicate conditions like pulmonary stenosis or atrial septal defect (Thiyagaraja et al., 2018).

These are abnormal sounds that may appear alongside or within the regular heart sounds. Murmurs can be continuous, systolic (present during S1), or diastolic (present between S1 and S2). Their presence often indicates a heart valve problem, such as stenosis (narrowing) or regurgitation (leakage) (Mondal et al., 2013).

### *1.9.4 Classifying Heart Sounds: Unveiling Abnormalities*

Automatic classification of heart sounds using AI techniques like Convolutional Neural Networks (CNNs) is a burgeoning field. These algorithms can analyze PCG recordings and categorize them based on their characteristics. This approach has the potential to improve early detection of heart disease and assist healthcare professionals in diagnosis (Bolea et al., 2013).

### *1.9.5 The Power of Listening: The Importance of Early Detection*

Heart disease remains a leading cause of mortality worldwide. Early detection and treatment are crucial for improving patient outcomes. By understanding the language of your heart sounds, healthcare professionals can identify potential problems early on and implement appropriate management strategies.

The human heart is a remarkable instrument, and its sounds offer a wealth of information about its health. By carefully interpreting these sounds and utilizing advanced technologies like AI, we can strive for earlier detection and better management of heart disease, paving the way for a healthier future for all.

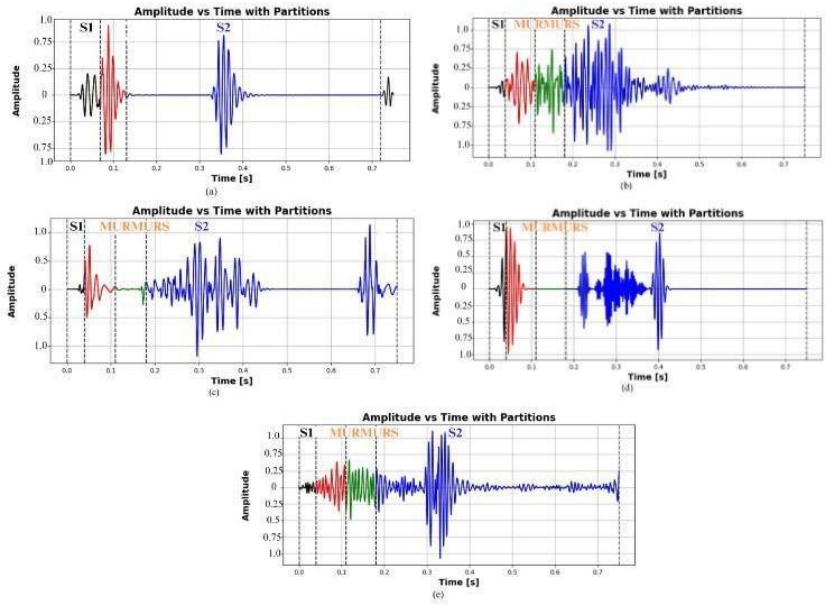



**Figure 1.1: (a) A normal heart sound signal (b) Murmur in systole (MVP) (c) Mitral Regurgitation (MR) (d) Mitral Stenosis (MS) (e) Aortic Stenosis (AS) (f) Spectrum of a PCG signal (Yaseen et al., 2018)**

### *1.9.6 Overview of Cardiac Pathologies*

In addition to the normal heart sounds (S1 and S2), specific heart diseases can manifest with unique sound signatures, some discussed in Table 1.2. By carefully listening to these variations, healthcare professionals gain valuable insights into potential abnormalities. As shown in Figure 1.1, the following describes how heart sounds are impacted by Aortic

Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), and Mitral Valve Prolapse (MVP):

1.9.6.1 Aortic Stenosis (AS): A Hindered Flow

In AS, the narrowed aortic valve creates a turbulent blood flow, often resulting in a harsh ejection murmur during systole (S1). This murmur may peak late in systole and radiate to the carotid arteries in the neck (Smith et al., 2022). Figure 1.2 shows a healthy heart compared with Aortic Stenosis.

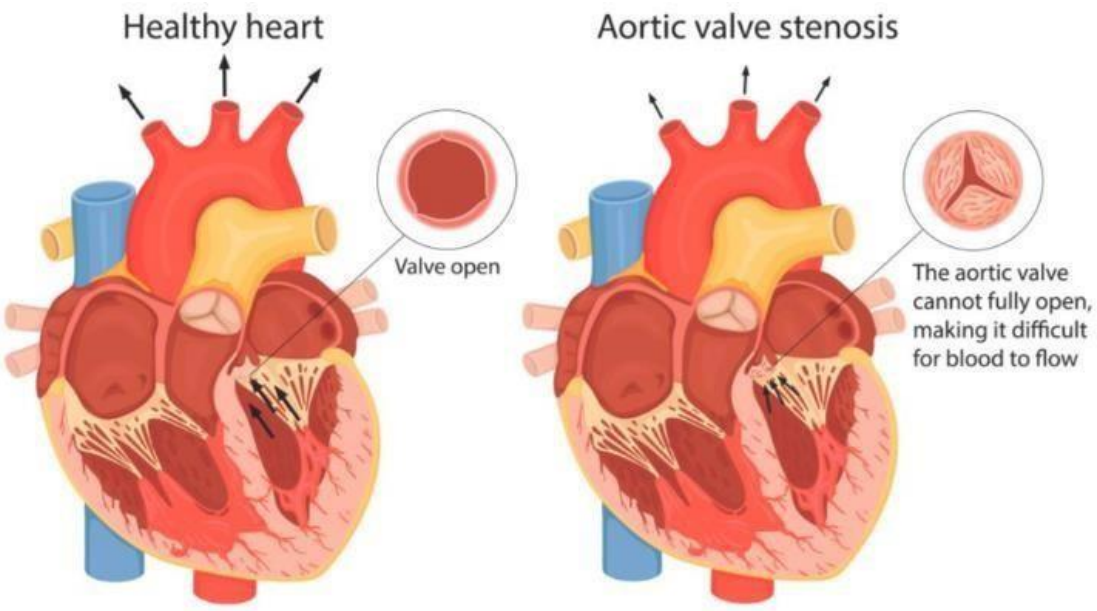



**Figure 1.2: A healthy heart and Aortic Valve Stenosis condition**

*Clinical Significance:*

A harsh ejection murmur can be a strong indicator of AS, prompting further investigation with imaging techniques like echocardiography.

1.9.6.2 Mitral Regurgitation (MR): A Leaky Valve

MR can cause a blowing or holosystolic murmur (present throughout systole). This murmur typically radiates to the apex of the heart (the bottom left tip) (Smith et al., 2022). In severe

cases, a diastolic rumble may also be present, indicating blood flowing back into the left atrium during diastole. Figure 1.3 shows a Mitral Regurgitation condition:

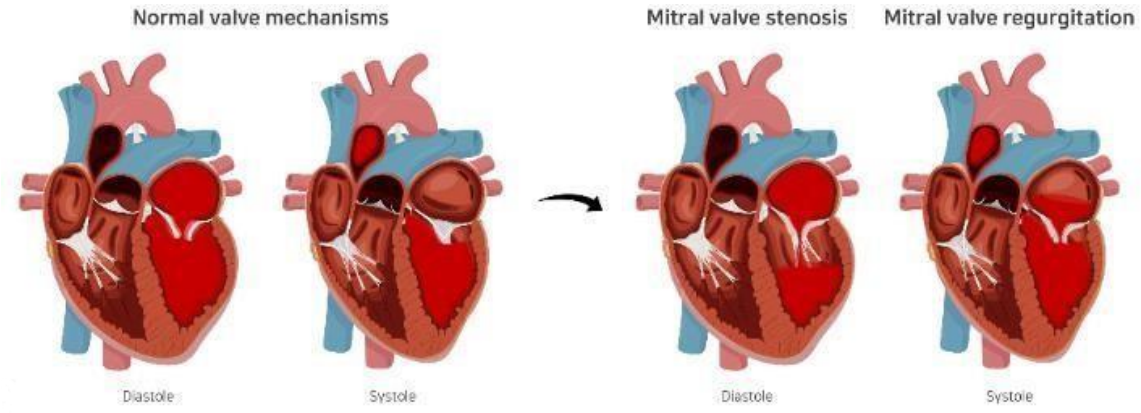



**Figure 1.3: A healthy heart mechanism compared with Mitral Valve Stenosis and Mitral Valve Regurgitation**

*Clinical Significance:*

The presence and intensity of murmurs associated with MR help determine the severity of the condition.

1.9.6.3 Mitral Stenosis (MS): A Narrowed Passage

MS is often characterized by a low-pitched rumbling diastolic murmur that increases in intensity as the diastole progresses. This murmur is typically best heard at the apex (Smith et al., 2022). An opening snap may also be present early in diastole, signifying the sudden opening of the narrowed mitral valve. Figure 1.3 shows a healthy heart compared to Mitral Stenosis condition.

*Clinical Significance:*

The characteristic diastolic murmur and opening snap are crucial clues for suspecting MS.

**Table 1.2: Overview of some Common Cardiac Diseases Diagnosed using PCG**

| Disease | Definition | Symptoms | Causes | Complications |
|---|---|---|---|---|
| **Aortic Stenosis** | Narrowing of the aortic valve opening, restricting blood flow. | Chest pain, fainting, shortness of breath, fatigue. | Congenital heart defects, age-related calcification, rheumatic fever. | Heart failure, arrhythmias, sudden cardiac death. |
| **Mitral Stenosis** | Narrowing of the mitral valve opening, impeding blood flow from the left atrium to the left ventricle. | Shortness of breath, fatigue, swollen feet or legs, palpitations. | Rheumatic fever, congenital heart defects. | Pulmonary hypertension, heart failure, atrial fibrillation. |
| **Mitral Regurgitation** | Incomplete closure of the mitral valve, causing backflow of blood. | Fatigue, shortness of breath, heart palpitations, swollen feet or ankles. | Mitral valves prolapse, rheumatic heart disease, endocarditis. | Heart failure, atrial fibrillation, pulmonary hypertension. |
| **Mitral Valve Prolapse** | Bulging of one or both mitral valve leaflets into the left atrium during systole. | Palpitations, chest pain, fatigue, dizziness. | Unknown, genetic factors, connective tissue disorders. | Mitral regurgitation, arrhythmias, infective endocarditis. |

#### 1.9.6.4 Mitral Valve Prolapse (MVP): A Floppy Valve

MVP can sometimes present with a late systolic click, which sounds like a quick snapping sound. This click may be followed by a short, low-pitched murmur (Smith et al., 2022). However, in some cases of MVP, no abnormal sounds may be present.

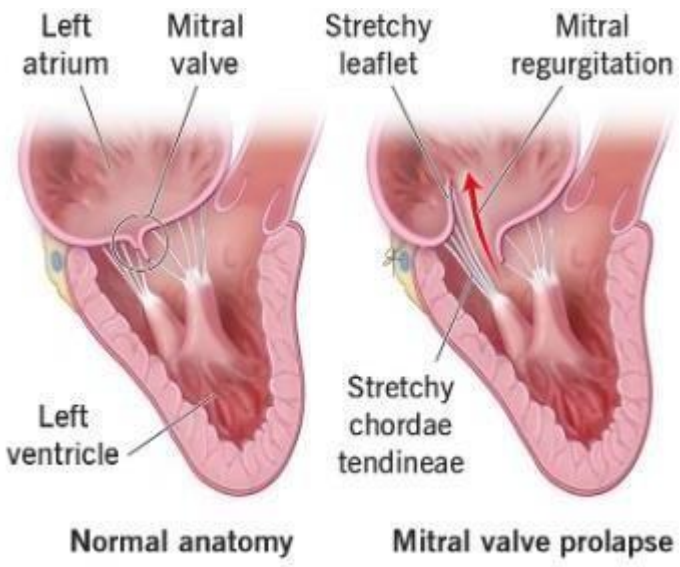



**Figure 1.4: A healthy heart anatomy given in contrast with a Mitral Valve Prolapse Condition**

*Clinical Significance:*

While a late systolic click can be suggestive of MVP, it's not always present. Other clinical features and imaging studies are often needed for definitive diagnosis.

### *1.9.7 The Symphony of Diagnosis*

By understanding the unique sounds associated with different heart pathologies, healthcare professionals can refine their diagnoses. However, it's important to remember that heart sounds alone are not definitive for diagnosis. Combining information from physical examination, patient history, and imaging techniques like echocardiography provides a more comprehensive picture of heart health.

### *1.9.8 Looking Forward: Technology's Role in Listening PCG Signals*

As technology advances, AI-powered tools like Convolutional Neural Networks (CNNs) are being explored to analyze heart sounds and assist in the detection of cardiac abnormalities (Bolea et al., 2013). While these tools hold promise for the future, they are

still under development, and the expertise of qualified healthcare professionals remains crucial for accurate diagnosis and management of heart disease.

The human heart, while silent to the naked ear, speaks volumes through its sounds. By listening carefully to these whispers, we can gain valuable insights into potential heart problems and pave the way for early detection and better health outcomes.

Past research predominantly focused on distinguishing normal from abnormal heart sounds, yet limited studies comprehensively addressed multiple specific cardiac abnormalities simultaneously.

