Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

(Bioinstrumentation Engineering)

A Master's Thesis Report

On

Classification of Lung Diseases using Multiresolution Analysis of Lung Sounds for Improving the Efficacy of Auscultation Diagnosis

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

(Bioinstrumentation Engineering)

By: Biruk Abera Tessema

December 22, 2021

Jimma, Ethiopia.

Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

(Bioinstrumentation Engineering)

Classification of Lung Diseases using Multiresolution Analysis of Lung Sounds for Improving the Efficacy of Auscultation Diagnosis

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

(Bioinstrumentation Engineering)

By: Biruk Abera Tessema

Advisor: Gizeaddis Lamesgin (Ph.D.)

Co-Advisor: Hundessa Daba (M.Sc.)

December 22, 2021

Jimma, Ethiopia.

Declaration

I, the undersigned, declare that this thesis work entitled as **"Classification of Lung Diseases using Multiresolution Analysis of Lung Sounds for Improving the Efficacy of Auscultation Diagnosis."** is my original work and has not been presented by any other person for a degree in this and any other university. All the sources of materials used for a successful completion of this research have been dully acknowledged.

Candidate: Biruk Abera Tessema

Signature_____ Date_____.

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

Advisors

Name: Gizeaddis Lamesgin (Ph.D.)   Signature_____ Date_____.

Name: Hundessa Daba (M.Sc.)   Signature_____ Date_____.

External Examiner

Name: __Million Meshesha (PhD)__ Signature_~Million~_ Date__Jan 11, 2022__.

Internal Examiner

Name: _____Signature_____ Date_____.

Chair

Name: _____Signature_____ Date_____.

_____

Abstract

Stethoscope-based auscultation is the most efficient, non-invasive and inexpensive technique for assessing lung conditions on the base of lung sounds analysis. However, it provides a subjective perception of lung sounds. Additionally, the stethoscope recording is highly vulnerable to different noises which could mask the important features of lung sounds and may lead to misdiagnosis. Moreover, given the non-stationary nature of lung sounds, the exclusive time or spectral domain analysis is not effective for analysis. Hence, the time-frequency analysis of lung sounds is paramount. In this research, a method for efficient analysis of lung sounds used for further classification of lung diseases has been proposed. An electronic stethoscope has been constructed for signal acquisition. The lung sound signals used in the study were normalized at a uniform sampling frequency of 44.1KHz, and16 bit-depth. We employed wavelet multiresolution analysis technique to analyze the lung sound signals. A six-level DWT was performed for better analysis of lung sound signals by decomposing them into details and approximation. Wavelet denoising technique was used for the pre-processing task. Four wavelet functions (*Db4, Db10, Sym5, and Sym13*) with soft and hard thresholding methods and four different threshold selection rules were used to analyze the required performance of denoising the lung sound signals. *Sym13* Wavelet function with soft thresholding method outperforms all other wavelet functions in denoising lung sound signals. In the feature extraction task, a total of 16 features have been extracted by following a DWT-based feature extraction procedure. Moreover, one-way ANOVA was applied for the feature selection task to select the most relevant features. After feature selection, a total of 13 features were used for final classification of data. A comparison among the classification accuracies of the different machine learning classifiers was performed and Fine Gaussian SVM has been selected due to its higher classification accuracy. At last, we optimized the selected model using Bayesian optimization technique and an accuracy, specificity, and sensitivity of 99%, 99.2%, and 99.04%, respectively, has been achieved on the unseen data. The proposed method delivered a considerable improved result for classification of lung diseases as asthma, pneumonia, COPD, URTI, LRTI, bronchiectasis, bronchiolitis or healthy. Furthermore, as a future work, the proposed method can be further improved using deep learning techniques especially for larger lung sound data.

Key words: Auscultation, Denoising, Feature Extraction, Optimization, SVM, Wavelet MRA.

_____

Acknowledgment

First of all, I would like to thank St. Mary and her son God, lord of all the worlds, for his unlimited help. I always asking God to forgive and direct me to the right road. Secondly, I would like to acknowledge Haramaya University, College of Health and Medical Sciences for providing me the opportunity to study master program.

Next, my deepest appreciation goes to my advisor Dr. Gizeaddis Lamesgin and to my co-advisor Mr. Hundessa Daba for their continual and strict supervision during the entire work of my research. I would also extend my special thanks to Dr. Bheem for his valuable guidance and helping me in providing sensors required for my research. Additionally, I would like to say thank you for all my classmates for sharing ideas and suggestions in doing my research. Moreover, I would also like to thank my clinical collaborators such as Prof. Yoo, Dr. Assefa, Dr. Bekela, and Dr. Aneso, who are the pulmonologists working in Jimma university medical center (JUMC), their ideas and directions were of a great importance for the successful completion of my research.

At last, but not least, I would like to express my gratitude to my father Mr. Abera Tessema and my mother Mrs. Alemieneshe Getnet, who always pray for me and made everything possible to succeed in my life.

Table of Contents

_____

Jimma University, Jimma Institute of Technology, School of BME, Bio-Instrumentation Stream

