Automatic Detection of Paediatric Congenital Heart Diseases (CHDs) from Phonocardiogram (PCG) Signals

Submitted by

Muhammad Zeeshan	18PWELE5298
Fazal-e-Haq	18PWELE5351
Zafar Ahmad	18PWELE5359
Muhammad Haris	18PWELE5367

Supervisor

Prof. Dr. Syed Waqar Shah

Co-Supervisor

Dr. Muhammad Salman Khan

Department of Electrical Engineering

University of Engineering & Technology, Peshawar, Pakistan

Automatic Detection of Paediatric Congenital Heart Diseases (CHDs) from Phonocardiogram (PCG) Signals

Submitted by

Muhammad Zeeshan	18PWELE5298
Fazal-e-Haq	18PWELE5351
Zafar Ahmad	18PWELE5359
Muhammad Haris	18PWELE5367

A thesis

presented to the University of Engineering and Technology, Peshawar

in partial fulfillment for the degree requirement of

Bachelor of science

in

Electrical Engineering

Session

2018-2022

Peshawar, Khyber Pakhtunkhwa, Pakistan

UNIVERSITY OF ENGINEERING AND TECHNOLOGY PESHAWAR
DEPARTMENT OF ELECTRICAL ENGINEERING

We hereby recommend that the thesis prepared under our supervision by
Dr. Muhammad Salman Khan, Prof. Dr. Syed Waqar Shah entitled:

"Automatic Detection of Paediatric Congenital Heart Diseases (CHDs) from Phonocardiogram (PCG) Signals"

be accepted in the partial fulfillment of the requirements for the Degree of **Bachelor of Science in Electrical Engineering.**

Prof. Dr. Syed Waqar Shah
(Project Supervisor and Chairman of Department)
Recommendation occurred in:

ABSTRACT

Among many birth diseases, cardiac defect which generally called "congenital heart diseases" occur at common rate in new born babies. Globally, specifically in US, CHD occurs approximately 1%. Listening to the internal body sounds (auscultation) is one of the oldest techniques in medicine to diagnose heart and lung diseases. Cardiologists auscultate the patient using a stethoscope and then confirm the results through echocardiography, however doing so, cardiologist faces many difficulties. Nowadays commonly used diagnostic technique is Phonocardiography (PCG), which is a graphical plot of heart sounds. During auscultation, medical expert and doctors use E-stethoscope for data collection, the data contains either normal or abnormal heart signal. The defective heart sound basic cause the dysfunction of heart valves and turbulent blood flow which are called murmurs. Murmurs produce extra sound along with the heartbeat S1 and S2. Researchers have used different methods for heart signal classification. In the proposed research study, CNN 1D algorithm is analyzed for binary classification; normal and abnormal heart sounds. The model is trained on both local and publicly available datasets. The local dataset has been taken through an electronic stethoscope and divided into two groups, school children (normal) and pediatric patient (abnormal) that are confirmed through echocardiography or by any other expert cardiologist. According to the literature review, all the classification models didn't predict the real-time normal and abnormal heart sounds accurately. It is because in real-time some background noises disturbed our interested region. To deal with this problem trained model on both local and publicly available datasets. Both the datasets have been filtered through low and high pass filter pass-band frequency 60Hz, stop-band frequency 650 Hz, and roll-off 48dB to minimize the background noise. The mean/average of the initial results that were obtained: Accuracy 95.56%, Precision 96.22%, F1 Score 95.72%, Specificity 0.96, and Sensitivity 0.95%. These results seem fearful but there was significant random variance which is not desirable. Details analysis was done to properly investigate the cause of this variance, the most prominent causes were insufficient data and the noise in the local data, similarly the third most probable cause was the difference between the sampling rate of the local data and public data. So, the methodology of this proposed study includes resampling, chunking and padding, data-augmentation, and more specifically pitch-shifting techniques were explored. The final results that obtained: Accuracy 98.56, Precision 98.56%, F1 Score 98.55, Specificity 0.99, Sensitivity 0.98. Cardiac signals are sensitive and in this research study murmurs are considered, accuracy but most importantly high sensitivity.

Acknowledgements

We are thankful to **Dr. Arshad Sohail** and RMI echo Room staff for providing the opportunity to collect data and teaching some of the basic knowledge and medical references that were the building blocks of that novel research. We are grateful to **Ma'am Misha Urooj Khan**, **and Syed Shahid Hussain** for their mentoring and guidance regardless of their preoccupations and busy schedules. We are pleased with the NCAI Healthcare lab for providing facilities throughout this research study. We are also thankful to the research board of RMI and LRH for providing the setting and data collection opportunities, without their cooperation it would not be that much easy.

Dedication

This novel research is dedicated to our parents and brothers who groomed us to this stage so that we can accomplish our goal; they were always there and helped us a lot in getting the such best education. Their blessings and motivation were a great source for us to accomplish this novel research.

Table of Contents

List of Figures

List of Tables

List of Acronyms

PCG	Phonocardiogram
ECG	Electrocardiogram
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
MR	Mitral Regurgitation
MS	Mitral Stenosis
AR	Aortic Regurgitation
VSD	Ventricular Septal Defect
ASD	Atrial Septal Defect
TOF	Tetralogy of Fallot
PDA	Patent Ductus Arteriosus
MVP	Mitral Valve Prolapse
NASA	National Aeronautics and Space Admiration Agency
CHDs	Congenital Heart Diseases
HR	Heart Rate
PS	Pulmonary Stenosis
DNN	Deep Neural Network
LSTM	Long Short-Term Memory
ANN	Artificial Neural Network
SVM	Support Vector Machine
CNN	Convolutional Neural Network
AUC	Area Under Curve
KNN	k Nearest Neighbor

CHAPTER 1 INTRODUCTION

1.1 Overview

In this chapter, brief overview of this prosed research study is presented. Heart defects is one of the significant causes of death in children and adults. These diseases may be presented by birth (Congenital heart diseases CHD's), or acquired after birth at any age (acquired heart disease). For instance, Stenosis of heart valves by Calcium Storage, leakage of valves, etc. These diseases may come out with birth (Congenital Heart Diseases CHD's), or due to some other factors in adulthood, for instance, blockage of heart valves by Calcium storage, leakage of blood in between valves, blood clots in the heart, etc. According to the research of Michigan University, globally, cardiac defects in newborn babies is commonly occur with 1 in 100 babies statistically [1].While in case of cyanotic CHDs, Tetralogy of Fallot (TOF) is generally the most common heart defect [2]. Out of 143 paeds sounds collected locally from Rehman Medical Institute (RMI), VSD has the higher frequency of occurrence which is 48% followed by PDA which has a frequency of 24%. Similarly, ASD and Pulmonic Valve Diseases were diagnosed 10% each in the collected data. There are several types of CHDs such as PDA, ASD, VSD, Pulmonic Valve Diseases, and Aortic Valve Diseases. Study shows that those babies that were born with heart defect significantly changed in the few decades, due to not proper care and prevention died out on large numbers, however, this number is potentially reduced due to early caring system and prevention and now the survival rate is increased to 96% [3]. The two most pioneering work in 1938, one is having children hospital in Boston. Robert Gross" develop a surgical solution for PDA [4], fand the other is "Alfred Blalock and Helen Taussig" in the relief of "blue babies" with tetralogy of Fallot in 1944 [5]. Nowadays a lot of other success stories are circulating.

These surgical initiatives achievements were indeed necessary to decrease the death rate that occurred due to CHDs; however, they were not alone in this area, some other scientists and engineers develop cardiac arteriography and echocardiography in the 1950s required further time to be fully developed and become clinically established in the sixties to seventies, leading to further preciseness in diagnostic systems, real-time imaging, and follow-up of the heart [3].Development in cardiac care didn't stop there, 1980s improvements in CPB (cardiopulmonary bypass) microcirculation hardware, with the development and the advent of the necessary preventive catheter-based cardiac intercede, and further processing in anatomical and biological understanding of single ventricle defects, this was the first successful step toward the multidisciplinary treatment of CHD. At this stage, guidelines were established for the single CHD defects to treat it with high precision near 100% survival and freedom from renovate or resection.

With the advent of modern techniques and technologies, Artificial intelligence, ML, and DL revolutionized almost every end of our life. People and industries move toward autonomous systems. These technologies are cheap and fast. However, it requires a large amount of data, because AI, DL, and ML learn from patterns and features. The terms AI, ML, and DL are based on different computational-algorithms such as CNN, ANN, DT, NB, K-mean, Support vector machine, KNN, and many more. The selection of the algorithm is completely based on the dataset type and accuracy constraints. AI, ML, and DL engineers will take everything into account before selecting and finalizing an algorithm.

Currently, medical specialists and experts use traditional diagnostic techniques to investigate cardiac defects in paeds are expensive and need significant time. In this research study, an AI model that will classify CHD's is presented. Real-time Phonocardiogram signals are imported, after execution, the model will predict as normal or abnormal. The aim of the study is not to replace humans with computers, it will enhance the current diagnostic system by bringing cutting-edge technology into use.

In science and technologies, it is generally true that it brings easiness to human life. However, it has limitations and constraints. AI in healthcare requires almost 100% accuracy and precision because in this research study human life is undertaken, which is sacred. Paediatric dataset is considered which is a combination of a local and public dataset. This study can be extended to multi-class classification when more and more data are available.

1.1.2 The Problem Statement

The current Diagnostic System is expensive and needs time to correctly investigate heart diseases. In manual auscultation, proper time and experience of doctor needs to detect a heart disease precisely and accurately is very delicate, the system can be enhanced by bringing the cutting-edge AI technology into use and developing a system that will automatically detect an abnormality.

1.1.2 Aim and Objectives

- Phonocardiogram Signal Enhancement: Signal Preprocessing, Filtering.
- To develop an AI based model for binary classification of normal and abnormal heart sounds.
- To explain and interpret the decision made by the model.
- To evaluate and compare the performance of the proposed model with other competing models.

1.1.3 Scope/Significance of the study

Currently, AI is the dominant technology where most of the research and development is going on. It drastically changes the way of our daily life. Every action in communication, energy sector, healthcare, advanced manufacturing industries, agriculture, and many others is going toward an autonomous system. Similarly, AI in health care rapidly changes the conventional diagnostic method. Nowadays, scientists develop a nano-type smart robotic chip that is inserted into the human body and is manually controlled by scientists using computer vision. For instance, to check infection in the small intestine or stomach, this robotic chip is sent and thoroughly checks human internal structure before doing any surgical operations. Similarly, nowadays researchers use x-ray image analysis to detect Covid-19. Moreover, scientists use image analysis to detect tumors in the brain. Similarly, there are multiple other research going on heart analysis using AI. Early detection of CHDs will save lives if the proposed methodology is adopted at primary healthcare units. Based on the time constraints and data availability, Paediatric side is focused and for that, analysis was done to develop an AI model that can detect an abnormality. So hopefully, this study will enhance the diseases detection system. The tools used in this algorithm development could be utilized in other one-dimensional signal analysis. Further research in this area will help in understanding of CHDs, local distribution of different types CHDs in Peshawar.

1.1.1 Frequencies wise distribution of heart sounds

At the lower end of this spectrum, between 20 and 500 Hz, are the fundamental frequencies of cardiac murmurs. 1,2 The third and fourth heart sounds as well as the diastolic murmur of mitral stenosis are examples of low-frequency sounds, which are those whose main frequencies are less than 100 Hz. Because the human ear senses lower frequencies considerably less well than higher frequencies, these noises are typically difficult to hear. Aortic regurgitation, whose predominant frequencies are around 400 Hz, is the murmur with the highest frequency sound. Other sounds and murmurs typically range in frequency from 100 to 400 Hz [7].According to Rennert et al. [38], the heart sounds fall between 20 and 650 Hz and are best picked up by the bell of a stethoscope. According to Spencer and Pennington [38],S1 and S2 occur between 50 and 500 Hz, S3 and S4 appear between 20 and 200 Hz, with S3 occurring at the lowest frequency. S1 and S2 appear at the highest frequency. At a frequency of around 300 Hz, other cardiac and pulmonary sounds such murmurs, ejection clicks, and crackles can be heard [8].

1.2 Background

The left atrium and right atrium (upper chambers), left ventricle, and right ventricle are the four chambers of the human heart (lower chambers). According to Figure 1.1 the atria are the upper two chambers and the ventricles are the lower two. The septum is a tissue wall that divides the chambers. Four cardiac valves help the process of pumping blood through the chambers. To allow just one direction of blood flow, the valves open and seal. At the top of the left corner of the heart, there is an electrical system that generates a millivolt pulse, this pulse is responsible for the heart beats and then relaxes muscle to pump the blood flow through the valves. According to medical experts, sometimes there exists two electrical systems that generate two irregular spikes which then create heart problems, so medical doctors disable one of the electrical systems to eliminate the irregularity in the heart muscle and eventually the heart rate (HR).

1. The tricuspid valve, situated between the right chamber and the right ventricle;
2. The pneumonic valve, between the right ventricle and the aspiratory vein;
3. The mitral valve, between the left chamber and left ventricle; and
4. The aortic valve, between the left ventricle and the aorta.

Every valve has a bunch of "folds" (likewise called handouts or cusps). The mitral valve regularly has two folds; the others have three [6].

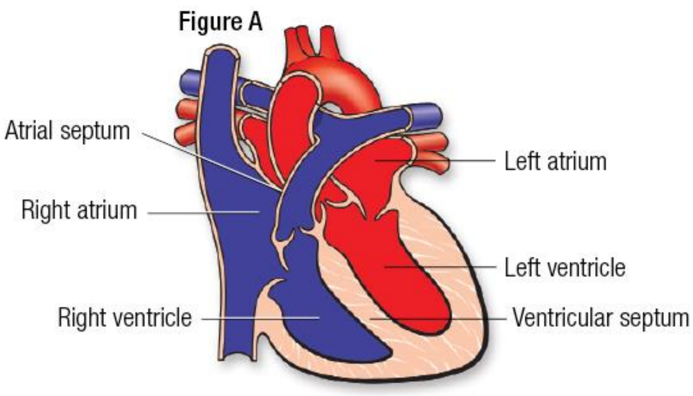

Figure 1.1: Four fundamental heart valves [6]

1.3 Normal heart blood flow patterns.

Clean blood (oxygenated Blood) flows from the lungs to the heart where it pulls to other parts of the body, similarly deoxygenated blood pulls by the heart toward the lungs, so it seems like it follows the pattern body-heart-lungs-heart-body.

1.3.1 Blood flow pattern from the body to heart

Figure 1.2 shows dark bluish color blood, the amount of oxygen is low, flowing back to the heart after circulating through the body. It returns to the heart through veins and enters the right atrium. This chamber empties blood through the tricuspid valve into the right ventricle.