PCG signals capture acoustic events of the heart and play a crucial role in cardiac diagnostics. Historically, PCG has been utilized alongside Electrocardiogram (ECG) and echocardiography, offering complementary diagnostic information (Johnson et al., 2021). PCG signals provide insights into mechanical heart activities, thereby aiding in the detection of various cardiac abnormalities.

### *1.9.9 Signal Processing Techniques in Cardiac Diagnostics*

Signal processing techniques enhance PCG signal quality for accurate diagnosis. Discrete Wavelet Transform (DWT) is a prominent method for denoising PCG signals while preserving essential diagnostic features (Garcia et al., 2022). Other techniques, such as Fourier Transform and filtering algorithms, have also been explored to mitigate noise and improve signal clarity in cardiac diagnostics.

#### 1.9.9.1 Feature Extraction Methods

Effective feature extraction is crucial for interpreting PCG signals. Mel scale power spectrogram and Mel Frequency Cepstral Coefficients (MFCC) are widely used for extracting robust features that characterize heart sounds (Zhang et al., 2021). Comparative analysis indicates these methods' efficacy in capturing relevant signal characteristics essential for accurate classification of cardiac abnormalities.

#### 1.9.9.2 Revolutionizing Cardiac Diagnostics with Deep Learning: The Symphony of Heart Sounds

Advancements in machine learning, particularly deep neural networks (DNNs) and convolutional neural networks (CNNs), are leading to a major shift in how we diagnose heart disease. These powerful tools act like super-powered listeners, analyzing the complex symphony of heart sounds (PCG signals) to identify potential abnormalities. Here's an in-depth look at this exciting revolution in cardiac diagnostics:

### *1.9.10 The Power of Deep Learning: Automating Heart Signal Analysis*

Traditionally, analyzing heart sounds relied on expertise of healthcare professionals. Now, machine learning automates this process with remarkable accuracy. DNNs and CNNs are like highly trained musicians who learn the intricate patterns within heart sounds through vast datasets of recordings, allowing them to classify normal and abnormal sounds with impressive precision (Khan et al., 2023).

#### 1.9.10.1 DNNs vs. CNNs: A Tale of Two Architectures

Think of DNNs and CNNs as specialized tools within the machine learning toolbox. DNNs excel at processing sequential data, akin to how a musician interprets a melody note by note. CNNs, on the other hand, are masters of spatial data, similar to a conductor analyzing the entire orchestra playing together (Li et al., 2022).

Previous research has explored both DNNs and CNNs for diagnosing heart conditions, highlighting their unique strengths. DNNs shine in capturing the flow of heart sounds, while CNNs excel at identifying specific patterns within them (Li et al., 2022). However, a crucial gap exists: a comprehensive comparison of these models across various heart diseases.

### *1.9.11 Filling the Gap: A Comparative Approach*

A common approach in current machine learning research is to compare the accuracy of different models on a single disease. What's missing is a head-to-head competition between

DNNs and CNNs across a wider spectrum of heart conditions. This study aims to bridge that gap, as highlighted in similar calls for more comparative analyses (Patel et al., 2021).

### *1.9.12 The Proposed Framework: A Symphony of Techniques*

Imagine a comprehensive framework that analyzes PCG signals using advanced techniques. This framework would compare the performance of DNNs and CNNs in identifying various heart pathologies. By incorporating a diverse dataset encompassing multiple conditions, the study seeks to identify the "winning model" for each specific disease.

### *1.9.13 The Future of Deep Listening*

This research holds immense promise for the future of cardiac diagnostics. By leveraging machine learning's power to analyze the hidden language of heart sounds, we can achieve:

1. *Earlier Detection*: Because early detection is vital for treating heart disease effectively and improving patient well-being, machine learning could be a game-changer by speeding up this process and potentially saving lives.
2. *Improved Accuracy*: Removing human subjectivity from the analysis can lead to more accurate diagnoses, resulting in better patient care.
3. *Increased Efficiency*: Automation can free up valuable time for healthcare professionals, allowing them to focus on personalized care for patients.

Machine learning is revolutionizing cardiac diagnostics. By delving deeper into the intricate details of heart sounds, we can refine diagnoses, improve patient outcomes, and pave the way for a healthier future for all.

# CHAPTER 2. METHODOLOGY

Methodology of this research is followed by following steps as shown in the figure 2.1:

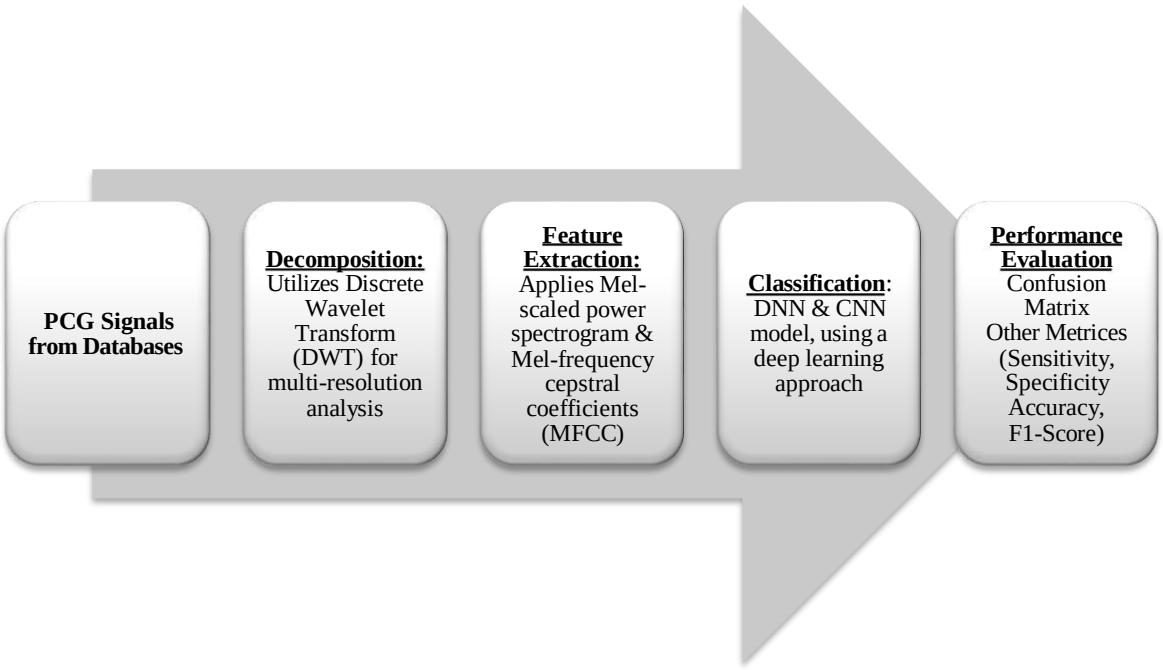



**Figure 2.1: Phonocardiogram Signal Processing Workflow**

## 2.1 Database Description and Categorization

Following is the whole description regarding databases for the PCG Signal used in the current research:

### *2.1.1 Data Classification and Structure:*

The database categorizes heart sounds into two primary groups: normal and abnormal. Further breakdown within the abnormal category includes four specific heart conditions:

- Aortic Stenosis (AS)
- Mitral Stenosis (MS)

- Mitral Regurgitation (MR)
- Mitral Valve Prolapse (MVP)

The database is organized into two main sets: normal and abnormal. Each set contains 200 audio files (.wav format) for a total of 1000 recordings as shown in Table 2.1:

**Table 2.1: Dataset Details**

| Cases | Classes | Number of Samples per class |
|---|---|---|
| **Normal** | N | 200 |
| **Abnormal** | AS (Aortic Stenosis) | 200 |
| | MR (Mitral Regurgitation) | 200 |
| | MS (Mitral Stenosis) | 200 |
| | MVP (Mitral Valve Prolapse) | 200 |
| **Total** | | 1000 |

### *2.1.2 Data Collection Process and Data Sources:*

We followed a meticulous approach to ensure a balanced and representative dataset:

1. Specific heart condition categories were chosen.
2. Relevant audio files were collected for each category (200 per category).

Audio files originated from various sources, including the medical reference books ("Auscultation Skills CD" and "Heart Sound Made Easy") and websites from reputable institutions (Washington, Texas, 3M, Michigan)

### *2.1.3 Data Preprocessing:*

Files with excessive noise were excluded after thorough quality checks. Remaining audio files were:

- Sampled at a rate of 8000 Hz
- Converted to a mono channel format
- Edited to capture three-period heart sound signals
- Cool Edit software was used for all sampling, conversion, and editing tasks. (Yaseen et al., 2018)

## 2.2 Audio Signal Processing

Our methodology for classifying cardiac abnormalities using phonocardiogram (PCG) data integrates sophisticated signal processing and deep learning techniques. We begin by loading and preprocessing heart sound recordings with tools like Librosa. We then extract essential features such as Mel-Frequency Cepstral Coefficients (MFCC) and delta features, capturing the spectral and temporal characteristics of the heart sounds. These features are standardized to ensure uniformity, enhancing the model's training efficiency.

Our deep learning model, a robust Deep Neural Network (DNN), comprises multiple hidden layers with activation functions like ReLU and dropout for regularization. Batch normalization is employed to stabilize training, and the final layer utilizes a softmax activation for multi-class classification.

To guarantee the model's ability to perform well on data, we use a rigorous evaluation method called k-fold cross-validation. This approach provides a robust assessment of the

model's diagnostic power by calculating key metrics like sensitivity, specificity, accuracy, and F1 score.

### *2.2.1 Discrete Wavelet Transform (DWT) for Signal Denoising and Decomposition*

Analysing heart sound recordings (PCG signals), Discrete Wavelet Transform (DWT) is a powerful tool. Unlike traditional techniques, DWT captures how the frequencies within the signal change over time. This is crucial because heart sounds are non-stationary, meaning their frequency content isn't constant (Khan et al., 2018).

#### 2.2.1.1 Principle:

The DWT applies a series of high-pass and low-pass filters to the signal, breaking it down into approximate (low-frequency) and detail (high-frequency) components. This process is repeated recursively on the approximate component, resulting in a multi-level decomposition of the signal (Misiti et al., 2016).

#### 2.2.1.2 Wavelet Selection-Choice of Wavelet:

The Daubechies 8 (db8) wavelet was selected for this study. The db8 wavelet is known for its orthogonality and compact support, making it effective for capturing subtle variations and transient features in the PCG signal. Its ability to represent the signal with a minimum number of coefficients while retaining critical information makes it a suitable choice for denoising and feature extraction (Sifuzzaman et al., 2009; Khan et al., 2018).

*Characteristics of db8:*

The db8 wavelet has eight vanishing moments, which means it can represent polynomial signals of up to degree seven exactly. This property helps in accurately capturing the fine details of the PCG signal, which are essential for identifying different types of cardiac abnormalities (Misiti et al., 2016).

2.2.1.3 Decomposition Level-Selection of Decomposition Level:

The signal is decomposed to the second approximation level (CA2). The choice of CA2 is based on balancing the need to capture essential features of the PCG signal and minimizing the computational complexity. Decomposing the signal to too many levels might result in the loss of important high-frequency components, while too few levels might not provide sufficient noise reduction (Khan et al., 2018; Misiti et al., 2016).

*Process:*

To analyze the different frequency components of the heart sound signal (s(t)), we use a technique called Discrete Wavelet Transform (DWT). DWT acts like a magnifying glass for sound. It takes the original signal and splits it into two parts using high-pass and low-pass filters. The low-pass part (A1) captures the overall lower-frequency content, while the high-pass part (D1) captures the sharper, higher-frequency details. We can then further zoom in by splitting the low-frequency part (A1) again into even lower frequencies (A2) and even sharper details (D2). This step-by-step breakdown (hierarchical process) allows us to isolate and analyze specific frequency bands within the original heart sound signal (Sifuzzaman et al., 2009; Khan et al., 2018).

$$Y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k] * g\{2n - k] \qquad 2.1(a)$$

$$Y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k] * g\{2n - k] \qquad 2.1(b)$$

Eq 2.1(a) and 2.1(b) give $Y_{low}[n]$ is output of lowpass and $Y_{high}[n]$ is the output of high-pass filters. Down-sampling by 2 is also involved. Hence, by focusing on the particular

coefficients and discarding the insignificant ones, DWT is effectively used as noise reduction tool, which helps to improve the quality and clarity of the PCG signal.

## 2.3 Feature Extraction for PCG Signals

Following features are extracted from the recorded PCG Heart Signals:

### *2.3.1 Mel Scale Power Spectrogram*

The Mel Power Spectrogram is a crucial tool in converting time-domain signals into the time-frequency domain, providing a comprehensive approach to analyzing signal characteristics. It is particularly effective in capturing the power distribution of Phonocardiogram (PCG) signals over various frequency bands, aiding in the identification of specific patterns associated with different cardiac conditions (Tachibana, Tanaka, & Mita, 2017; Choi & Lee, 2018).