List of Figures

List of Tables

List of Abbreviations

ANN………………. Artificial neural network

ANOVA………………. Analysis of variance

ARMA………………. Autoregressive moving average

CNN………………. Convolutional neural networks

COPD………………. Chronic obstructive pulmonary disease

CORSA………………. Computerized respiratory sound analysis

CWT………………. Continuous wavelet transforms

DCNN………………. Deep convolutional neural network

DWT………………. Discrete wavelet transforms

ELM………………. Extreme learning machine

EMD………………. Empirical mode decomposition

FFT………………. Fast Fourier transform

ICBHI………………. International conference on biomedical and health informatics

IDWT………………. Inverse discrete wavelet transforms

JUMC………………. Jimma university medical center

LRTI………………. Lower respiratory tract infection

MFCC………………. Mel frequency cepstral coefficient

MRA………………. Multiresolution analysis

PSD………………. Power spectral density

STFT………………. Short time Fourier transform

SVM………………. Support vector machine

URTI………………. Upper respiratory tract infection

_____

CHAPTER ONE

INTRODUCTON

1.1.   Background of the Study

Lung diseases include many disorders affecting the lungs. Respiratory problems caused by the lung disease may prevent the body from getting enough oxygen. There are many lung diseases affecting the lungs. These include COPD, asthma, pneumonia, URTI, LRTI, bronchiectasis, bronchiolitis, and others. COPD is a common chronic inflammatory lung disease [1,2,3]. Smoking is the main cause of COPD [1]. COPD causes breathing difficulty, cough, production of mucus (sputum), and wheezing [2,3]. These symptoms are also common in asthma disease [4]. Asthma is also one of the most common widespread lung diseases among children as well as adults and it associated with airway obstruction in the lungs [4,5,6].It causes recurrent episodes of wheezing, breathlessness, chest tightness, and coughing particularly at night or early in morning [5,6]. Pneumonia is an infection which can be caused by bacteria, viruses, and fungi and inflames the air sacs called alveoli in one or both lungs [7]. It causes coughing that may cause production of mucus, fever, chest pain, shortness of breath, and formation of adventitious lung sounds like rhonchi and crackles. URTI is a respiratory illness that occurs commonly both in children and adults and is a major cause of mild morbidity [8]. It is caused by several families of virus such as rhinovirus, coronavirus, parainfluenza, respiratory syncytial virus (RSV), adenovirus, human metapneumovirus, influenza, enterovirus, and recently discovered bocavirus [8]. URTI will affect upper respiratory tract including nose, sinuses, pharynx, or larynx [9,10]. Symptoms of LRTI vary and depend on the severity of the infection and it is the main cause of pediatric mortality and morbidity in low- and middle-income countries [10]. The less severe infections can have the same symptoms as bronchiolitis or bronchiectasis. Bronchiectasis is a long-term condition where the airways of the lungs become abnormally widened [11]. Bronchiectasis can make the lungs more vulnerable to infection by leading to a build-up of excess mucus. It causes a persistent cough which usually brings up phlegm (sputum), and breathlessness. Bronchiolitis is a common lung infection among infants, happens when small breathing tubes (bronchioles) become infected.

Generally, lung diseases cause a significant burden for public health systems. Currently, lung diseases are the third leading cause of death worldwide and cause an immense health, economic, and social burden [12]. Diagnosis of lung diseases, which are the third most common cause of death worldwide, is of great importance in the medical field. As a result, plenty of research efforts have been dedicated for improving early diagnosis and monitoring of patients with different respiratory diseases to allow for timely innervations [12]. Most recently, a number of researches has been focused in the auscultation and characteristics of lung sounds.

Auscultation is the act of listening sounds within the body via the help of stethoscopes. Stethoscopes are valuable tools used for clinical lung auscultation diagnosis for assessing patient's respiratory conditions using lung sounds [13]. Auscultation is an inexpensive and efficient technique for assessing respiratory conditions using lung sounds. Because, it provides direct information about the function of the lung when the pathological changes of the lung produce the characteristic sounds. Lung sounds are important indicators of respiratory health and disorders [14,15]. Moreover, the symptoms of all the above-discussed lung diseases are very common. This will cause a wrong diagnosis to be done by doctors. Hence, it is very useful to be able to identify or diagnose the disease using lung sounds without taking into account the rest of the symptoms.

Due to these advantages of auscultation diagnosis, improving diagnosis of pulmonary diseases using lung sound signals is becoming the most common research hotspots in the present time. In this study, an attempt has been done to improve the efficacy of auscultation diagnosis of pulmonary diseases. Therefore, multiresolution analysis of lung sounds has been done in this study for further classification of pulmonary diseases. A DWT based multiresolution decomposition was performed for better analysis of the non-stationary lung sound signals. Following signal acquisition, pre-processing, and extraction of features from the lung sound signals, important features were selected and given as input to different machine learning classifiers. Finally, after comparing classification accuracy of the different machine learning classifiers, SVM algorithm particularly fine gaussian SVM has been chosen for final classification of lung diseases as COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma or healthy.

1.2.   Motivation of the Study

At large lung diseases are among the leading cause of death worldwide. Diagnosis of lung diseases is often limited by the devices available to diagnose these disorders in low resource settings of the developing nations including our country Ethiopia. Devices such as chest x-ray (CXR), computed tomography (CT) scans, magnetic resonance imaging (MRI), spirometer, arterial blood gas analyzer, and auscultation with stethoscope can be used for diagnosis of lung diseases. The Imaging modalities (CXR, CT scans, and MRI) are highly accurate to diagnose the lung diseases [16]. However, their cost provides a high barrier for many patients of the third world countries with certain financial limitations. Such radiographic devices are not readily available in most health care settings of developing nations.

Moreover, it is difficult to use them for a day-to-day assessment of lung problems due to the risk of repeated dose of radiation. Therefore, the risk of repeated dose of harmful radiation, cost of machines, and inconvenient to deploy in many healthcare settings are some challenges to get such diagnosis approaches. Spirometer is another commonly used device to perform lung function tests by measuring the amount of air entering and leaving the lungs before and after the use of inhaled bronchodilator. However, spirometer requires skilled operator, it is costly, it is not efficient to detect a mix of obstructive-restrictive defect, and it needs multiple number of patient's breathing maneuvers [17]. Above all it needs forceful breathing which is difficult to apply for patients who are unable to breath forcefully.

Arterial blood gas analyzer (ABG) is also used to assess the lung problems by taking blood sample from the patient with the help of a syringe and a thin needle [18]. ABG test is too expensive and invasive leads to patient discomfort and prone to misdiagnosis. Due to the problems or challenges associated with the above-mentioned diagnostic approaches, stethoscope is the most commonly used valuable tool for assessing the lungs status through chest auscultation diagnosis [19]. Despite the rapid and continuous developments of technology in the area of chest disease diagnosis approaches, auscultation is still the most commonly used and indispensable diagnostic method [19]. It is a useful, simple, cost-effective (cheap) and fully non-invasive diagnostic device used to diagnose the lung diseases on the base of lung sounds.

Therefore, the goal of this research was to take these advantages of stethoscope and integrate it with computer in order to easily acquire and analyze the lung sounds for further classification of pulmonary diseases. Most importantly, it is highly useful to be able to classify the lung diseases using lung sounds without taking into account the rest of similar symptoms of the diseases. These were the concrete reasons, which motivate us to do our research on classification of most common lung diseases based on lung sound analysis for improving the current existing stethoscope-based auscultation diagnosis of lung diseases.

1.3.   Statement of the Problem

Lung diseases challenge millions of lives globally and annually. The economic impact of these respiratory problems is also very high. For example, in the UK, respiratory diseases cost the National Health service approximately 3 and 2 million GBP for asthma and chronic obstructive pulmonary disease (COPD) respectively [20]. Furthermore, lung diseases were among the top 10 causes of death worldwide (ranked first for low-income countries and ranked fifth for high-income countries) in 2016 [20]. Chronic respiratory diseases such as COPD and asthma affect over 15% of the world population. It is estimated that the prevalence of COPD is more than 250 million cases reported annually, resulting in over 3 million deaths globally [21]. It is also estimated that more than 235 million people suffer from asthma worldwide, with the disease causing in excess of 300,000 deaths per year [21]. Based on the information provided by United Nations Children's Fund (UNICEF), half of the morbidity from the recorded 5.9 million under-five deaths has caused by infectious diseases in which pneumonia lies at first rank in 2015 [22].