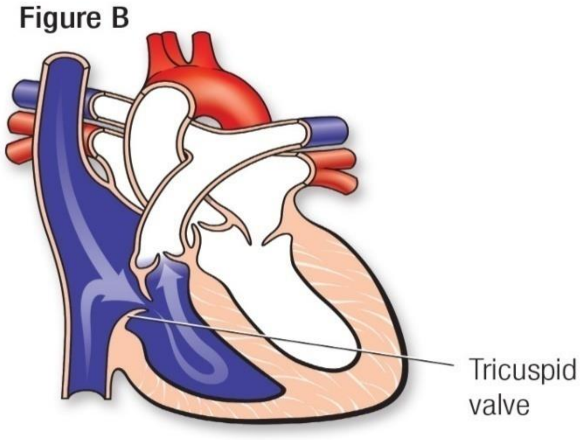

Figure 1.2: Blood flow pattern from the body to heart [6]

1.3.2 Blood flow pattern from the heart to the lungs

Figure 1.3 shows the right ventricle siphons the blood under low tension through the aspiratory valve into the pneumonic supply route. From that point the blood goes to the lungs where it gets new oxygen [6].

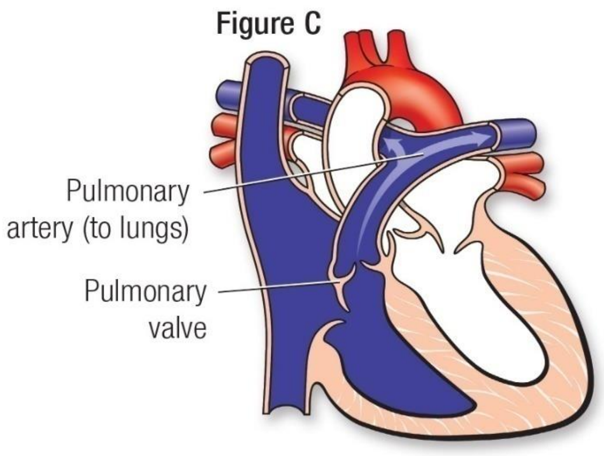

Figure 1.3: Blood flow pattern from the heart to the lungs [6]

1.3.3 Blood stream design from the lungs to the heart

Figure 1.4 shows that after the blood is oxygenated, it's dazzling red. Then, at that point, it gets back to the passed-on heart through the aspiratory veins to the left chamber. From that point it goes through the mitral valve and enters the left ventricle [6].

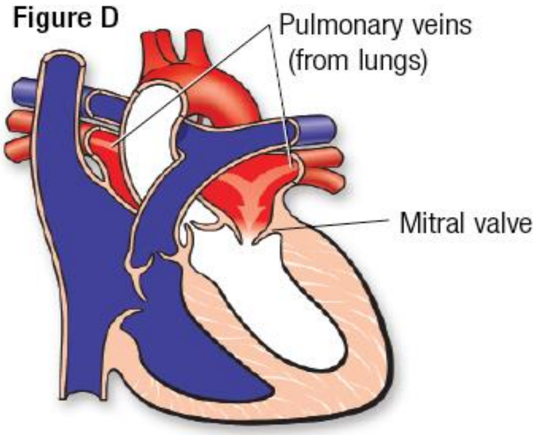

Figure 1.4: Blood flow pattern from the lungs to the heart [6]

1.3.4 Blood stream design from the heart to the body

Figure 1.5 shows that the left ventricle siphons the red oxygen-rich blood out through the aortic valve into the aorta. The aorta takes blood to the body's overall dissemination. The circulatory strain in the left ventricle is equivalent to the tension estimated in the arm [6].

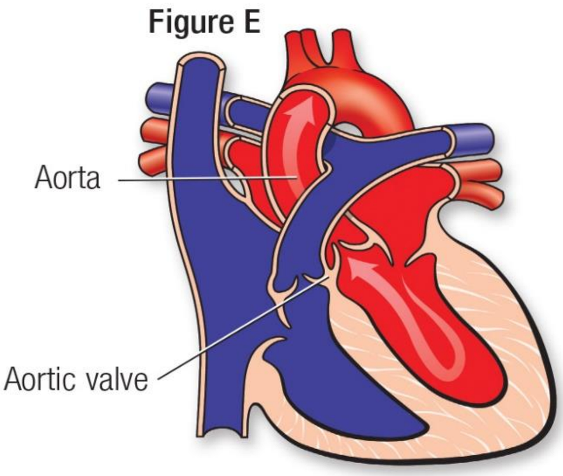

Figure 1.5: Blood stream design from the heart to the body [6]

1.4 Paediatric Heart Murmurs types (the most common ones)

CHDs results when the heart, or veins close to the heart, doesn’t grow regularly before birth. Following are a few distinct sorts of CHDs.

1.4.1 Atrial Septal Defect

Figure 1.6 shows the heart anatomical ASD. It happens when there is an opening or opening which is known as an imperfection in the wall (for short it is called septum) that isolates the best two offices of the heart (likewise called atria).ASD permits oxygenated blood to spill into the deoxygenated blood chambers in the heart. ASD is an imperfection in the septum between the heart's two upper chambers. The septum is a wall that isolates the heart's left and right sides.

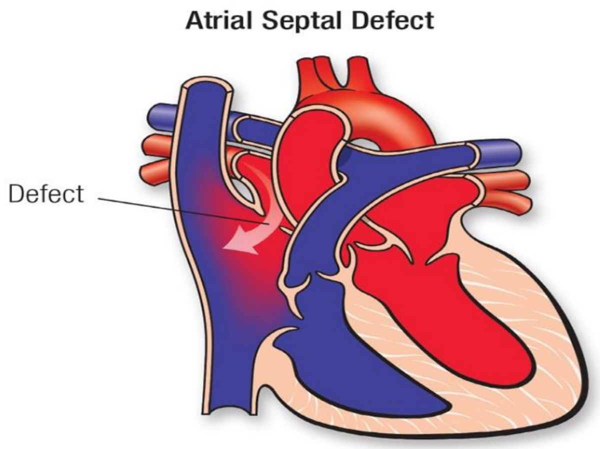

Figure 1.6: Heart anatomical ASD defect [6]

1.4.2 Patent Ductus Arteriosus

Figure 1.7 shows the heart anatomical PDA. At the point when the opening in the fundamental body supply route aorta isn't as expected shut. Before when a child is conceived, the unborn child's blood doesn't have to go to the lungs to get cleaned and oxygenated. The opening that is called ductus arteriosus, permits the blood to avoid the dissemination to the lungs. Be that as it may, when the child is conceived, the blood should be oxygenated in the lungs and this opening should close. On the off chance that the ductus arteriosus stays open, the blood might avoid this indispensable step of dissemination. This open opening is known as the patent ductus arteriosus.

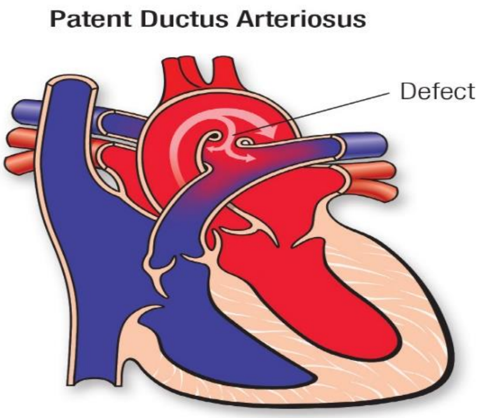

Figure 1.7: Heart anatomical PDA defect [6]

1.4.3 Tetralogy of Fallot

Figure 1.8 shows the heart anatomical TOF.A coronary illness that incorporates the accompanying issues;

- An opening between the lower offices of the heart.
- An impediment from the heart to the lungs.
- The aorta (vein) lies over the opening in the lower chambers.
- The muscle encompassing the lower right chamber turns out to be excessively thickened.

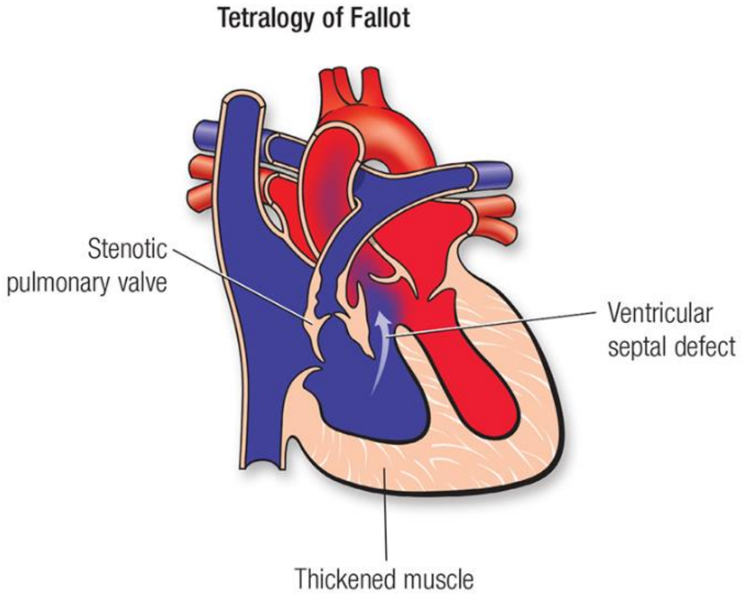

Figure 1.8: Heart anatomical TOF defect [6]

1.5 Basics of PCG (Phonocardiogram)

The phonocardiogram concept was developed by Robert Hooke (1635-1703), at no proper equipment was available until the 1930s. In 1950, the phonocardiogram framework was from that point utilized on one of the Project Gemini space flights, NASA's second human spaceflight program, to screen the pulses of astronauts [7].

Phonocardiography started as an endeavor to portray the event of heart sounds in the cardiovascular cycle with the mechanical movement of the heart, as kept in the pinnacle thump or blood vessel pulsations [37]. Figures 1.9 and1.10 shows the ideal or rough sketch and real-time signal plot respectively.

1.6 Thesis Layout

Chapter 1

In this chapter, brief overview of the proposed research study is presented. It explains what CHD is and what their different types are. It also explains how these diseases were treated at the very early stage of scientific era. Different AI algorithm was analyzed and what were the most successful results. Moreover, aims and objectives of this study are discussed. In addition to that, heart anatomy and different of CHDs are briefly discussed.

Chapter 2

In this chapter, a detail literature review is presented. It highlighted the achievements and results that other researcher have achieved. It also presented the algorithms detail that by using it, the researchers have achieved those results. Finally, a table of literature review is presented.

Chapter 3

In this chapter, the algorithm that is explored in this proposed study is 1D convolutional neural networks. Further, its mathematical background is highlighted with flow diagrams. In addition to that, it is further explained how it works.

Chapter 4

In this chapter, the methodology of this proposed study is explained in details. It briefly explains how the data were collected. Similarly, it also highlights the data pre-processing techniques. It also presents the filter specification, the pitch-shifting technique. Finally, in the last few paragraphs, the effect of pitch-shifting is shown in spectrogram.

Chapter 5

In this chapter, results and discussion of this study is presented in detail. It also explains the issues and challenges that were overcome during this study. It represented very first result that were achieved when the data were noisy, it also highlighted the mean result of twelve-time simulations. Different model evaluation parameters along with their mathematical equations are presented. Finally, it shown the final results which were achieved in this study.

Chapter 6

In this chapter, conclusion and future work were discussed. Finally, limitation of this novel scientific study is briefly highlighted.

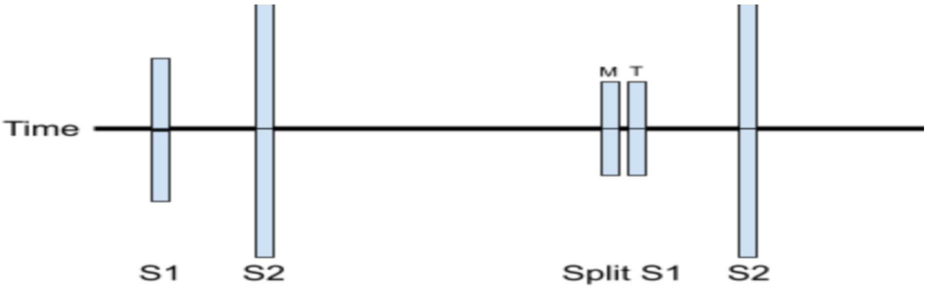

Figure 1.9: Rough sketch of PCG [6]

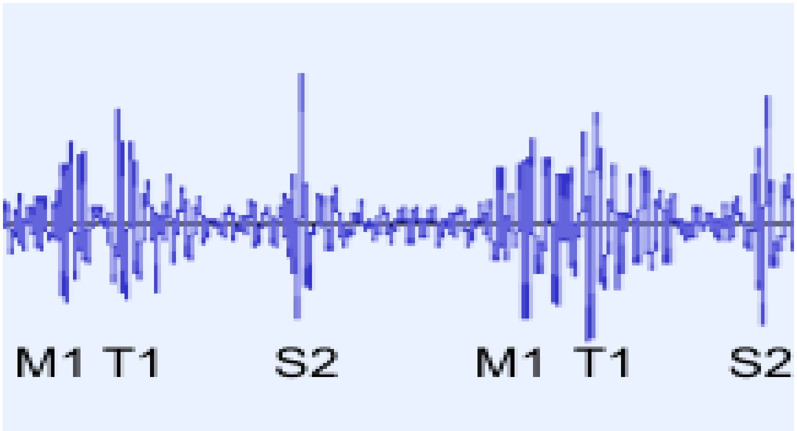

Figure 1.10: Real time plot of PCG [6]

CHAPTER 2 LITERATURE REVIEW

Although PCG signals display secret information about the functioning and anatomy of the human heart, they are frequently used in heart sound modeling approaches. The primary cardiac sound components including S1, S2, S3, also S4 sounds, have been used by researchers in recent years to distinguish between normal and pathological heart sounds.

A recent development in the field of heart sound analysis research is the combination of machine learning and processing methods. In order to identify pathological cardiac diseases, there is a need for sophisticated intelligent heart sound assessment and health monitoring system that can transform heart sound data into practical clinical informatics tools. This could help with early discovery and, as a result, early medicine, which could be vital in preserving a life.

Yaseen *et. al.* [11] focused on the classification of heart sound signals by using the Mel Frequency Cepstral Coefficient (MFCCs) as well as Discreet Wavelets Transform (DWT) to extract characteristics and used three distinct classifiers, including the centroid displacement based k closest neighbor technique, deep neural network, and support vector machine for the classification and learning of heart sound signals with the sample size of 1000 (1 normal class=200 samples, 4 abnormal classes each category consists of 200 samples ,i.e., AS, MR, MS, and MVP) whereas one-fifth of the data sample from each category was employed for testing, as well as fourth-fifth of the sample data from each category was utilized for training. Limitations and research gaps in this article were, that they focused on a very small dataset also the dataset is noiseless, it can be elevated to a large dataset. In real-time auscultation, there is background noise along with heart sound signal for this we need to build a model with a local dataset to overcome the real-time auscultation problems.