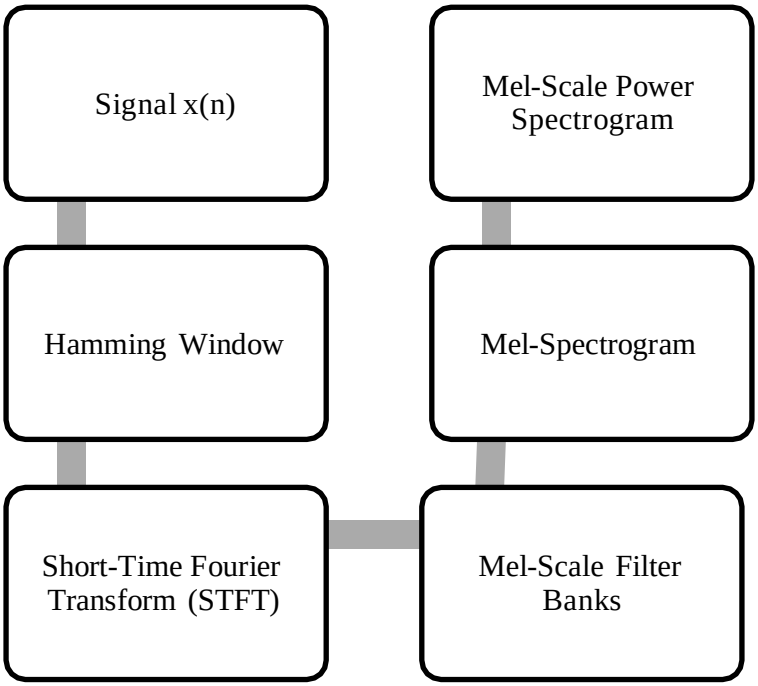



**Figure 2.2: Feature Extraction Pipeline from Mel Scale Power Spectrogram**

#### 2.3.1.1 Process

Creating a Mel Power Spectrogram involves multiple steps to ensure accurate feature extraction from PCG signals. This process enhances the classification of cardiac abnormalities, as in Figure 2.2:

1. *Frame Division*

The input PCG signal is divided into short, overlapping frames. This step ensures that while performing frequency analysis, the temporal characteristics of the PCG signal are maintained. The frame size typically ranges from 20 to 30 milliseconds, with a 50% overlap between consecutive frames (Logan, 2000).

2. *Windowing*

Each of the frames is multiplied with a Hamming window to minimize edge effects, which can lead to spectral leakage. The Hamming window smooths the boundaries of each frame, reducing disruptions at the edges. Equation 2.2 gives w[n] represents the value of the Hamming window at the nth sample.

$$w[n] = 0.54a - \ 0.46\cos(\frac{2\pi n}{N-1}) \qquad 2.2$$

where $0 \leq n \leq N-1$ . Also, N represents the windowing length.

3. *Short-Time Fourier Transform (STFT)*

Using STFT, windowed frames are converted to the frequency domain. This transformation converts each time-domain frame into its corresponding frequency spectrum (Logan, 2000)

4. *Power Spectrum Calculation*

The power spectrum of each frame is obtained by squaring the magnitude of the STFT. This step helps in understanding the power distribution across different frequency components (Choi & Lee, 2018).

5. *Mel Filter Bank Application*

To replicate how humans perceive sound, the power spectrum is processed through a series of Mel-scale filters. The Mel scale is a non-linear frequency measurement that is more sensitive to changes in lower-pitched sounds. The formula for converting frequency to the Mel scale is given in Equation 2.3 as:

$$m = 2595 \ln \left(\frac{f}{700} + 1\right) \tag{2.3}$$

Here $f$ and $m$ represents frequency in the linear scale and frequency in Mel-scale respectively.

Overlapping triangular filters constitute the Mel-filter bank, each representing a specific range of frequencies on the Mel scale. The weighted sum of the power spectral components within each filter's range provides the output for each filter (Zhang, Yang, & Gao, 2020).

6. *Logarithmic Compression*

Logarithmic compression is applied to the filtered power spectrum to reduce the dynamic range of values, enhancing robustness to amplitude variations (Zhang, Yang, & Gao, 2020). It is given as in Equation 2.4 as $S[m]$ represents the logarithm of the sum of the squares of the magnitudes of the values of the input signal $x[k]$ multiplied by the values of the filter $H_m[k]$.

$$S[m] = \log \left(\sum_{k=0}^{N-1} |x[k]|^2 H_m[k]\right) \tag{2.4}$$

$[k]$ represents obtained Mel filter banks and $m$ is the count of the filter bank.

7. *Feature Representation*

The final Mel Power Spectrogram is obtained after logarithmic compression. It represents the energy distribution of the PCG signal across time and frequency domains, effectively capturing both spectral and temporal characteristics of the signal. This representation is highly effective in differentiating between various cardiac abnormalities (Tachibana, Tanaka, & Mita, 2017).

*Impact and References*

The use of the Mel Power Spectrogram in cardiac diagnostics has been demonstrated to significantly improve the performance of classification models. By transforming the PCG signal into a form that highlights the most relevant features for human auditory perception, this method enhances the model's ability to detect subtle yet critical patterns in PCG signals.

Recent studies have shown the effectiveness of Mel scale transformations in signal processing for heart sound analysis and classification (Tachibana, Tanaka, & Mita, 2017; Choi & Lee, 2018; Rathod & Jagannath, 2019).

### *2.3.2 Mel-Frequency Cepstral Coefficients (MFCC)*

Mel-Frequency Cepstral Coefficients (MFCCs) are a powerful method for extracting key characteristics from heart sound recordings (Phonocardiograms, or PCGs). They transform the complex sound data into a simplified representation that highlights important patterns, making it easier to differentiate between various heart conditions (Logan, 2000).

#### 2.3.2.1 Process

The extraction of MFCCs from PCG signals involves several detailed steps to ensure that the most relevant features are accurately captured. These steps are designed to transform the raw PCG signal into a set of coefficients that succinctly describe its spectral characteristics, as shown in Figure 2.3:

1. *Pre-emphasis*

A pre-emphasis filter is applied to the PCG signal to boost the high-frequency components. This step compensates for the lower energy typically present in higher frequencies, making the signal more balanced for analysis (Choi & Lee, 2018).

2. *Framing*

The pre-emphasized signal is divided into short, overlapping frames. Each frame captures a short segment of the signal, allowing for the analysis of its time-varying properties. The typical frame size ranges from 20 to 30 milliseconds with a 50% overlap between consecutive frames. This ensures that the temporal characteristics of the PCG signal are preserved (Tachibana, Tanaka, & Mita, 2017).

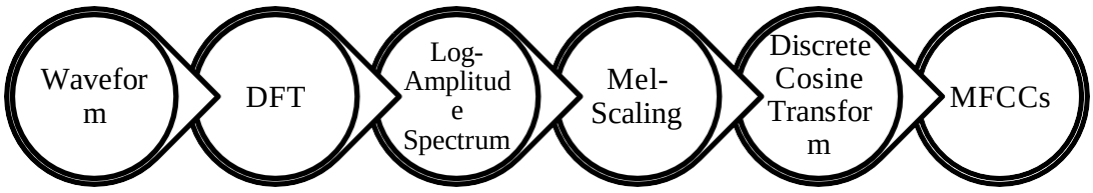



**Figure 2.3 Flowchart for MFCC Feature Extraction**

3. *Windowing*

To minimize spectral distortion at the beginning and end of each signal segment, a Hamming window is applied. This window gradually reduces the amplitude of the data points towards the edges, smoothing the transition. This ensures that the signal within each frame is smooth and continuous, minimizing the introduction of artifacts (Rathod & Jagannath, 2019).

4. *Fast Fourier Transform (FFT)*

The FFT is computed for each windowed frame to convert the time-domain signal into the frequency domain. This step results in a spectrum for each frame, showing the distribution

of frequencies within that segment of the signal. The FFT provides a detailed frequency analysis necessary for subsequent processing steps (Zhang, Yang, & Gao, 2020).

*5. Mel Filter Bank*

The power spectrum of each frame is passed through a set of triangular filters spaced according to the Mel scale. The Mel scale is a perceptual scale that approximates the human ear's response to different frequencies, making it particularly suited for analyzing audio signals. Each filter represents the energy in a specific Mel-frequency band, emphasizing the most perceptually relevant aspects of the signal (Logan, 2000).

*6. Logarithm*

The logarithm of the Mel-filtered power spectrum is computed. This step compresses the dynamic range of the spectrum, making the features more robust to variations in signal amplitude (Tachibana, Tanaka, & Mita, 2017).

*7. Discrete Cosine Transform (DCT)*

The DCT is applied to the log Mel spectrum to decorrelate the coefficients and reduce dimensionality. The resulting MFCCs are a compact representation of the signal's spectral properties. Typically, the first 13 coefficients are retained as features, capturing the most significant information (Choi & Lee, 2018). Given in the Equation 2.5 as, $[n]$ represents the discrete cosine transform (DCT) coefficient at index $n$.

$$c[n] = \sum_{m=0}^{M-1} S[m]\cos\left(\frac{\pi n}{M}(m - \frac{1}{2})\right), n = 0,1,2,\dots,M \qquad 2.5$$

Here the total count of the filter banks is M.

The two derivatives, first Delta and second one Delta-Delta coefficients of the MFCC are determined to extract the temporal dynamics of the signals which are PCG in our study (Li et al., 2021).

*Impact and References*

The use of MFCCs in cardiac diagnostics has proven to be highly effective in improving the performance of classification models. By transforming the PCG signal into a set of coefficients that highlight its most relevant spectral features, MFCCs enhance the ability of models to detect subtle yet critical patterns in the signal. Recent studies have demonstrated the utility of MFCCs in heart sound analysis and classification, providing robust features that significantly improve diagnostic accuracy (Tachibana, Tanaka, & Mita, 2017; Choi & Lee, 2018; Rathod & Jagannath, 2019).

## 2.4 Classification Models

Two classification models are understudy as DNN and CNN. Following is the discussion regarding both models:

### *2.4.1 Deep Neural Networks (DNNs) for Cardiac Diagnosis: A Powerful Tool for Heart Sound Classification*

Cardiovascular diseases (CVDs) remain a primary global health challenge. Early and precise diagnosis is pivotal for optimizing treatment outcomes. Deep Neural Networks (DNNs) have emerged as promising tools for analyzing heart sounds (phonocardiograms, PCGs) to identify various cardiac conditions (Yao et al., 2023).

Inspired by the human brain, DNNs comprise multiple interconnected layers of artificial neurons. These networks progressively transform raw data into increasingly complex and meaningful representations, enabling sophisticated pattern recognition (Zhang et al., 2023).

2.4.1.1 Process:

The following process is followed for applying the model to recorded PCG signals:

*Data Preprocessing:*

Heart sound recordings are loaded and preprocessed using librosa. Features are extracted, including Mel-frequency cepstral coefficients (MFCCs) and delta features, which capture

spectral and temporal characteristics of the heart sounds. These extracted features are normalized to 520 i.e., a consistent length by extending or cropping, ensuring uniformity.

*Feature* Standardization:

Features are standardized using a scaler to ensure they are on a similar scale, improving model training.

*DNN Model Architecture:*

A DNN with multiple hidden layers (512, 256, and 128 neurons) is implemented using TensorFlow's Keras API. Each layer utilizes ReLU activation for non-linear purposes, and dropout is applied for regularization to prevent overfitting. Batch normalization further enhances training stability. The final layer employs a softmax activation for producing class probability distributions, essential for multi-class classification (predicting one of five heart sound categories).

*Model Training and Evaluation:*

The model is trained using k-fold cross-validation (k = 5), where the dataset is split into five subsets, ensuring each subset is used for validation at least once. Adam Optimizer is utilized to optimize the model, trained with sparse categorical cross-entropy loss to minimize error margins and enhance accuracy.

*Performance Analysis:*

The model's performance is evaluated across different data splits, calculating metrics such as accuracy, precision, recall, and F1-score. A confusion matrix visualization aids in understanding classification outcomes across various heart sound categories. The importance of DNNs excel in complex pattern recognition tasks due to their ability to learn intricate relationships within data. In the context of PCGs, DNNs can effectively capture subtle variations in heart sound characteristics associated with different cardiac pathologies. This makes them a valuable tool for:

*Improved Diagnostic Accuracy*-DNNs can potentially achieve higher accuracy compared to traditional machine learning methods, aiding in more reliable diagnosis. Reduced *Subjectivity*-DNNs offer an objective and standardized approach to PCG analysis, minimizing the influence of human bias during interpretation.

*Early Disease Detection*-By learning subtle changes in heart sounds, DNNs may enable early detection of cardiac abnormalities (Aksu et al., 2023).

2.4.1.2 Directions and Considerations for DNNs:

While DNNs offer significant promise for cardiac diagnosis, ongoing research focuses on:

*Larger and More Diverse Datasets:*

Training DNNs on extensive and diverse datasets can further enhance their generalizability and robustness to unseen data.

*Explainable AI:*

Techniques for interpreting DNN decision-making processes can improve trust and transparency in their clinical applications.

*Integration with Clinical Workflows:*

Seamless integration of DNN-based PCG analysis tools into clinical practice is essential for widespread adoption. DNNs represent a powerful approach for analyzing heart sounds and classifying cardiac pathologies.