Acute respiratory infections are also the major problem in Ethiopia, accounting for about 10% of under-five deaths each year [23]. There are different clinical alternatives for diagnosing lung or respiratory diseases. Stethoscope-based auscultation is one of the simplest and a primary diagnostic approach used to evaluate lung diseases using lung sounds analysis. Stethoscopes are highly valuable tools used for clinical lung auscultation diagnosis. However, stethoscope-based auscultation provides limited and subjective perception of the respiratory sounds. The subjectivity results inconsistency during interpretation of lung sounds by the different medical experts. Such subjectivity and inconsistency are due to the physician's hearing capacity, internal listening variability as well as their ability or experience to differentiate and characterize different sound patterns.

So, stethoscope-based auscultation technique requires highly advanced expertise to correctly analyze the sounds and accurately identify the associated lung diseases. In addition, the stethoscope recording is highly vulnerable to different noises. The noise signals will mask the important features of lung sound signals and lead to wrong diagnosis of lung diseases.

Nowadays, the field of computer technology and the emerging of digital signal processing have made tremendous advancements in the early and rapid diagnosis of lung diseases from lung sounds. To overcome the limitations associated with stethoscope-based auscultation, improvements had been done following the development of digital signal processing techniques. Gogus, et al. [24], performed a comparative study for identification of pulmonary disorders using different spectral analysis methods such as fast Fourier transform (FFT), autoregressive (AR) and autoregressive moving average (ARMA). They estimated the power spectral density (PSD) of lung sounds using these spectral methods. Finally, they fed the features extracted from the PSD of lung sounds as inputs for the classifiers for further classification pulmonary diseases. Similarly, Islam, et al. [25], used Welch spectral method for estimation of PSD of lung sounds. They extracted different features from the PSD and fed to the classifiers to perform a three-class classification of lung diseases. Moreover, Kurt, et al. [26], did detection of lung abnormalities using Mel-frequency cepstral coefficient (MFCC) features of the lung sounds. Furthermore, Haider, et al. [27], also conducted respiratory sound-based classification of pulmonary diseases using median frequency and linear predictive coefficients of lung sound signals. It can be observed that the above-mentioned research works implemented spectral analysis techniques for analyzing the lung sound signals. However, Spectral analysis techniques couldn't provide temporal information of signals. They only provide spectral information of signals. Moreover, spectral analysis techniques provide better results in the analysis of stationary signals. However, lung sounds signals are non-stationary signals whose statistical characteristics changes with time. Hence, given the non-stationary property of lung sounds, the spectral analysis techniques are not efficient for analyzing such signals. Therefore, to get important diagnostic information (such as presence or absence of adventitious sounds, and lung diseases types) from the non-stationary lung sounds, the time-frequency analysis of lung sounds is highly needed. It is therefore that the aim of this study is to undertake a multiresolution analysis of lung sounds for improving the efficacy of auscultation diagnosis of pulmonary diseases.

1.4. Research Questions

- Does wavelet-based signal pre-processing method is suitable for denoising the non-stationary lung sound signals?
- How to extract and select relevant features of lung sound signals which can be further used for classification of pulmonary diseases?
- How to construct an electronic stethoscope by modifying the available acoustic stethoscope?
- Does time-frequency multiresolution analysis is the best suited method for analyzing the non-stationary lung sound signals?

1.5. Objectives of the Study

1.5.1. General objective

The main objective of this research is to construct a model for lung diseases classification based on multiresolution analysis of lung sounds towards improving the efficacy of auscultation diagnosis.

1.5.2. Specific objectives

This research strived to achieve the general purpose of the study by solving the following specific objectives, which are directed towards answering the research questions.

1) To construct an electronic stethoscope by modifying the existing traditional acoustic stethoscope.
2) To acquire lung sound signals from COPD, URTI, LRTI, bronchiectasis, bronchiolitis, pneumonia, asthma and healthy subjects.
3) To decompose and denoise the lung sounds using wavelet denoising technique and reconstruct the signals using IDWT.
4) To extract and select the most relevant features from the lung sound signals.
5) To train and compare different classical machine learning classifiers and select the best for classification of lung diseases by feeding the extracted and selected features as inputs.

1.6. Significance of the Study

Most recently, a number of researches has been focused in the auscultation and characteristics of lung sounds. This is due to the fact that the lung sounds are directly related to movement of air, changes within the lung tissues, and position of secretions within the tracheobronchial tree. Hence, lung sounds provide essential information about the health of the lungs. This is also due to the fact that the lung sounds can provide direct information about the function of the lung when the pathological changes of the lung produce the characteristic sounds. All these makes the lung sounds are valuable indicators of respiratory health and disorders. Acquisition and wavelet multiresolution analysis of lung sounds for accurate classification of most common lung diseases has been done in this study. This research can provide lots of advantages for the physicians in terms of storage, analysis and communication of sounds, graphical presentation of features of importance, correlation of lung sound with other physiological signals, comparison of lung sound obtained at different times during the progression of respiratory diseases or during the time of treatment. Most importantly, the study can be used for monitoring of lung sounds of patients who are in critical care units as well as for children who found difficulty to blow hard several times in the course of lung function tests. Moreover, this study will have a good potential to help researchers who are dedicated to study lung diseases classification on the base of lung sounds analysis.

1.7. Scope and Limitation of the Study

The study mainly covers construction of electronic stethoscope, signal acquisition, signal pre-processing, feature extraction and selection, feature normalization, data splitting, model training and evaluation, and model optimization. Head of traditional acoustic stethoscope, condenser microphone sensor, poly vinyl chloride tubing (medical grade tubing), audio cable and computer installed with audacity software were used for construction of our electronic stethoscope. We recorded the lung sounds using our electronic stethoscope under quite supervision of the clinical collaborators. The local lung sound files were collected from Jimma University Medical Center (JUMC), Department of Internal Medicine, Pulmonology Unit. Additionally, we used some annotated lung sound records available online. Moreover, a discrete wavelet transform (DWT)-based denoising of the lung sound signals which involves the three important steps, such as decomposition, detail coefficients thresholding, and reconstruction have been done.

Four wavelet functions (*Db4, Db10, Sym5,* and *Sym13*) and four different threshold selection rules (*Rigrsure, Sqtwolog, Heursure,* and *Minmax*) were used to analyze the required performance of denoising the lung sound signals. Moreover, a DWT-based feature extraction procedure involving two important steps (decomposition and reconstruction) was done to extract relevant features from the lung sound signals. After pertinent feature extractions, feature selection was done by implementing one-way ANOVA feature ranking method using $minmax$ normalization scheme to select the most discriminative features. Following feature extraction and selection, normalization of features has been done by subtracting the mean and dividing the standard deviation of each column ($minVar$ approach). The final step was data splitting to prepare the data for model training. In this research, first the dataset has been split using the hold-out method. Once we create a cross-validation partition for data, an 8-fold cross-validation has been applied to the training dataset. The classification performance of different classical machine learning algorithms was compared to select the one with the highest classification accuracy. Finally, we optimized the selected model using Bayesian optimization technique to improve its accuracy. MATLAB software (R2019b) and the latest version of audacity (audacity 2.4.2) software materials were used for the successful completion of this research.