Ghaffari *et. al.* [12] concentrated on, supporting an automated PCG diagnostic system for the prognosis of ventricular septal defect (VSD) in infants as well as in adolescents, that dataset which was used in this paper consists of 55 children their age in the middle of 6 months to 2 years (normal=5 children, VSD=20, ASD=7, TOF=4, AS=10, PS=5, and MR=4) by using a variety of techniques for digital signal processing, among them the segmentation, short-time Fourier transform, and autocorrelation, used MFCC as well as its derivatives for feature extraction then classified these features by employing the K-Nearest Neighbors. The accuracy achieved was 93.2 percent, this accuracy could be improved if a large dataset would use, with using the advance features extraction and classification techniques and further this can be extended to classify other types of CHDs.

Chen *et. al.* [13] pay attention to real-time automated analysis of healthy and unhealthy heart sounds to minimize the analysis time of PCG signals without the use of PCG pre-segmentation by using a 1D-CNN as well as long short-term memory network, for the building of this model Physionet Challenge 2016 dataset (CinC)(3099 samples of heart sound were used) is used to train and validate the PCG classification model. Moreover, there are some limitations the author did not classify the specific heart disease, also prior identification and diagnosis of heart abnormality are very handy when dealing with a newborn, where recordings of ECG and other approaches are hard to carry out.

Akbari *et. al.* [14] aimed to improve the classification of PCG signals by using different classifiers, including SVM, kNN, multilayer perceptron (MLP), and maximum likelihood (ML), were combined with various mathematical techniques, such as power spectrum estimation, wavelet transform (WT), and MFCC used for the extraction of features were used on that dataset which comprised of 90 PCG and ECG signals (included normal subjects=40, PS=20, ASD=21 and VSD=12) from the subjects aged among 1 month to 8 years, it was noticed that with the classifier and feature fusion rule the accuracy improved significantly up to 93.2 percent. Furthermore, the accuracy could be improved if a huge dataset was used, also it may be extended to a wide range of pathologies.

Roy *et. al.* [15] worked on to find the best-fitted classifier for valvular heart diseases classification by using three different datasets, first PCG dataset comprised one thousand recordings of PCG (one normal category=200 sample, four abnormal categories each category consists of 200 samples ,i.e., AS, MR, MS, and MVP), second heart sound dataset composed of 461 samples of heart sounds (normal=320, murmur=95, and extra-systolic=46), and third heart sound dataset was Physio Net dataset comprised with 5 training databases composed of 3126 samples of the heartbeat are used in total. It was witnessed in his article that SVM and Random Forest Classifiers are the prime classifiers of machine learning. However, there are certain complications if the amount of training data is quite huge. The best AI classifier for the heart sound is a CNN-based customized deep learning exception network.

Nassralla *et. al.* [16] focused on the classification of regular and irregular PCG recordings using the PhysioNet datasets (3126 heart sound recordings), for this approach they extracted different time and frequency features (Wavelet entropy, MFCCs, as well as power spectrum) from heartbeat signals and given this feature to Random Forest classifier for training to classify the PCG recordings. The proposed approach included the preprocessing, segmentation, feature extraction, and Random Forest learning stages to correctly identify the heart sound signals. Finally, they got an accuracy of 92% on the following PhysioNet datasets. Future study can expand on this work by utilizing long deep learning models (DL) that can classify time series data.

Boulares *et. al.* [17] aimed of the author is to provide an authentic CVD identification model constructed on using CNN. The presented approach is assessed on publicly accessible PASCAL dataset which contained of 425 samples of heart sounds in total (Normal=231, Murmur=129 Extra-systole=65), and the Concerning PhysioNet [89] dataset, which is composed of 665 normal samples of PCG signals, as well as 2575 abnormal samples of PCG signals. They established a segment choosing approach that allows the automatic picking of the most correlated sections following the usage of infinite impulse response filters in the pre-processing stages. Ultimately, to train the pre-trained model on the heartbeat cycles MFCC spectrogram pictures, they made necessary adjustments. For the PASCAL dataset obtained accuracy was 0.87% and, for PhysioNet dataset achieved 0.97% accuracy. In the future, CNN models and mask RCNN for object detection will be merged to improve classification outcomes according to model voting.

Aziz *et. al.*[18] proposed a model which included the following steps; classification, extraction of features, as well as pre-processing. PCG sounds collected from various patients enrolled in Pakistan's Rawalpindi Institute of Cardiology. Dataset consists of a total of 240 signals (ASD=85, VSD=55, and normal=140 signals) that were collected in a hospital environment, for denoising raw PCG signals empirical mode decomposition was utilized. For the extraction of the various features from the denoised PCG signals one-dimensional local ternary patterns as well as MFCCs being used. Finally, SVM classifier was fed the fused feature vector of one-dimensional local ternary patterns as well as Mel Frequency Cepstral Coefficient using 10-fold cross-validation. This methodology achieved a 95.24 percent mean accuracy in recognizing normal subjects, VSD, as well as ASD. Due to its ability to provide cardiologists with a quicker and more objective assessment of PCG signals, this model may be a viable alternative. The presented model can be enhanced by making a large dataset of PCG. Implementing feature minimization and fusion techniques could even reduce the size of the characteristic vectors and improve the accuracy of the system.

Humayun *et. al.* [19] presented innovative Conv layers as teachable FIR filter banks for stethoscope invariant heart sound anomaly detection using a branched 1D-CNN model. The outcomes show how well the learnable filter bank CNN architecture described here achieves robustness against sensor/domain variability for PCG signals, by using publicly available multi-domain datasets, PhysioNet 2016 PHSDB dataset (3153 heart sound samples from 764 subjects were gathered, and 2488 abnormal heart sounds as well as 665 normal heart sound samples), and 2018 INTERSPEECH Compare Dataset (HSSDB) (845 samples from 170 various patients were used). The proposed method will help in deploying automated cardiac diagnostic systems. The top-performing systems for the binary classification of heart sounds are outperformed by the proposed approach. In comparison to state-of-the-art techniques, they were able to achieve relative improvements in Macc of up to 11.84 %.

Khan *et. al.* [20] presented an automatic classification of heart sounds from segmented and unsegmented PCG signals using frequency and time domain parameters was the author's main focus. This suggested strategy made use of the PhsioNet/challenge 2016 publicly accessible database (consist of total recordings of 3126, test recordings of 625, as well as Training recordings of 2501). The methodology used consists of the following stages; heart sound recordings, segmented/unsegmented, pre-processing, extraction of features, finally then classification. Depending on the PCG signal used, this task was split into two types for the extraction of features from PCG signals. Using segmented PCG signals, time and frequency information are obtained in the first kind. In this study, a variety of classification algorithms including support vector machine, kNN, decision tree, ANN, LSTM networks, and ensemble classifier were used to compute MFCC utilizing unsegmented PCG samples for the second type. Long short-term memory classifier and Mel Frequency Cepstral Coefficient characteristics of unsegmented data are used to obtain a better score. This combination outperformed other algorithms, scoring 91.39 percent on the AUC scale. Ultimately, all classifiers performed better than LSTM when using different time-frequency characteristics of segmented data, as well as the ensemble classifier outperformed alternative classifiers by achieving an AUC result of 87.92 percent. CNN and bidirectional LSTM were two deep learning classification techniques that the author plans to use in the development of a PCG signal database. In addition, novel features including power spectral density, discrete wavelet transforms, as well as the Hilbert transforms will be explored.

Dastagir *et. al.* [21] proposed research modified technique for automated classification of phonocardiogram (PCG) signals is defined. Among all the classification machine learning algorithms, four of them were used in this research which includes DT, SVM, KNN, and Tree Bagger (TB). An imbalance data of sample size 3222 was used. Which is further divided into normal sounds of 2563 and abnormal sounds of 659. The dataset is taken from Physio-Net 2016. They have used a different percentage of the training data to train our model. By extending training data from 80% to 95%, SVM has been improved from 87.37% to 94.20%. Similar improvements were made in the final scores of KNN, DT, and TB, which went from 73.64% to 86.50%, 82.90% to 89.07%, and 83.56% to 94.07%, respectively. While utilizing KNN and TB, they were able to get better specificities of 98.4%, and 99.17%, and the sensitivities of mediocre are 72.97%, 65.85%. Similar to how Discreet T fared well improving so accuracy will be 95.31% and Sensitivity of 98.15% at the price of Specificity, which was only 80%. With a Final Sensitivity of 92.30%, Accuracy of 95.31%, and Specificity of 96.08%, SVM scored quite well overall. To strengthen the system and provide better results, the scientists want to include learning-based classification approaches such as CNN and LSTMs in this study in the future. Additionally, several new features, including the power spectral density and the Hilbert transform, will be included.

Jiaming Wang *et. al.* [22] proposed that heart auscultation is being developed to be an intelligent tool utilized in online medicine and is a practical tool for the early identification of heart problems. Few research has been done on the accurate detection of children murmurs caused by congenital heart disease (CHD). Using an electronic stethoscope, 86 PCG signals from 24 children with normal heart sounds and 62 children with (CHD) murmurs were recorded for this study (Shanghai Tuoxiao Intelligent Technology Co, ChildCareG-10, Ltd, China, Shanghai. The maximum duration of recorded sounds was the 20s. First, they differentiate the S1, S2from the PCG signals, closed semilunar valve (CSV), closed atrioventricular valve (CAV), and. The categorization method was then updated to include the CAV and CSV. to classify the CHD's murmur. The DWT technique was used for de-noising the PCG signals with Daubechies 6 as a wavelet basis function. In the classification algorithm, 86 artificial neural networks (ANN) that made up the classification system each had a network of 3 layers structure with 1 neuron in the output layer and 10 neurons each in the hidden and input layers. 10 retrieved characteristics are inputted into each artificial neural network, which then outputs either 0 or 1, showing the murmur of a CHD and the typical heart sound, respectively. The learning rate started at 0.01 and reached its maximum after 1000 iterations, and a goal of 106. These were the parameters chosen for the AI network. With regard to the activation of the hidden layer and output layer, respectively, were linear and hyperbolic tangent functions. The appropriate bias settings for Bias1 and Bias2, as well as the ideal weight parameters for W1 and W2, were discovered using the Liebenberg-Marquardt optimization method. The squared error loss was used to calculate the discrepancy between the projected value and the actual output. The Jack-Knife method was used to evaluate the generalization ability of the classification system. The classifier that had been trained using the other data was validated iteratively by excluding one example at a time. To create a classification system, there were merged 86 artificial neural networks. A threshold of 0.5 was used to quantify the expected value of the classification system to 0 (0.5) or 1 (>0.5). As a result, 85 samples were used to train 86 artificial neural networks, and the one missing sample was used to confirm the results. They attained 93 percent accuracy, 93.5 percent sensitivity, and 91.7 percent specificity. of the heart murmurs.

Xiao *et. al.* [23] focused on utilizing two brand-new lightweight convolution neural networks, they create a computer-aided approach for Pediatric CHDs diagnosis based on deep learning (CNNs). Research studies state that neural networks need a large number of datasets for better prediction and performance. In this paper, they collect their dataset of 528 fine quality recordings of heart sounds containing 193 normal signals from 137 patients aged 12 months to 12 year old. The duration of heart sounds ranges from 3 to 249 seconds. They use 1-D time series data as an input for the 1-D CNN algorithm. To reduce background noise and compress the inputs for future training, 2000 Hz, the original heart sounds are resampled. They separated the model into the dense block-based sub model and the model and the clique block-based model. After implementation, it was observed that the dense block-based method shows better performance than other proposed methods. The achieved performance of the dense block-based model with a sensitivity of 95.36%, specificity of 98.08%, and accuracy of 96.21%.

Zhang *et. al.* [24] proposed a method of two noise reduction algorithm namely called wavelet soft threshold algorithm and variational modal decomposition. The noise reduction is specially done for removing children crying from the recordings. Noise in the recordings mixing irregularly and intermittently with heart sounds can affect the precision of PCG feature analysis and extraction, or by the noise that exhibits bandwidth characteristics or plasticity. They also discussed different de-noising methods, blind source separation algorithm, Wavelet-based threshold noise reduction and empirical mode decomposition-based adaptive noise reduction. In this paper they trained the model on 103 PCG signals which are collected from the hospitals, their dataset contains 37 normal PCG recordings and 66 abnormal heart sounds. All the recordings were then down-sampled to 2000 Hz. They use the frequency range of 30 to 500 Hz for de-noising due to their electronic stethoscope filter. After de - noising, they classified data using an artificial neural network (ANN). The model's performance was reached with 92.23 percent accuracy, sensitivity of 92.42%, and specificity of 91.89%. It has been discovered that the PCG de-noised by that approach increases the performance of the classification system based on the VWG approach and increases the accuracy of the intelligent diagnosis of congenital heart diseases.

Mahmood *et. al.* [25] proposed a method which is called Cardio help, which predicts heart-related fatalities using the deep neural network CNN. They use one of the regression techniques which is called least absolute shrinkage, also known as LASSO, which is useful for reorganizing and choosing variables to increase the predictability and accuracy of the generated model. Dataset has been taken from the UCI standard repository [27]. The LASSO approach aids in the removal of parameters and the choice of variables by condensing the data values to a single central point, or Pc. Regression methods of this kind work well with extremely multicollinear models. They performed both the classification binary and multiclass. They designed the CNN by taking the input from the LASSO algorithm to a fully connected dense layer with 64 neurons. Similarly, following the other layers which are discussed in detail in the recent research paper. They use the learning rate of 0.005 with 0.15 dropout. The achieved performance for binary classification with accuracy of 97%, F1 scores of 96.7%, the precision of 97.06%, and recall of 96.35%. The overall performance of multiclass classification with a precision of 86.69%, accuracy of 86.67%, recall of 81.74.35%, F1 scores of 84.14.7%. In the proposed paper sensitivity and specificity are not mentioned, also they didn't mention the dataset details.

As nowadays artificial intelligence is an emerging technology and its models need a high-performance system to train the models. Because of using a large number of parameters, the preprocessing and complex computation takes a long time to overload the systems. For this purpose, one of the proposed methods related to the fewer parameter consuming model presented in the given research paper by Yunqiu*et. al*. [26]. The purpose of the research is to classify heart-related abnormalities with a fewer number of parameters. They proposed the novel 1-D CNN model for classification which predicts the output based on the patch-level results. The suggested method's primary pipeline may be separated into three sections: employing a unique 1-D deep CNN for PCG patch preprocessing and automated classification, integrating patch-level estimates using a majority vote decision method with recording-level outcomes. In the proposed method 1-D raw signals are imported into the model using a 2000 Hz sampling rate. They use the band filter for noise reduction. The PCG recordings are divided into many patches that are each 3 seconds long and have a 1-second stride. More significantly, PCG segmentation can expand the training set, which is crucial for systems based on deep learning. The models have been trained on the dataset taken from the Physio-Net 2016 with 665 abnormal PCG signals and 2488 normal signals. To assess the classification performance, 10-fold cross-validation is used. With no data augmentation, we train the CNN model from scratch A weighted cross entropy with a rate of 0.25 to 1 (normal to abnormal) is selected as the loss function after taking into account the class imbalance. The stochastic gradient descent (SGD) optimizer is chosen, with a mini-batch size of 64 and a 0.9 Nesterov momentum. Dropout layers with a rate of 0.1 and a weight decay of 10-4 are also utilized to prevent over fitting. The model is trained using our method for 40 epochs with an early stop. The starting learning rate is 0.1, and at epochs 20 and 30, the learning rate decline of 0.1 is applied. The output prediction is based on the majority voting in which predicted patches are counted. Achieved performance with a sensitivity of 85.81 percent, specificity of 95.12 percent, and accuracy of 93.21 percent and a score of 90.46% with only 0.19M parameters. The accuracy of the model is not improved but they decrease the number of parameters and improve the sensitivity and score of the model. Compared to MFCC-CNN [28] the suggested model employs 65 times fewer parameters.