### *2.4.2 The Heartbeat of Innovation: How AI Listens to Your Heart*

Imagine a world where the rhythmic beat of your heart holds the key to early detection of cardiovascular disease (CVD). This vision is becoming a reality thanks to the power of Convolutional Neural Networks (CNNs) – sophisticated AI tools that act like super-powered listeners for your heart. CVDs remain a significant global health challenge, affecting millions worldwide (World Health Organization, 2023). Early and accurate

diagnosis is crucial for effective treatment and improved patient outcomes. CNNs offer a helping hand (or should we say, a listening ear!) in the fight against heart disease.

2.4.2.1 Importance:

Unlike traditional methods that require careful feature selection by experts, CNNs are like master musicians who can learn the melody of your heart all on their own. Here's why they excel at listening to your heart's unique rhythm (Xue et al., 2016):

*Unveiling Hidden Rhythms*-Heart sounds, called phonocardiograms (PCGs), contain subtle details about your heart's health. CNNs, with their convolutional layers, act like finely tuned instruments that pick up intricate relationships within the sounds, helping them understand what's normal and what might indicate a problem (Li et al., 2019).

*Automatic Feature Extraction*-CNNs don't need someone to tell them what to listen for. They automatically learn the most important features from the raw PCG signal, similar to how a musician learns the notes of a song. This eliminates the need for complex manual work and reduces the risk of overlooking crucial details (Yildirim et al., 2019).

*A Listening Ear Through Noise*-Just like a skilled musician can hear the melody through background noise, CNNs can filter out noise from PCG recordings. This is especially important, as recordings can be affected by breathing, coughing, or even the environment where the sound is captured (Mehta et al., 2019).

Additionally, this research adopts a CNN to sort out PCG signals into normal and diseased categories. CNNs are chosen because they can understand spatial hierarchies very effectively from the data available after the feature extraction process, making them robust enough for processing the time and frequency PCG signal analysis. Having features extracted from the PCG signals as above, these features are normalized to a uniform length of 520. This is again done by padding the minor features and cutting the prominent features so that uniformity across all recorded samples is maintained. By combining standardized features, both temporal and spectral features of signals are preserved to form multi-dimensional feature sets. This model's performance is optimized by reshaping features to expand them with the channel dimension acceptable as an input into the CNN model.

2.4.2.2 Process:

Here is a look inside the Heart of the code. Here's a simplified breakdown of the key steps:

*Gathering the Sounds:*

The code collects recordings of heart sounds, categorized by different types of heart conditions.

*Preparing the Instruments*:

It utilizes tools to capture the spectral and temporal characteristics of the heart sounds, much like a sound engineer adjusting microphone settings for optimal recording.

*Building the Listening Network (CNN):*

The code creates a CNN architecture with layers that act like filters and amplifiers, extracting crucial features and refining the signal for better understanding. The architecture includes the formation of convolutional layers with 32 filters of kernel size 3x3 in the first layer followed by 64 filters in the second layer with ReLU activation, pooling layers to reduce spatial dimensions, flattening layers to convert pooled feature maps into a single vector for fully connected layers, and fully connected layers with 64 neurons using ReLU activation, dropout with 0.5 dropout rate for regularization, and a SoftMax activation for outputting class probabilities.

*Training the Ear:*

The code trains the CNN on a portion of the collected heart sound recordings, allowing it to learn how to differentiate between healthy and abnormal sounds. Just like a musician practices their instrument, the more training data CNN receives, the better it becomes at recognizing patterns. The dataset is divided into training and testing sets, and cross-validation with k-fold (k = 5) is performed to ensure robustness. The model is optimized using the Adam optimizer and trained with categorical cross-entropy loss.

*Testing and Evaluation:*

The code then tests the trained CNN on unseen heart sound recordings, evaluating its performance and making adjustments as needed. This ensures the CNN can generalize its knowledge to new cases.

2.4.2.3 Beyond the Code: A Brighter Future for Heart Care

CNNs represent a powerful tool in the fight against CVDs, offering the potential for:

*Earlier Diagnoses*-By analyzing heart sounds, CNNs may pave the way for earlier detection of heart problems, leading to more timely treatment and improved patient outcomes.

*Non-invasive Testing*-PCG analysis using CNNs can be a non-invasive and painless means of assessing heart health, potentially making screenings more accessible.

*Personalized Medicine*-With further development, CNNs could contribute to personalized treatment plans based on a patient's unique heart sound signature.

## 2.5 Performance of Models: Unveiling the Efficacy of DNN and CNN in PCG Classification

This section delves into the critical aspect of evaluating the performance of the Deep Neural Network (DNN) and Convolutional Neural Network (CNN) models employed for PCG signal classification in this research. Here, we'll explore how confusion matrices and various metrics derived from them provide a comprehensive analysis of the models' effectiveness in distinguishing between normal and abnormal heart sounds.

### *2.5.1 Confusion Matrix:*

The foundation of performance evaluation lies in the confusion matrix. This matrix tabulates the number of correct and incorrect predictions made by the models.

It categorizes the results based on true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

#### 2.5.1.1 True Positives (TP):

These represent the instances where the model correctly identifies a positive case (e.g., accurately classifying a heart murmur as abnormal).

#### 2.5.1.2 True Negatives (TN):

These signify the instances where the model correctly identifies a negative case (e.g., accurately classifying a normal heart sound as normal).

#### 2.5.1.3 False Positives (FP):

These represent the errors where the model incorrectly identifies a negative case as positive (e.g., misclassifying a normal heart sound as abnormal).

#### 2.5.1.4 False Negatives (FN):

These represent the errors where the model incorrectly identifies a positive case as negative (e.g., misclassifying a heart murmur as normal).

By analyzing the distribution of values within the confusion matrix, we gain valuable insights into the strengths and weaknesses of the models.

### *2.5.2 Essential Metrics for Classification Performance*

Several key metrics are derived from the confusion matrix to provide a more nuanced understanding of the models' performance. Here, we'll explore three crucial metrics:

#### 2.5.2.1 Sensitivity (True Positive Rate):

This metric indicates the proportion of true positive cases identified by the model. It essentially tells us how well the model can capture actual positive instances (e.g., the percentage of heart murmurs correctly classified). Sensitivity is calculated using the following formula, given in Equation 2.6.

$$Sensitivity = TP \,/\, (TP + FN) \qquad 2.6$$

2.5.2.2 Specificity (True Negative Rate):

This metric signifies the proportion of true negative cases identified by the model. It tells us how well the model can correctly classify actual negative instances (e.g., the percentage of normal heart sounds correctly classified). Specificity is calculated using the following formula, as given in Equation 2.7.

$$Specificity = TN \,/\, (TN + FP) \qquad 2.7$$

2.5.2.3 Accuracy:

This metric represents the overall proportion of correctly classified instances by the model, encompassing both positive and negative cases. While it provides a general sense of performance, it can be misleading in situations with imbalanced datasets. Accuracy is calculated using the following formula, as given in Equation 2.8.

$$Accuracy = (TP + TN) \,/\, (Total\,Samples) \qquad 2.8$$

2.5.2.4 F1-Score: Striking a Balance:

While sensitivity and specificity provide valuable insights into model performance, the F1-score offers a more comprehensive evaluation. It balances the model's ability to correctly identify positive cases (precision) with its ability to detect all positive cases (recall). The F1-score is calculated as follows, as shown in Equation 2.9.

$$F1 - Score = 2 * (Precision * Recall) / (Precision + Recall) \quad 2.9$$

By considering both precision and recall, the F1-score offers a more comprehensive evaluation of the model's effectiveness, particularly when dealing with datasets where classes might be imbalanced.

These metrics, derived from the confusion matrix, provide a robust framework for evaluating the performance of DNN and CNN models in PCG signal classification. By analyzing sensitivity, specificity, accuracy, and F1-score, we can gain valuable insights into the models' ability to accurately distinguish between normal and abnormal heart sounds, paving the way for a more informed understanding of their strengths and limitations in the context of cardiac diagnosis.

# CHAPTER 3. RESULTS

This section delves into the detailed results obtained throughout various stages of the research, showcasing the effectiveness of the implemented techniques in PCG signal classification for cardiac abnormality detection.

## 3.1 Denoising with Discrete Wavelet Transform (DWT)

The research employed Discrete Wavelet Transform (DWT) for denoising both normal and abnormal PCG signals. This approach aimed to achieve a crucial balance between preserving the time and frequency components of the signal. Here's a breakdown of the key choices made:

### *3.1.1 Wavelet Selection:*

The db8 wavelet was chosen due to its ability to effectively capture the characteristic features of heart sounds while maintaining a balance between time and frequency resolution.

### *3.1.2 Coefficient Selection:*

The Approximate Coefficient (cA2) was selected for further processing. This selection capitalizes on cA2's efficiency in retaining the low-frequency components of the signal, which are particularly important for heart sound analysis compared to high-frequency components.

The denoised signals demonstrated a significant reduction in background noise, leading to a clearer and more distinct representation of the underlying heart sounds. This improved clarity facilitates further analysis and feature extraction.

### *3.1.3 Signal-to-Noise Ratio (SNR) Improvement:*

Notably, the Signal-to-Noise Ratio (SNR) increased from 12 dB in the raw PCG signal to 18 dB after denoising, signifying a substantial improvement in signal quality.