The lung sound records used in this study were collected from only eight classes (COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma and healthy subjects). Since the novel coronavirus pandemic is one of the respiratory cases, the respiratory medicine department of JUMC was under lockdown for many times. Thus, only limited local lung sound records were used in this study. The other limitation of the study is that the lung sound signals were acquired using single channel data acquisition system. Multichannel data acquisition system (which requires high cost) will be helpful to gather more adequate information about the lungs pathology. Moreover, we only compared the classification performance of classical machine learning classifiers. Meaning, the study didn't use the state-of-the-art of deep learning techniques. Deep learning techniques involve use of large amount of data and complex algorithms which require powerful computation hardware (i.e., computer).

1.8. Methodology of the Study

1.8.1. Research Design

This study followed an experimental research design to achieve its objective. Experimental research is a study carried out through a scientific approach using two or more variables [28]. It includes a collection of research designs that utilize manipulation and controlled testing to understand the causal process. In the experimental research design, one or more variables are manipulated to determine their effect on a dependent variable. Algorithms and datasets used in the study are the independent variables. While the performance and the parameters are the dependent variables.

1.8.2. Data Collection and Preparation

The major tasks which were performed in this stage were collection of lung sounds and preparation of the data for further processing. Both online and local lung sound records were used to achieve the study. The online lung sound records were collected from the international conference on biomedical and health informatics (ICBHI) respiratory sound database. The local lung sound records were acquired from patients found in Jimma University Medical Center (JUMC), Department of Internal Medicine, Pulmonology Unit. Following data collection, a wavelet-based denoising technique which involves three basic steps (i.e., decomposition, detail coefficients thresholding, and reconstruction) was conducted to denoise the lung sound signals.

The denoising performances of the four wavelet functions (*Db4*, *Db10*, *Sym5*, and *Sym13*) with soft and hard thresholding methods along with the four threshold selection rules *Sqtwolog, Rigrsure, Heursure, and Minmax*) were investigated by calculating the SNR values. Additionally, the DWT feature extraction algorithm which involves series of steps has been implemented to extract important features from the lung sound signals. After pertinent feature extraction, one-way ANOVA was implemented for selection of the most relevant features.

1.8.3. Implementation Tools for Modeling

The study followed key steps to construct a model used for classification of lung diseases based on multiresolution analysis of lung sounds. These steps include lung sound data collection, data preparation, model training, optimization (parameter tunning), and performance evaluation. All the steps were performed using MATLAB software.

MATLAB is an ideal a multi-paradigm programming language and numeric computing environment designed for analyzing data, developing algorithms, creating models, and for different engineering, science, and economics tasks. It includes several new improvements, tweaks, and fixes with noticeable features which results in better workflow and improved working environment. Therefore, we utilized different MATLAB (R2019b) software toolboxes to construct the model. Signal processing toolbox and wavelet toolbox were used for pre-processing and feature extraction stages. While statistics and machine learning toolbox were used for constructing a model, we used the classification learner application to interactively train and compare the classification accuracies of different classifiers. Moreover, Bayesian optimization technique was implemented to automatically optimize the model.

1.8.4. Performance Evaluation

The performance evaluation matrices such as sensitivity, specificity, and accuracy were used to evaluate the final optimized model. Sensitivity is a performance metric which represents the correctly classified positive samples to the total number of positive samples [29]. It is used to evaluate a model's ability to predict true positive rates of each available category. Specificity is a performance metric which represents the correctly classified negative samples to the total number of negative samples [29]. It is used to evaluate a model's ability to predict true negatives. Moreover, accuracy is one of the most commonly used performance measures used to evaluate a model's ability to correctly differentiate observations from different category. It is defined as a ratio between the correctly classified samples to the total number of samples [29].

1.9.   Organization of the Thesis

To understand the work easily, this thesis has been divided into five chapters. **Chapter One** introduces the background of the study, and subsequently followed by motivation of the study, the base problem, objectives, significance of the study, scope and limitation of the study, methodology, and organization of the study. **Chapter Two** is literature review. This chapter is divided into two sections. The first section holds the conceptual literature review. The second section holds a review of related works. In the second section, other lung sound related research works conducted by various researchers are clearly reviewed. The signal processing techniques and the classification methods they used for the classification of lung diseases via processing of lung sounds have been discussed. **Chapter Three** is methods and materials section. In this chapter, the different methods used for achieving the different tasks of this research are clearly described. Moreover, the materials used for the successful completion of the study are also described. **Chapter Four** is the results and discussion section. It summarizes all the results found in the study. Moreover, the interpretation of the results is found in this chapter under the discussion section. At last, **Chapter Five** draws conclusion and recommendation. This chapter holds the overall conclusion of the study and leaves a recommendation to be addressed in future.

CHAPTER TWO

LITERATURE REVIEW

2.1. Lungs

Lungs are the primary organs of the human respiratory system which allow us to breath. Respiration or breathing is the main function of lungs. During respiration, oxygen from the incoming air enters in the blood and carbon dioxide (a waste gas) from the metabolism leaves the blood [30]. Figure 2.1 shows gas exchange between alveolus wall and capillary wall. The lungs which are the primary organs of the human respiratory system can be affected by different lung diseases. There are different clinical alternatives for diagnosis of pulmonary diseases. These includes chest x-ray, CT scans, MRI, spirometer, sputum culture, sputum cytology, rigid and flexible bronchoscopy, lung biopsy, and arterial blood gas analyzer can be used for diagnosis of lung diseases. Moreover, lung diseases can be identified on the basis of lung sounds through lung auscultation diagnosis. Stethoscope-based auscultation technique is the most commonly used primary diagnostic device used for assessing the lungs status. It is the simplest, cheapest, and non-invasive diagnostic device used to diagnose the lung diseases on the base of lung sounds. The goal of this research was to take these advantages of stethoscope and integrate it with computer in order to easily analyze the lung sounds and classify the most common lung diseases.
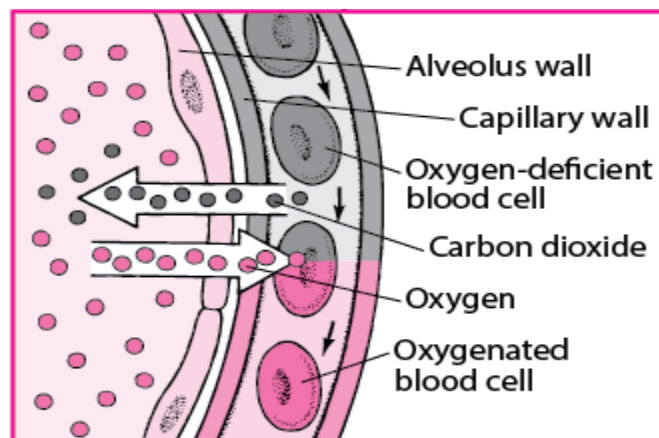
Figure 2.1: Gas exchange between alveolus wall and capillary wall [30].