Ahmad *et. al.* [29] focused on a novel approach for the classification of PCG signals (that were taken from the e-general medical having 300 samples which were sampled at 11025 Hz) based on statistical properties using a fuzzy inference system that removes absolute boundaries and assigns a degree of association to every segment of the signal. Before the analysis they filtered low pass using a filter called Chebyshev having Fc 882 Hz. Certain Rules for classification are created and output is computed according to the methods of 1. Envelogram Computation 2. Feature Extraction 2a. The amplitude of Envelopes 2b. The energy of Envelopes 2c. Duration of Envelopes 2d. Duration of zero segments. So, by using this classifier average accuracy to identify different segments of the PCG signal is 97%. The final classification of the output sequence of normal and nearly normal cases is highly accurate up to 100 percent using this technique.

Redlarski *et. al.* [30] proposed that the difficulties in classification of PCG signals is very large, this research focuses to use spectral analysis (used in the LPC algorithm). The methodology in this research is 1. Modified Linear Predictive Coding Algorithm, 2. Support Vector Machine Classifier 3. Modified Cuckoo Search. Overall flow chart of this research is (1) LPC signal processing(a) Acquisition of heart tones (b) signal segmentation (c) Estimation of filter coefficients by the LPC algorithm). (2) SVM-MCS classifier: (a) Creating initial swarm population, (b) Training SVM with initial solutions and fitness function evaluation (c) searching for new solutions by swarm optimizer, (d) Training SVM with new solutions and fitness function evaluation, saving best solution, (f) maximum no of iterations reached. This system is simulated in MATLAB. The 8 types of pathological sounds were created by the LPC algorithm and 4 normal sounds which contains S1, S2, S3, S4. For comparison, three types of Support Vector classifiers (Machine) were created. For all the cases our accuracy reached up to 93%.

DeGroff *et. al.* [31] presented a model to classify between pathological and normal sounds of heart they use ANN because in the field of classification and complex pattern ANN is the perfect tool and this is the purpose of this study. From 69 patients having 37 pathological and 32 normal they took the whole recordings using smart stethoscope each recording takes 5 to 10 minutes. 44.1 kHz sampling rate is used in ANN. FFT is used to obtained the normalized energy spectrum having the frequency ranges from 0 to 90, 150, 210, 255 and 300 respectively.

Nizam *et. al.* [32] presented to classify multiple cardiac abnormalities in time domain from the heart sounds using Hilbert envelope extraction technique is proposed in this research. Upon the presence of SNR (0 to 15dB) and respiratory noise heart sounds was taken from the two states of heart, total of 1000 recordings were taken for abnormal (AS, MR, MS, MVP) sounds of heart and 200 recordings for the normal sounds. All the data contains three cardiac cycles. The dataset is sampled at 8 kHz and each sound is about 2 to 3 sec lengths. An experiment was performed using the following methodology (a) Pre-processing (b) Segmentation (c) Extraction of SCC (d) Hilbert-envelope extraction Envelope feature aggregation. The accuracy, specificity and sensitivity with 0dB SNR of this model is 94.78%, 96.87%, and 87.48% respectively. There are some limitations of this classification that if the first and second sound is not identified by the model effectively so this extraction process may not work and may not classify your input data.

Pedrosa *et. al.* [33] proposed two new algorithms for murmur identification and computational analysis of heart sounds. One is used to divide cardiac sounds into heart cycles, and the other is used to look for heart murmurs. Recordings of 3200 of ages containing 6 to 17 years were taken using the Littman72 signals in total each having lengths< 1-minute from all four action sounds spots. The test sets and the train, and a division according to the number of patients is identified by murmur identification algorithm. A periodic component of the signals is identified by feature extraction method which is based on segmentation algorithm having the sensitivity of 89%. The first gave results with a specificity of 97.21%, minimum error of 2.19%, and sensitivity of 98.42% respectively. With a minimal inaccuracy of 33.65 percent. By adjusting the proportion of segments that must be categorized as noise (murmurs) to be made, the operational point or threshold was determined to be at a specificity of 46.91 percent and a sensitivity of 69.67 percent for a total error of 38.90 percent. Our strategy consists of two stages:1) The signal is divided into systolic and diastolic segments when the primary heart sounds are recognized; and 2) classifier is used to figure out murmur in each portion 3) Murmur Detection: Using the first method, a random train-test division, the minimal error was 4.65 percent for the entire feature set and 2.19 percent for the best subset of 167 features. The error is 35.5 percent for the entire collection of features and 33.65 percent for the ideal subset in the second strategy, which split train-test sets at the patient age. 52.38 percent and 79.40 percent, respectively, are the sensitivity and specificity. The operational point was altered to have a sensitivity of 69.67 percent and a specificity of 46.91 percent to detect more murmur cases at the expense of raising the overall error to 38.9 percent. The overall accuracy rate was 79.2 percent.

Singh *et. al.* [34] proposed an abnormality detection and classification on an unbalanced PCG is a laborious task. The focus of author is totally on that model to classify PCG signals automatically without any segmentation of sounds for detecting any abnormality in heart sounds. This proposed method extracts 5 different time features from the dataset of PCG signals. Our architecture uses traditional classifiers called linear discriminant analysis, vector machine, ensemble, and KNN because PHYSIO-NET 2016 dataset uses KNN. There is linear and nonlinear wavelet but our model uses nonlinear wavelet and achieves sensitivity of 95.04%, specificity of 98.72%, and sensitivity of 97.82%. In comparison to other previously published works of literature, our focus in this research is Paediatric CHDs. We want to enhance the early diagnostic system using advanced AI computation and algorithms

Table 2.1: Literature review details

Paper Reference	Year	Dataset (Normal And Abnormal)	Features	Classification algorithm	Accuracy
Yaseen *et. al.*[11]	2018	1000 samples (normal=200 and abnormal=8000)	MFCC, DWT, MFCC+DWT	SVM, DNN, and k-NN	(MFCC+DWT) KNN acc=97.4% SVM acc=97.9% DNN Acc=92.1%
Ghaffari *et. al.*[12]	2017	55 children samples (normal=5, VSD=20, ASD=7, TOF=4, AS=10, PS=5, MR=4)	MFFC	KNN	93.2%
Chen *et. al.*[13]	2022	Physionet Challenge 2016 dataset (CinC) (3099 samples)	Detect abnormal PCG from unsegmented heart-sound signals without any feature-engineering process	1D-CNN, LSTM, and Conv1D +LSTM	(1D-CNN + LSTM) Sensitivity=87% Specificity=82% MAcc=86% F1-score=91% AUC=92%
Akbari *et. al.*[14]	2019	90 PCG and ECG signals	power spectrum estimation (PSE), wavelet transform (WT), (MFCC)	SVM,k-NN, multilayer perceptron (MLP) and maximum likelihood (ML)	92.5%
Roy *et. al.*[15]	2021	Heart Sound Dataset A (1000 samples). Pascal Challenge Dataset B (461 samples) Physio Net Challenge training set comprised of 5 training databases (A through E) composed a total of 3126 heart sound samples	Root Mean Square (RMS), Signal Energy and Power, Zero-Crossing Rate (ZCR), Total Harmonic distortion (THD), and Skewness and Kurtosis	SVM, KNN, Random Forest, Naïve Baye and ANN	SVM=0.99% KNN=0.76% Random Forest=0.99% Naïve Bayes=0.96% ANN=0.65%
Nassrallaet. *al.*[16]	2017	PhysioNetdatasets (3126 heart sound recordings)	Time features and frequency features (MFCCs, wavelet entropy, and the power spectrum)	Random forest algorithm	92%
Boulares *et. al.*[17]	2021	Pascal dataset: Normal=231, Murmur=129 Extrasystole=65 PhysioNetdataset,	Mel-Frequency spectrum images	CNN	PhysioNet dataset= 0.87% Pascal dataset=

		normal=665 samples, abnormal=2575 samples			0.97%
Aziz *et. al.*[18]	2020	Dataset(local hospital data) Normal=140 samples 28 subjects, ASD=85 samples from 17 subjects, VSD=55 samples from 11 subjects.	Fusion of 1D-LTPs and MFCCs	SVM (using 10-fold cross-validation)	95.24%
Humayun *et. al.*[19]	2020	2016 PhysioNet Heart Sound Database (3153 HS recordings. 2018 INTERSPEECH CompareDataset (845 recordings from 170 different subjects).	1-D time series signals	tConv layers with a branched 1D-CNN model	Improvements of up to 11.84% in terms of MAcc, compared to state-of-the-art methods
Khan *et. al.*[20]	2020	PhysioNet/ challenge 2016 (3126 samples)	Time and frequency domain features (segmented) and MFCCs (unsegmented)	k-NN, decision tree, ensemble classifier, ANN, SVM, and LSTM	91.39%
J.Dastagir et. al. [21]	2021	PhysioNet/Challenge 2016	Multidomain ,i.e., frequency and Time	SVM, KNN, DT & TB	SVM outperformed Accuracy= 95.31%,sensitivity=92.30%, specificity= 96.80%
J.Wang et. al.[22]	2020	Locally collected (86 children, 24 normal and 62 CHD murmur)	Time and frequency	ANN	Accuracy=93%, sensitivity = 93.5%, specificity =91.7%
Xiao et. al. [23]	2019	Locally collected (MySQL)	1-D raw waveforms	Two 1-D CNN models (dense/clique)	Accuracy= 96.21%, sensitivity= 95.36%, specificity= 98.08%
Zhang et. al. [24]	2022	Locally collected (103 phonocardiogram samples)	Time and frequency	VMD, WST	Accuracy = 92.23%, sensitivity= 92.42%, specificity= 91.89%
Mahmood et. al.[25]	2020	UCI repository	--	CNN	Binary classification accuracy=97% and multiclass accuracy=

					86.67%
Yunqiu et al.[26]	2018	PhysioNet/Cinc2016	1-D raw waveforms	Noval 1-D CNN	Accuracy = 93.21%, sensitivity= 85.81%, specificity= 95.12%
Ahmad *et. al.*[29]	2014	e-general medical 300 samples	Time domain	fuzzy inference	97%
Redlarski G *et. al.*[30]	2014	Public dataset Locally collected	Time domain	SVM-MCS And Artificial Neural Network	93%
DeGroff C *et. al.*[31]	2001	69 sounds (37 pathological and 32 innocent murmurs)	Time and Frequency domain	ANNs	98%
Nizam *et. al.*[32]	2020	200 sounds	standard acoustic and Resnet	Hilbert envelope feature	Accuracy = 94.78% sensitivity = 87.48%
Pedrosa J *et. al.*[33]	2014	3200 sounds	Multi domain time and frequency	segmentation algorithm	sensitivity = 69.67% Accuracy = 79.40%
Singh et. al.[34]	2022	public PhysioNet	Time domain	KNN	97.82%,

CHAPTER 3 1D CONVOLUTIONAL NEURAL NETWORKS (1D-CNN)

3.1 Introduction

Over the previous ten years, for so many Computers Vision and Machine Learning processes Convolutional Neural Networks becomes the standard. Convolutional Neural Networks are called the feed forward artificial neural networks where the subsampling and convolution are alternated. It is capable of learning to a sizable dataset with proper labeling, deep 2D CNNs having millions of parameters, and numerous hidden layers that could understand complicated materials and shapes. Also with appropriate training, it can serve as the main tool for a variety of technological operations for 2-dimensional data, like pictures and video clips. Nonetheless, this could not be an effective alternative in many applications compared to 1-dimensional data, specifically if the data for training is limited. 1D CNNs have recently been presented as a solution to this problem and have already attained the contemporary level of performance in a variety of situations, having identification and the timely treatment of customized biomedical data, structural health monitoring, and anomaly detection. Due to the simple and inexpensive setup of 1D CNNs, which only perform 1D convolutions, another significant benefit is the feasibility of affordable and real-time circuitry installation [35]. In 1D CNNs, there are two categories of layers: Sub-sampling as well as 1D convolutions take place in CNN layers, and layers which are completely connected and resemble the output and invisible layers of a standard multi-layer perceptron (MLP).

3.2 Comparison of 1D and 2D convolution neural networks

According to certain studies, the mentioned factors make 1-dimentional CNNs preferable to those 2D equivalents when handling 1-dimensional data for a particular operation [35].

3.2.1 In terms of computational complexities

While the analogous 1D convolution (N and K both possess the same sizes) has the computational complexity of ~ O (NK), The complexity of computation of a picture having N x N sizes convolved with a K x K kernel is ~ O (N^2K^2). This implies that the 1D CNN's calculation complexity is considerably less from 2D CNNs under comparable circumstances (same arrangements, network, and hyperparameters). So, the ideal solution for the real-time as well as for small-budget applications is 1D CNN.

3.2.2 In terms of hardware

Deep 2D CNN training typically requires specific hardware (such as cloud computing). To the flip side training, every central processing unit to put into actions on common computer was possible as well as rather quick to tiny 1D CNNs including little invisible layers (such as two, or lower than 2) as well as neurons (such as lower than fifty).

3.2.3 In terms of parameters

As a general observation, particularly in light of recent studies, small configurations having one-two hidden layers of CNN and systems with less than 10k characteristics were used in the majority of 1D CNN operations, whereas almost every 2-dimensional CNN system utilized "deep" designs including over 1M (usually above 10M) parameters. Systems having shallow architectures were easier to create as well as operate.

3.3 1D-CNN overview

Compact 1D CNNs have shown improved performance in the abovementioned recent research for applications having large signal variations along with few labeled data obtained out of a plethora of sources (including PCG or ECG patients, structural, automotive, and aeronautical) [35].

Two unique layer types are suggested for 1D CNNs, as depicted in Figure 3.1. The first one is called "CNN-layers," in which sub sampling (pooling), activation functions, along with 1D convolutions all are used.2) The second one is called “Fully connected layers” which are referred to as "MLP-layers" because they are identical to the layers of a standard Multi-layer Perceptron (MLP). Following hyper parameters combine to make a 1D-CNN configuration:

1) Amount of concealed MLP and CNN layers (2 concealed MLP along with 3 CNN layers, in 1D CNN illustrated by Figure 3.1).

2) Size of the kernel in every CNN layer (as Illustrated in Figure 3.1, kernal size have fourty one in all concealed layers of CNN).