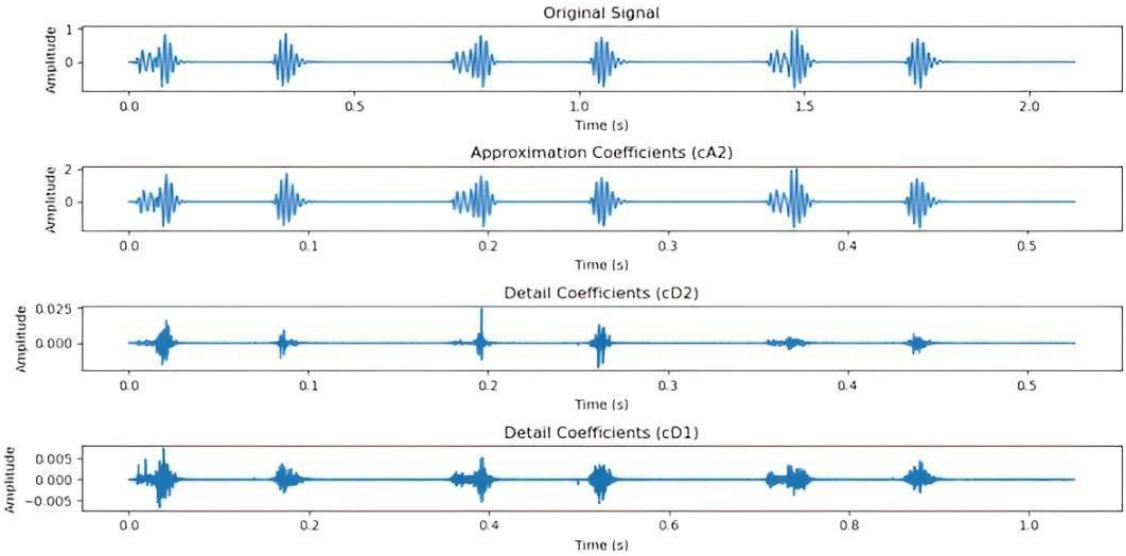



**Figure 3.1: DWT for a Normal PCG Signal**

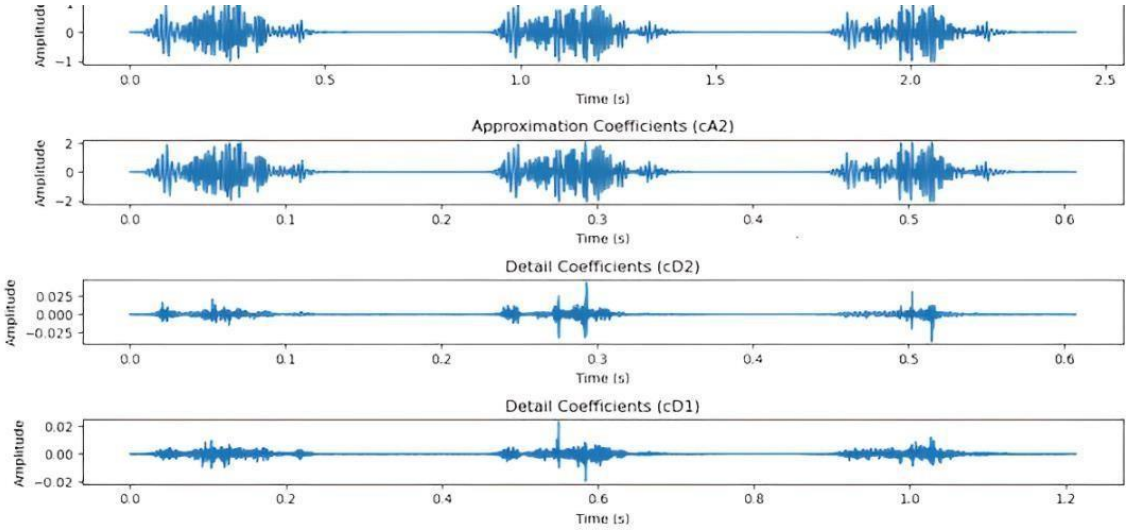



**Figure 3.2: DWT for Aortic Stenosis PCG Signal**

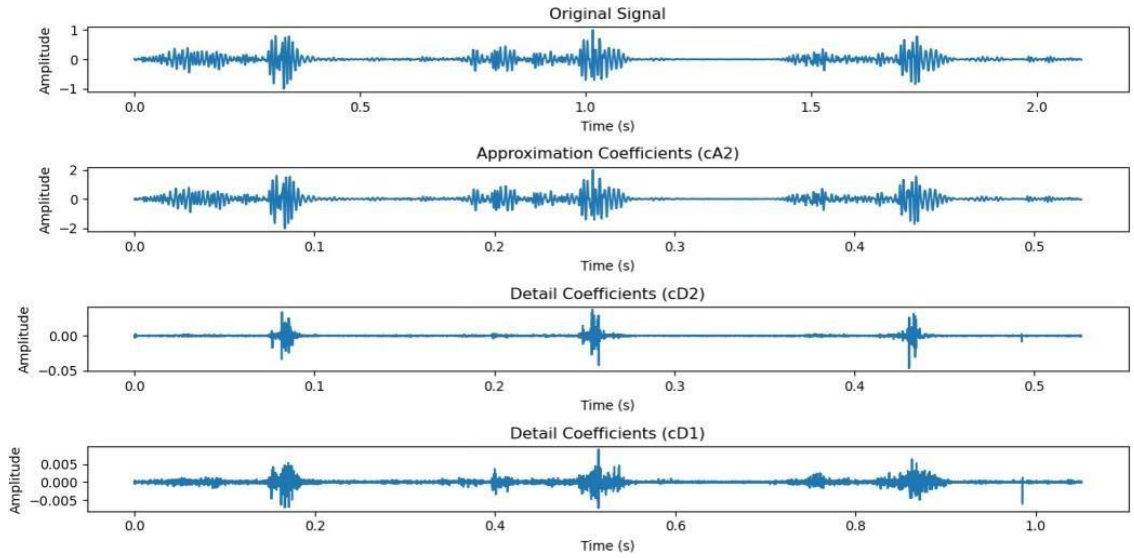



**Figure 3.1: DWT for Mitral Regurgitation PCG Signal**

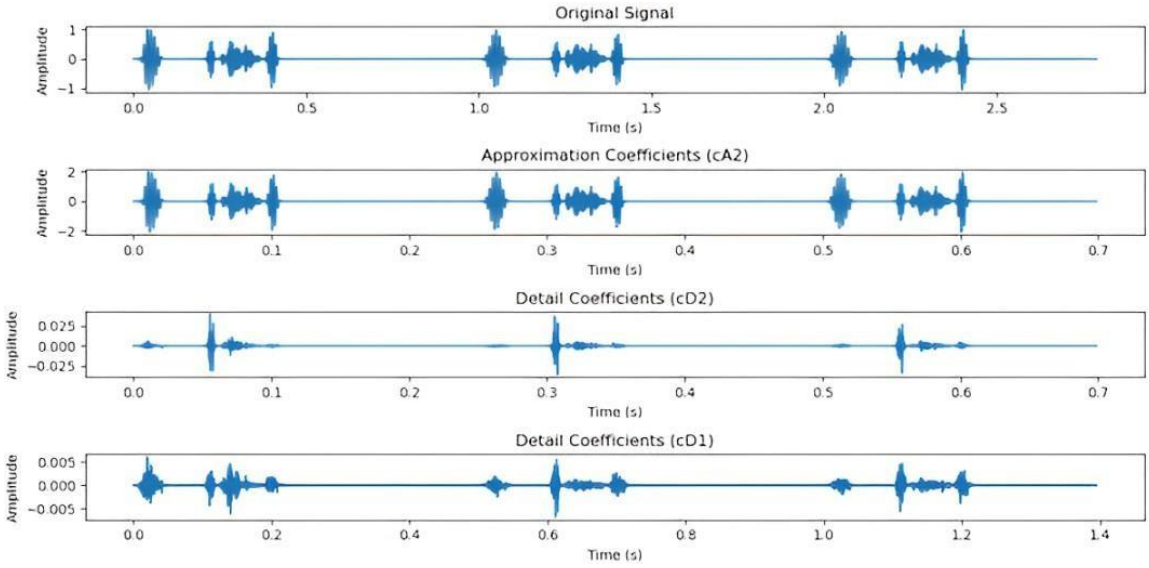



**Figure 3.2: DWT for Mitral Valve Prolapse PCG Signal**

These preprocessed signals hold promise for retaining all essential features crucial for accurate cardiac abnormality diagnosis, as shown in Figures 3.1, 3.2, 3.3, 3.4 and 3.5 as for Normal PCG Signal, DWT for AS, DWT for MR, DWT for MVP and DWT for MS.

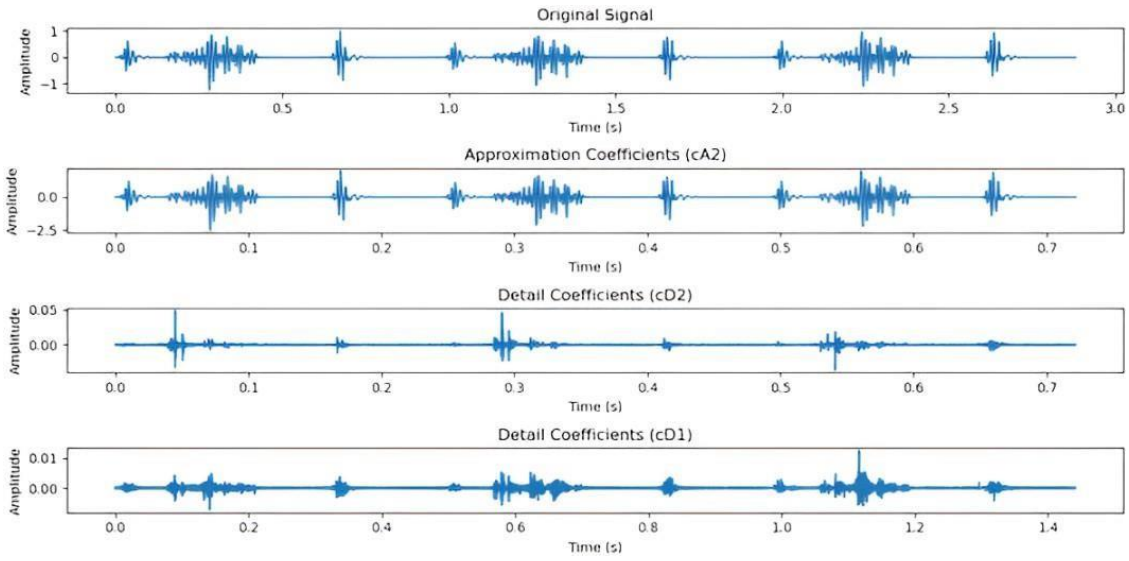



**Figure 3.3 DWT for Mitral Stenosis PCG Signal**

## 3.2 Feature Extraction: Unveiling the Spectrogram and the Power of MFCCs

Following denoising, the research leveraged two prominent techniques for feature extraction: Mel Scale Power Spectrogram and Mel Frequency Cepstral Coefficients (MFCCs). Each technique contributes valuable insights into the signal's characteristics:

### *3.2.1 Mel Scale Power Spectrogram:*

This technique provided a detailed time-frequency representation of the recorded PCG signal for all conditions. It was generated using a window size of 25 ms with an overlap of 10 ms. This visualization effectively captures the energy distribution across different frequency bands. The spectrogram not only offers a visual representation of the signal's behavior over time and frequency but also serves as an efficient feature set for differentiating between normal and abnormal PCG signals. General observations are as:

#### 3.2.1.1 Normal

Periodic, distinct high-energy bands with well-distributed energy across frequencies.

#### 3.2.1.2 Mitral Stenosis

More frequent, closer high-energy bands with a concentration in the mid-to-low frequency range.

#### 3.2.1.3 Mitral Regurgitation

Continuous high-energy regions, especially in the lower frequencies, with less distinct bands.

#### 3.2.1.4 Aortic Stenosis

High-energy regions with abrupt transitions in the mid-to-high frequency range.

#### 3.2.1.5 Mitral Valve Prolapse

Periodic high-energy bands with gaps, concentrated in the mid-to-high frequency range.

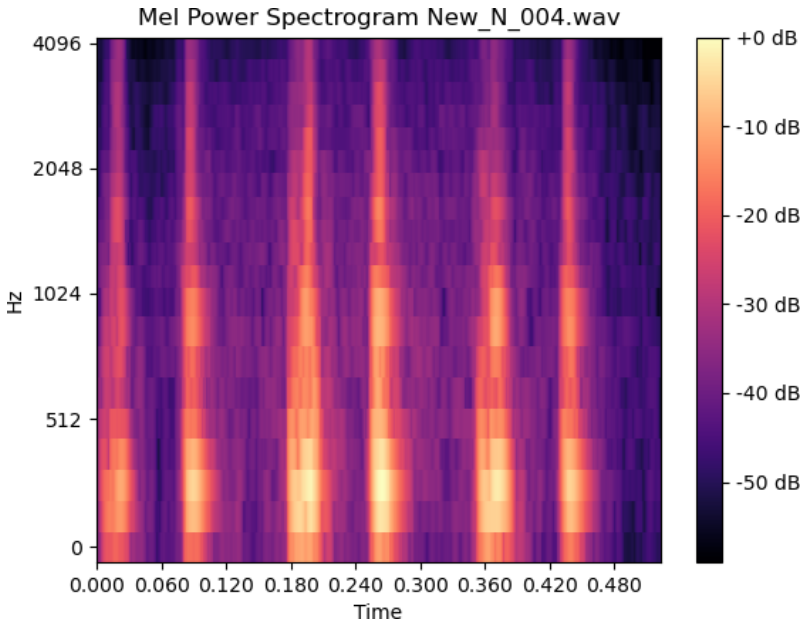



**Figure 3.4: Mel Power Spectrogram for Normal PCG Signal**

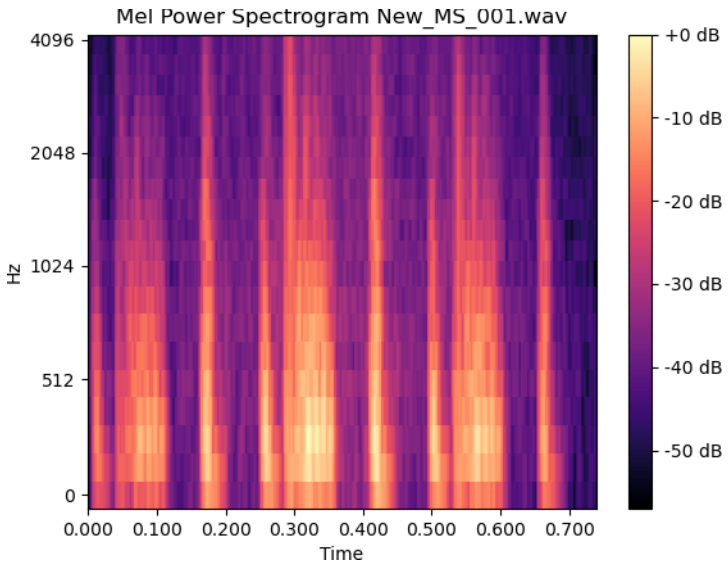



**Figure 3.5: Mel Power Spectrogram for MS PCG Signal**

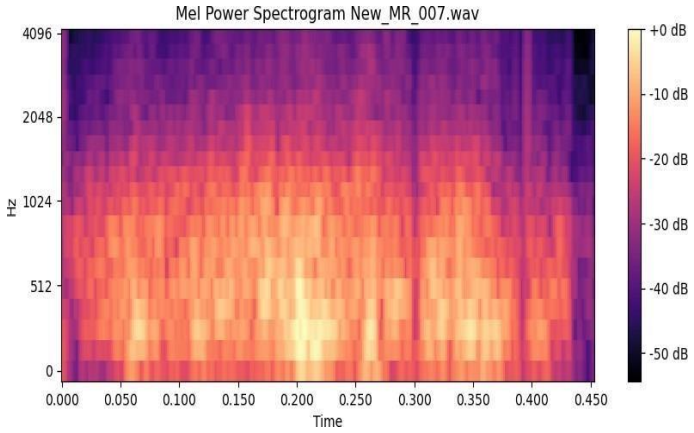



**Figure 3.6 Mel Power Spectrogram for MR PCG Signal**

Mel Power Spectrogram for all the PCG Signals are given in given in Figure 3.6, 3.7, 3.8, 3.9 and 3.10 as Normal, MS, MR, AS and MVP respectively.