2.1.1. Anatomy and Physiology of Lungs

Lungs are paired organs which are connected to the trachea through the right and the left bronchi. The lungs consist of right and left sides, each side has their lobes. The right lung has three lobes namely upper lobe, middle lobe and lower lobe while the left lung has two lobes namely upper lobe and lower lobe. Lobes are smaller units which composed each lung. The lungs are bordered by the diaphragm on the inferior surface. The diaphragm is the flat, large dome-shaped muscle that contracts and relaxes during breathing, also separates the chest and abdominal cavity. The heart sits in the mid chest extending into the left side. The cardiac notch is an indentation on the surface of the left lung and it allows space for the heart. Figure 2.2 shows gross anatomy of the lungs.
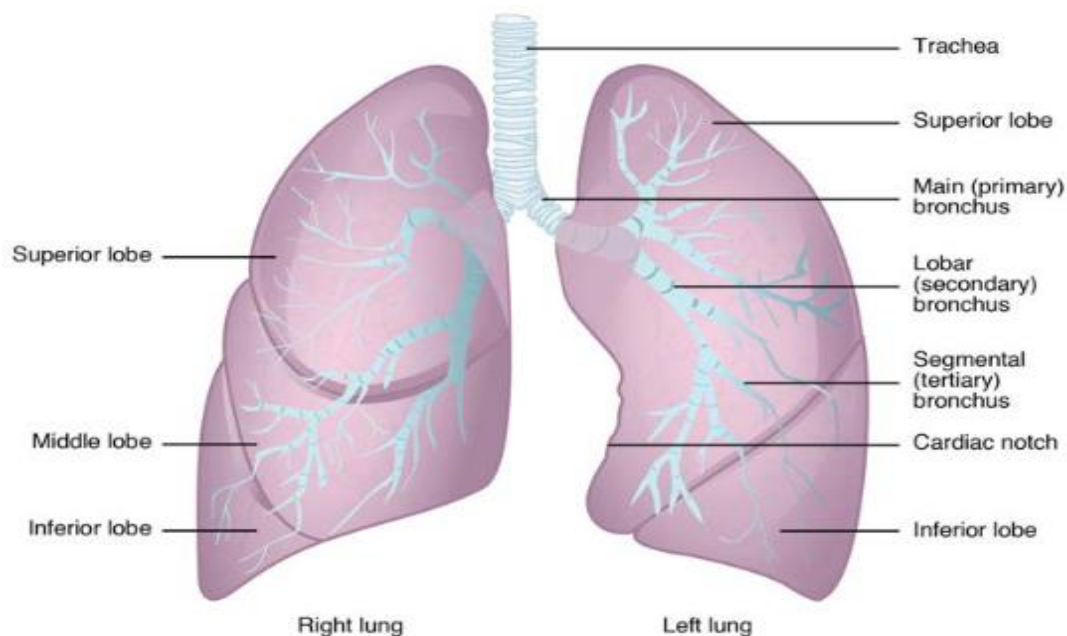
Figure 2.2: Gross Anatomy of the Lungs [31].

The functions of lungs include ventilation, diffusion, and transportation of gases such as oxygen and carbon dioxide [32]. Ventilation is the process of inhaling oxygenated air into the lungs and exhaling carbon dioxide. During a normal condition when the diaphragm contracts, it moves downward and increasing the diameter of the chest and elevating the lower ribs. This action decreases the amount of negative pressure within the alveoli, this change in pressure results in air being pulled into the lungs. When the lungs are now filled with air, the pressure within the lungs is greater than the pressure in the atmosphere, hence air is exhaled.

Ventilation includes the following four main functions of the respiratory systems such as supplying oxygen to the body, removing waste (carbon dioxide) from body tissues, maintain hemostasis (acid-base balance) of arterial blood, and maintaining heat exchange [32]. Any injury or disease process which impacts the lungs could impact one or all of these ventilation functions of the respiratory system. Diffusion is the exchange of gases from region of higher to region of lower concentration gradient. Diffusion occurs when the oxygen molecules move from the alveoli into the blood stream and the carbon dioxide molecules move from the blood stream into the alveoli. During inspiration, the oxygen in the lungs is at high concentration, moves across the respiratory membrane in the pulmonary arterioles where the oxygen concentration is rather low. At the same time, the high concentration of carbon dioxide moves out of the blood into the lungs. Transportation is the final step in respiration. It is the transport of oxygen and carbon dioxide to the cells. Oxygen can be transported in the blood in two ways such as by bounding to hemoglobin on the red cells and by dissolving in the plasma. Until the molecule is fully saturated, hemoglobin binds to dissolved oxygen in the blood. Finally, the amount of hemoglobin saturated with oxygen will be measured using pulse oximeter as SpO_2 and the remaining oxygen is dissolved in the plasma and can be measured by the arterial blood gas sampling as paO_2. The normal values of SpO_2 and paO_2 are about above 95% and 80 to 100 mm Hg (mercury) respectively [33].

One of the fundamental concepts used for understanding how oxygenation cooccurs is the oxyhemoglobin dissociation curve, shown using Figure 2.3. According to this curve, as the amount of oxygen in the atmosphere (external oxygen supplied to the patient) increases the paO_2 will increase but only up to a certain point. Generally, the oxyhemoglobin dissociation curve flattens out at about 60% oxygen and there will be a relatively little change in the oxygen saturation [34]. The predictability of the oxyhemoglobin dissociation curve relies on factors such as the patient being normothermic and the PH being within the normal range [34]. When one of these factors is impacted, the oxyhemoglobin dissociation curve will be impacted and an increase in oxygen will not have effect on the oxygenation of the blood.
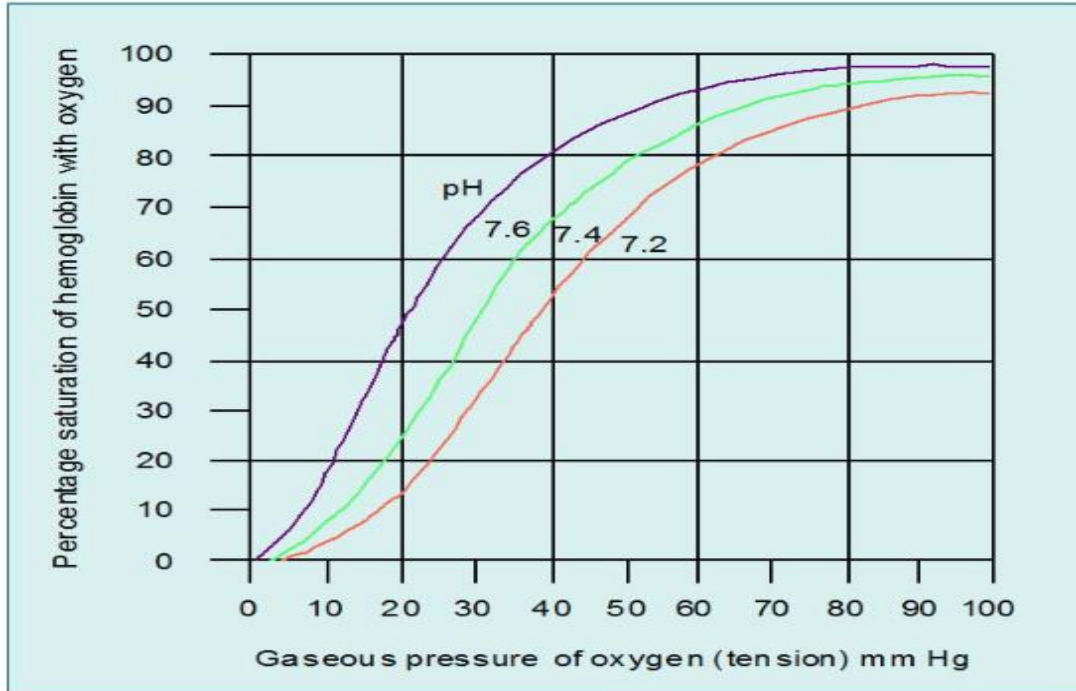
Figure 2.3: Oxyhemoglobin dissociation curve [34].