3) The sub sampling factor for each CNN layer (as depicted in Figure 3.1, which is 4).

4) Selection of the activation functions as well as pooling.

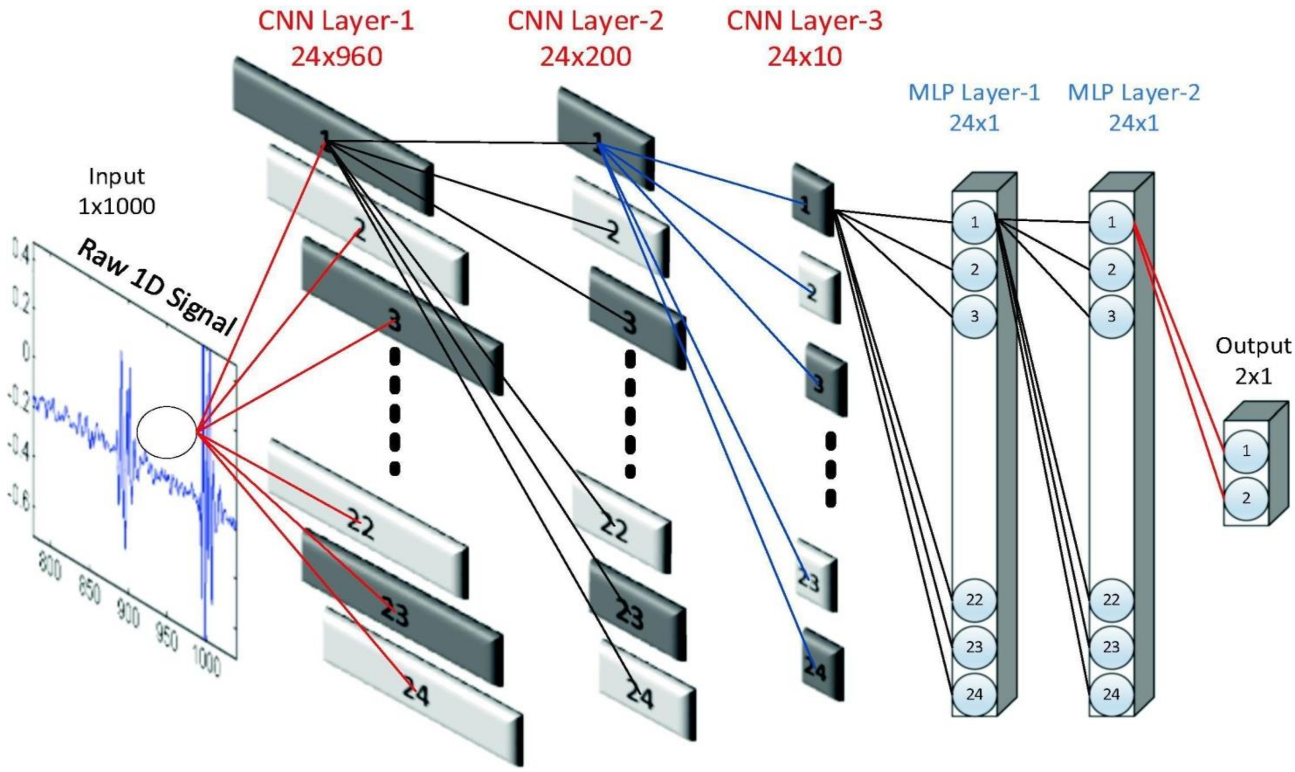

Figure 3.1: Sample 1D CNN arrangement with 3 CNN and 2 MLP layers [35]

Having a sub sampling factor of two along with a kernel size of three, the 3 sequential CNN layers are displayed in Figure 3.1. Only as an example, these network configurations were chosen.

Figure 3.2 shows three consecutive hidden CNN layers of a 1D CNN. Whenever convolution 1D is executed, the output characteristic array dimension in the leftmost layer (l-1) is 22, as well as in the following layer (l), the input characteristic size of the array is 20, where the size of the output feature array will be ten (10) with a sub sampling of factor 2.

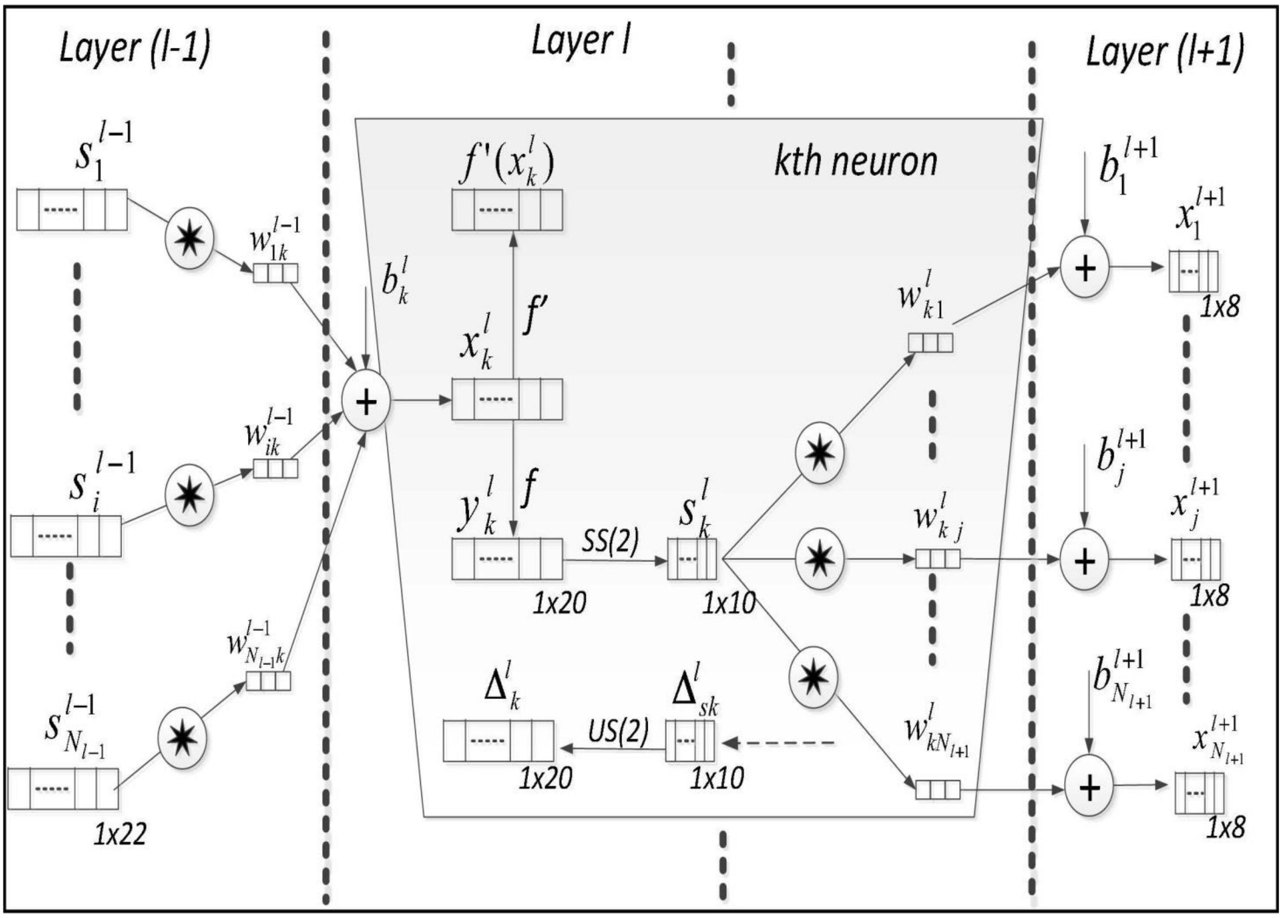

Figure 3.2: Three consecutive hidden CNN layers of a 1D CNN [9]

Another convolution with a kernel size of 3 will result in the input characteristics array at layer (l+1) having a final size of 10-(3-1) = 8. To show the adaptiveness of the proposed 1D-CNN, suppose layer l + 1 is the final CNN layer. The final array size, in this case will be 1, allowing for a direct connection to the deep MLP layers, and the sub-sampling factor will be flexibly tuned to eight to achieve this. As a result, the output size will always be 1, no matter how many CNN layers are present or how long the input signal is. This strategy aims to allow complete freedom in setting the network parameters, likely number of the CNN layers, subsampling factors, as well as kernel sizes.

In suggested method, 1D CNNs are configured in a straightforward and shallow way, with very few hidden layers and neurons. The network typically has only a few hundred parameters, and it can adjust to the length (resolution) of every PCG segment [36].

3.4 Forward- and back-propagation in CNN-layers

3.4.1 Forward Propagation

Get input data, interpret the information, and generate an outcome.

The 1D forward propagation is represented like this in each CNN layer:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv1D(w_{ik}^{l-1}, s_i^{l-1}) \quad (3.1)$$

Where the input is x_k^l, at layer l, the bias of k^{th} neuron is b_k^l, and at layer $(l-1)$, the output of the ith neuron is s_i^{l-1}. The kernel, designated as w_{ik}^{l-1} runs as from the i^{th} neuron at layer l-1 to the k^{th} neuron at layer l.

3.4.2 Backward Propagation

Update the network's parameters and compute error.

The error's BP begins from MLP final layer. Assume that l= 1 and l = L are the input and output layers accordingly. The final layer's error can be expressed as,

$$E = E(y_1^L, y_2^L) = \sum_{i=1}^{2}(y_i^L - t_i)^2 \quad (3.2)$$

Regarding the input vector p and its corresponding goal and output vectors, $[y_1^L, y_2^L]$ and $[t_1^L, t_2^L]$ respectively. The derivative of this error regarding a certain weight (connected to the neuron, k) w_{ik}^{l-1} , as well as the bias of that neuron k, b_k^l what we are interested in findingas a result, we can use the gradient descent approach to reduce the error appropriately. If the BP has calculated all of the delta errors in each MLP layer, then the weights as well as bias of every neuron could be modified using the gradient descent method. The delta inaccuracy of that k^{th} neuron at layer l, Δ_k^l, will specifically be used to update that neuron's biased and all of the weights of that neurons in the preceding layer connected to that neuron as follows:

$$\frac{\partial E}{\partial w_{ik}^{l-1}} = \Delta_k^l y_i^{l-1} and \frac{\partial E}{\partial b_k^l} = \Delta_k^l \quad (3.3)$$

Therefore, simply regular BP is carried out from MLP input layer towards final layer of CNN as follows:

$$\frac{\partial E}{\partial s_k^l} = \Delta s_k^l = \sum_{i=1}^{N_{l+1}} \frac{\partial E}{\partial x_i^{l+1}} \frac{\partial x_i^{l+1}}{\partial s_k^l} = \sum_{i=1}^{N_{l+1}} \Delta_i^{l+1} w_{ki}^l \quad (3.4)$$

Further back-propagation to input delta is possible, Δ_k^l, once the first BP has been carried out from that following layer l + 1, up to present layer, l. Let $s_k^l = up(s_k^l)$ be the zero order up-sampled map, then one can write:

$$\Delta_k^l = \frac{\partial E}{\partial y_k^l}\frac{\partial y_k^l}{\partial x_k^l} = \frac{\partial E}{\partial us_k^l}\frac{\partial us_k^l}{\partial y_k^l} f'\left(\mathrm{x}_\mathrm{k}^\mathrm{l}\right) = up\left(\Delta s_k^l\right)\beta f'(\mathrm{x}_\mathrm{k}^\mathrm{l}) \tag{3.5}$$

Where $\beta = (ss)^{-1}$. The Back Propagation of delta error $(\Delta s_k^l \xleftarrow{\Sigma} \Delta_i^{l+1})$ therfore be represented as,

$$\Delta s_k^l = \sum_{\mathrm{i=1}}^{\mathrm{N_{l+1}}}(\,\Delta_i^{l+1}, \mathrm{re}v(w_{ki}^l)\,) \tag{3.6}$$

In this case, reverses the array, as well as $conv1Dz$ executes a whole 1D convolution with K-1 zero padding. Last but not least, the weight and bias sensitivities are given as [36],

$$\frac{\partial E}{\partial w_{ki}^l} = conv1D\left(\Delta s_k^l, \Delta_i^{l+1}\right) = \sum_n \Delta_k^l(n) \tag{3.7}$$

CHAPTER 4 METHODOLOGY

In this research study, various researchers' works and techniques were explored as shown in the literature review Table 2.1. In this research study, a combined dataset containing both local and publicly available data was used. The overall research process of automatic Paediatric heart sound analysis and abnormality detection process is illustrated in Figure 4.1 and is detailed in the following subsections.

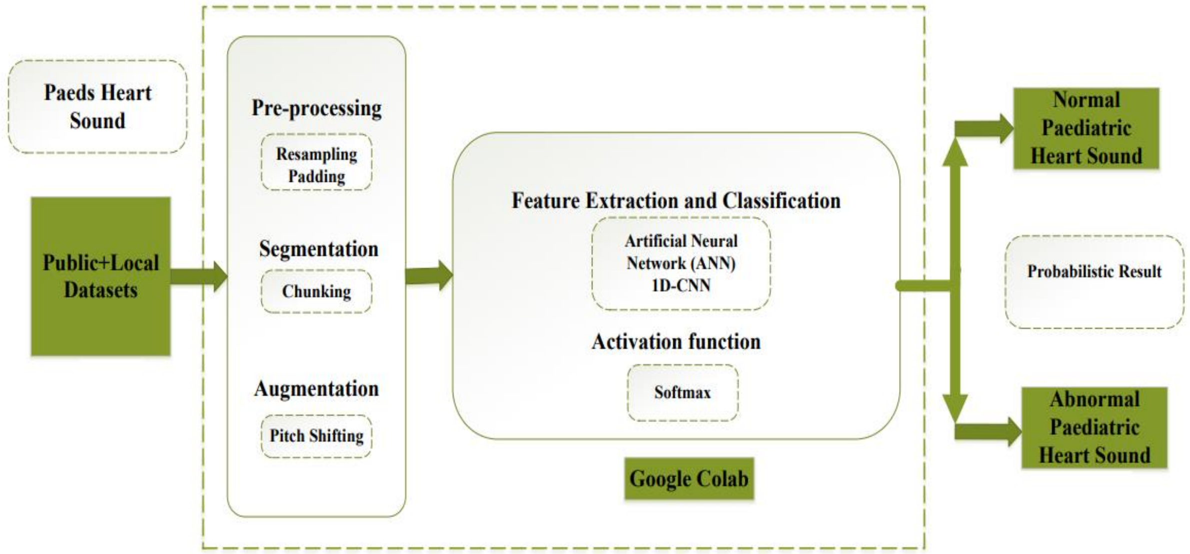

Figure 4.1: Block diagram of the proposed methodology

4.1 Datasets

Figure 4.2 shows the dataset statistics. The local data were collected from Rehman Medical Institute (RMI) Hospital Peshawar and Lady Reading Hospital (LRH) Peshawar while the remaining data were taken from the publicly available dataset.