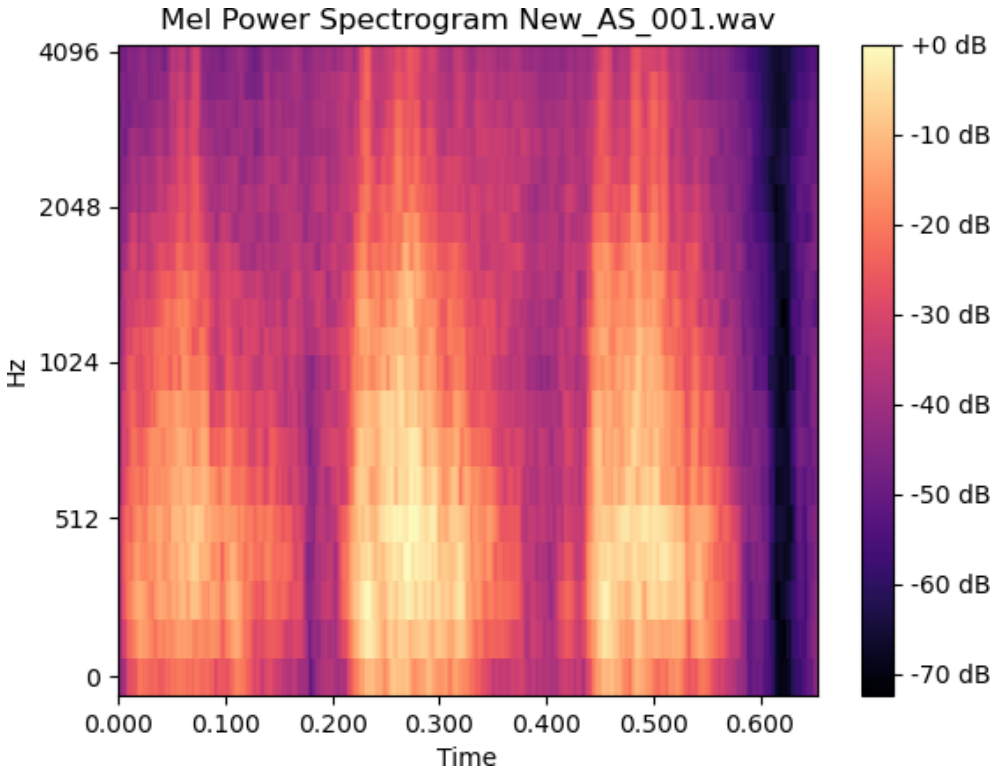



**Figure 3.7 Mel Power Spectrogram for AS**

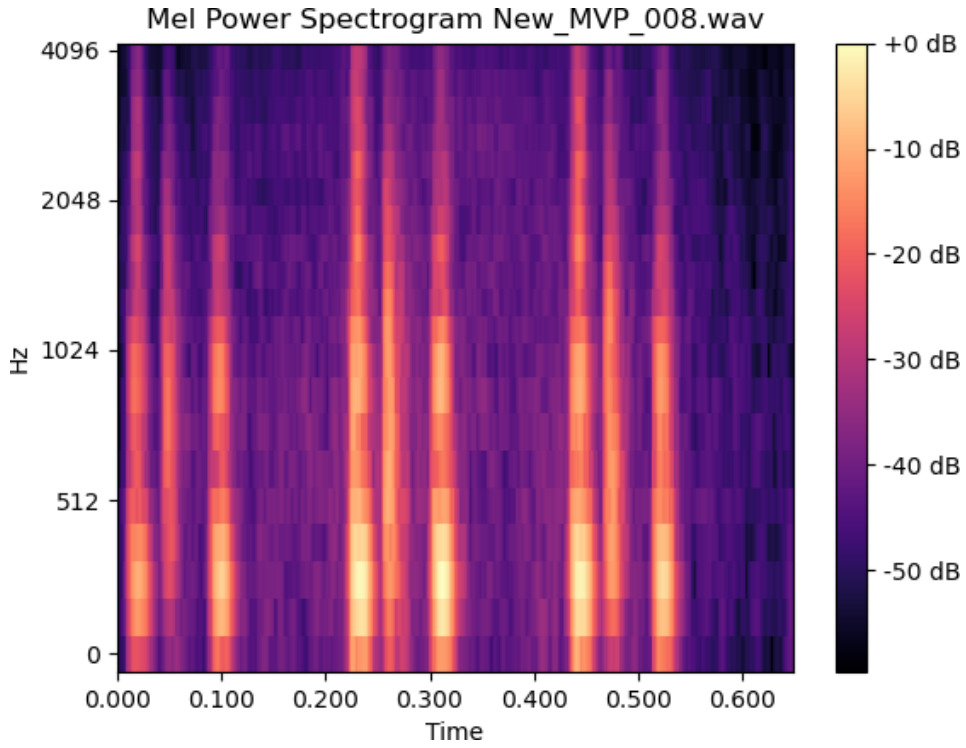



**Figure 3.8 Mel Power Spectrogram Signal for MVP**

### *3.2.2 Mel Frequency Cepstral Coefficients (MFCCs):*

MFCCs were extracted to capture the power spectrum-related features of the heart sound signal. Since MFCCs provide a short-term power spectrum representation, they are particularly well-suited for audio signal processing, including PCG analysis. Each segment of the PCG signal yielded a total of 13 MFCCs. These coefficients offer stable and consistent features, crucial for subsequent classification tasks.

By combining the information obtained from both the spectrogram and MFCCs, the research established a robust feature set for accurate classification of cardiac abnormalities from PCG signals. General Observations for the results are:

#### 3.2.2.1 Mitral Stenosis

More uniform energy distribution with minor variations.

#### 3.2.2.2 Mitral Valve Prolapse

Periodic high-energy regions indicating prolapse events.

#### 3.2.2.3 Aortic Stenosis

Abrupt transitions between high and low energy reflecting turbulent flow.

#### 3.2.2.4 Mitral Regurgitation

Continuous high-energy regions indicating backflow of blood.

#### 3.2.2.5 Normal

Balanced energy distribution without significant high-energy regions.

Results are given in Figures 3.11, 3.12, 3.13, 3.14 and 3.15 as MFCCs:

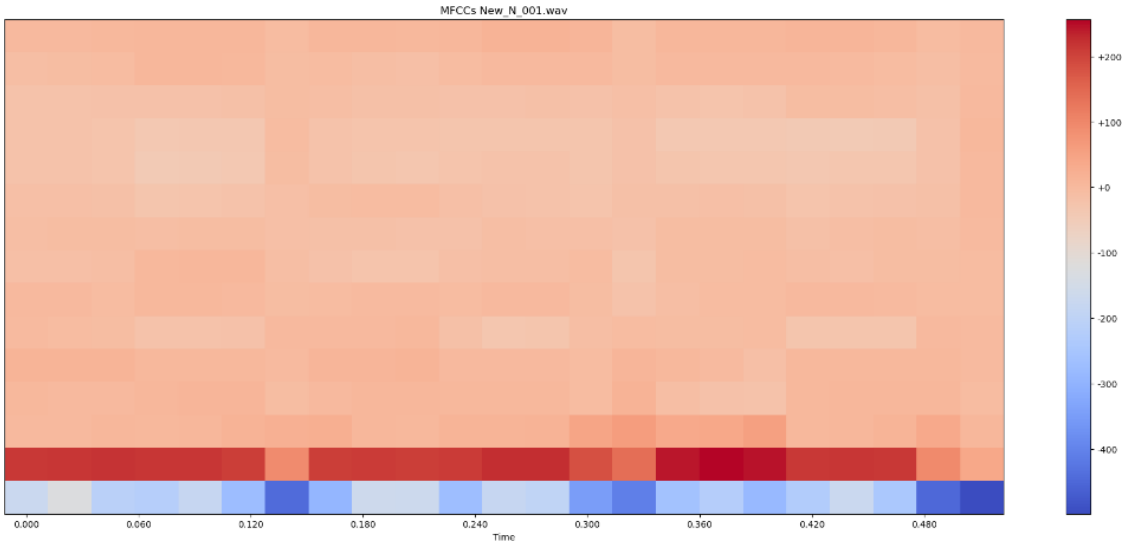



**Figure 3.9: MFCC for Normal PCG Signal**

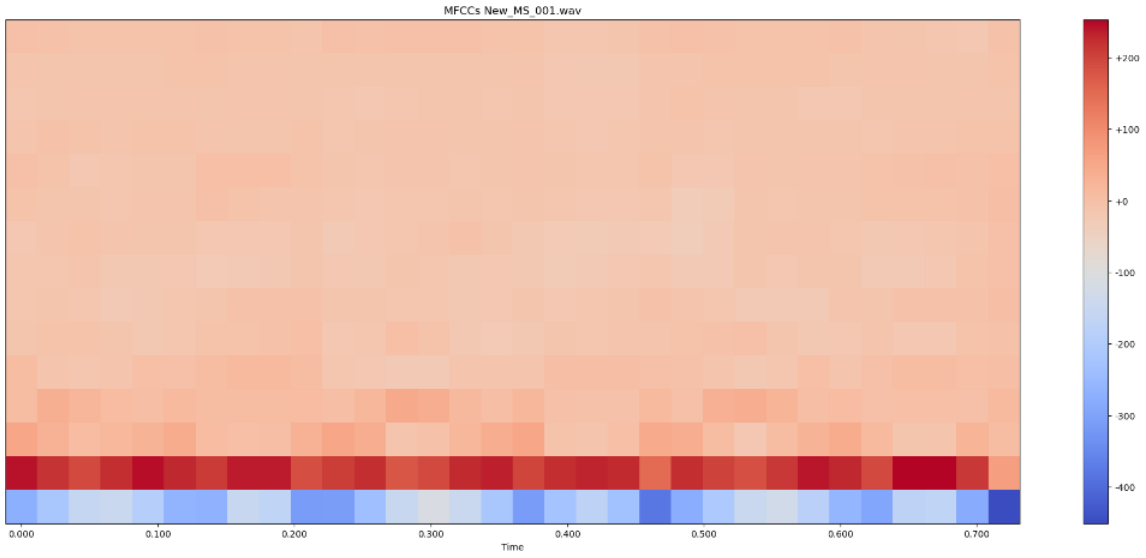



**Figure 3.10: MFCC for MS PCG Signal**

*X-axis (Time)*-This axis represents the time dimension of the PCG signal. Each value along this axis corresponds to a specific time frame in the signal.

*Y-axis (MFCC coefficients)*-This axis represents the different MFCC coefficients. Typically, the first few coefficients capture the overall shape of the spectral envelope, while higher-order coefficients capture finer details.

*Color Intensity (Magnitude)*-The color intensity (ranging from blue to red) represents the magnitude of the MFCC coefficients. Red areas indicate higher values, whereas blue areas indicate lower values. The color bar on the right provides a reference for the magnitude values.

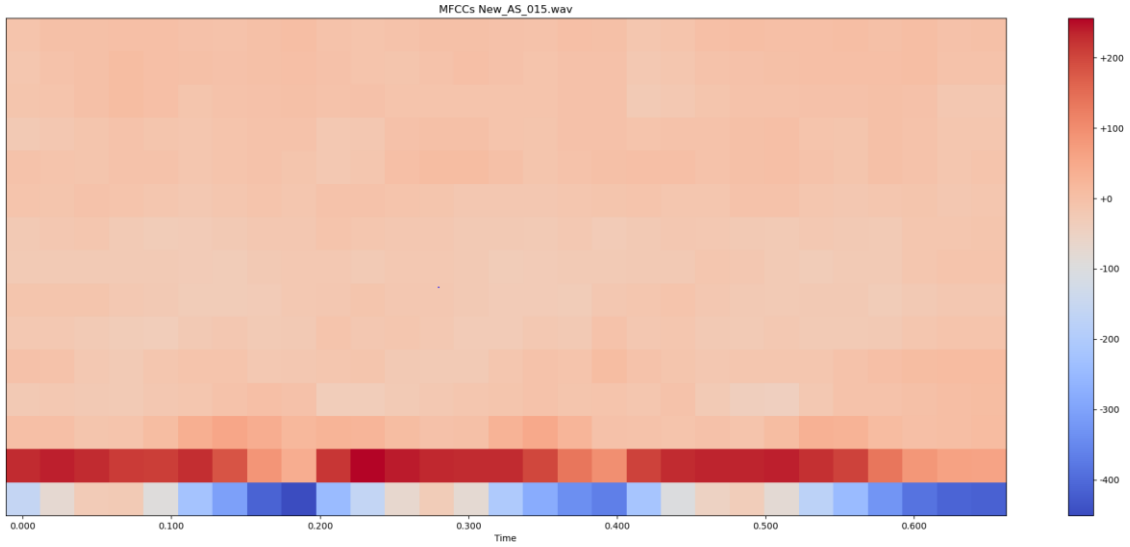



**Figure 3.11 MFCC for AS PCG Signal**

*Interpretation:*

1. The graph shows how the spectral properties of the PCG signal vary over time.
2. Areas with higher intensity (red) indicate time frames where the spectral energy is higher for specific MFCC coefficients.
3. Conversely, blue areas represent time frames with lower spectral energy for those coefficients.