The second key part of transportation is the elimination of carbon dioxide (the product of cellular wastes). Carbon dioxide can be transported in the blood in three ways such as combined with the hemoglobin molecules, dissolved in the plasma, and can also be carried in the form of bicarbonate. The lungs which are the primary organs of the human respiratory system can be affected by different lung diseases and lung sounds are important indicators of respiratory health and disorders. Meaning, diagnosis of lung problems can be simply done on the basis of lung sound analysis. The following section deals about lung sounds, physics of lung sounds, mechanism of lung sounds production, and overview of signal processing techniques.

2.2. Lung Sounds

Lung sounds are also called respiratory sounds, which covers all the respiratory sounds heard or detected over the chest wall. The lung sounds detected over the chest wall can be mainly classified into two major categories such as normal and adventitious or abnormal respiratory sounds. Normal respiratory sounds heard over the chest wall are characterized by a low noise during inspiration and hardly audible during expiration [35]. Over the trachea, normal respiratory sounds are characterized by a border spectrum of noise (for example containing high frequency components) and are audible both during inspiratory and expiratory phase [35,36].

_____

Adventitious respiratory sounds are abnormal sounds which can be detected over the chest wall. These sounds can be continuous, discontinuous or some of them like squawks may have both characteristics [35]. Some of the most known adventitious or abnormal respiratory sounds include wheezes, crackles, squawks, stridors, and rhonchi. Wheezes are continuous adventitious sounds having a musical character. Wheezes are high-pitched sounds and can be heard during inspiration, expiration, or both. Acoustically, wheeze sounds are characterized by periodic waveforms having a dominant frequency usually more than 100 Hz with a time duration of greater than or equal to 100 ms [36]. Crackles are discontinuous, explosive and transient adventitious sounds in nature. Such sounds are more frequently occurs in cardiorespiratory disease and generally appear during inspiratory phase. Acoustically, crackles are characterized by their specific waveform with a wide frequency content and time duration less than 20 ms [36,37]. Squawks are short inspiratory wheezes, that occur primarily in restrictive lung diseases like interstitial lung disorders; are mixed sounds with short musical component (short wheeze) proceeded by crackles, often begin with a crackle. Acoustically, their waveform looks like short wheezes but different from those of the wheezes in obstructive lung disease. The time duration of squawks may vary between 50 and 400ms [37,38]. Stridors are very low wheezes; most frequently heard during inspiratory phase and can appear during in whooping cough, laryngeal or tracheal stenosis [38,39]. Such sounds may be audible at a distance of mouth, trachea, and over the chest wall without the help of stethoscope. Stridors are characterized by a prominent peak at about 1000 Hz [38]. Rhonchi are a low- pitched, rattling and bubbling wheeze sounds. Acoustically this sounds are characterized by periodic waveforms having a time duration above 100 ms and a frequency range of below 300 Hz [39]. Such sounds may occur during inspiratory, expiratory or in both phases and can be found in patients having chronic bronchitis.

2.2.1. Physics of lung sounds

Breath sounds have two major characters such as frequency and amplitude (intensity) which enables us to differentiate sounds [40].

I.   Frequency and Pitch

Frequency is the measure of the number of sound waves or vibrations per a given amount of time. It is an objective measurement measured in hertz (Hz). It depends on the wavelengths.

When the wavelengths are shorter, there are a greater number of sound waves per second, and the frequencies will be higher and when there is a longer wavelength, the frequencies will be lower. Wavelength commonly designated by the Greek letter lambda (λ), is the distance from the peak of one sound or pressure wave to the peak of the next sound or pressure wave. It depends on the speed of the sound waves, the medium where the sound waves are traversing and the temperature of the medium. Pitch is a perceptual property of sounds; it is a subjective perception of the sound's frequency. Pitch depends on the frequency. The human ear can perceive sound waves in the range of frequencies ranging from 20 to 20K Hz.

II.   Amplitude or Loudness

Amplitude is related to the energy of the sound waves while loudness is the subjective perception of amplitude. Amplitude is measured from the height or the peak of sound waves to the center line position. The range of amplitude of sound waves is extremely wide, hence it is measured on a logarithmic scale and is expressed by decibels (dB). A decibel (dB) is defined as one tenth of a bel. The bel is an amplitude unit defined for sound as the log (base 10) of the intensity relative to some reference intensity as shown in Equation (2.1) [40].

$$\text{Amplitude}_{\text{in}_{\text{bels}}} = \log 1o \left(\frac{\text{Signal}_{\text{Intensity}}}{\text{Reference}_{\text{Intensity}}} \right) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.1)$$

The choice of reference intensity or power determines the particular choice of dB scale. Signal intensity, power and energy are always proportional to the square of the signal amplitude, shown in Equation (2.2) [40].

$$\text{Amplitude}_{\text{in}_{\text{bels}}} = \log 10 \left(\frac{\text{Amplitude}^2}{\text{Amplitude}^2 \text{ ref}} \right) = 2\log 10 \left(\frac{|\text{Amplitude}|}{|\text{Amplituderef}|} \right) \dots \dots \dots \dots .. (2.2)$$

Since decibel (dB) is one tenth of a bel, there are 10 decibels to a bel. Hence the amplitude is given in decibel (dB) using Equation (2.3) [40]:

$$\text{Amplitiude}_{\text{in}_{dB}} = 20\log10\left(\frac{|\text{Amplitude}|}{|\text{Amplituderef}|}\right) = 10\log10\left(\frac{\text{Intensity}}{\text{Intensityref}}\right)$$

$$= 10\log10\left(\frac{\text{Power}}{\text{Powerref}}\right)$$

$$= 10\log\left(\frac{\text{Energy}}{\text{Energyref}}\right)\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(2.3)$$