4.1.1 Local Dataset

In the summer of 2021, the collected data contains 143 abnormal Paediatric heart signals (PCG) having different congenital heart diseases (CHDs) from different local hospitals including RMI and LRH in Peshawar. Each signal recording time was 10 to 12 seconds; however, these signals were noisy and less in number. The sample size of our research was 800 signals; 400 abnormal heart signals and 400 normal heart signals. So, the alternative option was to use publicly available datasets and combine them.

4.1.2 Publicly Available Dataset

(Professional Skills Builder Heart Sounds Dataset [10])

This dataset contains 23 heart sounds of high quality with comparably less noise. The signal recording time ranges from 1 minute to 1 minute and 15 seconds. The overall summary of the diseases in the dataset is:

Arterial Septal Defect (ASD), Ventricular Septal Defect (VSD), coronary artery disease (CAD), Mitral Valve Prolapse, Acute Mitral Regurgitation, Pulmonary Valve Stenosis, Severe Aortic Stenosis, Mital Regurgitation, Mitral Stenosis, Aortic Regurgitation.

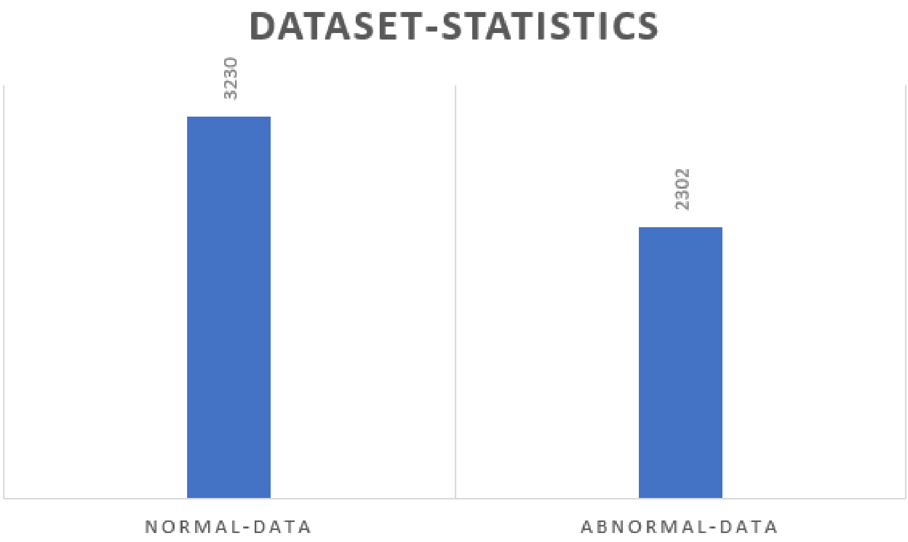

Figure 4.2: Dataset statistic

4.1.3 Splitting the Recorded Signals

Multiple research papers were analyzed to figure out the appropriate signal time and, in most papers, researchers used 8 to 12 seconds, so after confirmation split time for each signal of the dataset using the same standard was chosen.

4.2 Data Pre-Processing

The collected data was quite noisy and insufficient. The most dominant effect was ambient noise and more specifically, baby crying, public talking noise, stethoscope rubbing sound during recording, and other environmental noise. These all factors produce a common effective noise that can disturb the accuracy, specificity, and sensitivity and can often confuse the AI model, so proper filtering and preprocessing techniques were required.

4.2.1 Filtering

The ambient noise usually falls in the low-frequency region between 0 and 50 Hz. Similarly, when something rubs with the ground or a baby cry, then this noise falls in the frequency region above 800 Hz. So, a Bandpass filter was needed. In this research, Audacity software was used which is one of the most widely used software for audio signal processing. Filter specification is shown in the Table 4.1.

Table 4.1: Filter specification

Required Band (Hz)	Pass Band frequency (Hz)	Stop Band frequency (Hz)	Roll-off (dB per octave)
60-650	**60**	**650**	**48**

4.2.2 Resampling

The local data was collected at the sampling rate of 8000 Hz. However, the publicly available data was sampled at 44.1 kHz, so all the data were sampled to 8000 Hz. If the dataset was not balanced in terms of all parameters such as sampling rate, normal and abnormal samples, then it often shows poor and sometimes unacceptable results. So, the dataset must be balanced in all parameters.

4.3 Segmentation

4.3.1 Chunking and Padding

AI models required a huge amount of data, and models trained with insufficient data may not perform well during testing. The data used for the proposed research was not sufficient for keeping the importance of AI results on the medical side, so the dataset was chunked to 4 seconds per sample. Hundreds of audio samples were not the multiple of 4, so the technique that was used is called padding. The basic purpose of that technique was that it adds the previous part of the signal as heart signals are periodic in general. After doing this process, the data was significantly increased and now it was approximately suitable for the AI model.

4.4 Data Augmentation

AI model requires a large amount of data to learn from. Data requirements are completely based on the purpose for which the AI model is trained. For instance, for business-related purposes, accuracy might be prioritized. Similarly, for aerospace engineering, the AI model should be précised, however for medical and healthcare-related purposes, sensitivity is highly concerned. All the technological advancement is doing to facilitate human life is life. If AI trained model predicts normal out of thousand decisions, it means that a was actually an abnormal patient which AI model predict wrong, so it might not be recommended for further treatment and thus human life could be loss, so to avoid such cases, highly well-trained AI model is needing and this is only possible when we have a huge dataset. If in case, the dataset was insufficient, then it is need for to create synthetic data of slightly improved parameters such a process is called Data Augmentation. There are multiple approaches that researchers used for their purposes, such as Pitch-Shifting, Time Stretching, Time compressing, etc. In this proposed study, pitch shifting which was the best for tackling this issue was explored.

4.4.1 Pitch-Shifting

This technique is often used to change the high-frequency component in an audio signal. Heart sound has S1 and S2 spikes which have high amplitude and high frequency, so using this technique will increase the S1 and S2 spikes to some higher frequency magnitude. Doing this will not affect the original signal, but rather slightly modified signal of slightly improved frequency components achieved. Some resultant pictures of the original and synthetic signal are shown in Figures 4.3 and 4.4.

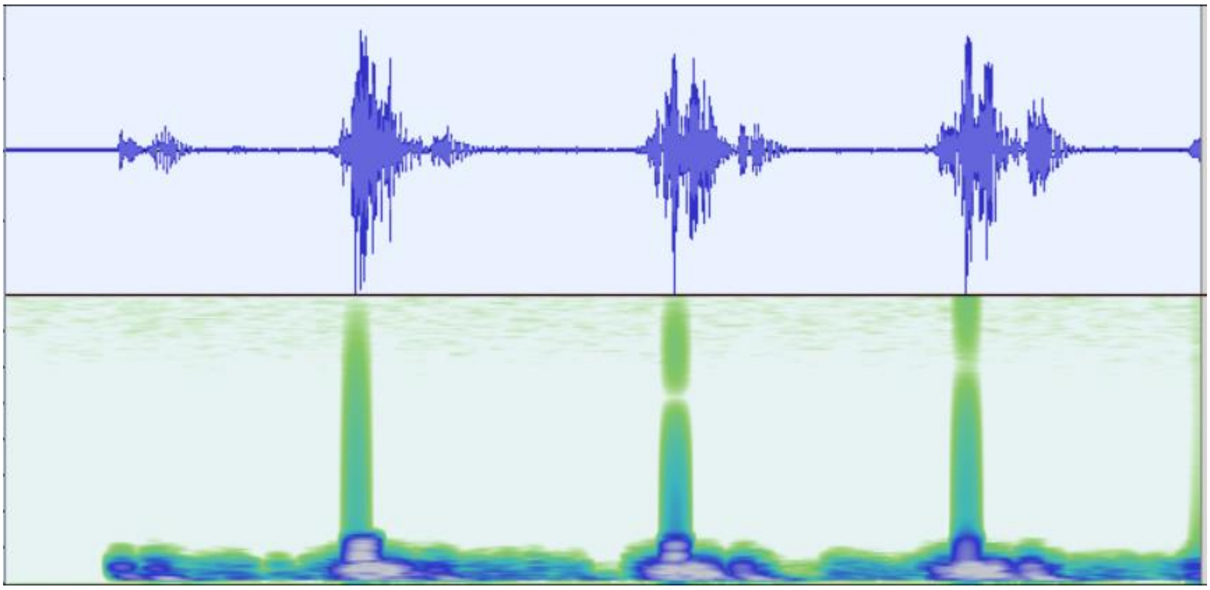

Figure 4.3: Original abnormal signal

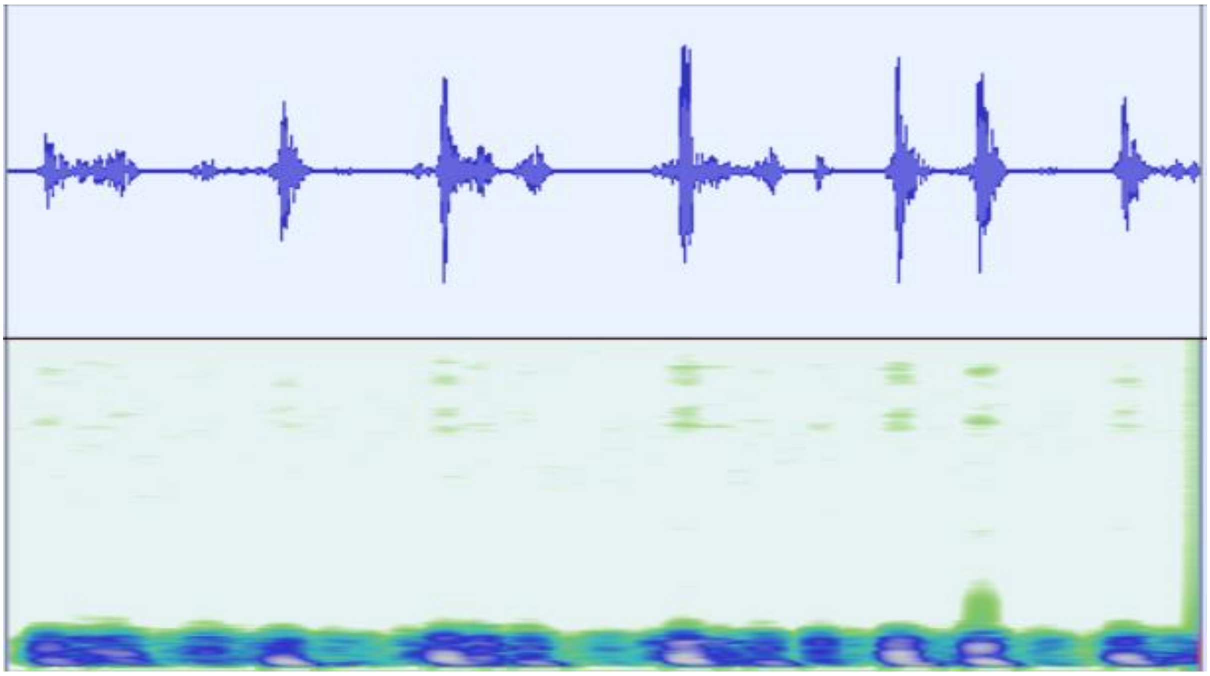

Figure 4.4: Synthetic abnormal signal

The upper portion in both cases shows the Phonocardiograph of the signals while theattached Figure 4.5 is the spectrogram. Original abnormal signal is shown in Figure 4.3. It can be observed in Figure 4.4 that the modified signal comparatively high-frequency magnitude. The filtered signal has a frequency band of 60 to 650Hz, after pitch shifting the band rises to 60 to 800Hz.

4.5 Spectrum Analysis

In Figure 4.5, the pitch-shifting is shown using a spectrum analyzer, spectrum analyzer gives us the overall band of frequencies that exist in the signal. It can be seen in Figures 4.5 and 4.6, that the pitch-shifted signal frequency band is increased from 250Hz to 350Hz and most importantly, high-frequency spikes improved.

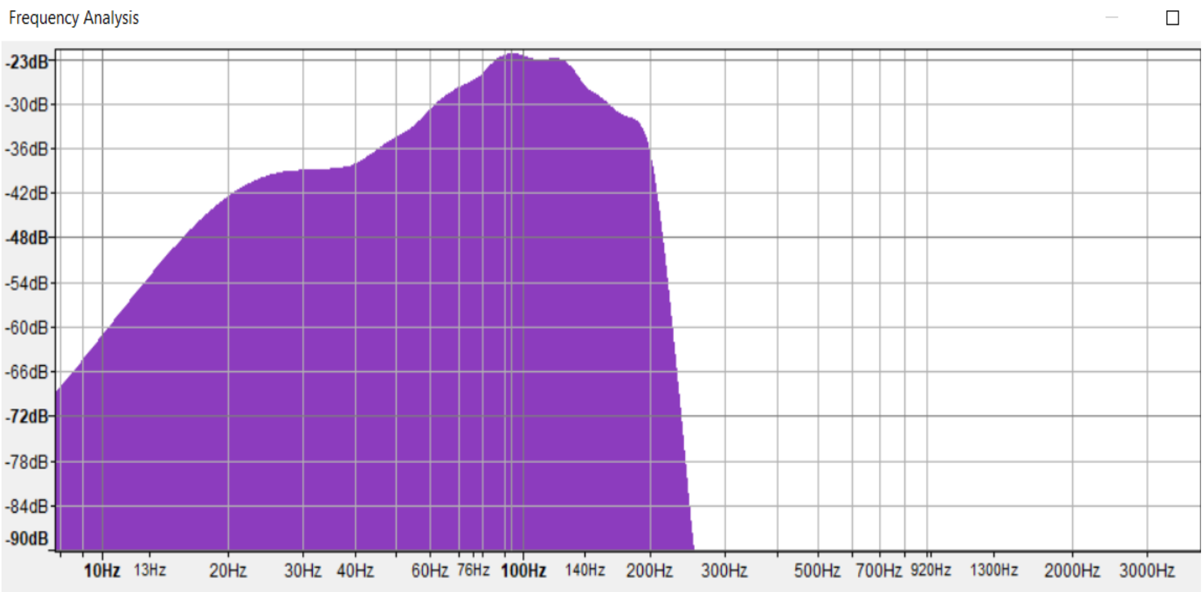

Figure 4.5: Spectrum of original signal

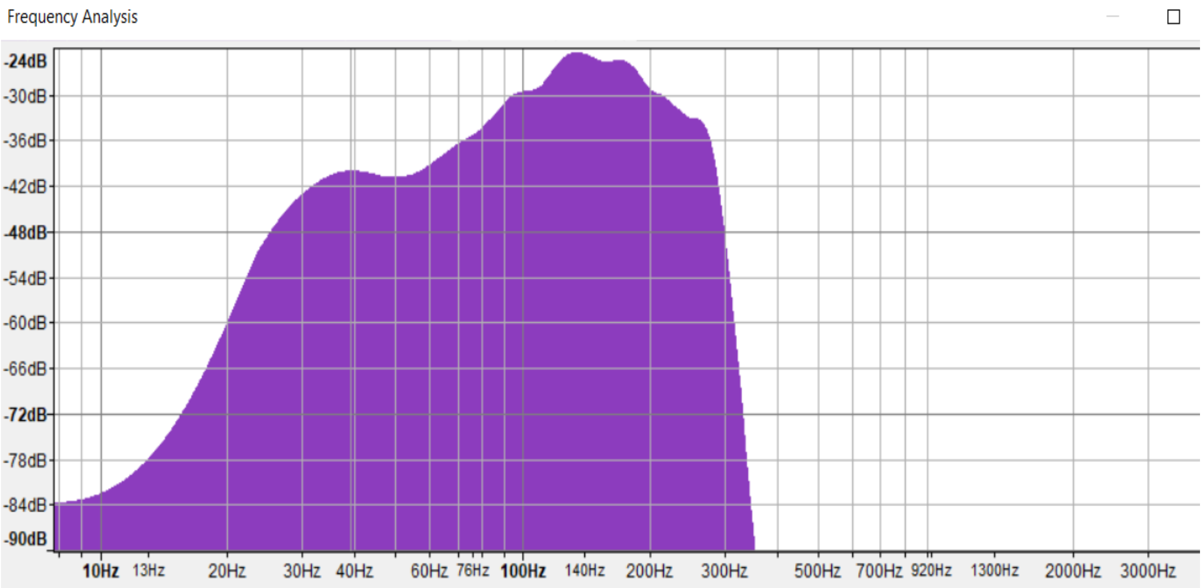

Figure 4.6: Spectrum of pitched signal

4.6 Feature Extraction and Classification

A convolution 1D algorithm was used for this research. Several other algorithms are used by other researchers worldwide. However, algorithm selection is based on the purpose, data type, accuracy, precision, sensitivity, and specificity requirements. The local dataset used for this proposed study was noisy, and this low-frequency noise is persistent. A study has shown that 1D CNN has good results on noisy data using audio signals.