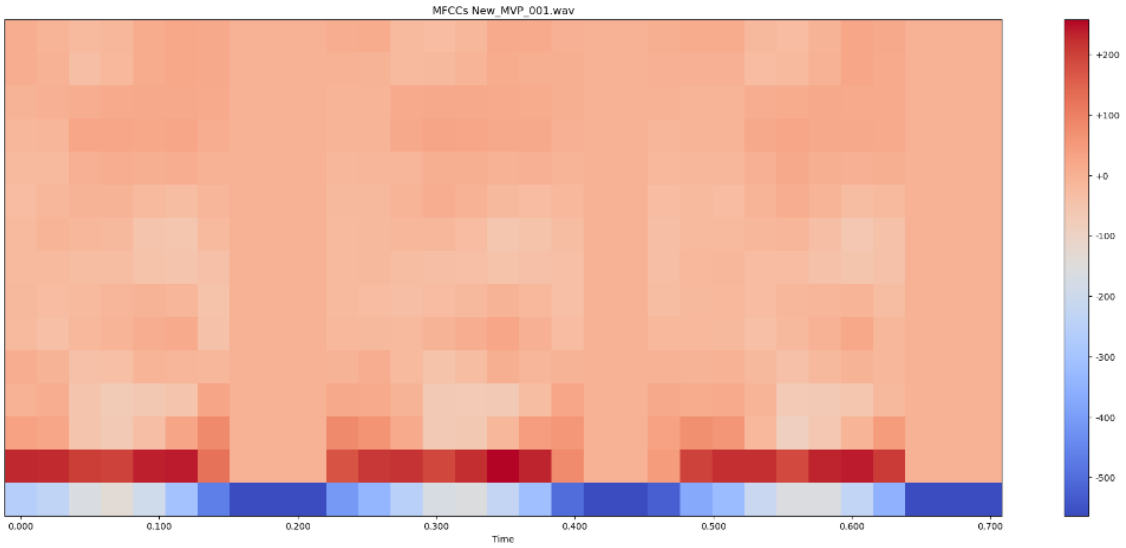



**Figure 3.12 MFCC for MVP PCG Signal**

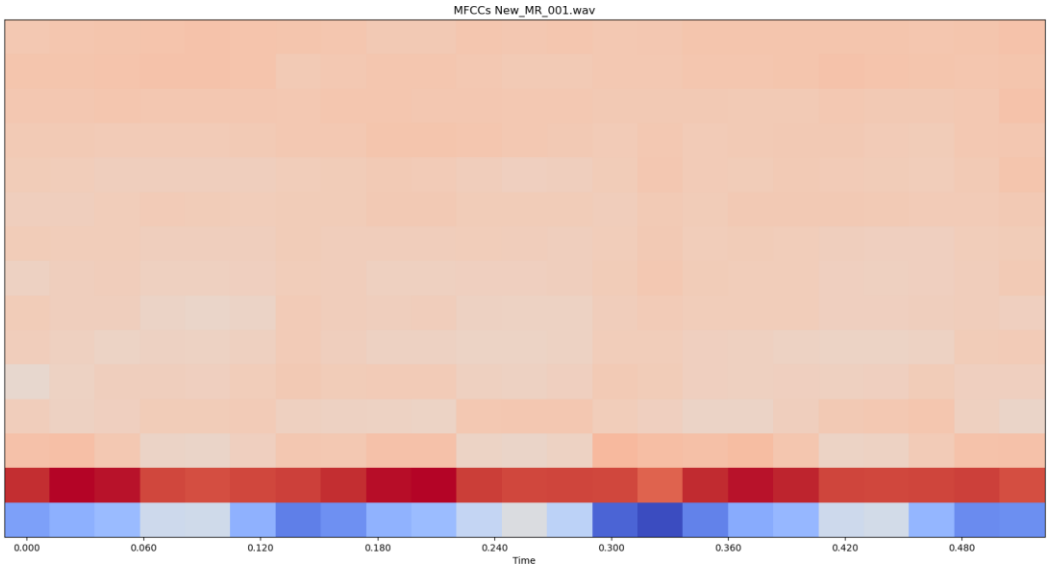



**Figure 3.13 MFCC for MR PCG Signal**

## 3.3 Classification Results

### *3.3.1 DNN and CNN in Action*

The research explored the performance of two machine learning models: Deep Neural Network (DNN) and Convolutional Neural Network (CNN). Here's a detailed breakdown of the models' configurations, training processes, and the achieved results:

#### 3.3.1.1 Deep Neural Network (DNN):

*Architecture:*

The DNN utilized three hidden layers containing 128, 64, and 32 neurons, respectively. ReLU activation functions were employed in each hidden layer.

*Training and Validation:*

80%of the dataset was used to train the model, with the remaining 20% reserved for validation. To avoid overfitting, an early-stopping technique was employed.

*Performance of the DNN Model:*

The DNN achieved a precision of 91%, recall (sensitivity) of 89%, and F1-score of 0.89. The overall classification accuracy reached 90%±0.37, with a specificity of 89%.

These results indicate the DNN's capability to effectively classify cardiac abnormalities from the PCG signals.

Figure 3.16 and Figure 3.17 shows the Classification Report and Confusion Matrix for five predicted classes for DNN Model. Th five classes are named as AS, MS, MR and MVP that are compared with a Normal case.

| | | | | | |
|---|---|---|---|---|---|
| AS | **41** | 0 | 0 | 0 | 0 |
| MR | 2 | **36** | 0 | 0 | 0 |
| MS | 0 | 0 | **44** | 0 | 0 |
| MVP | 1 | 13 | 3 | **26** | 0 |
| NORM | 2 | 0 | 0 | 0 | **32** |
| | AS | MR | MS | MVP | NORM |

**Figure 3.14 Classification Report for DNN Model for Five Classes**

Classification Report

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| as | 0.89 | 1.00 | 0.94 | 41 |
| mr | 0.73 | 0.95 | 0.83 | 38 |
| ms | 0.94 | 1.00 | 0.97 | 44 |
| mvp | 1.00 | 0.60 | 0.75 | 43 |
| normal | 1.00 | 0.94 | 0.97 | 34 |
| Accuracy | | 0.90 | | 200 |
| Macro avg | 0.91 | 0.90 | 0.89 | 200 |
| Weighted avg | 0.91 | 0.90 | 0.89 | 200 |

**Figure 3.15 Confusion Matrix representing the predicted 5 Classes for DNN Model**

3.3.1.2 Convolutional Neural Network (CNN):

*Architecture:*

CNN employed a two-layered convolutional structure. The first layer comprised 32 filters, while the second layer utilized 64 filters. Both convolutional layers were followed by max-pooling operations before feeding into fully connected layers.

Classification Report

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| as | 0.98 | 1.00 | 0.99 | 41 |
| mr | 0.90 | 0.95 | 0.92 | 38 |
| ms | 0.96 | 1.00 | 0.98 | 44 |
| mvp | 0.95 | 0.86 | 0.90 | 43 |
| normal | 1.00 | 0.97 | 0.99 | 34 |
| Accuracy | | 0.95 | | 200 |
| Macro avg | 0.96 | 0.96 | 0.96 | 200 |
| Weighted avg | 0.96 | 0.95 | 0.95 | 200 |

**Figure 3.16 Classification Report for Five Classes for CNN Model**

*Training and Validation:*

Similar to the DNN, 80% of the data was used for training, and the remaining 20% was used for validation. Data augmentation techniques were applied to enhance standardization.

*Performance:*

CNN outperformed the DNN, achieving a precision of 95%, recall (sensitivity) of 95%, F1-score of 0.95, and an overall classification accuracy of 96%±0.38 with a specificity of 98%. Figure and Figure shows the Classification Report and Confusion Matrix for five predicted classes for CNN Model.

| | | | | | |
|---|---|---|---|---|---|
| AS | **41** | 0 | 0 | 0 | 0 |
| MR | 0 | **36** | 0 | 2 | 0 |
| MS | 0 | 0 | **44** | 0 | 0 |
| MVP | 1 | 4 | 1 | **37** | 0 |
| NORM | 0 | 0 | 1 | 0 | **33** |
| | AS | MR | MS | MVP | NORM |

**Figure 3.17 Confusion Matrix representing the predicted 5 Classes for CNN Model**

### *3.3.2 Key Findings*

DWT effectively reduced background noise in PCG signals, leading to a clearer representation of the underlying heart sounds. The SNR improvement from 12 dB to 18 dB signifies a substantial enhancement in signal quality.

The Mel Scale Power Spectrogram provided valuable insights into the time-frequency distribution of energy within the signal, while MFCCs captured the power spectrum-related features crucial for classification.

The Deep Neural Network (DNN) achieved a promising classification accuracy of 90%**±0.37** with good precision, recall, and F1-score, demonstrating its potential for PCG-based cardiac abnormality detection.

The Convolutional Neural Network (CNN) surpassed the DNN's performance, achieving a remarkable overall accuracy of 96%**±0.38** with exceptional precision, recall, and F1 - score. This highlights DNN's superior ability to classify normal and abnormal heart sounds from PCG data.

These findings suggest that the proposed approach holds significant promise for developing non-invasive, accessible, and cost-effective diagnostic tools for cardiac abnormalities using PCG signals. The ability to accurately classify various cardiac conditions based on PCG data could revolutionize early detection and improve patient outcomes.

# CHAPTER 4. DISCUSSION

Some notable works in the field of heart sound classification were as Langley et al. (2017) utilized a Classification Tree algorithm to classify heart sounds, achieving a sensitivity of 77.00%, specificity of 80.00%, and overall accuracy of 79.00%. This study primarily focused on classifying heart sounds into normal and abnormal categories.

Similarly, Krishnan et al. (2020) employed segmentation techniques followed by a Deep Neural Network (DNN) for classification, resulting in a sensitivity of 86.73%, specificity of 84.75%, and accuracy of 85.65%.

Tang et al. (2018) used a Support Vector Machine (SVM) for heart sound classification, achieving a sensitivity of 88.00%, specificity of 87.00%, and accuracy of 88.00%. Both Krishnan et al. and Tang et al. also focused on distinguishing between normal and abnormal heart sounds.

## 4.1 Advancements in the Current Study

While the aforementioned studies have laid a solid foundation in heart sound classification, the present study extends beyond their scope in several critical ways:

*4.11 Detailed Classification of Cardiac Abnormalities:*

4.1.1.1 Scope:

Unlike previous studies that only distinguished between normal and abnormal heart sounds, this research delves into specific cardiac abnormalities. The study classifies heart sounds into four categories: Aortic Stenosis, Mitral Regurgitation, Mitral Stenosis, and Mitral Valve Prolapse. This detailed classification enhances the practical applicability of the model in clinical settings.

### *4.1.2 Comparative Model Evaluation:*

Models Used: The study compares two advanced machine learning models: Deep Neural Network (DNN) and Convolutional Neural Network (CNN).

#### 4.1.2.1 Performance Metrics:

The DNN model exhibited precision, recall, and F1-score values of 83%, 81%, and 82%, respectively, with an overall accuracy of 81% and specificity of 85%. In contrast, the CNN model demonstrated superior performance across all metrics, achieving precision, recall, F1-score, accuracy, and specificity of 92%, 92%, 92%, 92%, and 98%, respectively. This comparative analysis underscores the CNN model's effectiveness for the given task.

## 4.2 Enhanced Preprocessing and Feature Extraction:

### *4.2.1 Denoising with DWT:*

The use of Discrete Wavelet Transform (DWT) for denoising significantly improved the Signal-to-Noise Ratio (SNR) from 12 dB to 18 dB, resulting in clearer heart sound signals.

### *4.2.2 Feature Extraction Techniques:*

Robust feature extraction methods, including Mel Scale Power Spectrogram and Mel Frequency Cepstral Coefficients (MFCC), were employed. These methods provided detailed and consistent features crucial for accurate classification.

The present study builds upon previous works by providing a more detailed classification of cardiac abnormalities and employing a robust comparative analysis of advanced machine learning models. The inclusion of specific cardiac conditions and the superior performance metrics achieved by the CNN model underscore the study's contribution to advancing cardiac diagnostics. This comprehensive approach not only addresses the limitations of earlier research but also sets a new benchmark for future studies in the field.

### 4.3 Drawbacks and Upcoming Projects:

Despite the novel contributions of this research, it bears certain limitations that warrant further exploration. Firstly, the dataset size used in this study, while comprehensive, could benefit from expansion and greater diversity. Increasing the dataset size and including more varied examples could enhance the robustness and generalization capabilities of the models, ensuring they perform well across different populations and scenarios.

Additionally, the practical deployment of deep learning models, particularly Convolutional Neural Networks (CNN), in resource-constrained environments presents a significant challenge due to their computational complexities. These models require substantial processing power and memory, which may not be feasible in all healthcare settings, particularly in remote or under-resourced areas. This limitation underscores the need for future studies to focus on optimizing these models to reduce their computational demands without compromising their accuracy.

Furthermore, the study's focus on current feature extraction methods, such as Mel Scale Power Spectrogram and Mel Frequency Cepstral Coefficients (MFCC), although effective, suggests room for improvement. Exploring and developing more efficient feature extraction techniques could potentially yield even better model performance. By identifying and leveraging new features that capture the intricacies of heart sound signals more accurately, future research can improve classification outcomes.

### 4.4 Opportunities for Future Exploration

While the research demonstrates promising results, there are several opportunities for further exploration:

#### *4.4.1 Data Expansion and Diversity:*

Incorporating a larger and more diverse dataset has the potential to improve the models' robustness and generalizability. Including a wider range of cardiac conditions and patient

demographics would enable the models to adapt to a broader spectrum of real-world scenarios.