2.2.2. Mechanism of Lung Sounds production

Air flow along the trachea-bronchial tree is the prerequisite for the production of normal breath sounds [41]. However, not all air flow types can produce breath sounds. Only turbulent and vortices air flow types are responsible for the production of breath sounds [41]. Laminar flow is silent flow and occurs in low flow situation. It is the characteristic of small peripheral airways. For example, it is so slow that airflow in the alveoli comes to an end. In laminar air flow, the streams of air flow are parallel to the walls and are parabolic in shape as shown in Figure 2.4.
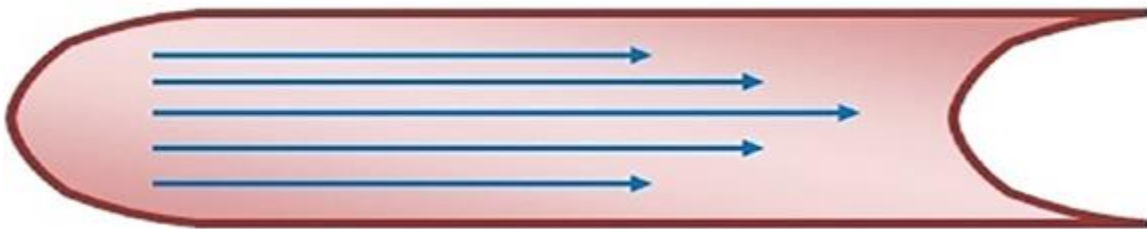
Figure 2.4: Laminar flow [42].

In laminar flow, air in the central layers moves faster than in the peripheral layers. The pattern of laminar flow follows the Poiseuille equation given in Equation (1.4) [42]. Laminar flow is directly proportional to the driving pressure (P), shown in Equation (2.4).

$$Q = \frac{\Delta P \pi r^4}{8\eta l}\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(2.4)$$

Where; Q is volume flow rate, P is the driving pressure, r is the radius, η is viscosity, and l is length.

Turbulent flow is the disorganized gas or liquid flow in which the fluid undergoes irregular fluctuations or mixing [42,43]. It is chaotic in nature and depends on the density of air more than viscosity. It is very rapid, complex and typical of large central airways such as trachea and major bronchi.

In turbulent air flow, the air molecules randomly collide against each other and onto the airway walls, hence this air movement is characteristically noisy. Turbulent air flow will occur when high velocity of flow passes through a large diameter air way such as through irregular airway walls [43]. As shown in Figure 2.5, for example, turbulent flows through bronchi and trachea or in the airway with sudden branching. Turbulent flow produces noise as the air molecules collide each other and with the air way wall.

Figure 2.5: Turbulent flow [42].

The development of vortices is another mechanism for the production of breath sounds [43]. Vortices also called whirlpools are developed when a stream of gas emerges from a circular orifice to a wider channel as shown in Figure 2.6. Vortices usually occurs between the fifth and thirteenth generations of bronchial tree. The vortices air flow is also called mixed or transitional flow, because it resembles both laminar and turbulent airflow, shown in Figure 2.6. It is a bit faster and typical of medium-sized branching airways. Branching separates airflows into different layers with different velocities. The interaction among the different layers generates eddies and makes vortices all noisy.
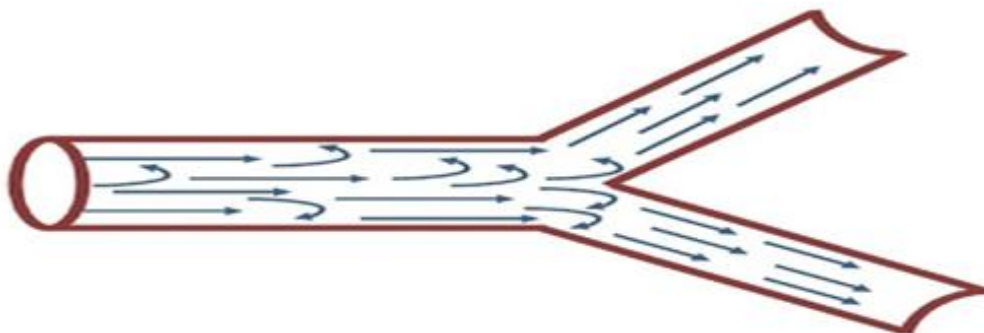
Figure 2.6: Transitional flow [43].

_____

Voice sounds are produced by the larynx. The lung sounds are different from the transmitted voice sounds generated by the larynx. Lung sounds also called respiratory sounds consists of normal breath sounds and adventitious or abnormal sounds heard or detected over the chest wall. The mechanism of production of normal breath sounds has been discussed above. The production or generation of abnormal or adventitious sounds associated with inflammation or infection of the small bronchi, bronchioles, and/or alveoli. The generation of abnormal or adventitious lung sounds are indication of the occurrence of lung diseases. Table 2.1 below shows adventitious respiratory sounds and their characteristics with the associated possible lung diseases.

2.3. Signal Processing Techniques

Signal analysis or processing is an area of electrical engineering, systems engineering, and applied mathematics which deals with analyzing, modifying, and synthesizing signals such as sound, images, and scientific measurements. There are various analysis techniques which can be used for analyzing signals using computers. These analysis techniques can be broadly classified into time-domain, frequency-domain and time-frequency domain analysis.

Table 2.1: Adventitious respiratory sounds and possible lung diseases [44].

Adventitious sound types	Characteristics	Possible lung diseases
Wheezes	Continuous high-pitched adventitious sounds. Characterized by periodic waveforms having a dominant frequency usually more than 100 Hz with a time duration of greater than or equal to 100 ms.	Asthma and cystic fibrosis.
Crackles	Discontinuous adventitious sounds. Characterized by their specific waveform with a wide frequency content and time duration less than 20 ms.	Pulmonary fibrosis, pneumonia, alveolitis, asbestosis, congestive heart failure, chronic bronchitis, bronchiectasis.
Stridors	Continuous very low wheeze sounds.	Tracheal stenosis, vocal cord

_____

	Characterized by a prominent peak at about 1000 Hz.	paralysis, laryngomalacia, laryngitis, tumors, airway inflammation following extubating.
Rhonchi	Continuous low- pitched, rattling and bubbling wheeze sounds. Characterized by periodic waveforms having a time duration above 100 ms and a frequency range of below 300 Hz.	Pneumonia, chronic obstructive pulmonary disease (COPD), tumors, and chronic bronchitis.
Squawks	The time duration of squawks may vary between 50 and 400ms.	Pulmonary fibrosis, interstitial fibrosis and allergic alveolitis.

2.3.1. Time-domain Analysis

Time domain analysis is the analysis of electronic signals, physical signals, biological signals like lung sound signals, mathematical functions, or time series of economic or environmental data in reference to time [45]. It is the process of analyzing data over a time period, the variable is always measured against time. For example, for electronic signals, the time domain analysis is usually based on the current-time plot or the voltage-time plot. Furthermore, a time domain graph shows how a signal changes with time while the frequency domain graph show how much of the signal lies within each given frequency band over a range of frequencies.