4.6.1 Softmax

The Softmax is a function often used in AI, ML, and DL models that converts a real number into the probability of K possible outcomes in generalized form. Moreover, it is a generalization of the logistic function to multi-dimensions and is used in logistic regression. Softmax converts the last results to normalized distributed for probabilistic results and it is based on Luce’s choice axiom.

CHAPTER 5 RESULTS AND DISCUSSION

The initial results were based on the combined dataset of local and public data; however, they were fear but not in the acceptable region and it was not fulfilling the criteria of medical standards. The initial investigation showed that the most prominent cause of the faulty results was an unbalanced and insufficient dataset. Secondly, the local data was quite noisy even some of the .wav files didn't hold any data.

5.1 Challenges and Issues

- Insufficient Data
- The available local data was very noisy
- The most suitable chunk time

5.1.1 Insufficient Data

This issue was hard and crucial because the AI model needs a large amount of data. Various techniques can be used for synthetic data creation and in this study pitch-shifting is used and its selection was based on the data sensitivity and result requirements.

5.1.2 The available local data was very noisy

The most difficult challenge was the ambient noise in the signal that affected the overall model performance, so the effective strategies that were brought into action in this proposed study a). all the data was thoroughly listened to and removed the unnecessary data files, b). After that, the signals were then passed through a band pass filter having the specification shown in the Table 4.1.

5.1.3 The most suitable chunk time

The recorded signals were of different time frames, so it was necessary to chunks all the data into the one-time frame, but how many seconds? This was one of the vital questions that must be referred to medical experts, after confirmation the chunks time was decided to be 4 seconds.

5.2 Google Colab

It is an online research lab that provides GPU, RAM, and Memory for scientists, engineers, and programmers to simulate AI models.

After solving the above challenges, the model was simulated in GoogleColab and the following initial result was obtained.

5.3 Initial Result

This result was simulated under the following conditions. In Figure 5.1, it can be seen a rapid fluctuation and no proper synchronization between the training accuracy curve and validation accuracy curve. It shows that the model was not properly trained during its training process. In other words, the model was confused during training. Similarly in Figure 5.2, the validation curve shows significantly high random variance, also the training loss and validation loss are not synchronized properly. So, this initial results glimpse on the weak zone of our data set. However, the initial results are unacceptable and can't fulfilling the medical standards. So, before changing and improving the dataset, our focused was to change the algorithm parameter and got the following results.

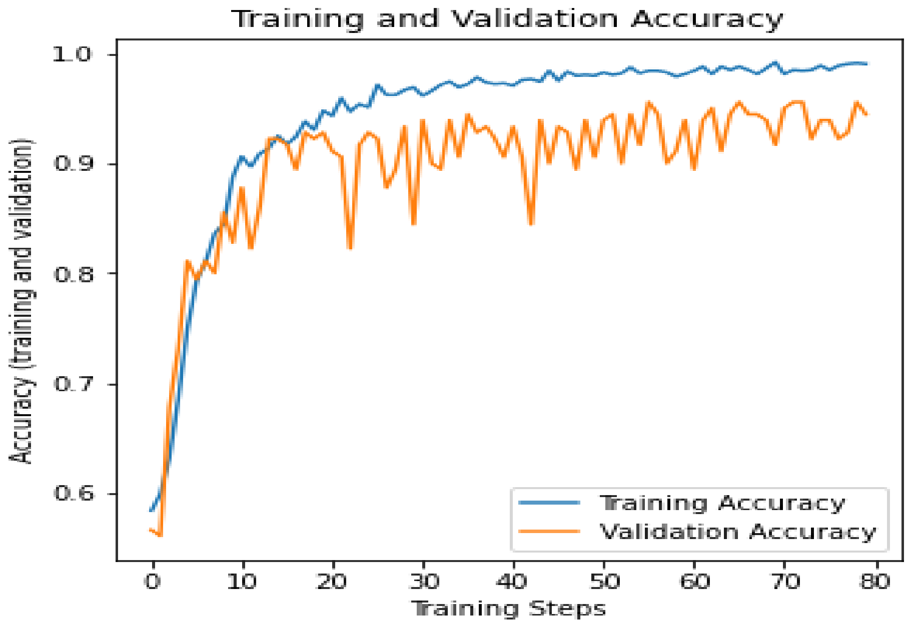

Figure 5.1: Training and Validation Accuracy

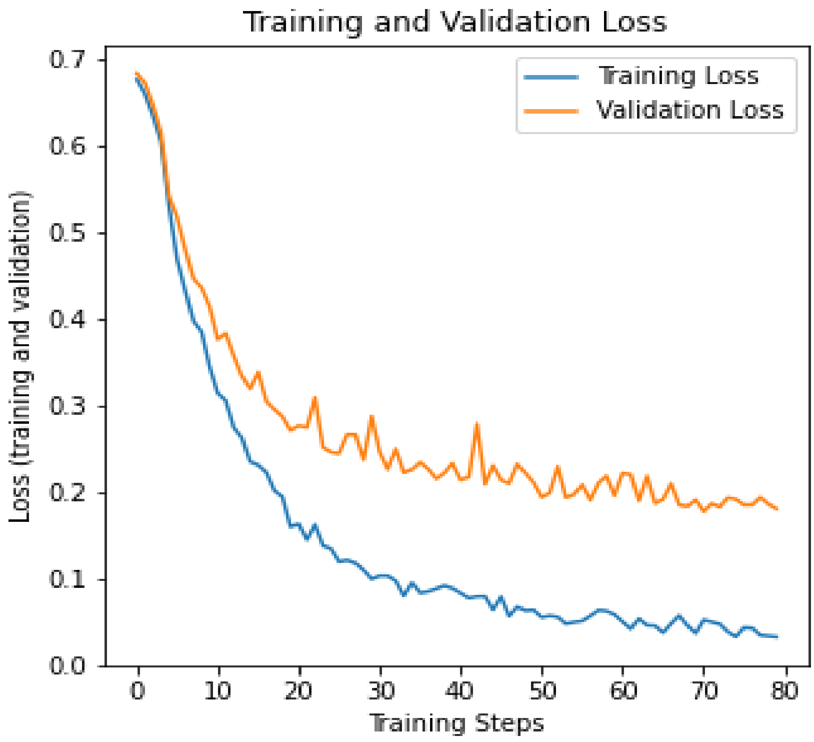

Figure 5.2: Training and Validation Loss

The initial results were analyzed thoroughly, the result could be improved if the dataset become noisy free, batch size could be changed, learning rate could be changed, number of epochs could be increased. After all these suggestions, the results were obtained are shown in Tables 5.1 and 5.2.

Table 5.1: Validation results of different simulations

Traning Data	Testing Data	Sampling Rate	Sample duration (sec)	Learning Rate	Batch Size	Epochs	Accuracy %	Precission %	F1 Score %	Specificity	Sensitivity
4729	277	8000	4	0.001	4	100	92.66	92.65	92.65	0.90666	0.9411
744	959	8000	4	0.0001	10	80	94.44	94.47	94.44	0.93	0.95
959	959	8000	4	0.01	15	80	96.11	96.47	96.14	1	0.93
744	959	8000	4	0.001	15	80	96.67	96.78	96.68	0.98	0.9
744	956	8000	4	0.001	8	80	90.31	91.54	90.22	0.98	0.81
744	959	8000	4	0.0001	8	100	94.9	94.92	94.9	0.95	0.93
744	959	8000	4	0.0001	8	150	93.37	93.41	93.36	0.94	0.91
744	959	8000	4	0.001	8	150	98.38	98.38	98.38	0.97	0.98
894	959	8000	4	0.0001	8	150	95.16	95.23	95.17	0.96	0.94
894	959	8000	4	0.0001	8	150	93.25	93.41	93.25	0.95	0.9
849	959	8000	4	0.0001	4	150	93.71	93.75	93.7	0.95	0.91
894	959	8000	4	0.001	4	150	89.76	91.79	89.83	0.82	1
4729	277	8000	4	0.0001	4	100	97.15	97.15	97.15	0.98	0.95

Table 5.2: Test results of different simulations

Traning Data	Testing Data	Sampling Rate (Hz)	Sample duration (sec)	Learning Rte	Batch Size	Epochs	Accuracy %	Precission %	F1 Score	Specificity	Sensitivity
4729	277	8000	4	0.001	4	100	94.03	94.14	94.05	0.95	0.93
744	959	8000	4	0.0001	10	80	95.12	95.13	95.12	0.93	0.96
959	959	8000	4	0.01	15	80	98.05	98.12	98.04	0.95	1
744	959	8000	4	0.001	15	80	95.61	95.64	95.6	0.93	0.97
744	956	8000	4	0.001	8	80	95.07	95.5	95.05	1	0.89
744	959	8000	4	0.0001	8	100	97.76	97.79	97.76	0.99	0.96
744	959	8000	4	0.0001	8	150	96.86	97.04	96.85	1	0.93
744	959	8000	4	0.001	8	150	98.97	98.99	98.97	1	0.98
894	959	8000	4	0.0001	8	150	97.25	97.33	97.25	0.99	0.95
894	959	8000	4	0.0001	8	150	94.85	94.86	94.84	0.95	0.93
849	959	8000	4	0.0001	4	150	95.88	96.17	95.86	1	0.91
894	959	8000	4	0.001	4	150	89.27	90.9	89.19	0.79	0.99
4729	277	8000	4	0.0001	4	100	98.56	98.57	98.56	0.98	0.99

The mean result of the all results is shown in Table 5.3 except the last one which is the final result of this proposed study as under:

Table 5.3: Mean of the first 12 results

Accuracy (%)	Precision (%)	F1 Score (%)	Specificity	Sensitivity
95.73	**96.22**	**95.72**	**0.96**	**0.95**

These results seem good but the random variance was quite high, after the results evaluation and analysis, pitch-shifting and re-filtering the dataset techniques were used.

5.4 Final result

Filtering the dataset, then applying pitch-shifting and increasing epoch to 150, the results were obtained are shown in Table 5.4. Figure 5.3 shows the accuracy and Figure 5.4 shows the loss respectively.

Table 5.4: Final test results

Accuracy (%)	Precision (%)	F1 Score (%)	Specificity	Sensitivity
98.56	**98.56**	**98.55**	**0.99**	**0.98**

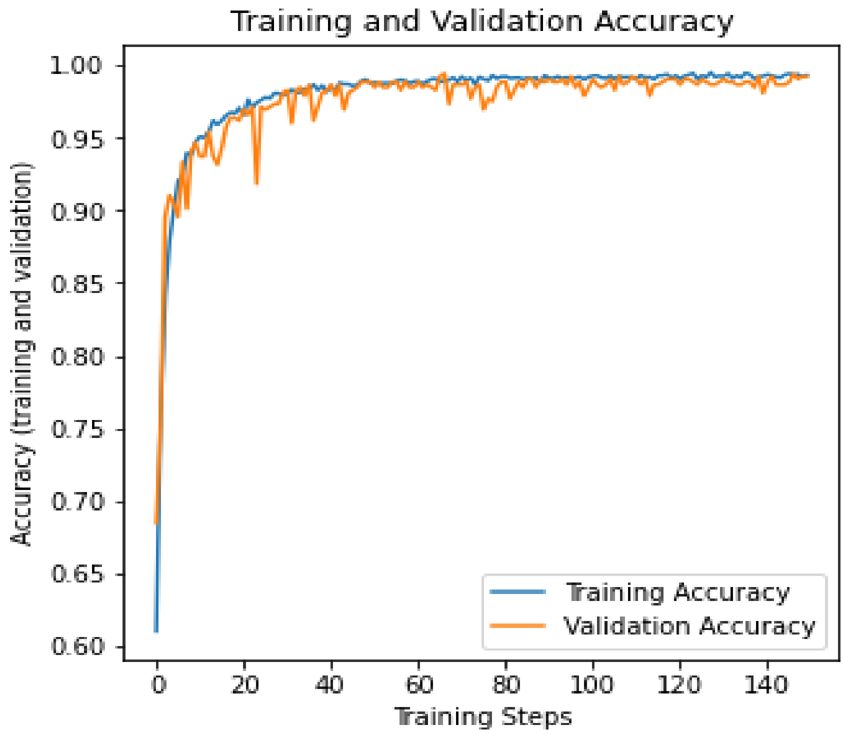

Figure 5.3: Accuracy (training and validation)

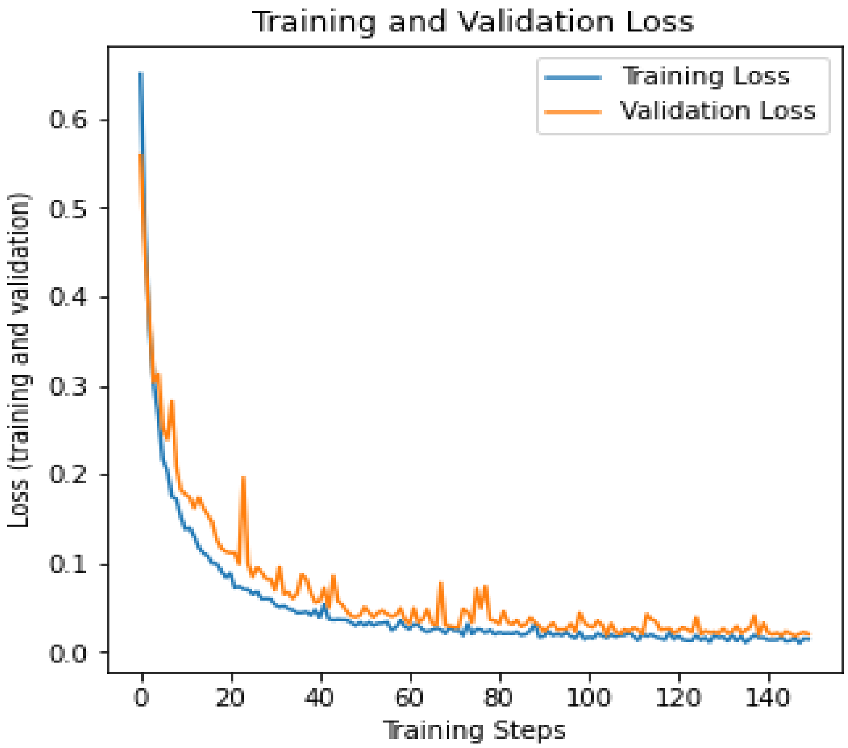

Figure 5.4: Loss (training and validation)

Now it sees that there is a good synchronization between the two curves however, there is still little random variance exists, but now it is acceptable. The most dominant reason of this variance is the noise in the local data. It can't completely diminish noise, it is a universal issue however, the noise can be reduced or control it to the desired limits. Most importantly, the point to be noted is that ConV 1D algorithm can tolerate somehow to this low frequency noise and it can give us quite good results.