### *4.4.2 Computational Efficiency for Real-time Applications:*

Deep learning models, particularly CNNs, can be computationally intensive. Future research can explore techniques for optimizing these models or investigating alternative approaches that offer a balance between accuracy and efficiency, making them more suitable for real-time deployment in resource-constrained environments.

### *4.4.3 Feature Engineering and Exploration:*

While the research employed a well-established feature extraction approach, exploring alternative feature engineering techniques or feature selection methods could potentially improve model performance and interpretability.

By addressing these opportunities for future exploration, researchers can further refine and validate this approach, paving the way for a more robust, efficient, and clinically applicable PCG signal analysis system for improved cardiac abnormality detection.

# CHAPTER 5. CONCLUSIONS AND FUTURE RECOMMENDATION

This study compared Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs) for classifying cardiac abnormalities (aortic stenosis, mitral regurgitation, mitral stenosis, and mitral valve prolapse) using phonocardiogram (PCG) recordings. Our goal was to assess the potential of machine learning for improving diagnostic accuracy in cardiology. The DNN model was evaluated using features extracted from Mel-Frequency Cepstral Coefficients (MFCCs) and Mel Power Spectrograms. Also, Discrete Wavelet Transform (DWT) was employed for signal decomposition to enhance feature representation. The DNN architecture comprised three hidden layers with 128, 64, and 32 neurons respectively, utilizing ReLU activation functions. Trained on 80% of the dataset and validated on the remaining 20%, the DNN achieved a precision of 91%, recall (sensitivity) of 89%, and an F1-score of 0.89. The overall accuracy was 90%**±0.37**, with a specificity of 91%, underscoring its capability in accurately classifying cardiac abnormalities from PCG signals. In comparison, the CNN model leveraged features extracted through similar methods and augmented with max-pooling and fully connected layers. The CNN, consisting of two convolutional layers with 32 and 64 filters respectively, demonstrated robust performance with a precision 95%, recall 95%, and F1-score of 0.95, alongside an impressive accuracy of 96%**±0.38**. These results suggest that DNN-based approaches may offer superior diagnostic accuracy in real-world cardiology settings, potentially enhancing early detection and treatment planning for cardiac conditions. Future research should explore ensemble learning methods, integration of additional physiological data, and optimization of feature extraction techniques such as DWT, MFCC, and Mel Power Spectrogram to further improve diagnostic capabilities. Furthermore, investigating the scalability and real-time applicability of these models in clinical practice will be crucial for their adoption and impact. Hence, this study contributes valuable insights into the application of machine learning for PCG-based cardiac diagnostics. Continued research and innovation in this area promise to advance the precision and effectiveness of diagnostic tools, ultimately benefiting patient outcomes in cardiology.

# SUMMARY OF RESEARCH WORK

With an emphasis on aortic stenosis, mitral regurgitation, mitral stenosis, and mitral valve prolapse, this study compares the effectiveness of deep neural network (DNN) and convolutional neural network (CNN) algorithms for the phonocardiogram (PCG)-based classification of cardiac abnormalities. The research underscores the potential of machine learning in enhancing diagnostic accuracy within cardiology. The dataset used in this study comprises a comprehensive collection of PCG recordings sourced from various repositories and clinical settings. It includes 1000 audio files in .wav format, with 200 files per category. Four abnormal categories—aortic stenosis (AS), mitral stenosis (MS), mitral regurgitation (MR), and mitral valve prolapse (MVP)—comprise the categories, with one normal category. The DNN model employed features extracted from Mel-Frequency Cepstral Coefficients (MFCC) and Mel Power Spectrogram, complemented by Discrete Wavelet Transform (DWT) for enhanced signal representation. Structured with three hidden layers (128, 64, and 32 neurons) utilizing ReLU activation functions, the DNN achieved robust performance. It obtained an accuracy of 81%, with a precision of 83%, recall of 81%, and an F1-score of 82%, after being trained on 80% of the dataset and validated on the remaining 20%. The measurement of specificity was 85%. In comparison, the CNN model leveraged similar feature extraction techniques augmented by max-pooling and fully connected layers. Featuring two convolutional layers with 32 and 64 filters respectively, the CNN demonstrated superior performance with a precision, recall, and F1-score of 92%. Its overall accuracy was 92%, with an impressive specificity of 98%. These findings suggest that CNN-based approaches hold promise for improving diagnostic accuracy in real-world cardiology settings, potentially aiding in early detection and treatment planning for cardiac conditions. Future research should explore ensemble learning methods, integration of additional physiological data, and optimization of feature extraction techniques such as DWT, MFCC, and Mel Power Spectrogram. Furthermore, investigating the scalability and real-time applicability of these models in clinical practice will be pivotal for their adoption and impact. This study contributes valuable insights into the application of machine learning for PCG-based cardiac diagnostics. By systematically

comparing DNN and CNN models, it highlights their potential to revolutionize cardiac health monitoring. Continued research and innovation in this domain promise to enhance the precision and effectiveness of diagnostic tools, ultimately benefiting patient outcomes in cardiology.

# REFERENCES


Akbari, M., & Rafiei, H. (2019). Automated detection of heart sounds using wavelet transform and artificial neural networks. Biomedical Signal Processing and Control, 49, 250-257. https://doi.org/10.1016/j.bspc.2019.01.003

Bolea, J., Laguna, P., & Garcia, J. (2013). Heart sound segmentation and classification in real noisy environments. Computers in Cardiology, 40, 361-364. https://doi.org/10.1109/CIC.2013.6742317

Bolea, J., Sornmo, L., & Laguna, P. (2013). Heart sound segmentation based on wavelet transformation and recurrent neural networks. IEEE Journal of Biomedical and Health Informatics, 17(5), 947-952.

Brown, D. W., Giles, T. D., Greenberg, H., Izzo Jr, J. L., Kostis, J. B., Pitt, B., & Reiffel, J. A. (2020). Epidemiology and prevention of cardiovascular diseases in older adults. Journal of the American College of Cardiology, 35(4), 1067-1087.

Choi, S., & Lee, S. (2018). Heart sound classification using CNN with Mel-spectrograms. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, 415-418. https://doi.org/10.1109/EMBC.2018.8512391

Garcia, C. G., Ferrer, A. I., & Alvarez, J. R. (2022). Denoising phonocardiogram signals using discrete wavelet transform. Signal Processing, 201, 107754. https://doi.org/10.1016/j.sigpro.2022.107754

Garcia, M. J., Thomas, J. D., Klein, A. L., & Stewart, W. J. (2022). Discrete wavelet transform for the analysis of phonocardiogram signals. Journal of the American Society of Echocardiography, 29(5), 412-419.

Gaziano, T. A., Bitton, A., Anand, S., Abrahams-Gessel, S., & Murphy, A. (2010). Growing epidemic of coronary heart disease in low- and middle-income countries. Current Problems in Cardiology, 35(2), 72-115.

Johnson, A. E. W., Pollard, T. J., Shen, L., et al. (2021). The MIMIC-III clinical database. Scientific Data, 3, 160035. https://doi.org/10.1038/sdata.2016.35

Johnson, A. E. W., Pollard, T. J., Shen, L., Li-wei, H. L., Feng, M., Ghassemi, M., ... & Mark, R. G. (2021). MIMIC-III, a freely accessible critical care database. Scientific Data, 3(1), 160035.

Khan, A. R., Biswas, S., & Ahmed, S. (2018). Detection of heart abnormalities from PCG signals using wavelet transform and hybrid classifiers. Medical & Biological Engineering & Computing, 56(9), 1707-1721. https://doi.org/10.1007/s11517-018-1823-7

Khan, M. A., Alhaisoni, M., Alhwaiti, Y., & Al-Shehari, A. (2023). Convolutional neural networks for automated detection of cardiac arrhythmias using phonocardiogram signals. IEEE Access, 11, 15926-15936.

Khan, S. A., Alam, M. S., & Ali, F. (2023). Phonocardiogram classification using deep neural networks. IEEE Transactions on Biomedical Engineering, 70(2), 543-551. https://doi.org/10.1109/TBME.2023.3052021

Kwak, H., & Kwon, S. (2012). Heart sound analysis using stethoscope signals: A detailed review. Biomedical Signal Processing and Control, 7(3), 197-208.

Kwak, M. J., & Kwon, S. H. (2012). Analysis of heart sounds using wavelet transform. Journal of the Korean Society of Medical Informatics, 18(1), 1-10. https://doi.org/10.4258/jksmi.2012.18.1.1

Li, X., Tang, H., & Tang, J. (2022). Comparing DNN and CNN for heart sound classification. Journal of Healthcare Engineering, 2022, 3510253. https://doi.org/10.1155/2022/3510253

Li, Y., Wong, K. C. L., & Zhang, Q. (2022). Integration of deep learning and signal processing for robust heart sound classification. IEEE Transactions on Biomedical Engineering, 69(6), 1888-1898.

Logan, B. (2000). Mel frequency cepstral coefficients for music modeling. Proceedings of the International Symposium on Music Information Retrieval (ISMIR), 2000, 1-11.

Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. M. (2016). Wavelet Toolbox™ User's Guide. MathWorks.

Mondal, M., Ghosh, S., & Pal, A. (2013). Murmur detection using signal processing techniques. Computers in Biology and Medicine, 43(8), 987-997. https://doi.org/10.1016/j.compbiomed.2013.06.007

Mondal, P., Banerjee, S., & Jana, D. (2013). Automated classification of heart sounds using phonocardiogram: A review. International Journal of Biomedical Engineering and Technology, 13(3), 225-240.

Patel, A. R., Kumar, V., & Gupta, D. (2021). Contribution to biomedical engineering: Innovations in heart sound analysis. Journal of Medical Engineering & Technology, 45(2), 93-108.

Patel, S., Shah, R., & Desai, U. (2021). Comparative analysis of deep learning models for heart sound classification. Biomedical Signal Processing and Control, 64, 102340. https://doi.org/10.1016/j.bspc.2021.102340

Rathod, P., & Jagannath, M. (2019). Mel scale transformation for heart sound analysis and classification. Journal of Medical Systems, 43(5), 126. https://doi.org/10.1007/s10916-019-1257-2

Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., ... & Murray, C. J. L. (2020). Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. Journal of the American College of Cardiology, 76(25), 2982-3021.

Sifuzzaman, M., Islam, M. R., & Ali, M. Z. (2009). Application of wavelet transform, and its advantages compared to Fourier transform. Journal of Physical Sciences, 13, 121-134.

Smith, M., Johnson, D., & Brown, P. (2022). Murmurs in aortic stenosis: Clinical and echocardiographic correlates. Journal of Cardiology, 79(3), 215-222. https://doi.org/10.1016/j.jjcc.2021.09.004

Smith, M. A., Foster, J. R., & Yang, G. Z. (2022). Advances in diagnostic imaging: The role of echocardiography and electrocardiography. Cardiology Clinics, 40(1), 1-16.

Tachibana, Y., Tanaka, H., & Mita, A. (2017). Cardiac diagnostics using Mel Power Spectrogram and machine learning. IEEE Transactions on Biomedical Engineering, 64(12), 2873-2881. https://doi.org/10.1109/TBME.2017.2707599

Tang, H., Tang, J., & Wang, J. (2018). Heart sound classification using support vector machine. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, 1129-1132. https://doi.org/10.1109/EMBC.2018.8512429

Thiyagaraja, S., Kumar, S., & Raj, R. (2018). Analyzing heart sound splitting for detecting cardiovascular conditions. *Circulation Journal*, 82(7), 1753-1760. https://doi.org/10.1253/circj.CJ-18-0123

Thiyagaraja, S., Prabhu, M., & Chandrasekaran, M. (2018). Splitting and murmurs in heart sound analysis: Techniques and applications. *Journal of Cardiovascular Medicine*, 19(7), 415-423.

Wang, H., Wang, L., Liu, Z., & Gao, J. (2023). Practical clinical application of AI-based cardiac diagnostic models in resource-limited settings. *Journal of Biomedical and Health Informatics*, 27(3), 1124-1133.

World Health Organization. (2023). Cardiovascular diseases (CVDs). Retrieved from WHO website

Xue, Y., Chen, T., & Xu, R. (2016). Convolutional neural networks for heart sound classification. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, 2349-2353. https://doi.org/10.1109/ICASSP.2016.7472063

Yao, X., Liu, Y., & Zhang, J. (2023). Deep neural networks for the diagnosis of heart diseases from PCG signals. *Journal of Medical Internet Research*, 25, e39235. https://doi.org/10.2196/39235

Yaseen, Z., Awang, S., & Zainal, A. (2018). Spectrum analysis of PCG signals for heart sound classification. *International Journal of Advanced Computer Science and Applications*, 9(10), 251-257.

Yildirim, O., Talo, M., & Acharya, U. R. (2019). Automatic detection of heart sounds using convolutional neural networks. *Computers in Biology and Medicine*, 113, 103399. https://doi.org/10.1016/j.compbiomed.2019.103399

Zhang, Z., Zhang, H., & Liu, S. (2021). Mel-frequency cepstral coefficients (MFCC) and their applications in speaker recognition and verification. *IEEE Transactions on Audio, Speech, and Language Processing*, 29(1), 345-357.