5.5 Result Evaluation Parameters

5.5.1 Accuracy

It shows the number of correctly classified data samples over the total number of data samples.

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \tag{5.1}$$

5.5.2 Precision

It is the ratio of correctly positive classified samples to that of a total number of samples classified as positive either correctly or incorrectly. In other words, precision reflects how reliable the model is in classifying samples as Positive.

$$Precision = \frac{TP}{TP + FP} \tag{5.2}$$

5.5.3 Sensitivity or Recall

The Sensitivity measures the model's ability to detect positive samples. The higher the sensitivity, the more positive samples detected.

$$\boldsymbol{Recall = \frac{True_{positive}}{True_{positive} + False_{negative}}} \tag{5.3}$$

5.5.4 F1 Score

The F1 score is the harmonic mean of precision and sensitivity or recall.

$$F1\ score = 2 * \frac{Precision\ * Recall}{Precision + Recall} \tag{5.4}$$

Generally, in deep learning, machine learning, and AI most researchers prefer “Accuracy” and “F1-score” for model evaluation. According to them, the higher accuracy model performs well in most areas such as autonomous systems, business, and agriculture. However, AI in health care is different, it is not only demanding high accuracy, but most importantly it is demanding high sensitivity. In this proposed study, heart data and murmur were analyzed. A signal is said to be normal if it follows the pattern as shown in Figure 1.6, now if a slightly different variant comes, the model will capture it and will be considered abnormal.

5.5.5 Sensitivity in the proposed study case

Cardiac data is sensitive, a slight change in the normal cardiac signal could change the whole story. A good AI model must be capable to capture the minimum change in the variation of the cardiac signal. In addition to that, the AI model must be capable to differentiate between ambient low-frequency noise and murmurs. These all property of the model refers to **model sensitivity**. The higher the sensitive data, the higher sensitivity will be required.

CHAPTER 6 CONCLUSION and FUTURE WORK

6.1 Conclusion

This proposed concentrate on presents an original calculation for irregularity identification of the PCG signal technique for consequently extricating the fundamental clinically significant heart sound and mumble boundaries and characterizing the sign into two classes; a parallel characterization. In this examination study, viability and adaptability of the proposed technique utilizing a blend of the neighborhood and public datasets of ordinary and unusual PCG signals. The public dataset of Pediatric information that were used for this exploration is the Open Michigan dataset [9] and the continuous PCG signals recorded utilizing our computerized stethoscope. Further evaluation of the heartiness of the proposed technique utilizing this consolidated dataset with two distinct information augmentation strategies, piecing and cushioning, and pitch-moving, and contrast the presentation and the underlying outcomes. Assessment results show that the technique accomplishes exactness of 98.56%, accuracy of 98.56%, a F1-score of 98.55%, a responsiveness of 0.98, and particularity of 0.99.

6.2 Limitation

Although the system looks like it will hopefully enhance the traditional diagnostic system; however, it still requires the knowledge and experience of medical experts. It can be depicted from the analysis that, the advancement of computer computational power it will make it easy for scientists and engineers to completely automates the process.

6.3 Future work

In this proposed study, a local dataset was combined with the public dataset of different sampling rates and noise. Local data was quite noisy. However, after proper treatment and filtering, the model trained on that data shows extremely good results. Because of the unavailability of the Paediatric dataset, this study was designed only for binary classification of normal and abnormal detection. In the future, when there is a large amount of local data, this study can be extended to Paediatric multiclass classification.

REFERENCES

[1]. Congenital Heart Disease (CHD) | CS Mott Children's Hospital | Michigan Medicine

[2].Chelu, A., Williams, S.G., Keavney, B.D. *et al.*, "Joint analysis of functionally related genes yields further candidates associated with Tetralogy of Fallot", *J Hum Genet* (2022). https://doi.org/10.1038/s10038-022-01051-y

[3]. Dodge-Khatami A.,"Advances and research in congenital heart disease",TranslPediatr. 2016 Jul;5(3):109-111. doi: 10.21037/tp.2016.05.01. PMID: 27709090; PMCID: PMC5035758.

[4]. Gross RE., "SURGICAL MANAGEMENT OF THE PATENT DUCTUS ARTERIOSUS: WITH SUMMARY OF FOUR SURGICALLY TREATED CASES",. Ann Surg. 1939 Sep;110(3):321-56. doi: 10.1097/00000658-193909000-00001. PMID: 17857455; PMCID: PMC1391376.]

[5]. TAUSSIG HB, BLALOCK A.,"The tetralogy of Fallot; diagnosis and indications for operation; the surgical treatment of the tetralogy of Fallot.", Surgery. 1947 Jan;21(1):145. PMID: 20279481.

[6]. How the Healthy Heart Works | American Heart Association

[7]. Understanding the clinical value of a PCG or Phonocardiogram | Eko (ekohealth.com)

[8]. Frequency Responses of Conventional and Amplified Stethoscopes for Measuring Heart Sounds - PMC (nih.gov)

[9]. T. Ince, S. Kiranyaz, L. Eren, M. Askar and M. Gabbouj, "Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks", in IEEE Transactions on Industrial Electronics, vol. 63, no. 11, pp. 7067-7075, Nov. 2016, doi: 10.1109/TIE.2016.2582729.

[10]. Professional Skill Builder | Heart Sound & Murmur Library (umich.edu)

[11]. Yaseen& Son, Gui-Young &Kwon, Soonil. (2018)., "Classification of Heart Sound Signal Using Multiple Features.", Applied Sciences. 8. 2344. 10.3390/app8122344.

[12]. M. Ghaffari, M. Ashourian, E. A. İnce and H. Demirel, "Phonocardiography signal processing for automatic diagnosis of ventricular septal defect in newborns and children", 9th International Conference on Computational Intelligence and Communication Networks (CICN), 2017, pp. 62-66, doi: 10.1109/CICN.2017.8319357.

[13]. Chen, D.; Xuan, W.; Gu, Y.; Liu, F.; Chen, J.; Xia, S.; Jin, H.; Dong, S.; Luo, J. ,"Automatic Classification of Normal–Abnormal Heart Sounds Using Convolution Neural Network and Long-Short Term Memory", Electronics 2022, 11, 1246. https:// doi.org/10.3390/electronics11081246

[14]. S. Akbari, H. Ghassemian and Z. Akbari, "Systolic Murmurs Diagnosis Improvement by Feature Fusion and Decision Fusion", IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 2019, pp. 6-11, doi: 10.1109/ICSIPA45851.2019.8977784.

[15]. Roy, Tanmay & Roy, Dr. Joyanta& Mandal, Nirupama. (2021).,"Classifier Identification Using Deep Learning and Machine Learning Algorithms for the Detection of Valvular Heart Diseases", SSRN Electronic Journal. 10.2139/ssrn.3975139.

[16]. M.Nassralla, Z. E. Zein and H. Hajj, "Classification of normal and abnormal heart sounds", Fourth International Conference on Advances in Biomedical Engineering (ICABME), 2017, pp. 1-4, doi: 10.1109/ICABME.2017.8167538.

[17]. Boulares M, Alotaibi R, AlMansour A, Barnawi A., "Cardiovascular Disease Recognition Based on Heartbeat Segmentation and Selection Process", Int J Environ Res Public Health. 2021 Oct 18;18(20):10952. doi: 10.3390/ijerph182010952. PMID: 34682696; PMCID: PMC8535944.

[18]. Aziz S, Khan MU, Alhaisoni M, Akram T, Altaf M.,"Phonocardiogram Signal Processing for Automatic Diagnosis of Congenital Heart Disorders through Fusion of Temporal and Cepstral Features", Sensors (Basel). 2020 Jul 6;20(13):3790. doi: 10.3390/s20133790. PMID: 32640710; PMCID: PMC7374414.

[19]. Humayun AI, Ghaffarzadegan S, Ansari MI, Feng Z, Hasan T.,"Towards Domain Invariant Heart Sound Abnormality Detection Using Learnable Filterbanks",. IEEE J Biomed Health Inform. 2020 Aug;24(8):2189-2198. doi: 10.1109/JBHI.2020.2970252. Epub 2020 Jan 31. PMID: 32012032.

[20]. Khan FA, Abid A, Khan MS., "Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features",Physiol Meas. 2020 Jun 3;41(5):055006. doi: 10.1088/1361-6579/ab8770. PMID: 32259811.

[21]. Dastagir, Jehan, Faiq Ahmad Khan, Muhammad Salman Khan, and Kaleem Nawaz Khan., "Computer-aided phonocardiogram classification using multidomain time and frequency features", In 2021 International Conference on Artificial Intelligence (ICAI), pp. 50-55. IEEE, 2021.

[22]. Wang J, You T, Yi K, Gong Y, Xie Q, Qu F, Wang B, He Z., "Intelligent Diagnosis of Heart Murmurs in Children with Congenital Heart Disease",JHealthc Eng. 2020 May 9;2020:9640821. doi: 10.1155/2020/9640821. PMID: 32454963; PMCID: PMC7238385.

[23]. Xiao, Bin, Yunqiu Xu, XiuliBi, Weisheng Li, Zhuo Ma, Junhui Zhang, and Xu Ma., "Follow the sound of children's heart: a deep-learning-based computer-aided pediatric CHDs diagnosis system",IEEE Internet of Things Journal 7, no. 3 (2019): 1994-2004.

[24]. Zhang A, Wang J, Qu F, He Z.,"Classification of Children's Heart Sounds With Noise Reduction Based on Variational Modal Decomposition", Front Med Technol. 2022 May 26;4:854382. doi: 10.3389/fmedt.2022.854382. PMID: 35693881; PMCID: PMC9178247.

[25]. Mehmood, Awais, Munwar Iqbal, Zahid Mehmood, Aun Irtaza, Marriam Nawaz, Tahira Nazir, and Momina Masood., "Prediction of heart disease using deep convolutional neural networks", Arabian Journal for Science and Engineering 46, no. 4 (2021): 3409-3422.

[26]. Yunqiu Xu, Bin Xiao, Xiuli Bi, Weisheng Li, Junhui Zhang, and Xu Ma., "Pay more attention with fewer parameters: A novel 1-D convolutional neural network for heart sounds classification",In 2018 Computing in Cardiology Conference (CinC), vol. 45, pp. 1-4. IEEE, 2018.

[27]. https://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease

[28]. Rubin, Jonathan, Rui Abreu, Anurag Ganguli, Saigopal Nelaturi, Ion Matei, and Kumar Sricharan., "Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients", In 2016 Computing in cardiology conference (CinC), pp. 813-816. IEEE.

[29]. Ahmad, Talha & Ali, Hussnain& Khan, Shoab. (2009).,"Classification of Phonocardiogram Using an adaptive Fuzzy Inference System", 2. 609-614.

[30]. Redlarski G, Gradolewski D, Palkowski A.,"A system for heart sounds classification",PLoS One. 2014 Nov 13;9(11):e112673. doi: 10.1371/journal.pone.0112673. PMID: 25393113; PMCID: PMC4231067.

[31]. DeGroff, C., Bhatikar, S.R., Hertzberg, J.R., Shandas, R., Valdes-Cruz, L.M., & Mahajan, R.L. (2001).,"Artificial Neural Network-Based Method of Screening Heart Murmurs in Children", *Circulation: Journal of the American Heart Association, 103*, 2711-2716.

[32]. "Hilbert-Envelope Features for Cardiac Disease Classification from Noisy Phonocardiograms",NusratBinta Nizam, Shoyad Ibn Sabur Khan Nuhash, Taufiq HasanmedRxiv 2020.11.17.20233064; doi: https://doi.org/10.1101/2020.11.17.20233064

[33]. Pedrosa J, Castro A, Vinhoza TT.,"Automatic heart sound segmentation and murmur detection in pediatric phonocardiograms", Annu Int Conf IEEE Eng Med Biol Soc. 2014;2014:2294-7. doi: 10.1109/EMBC.2014.6944078. PMID: 25570446.

[34]. SinamAjitkumar Singh, Ningthoujam Dinita Devi, Swanirbhar Majumder, "An Improved Unsegmented Phonocardiogram Classification Using Nonlinear Time Scattering Features", *The Computer Journal*, 2022;, bxac025.

[35]. Kiranyaz, Serkan& Avci, Onur &Abdeljaber, Osama & Ince, Turker &Gabbouj, Moncef & Inman, Daniel. (2021).,"1D Convolutional Neural Networks and Applications: A Survey. Mechanical Systems and Signal Processing", 151. 10.1016/j.ymssp.2020.107398.

[36]. Kiranyaz, Serkan &Zabihi, Morteza &Bahrami Rad, Ali & Ince, Turker & Ridha, Hamila&Gabbouj, Moncef. (2020)., "Real-time Phonocardiogram Anomaly Detection by Adaptive 1D Convolutional Neural Networks", Neurocomputing. 411. 10.1016/j.neucom.2020.05

[37]. Heart Sounds - an overview | ScienceDirect Topics

[38]. Alanazi AA, Atcherson SR, Franklin CA, Bryan MF.,"Frequency Responses of Conventional and Amplified Stethoscopes for Measuring Heart Sounds", Saudi J Med Med Sci. 2020 May-Aug;8(2):112-117. doi: 10.4103/sjmms.sjmms_118_19. Epub 2020 Apr 17. PMID: 32587492; PMCID: PMC7305673.