Time Domain

Input $X(t)$ → Impulse Response $h(t)$ → Output $Y(t) = h(t) * X(t)$
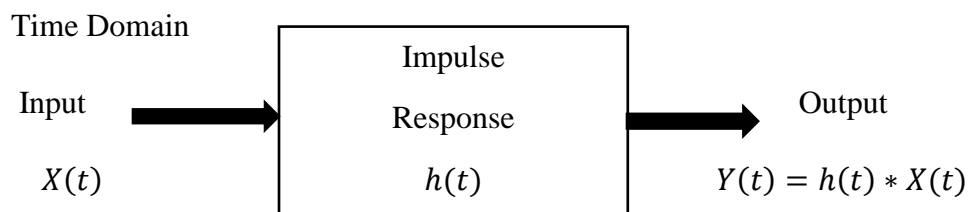
Figure 2.7: Time-domain approach [46].

When we analyze an audio signal using time domain approach, the X-axis is time. So, the value of the Y-axis depends on the change in signal with respect to time. The time-domain waveform representations of the lung sound signals of the COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma, and healthy subjects are shown using Figures 2.8, 2.9, 2.10, and 2.11.
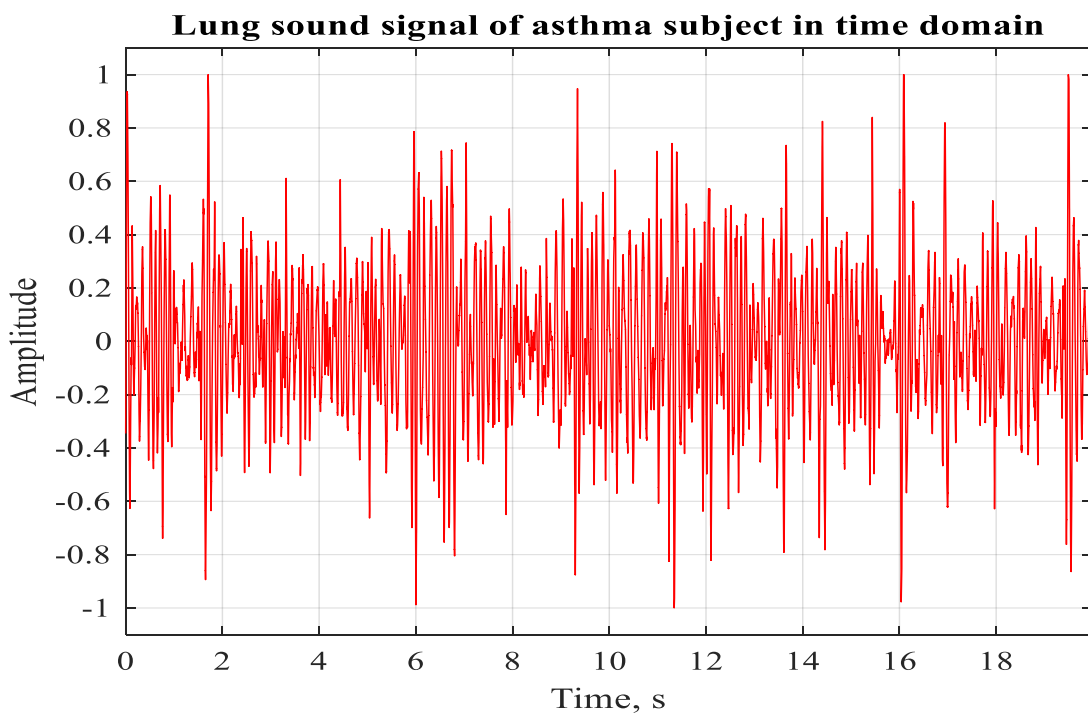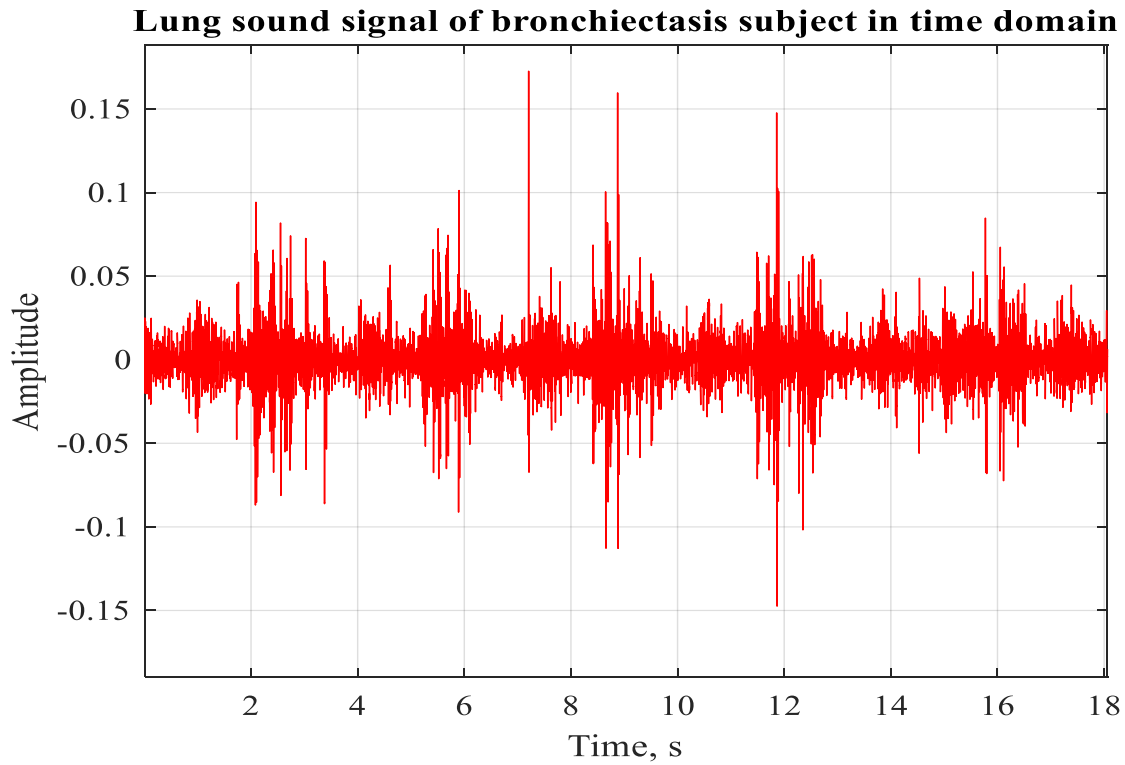
Figure 2.8: Time-domain waveform representation of lung sound signals; (top), bronchiectasis (bottom) asthma subjects.
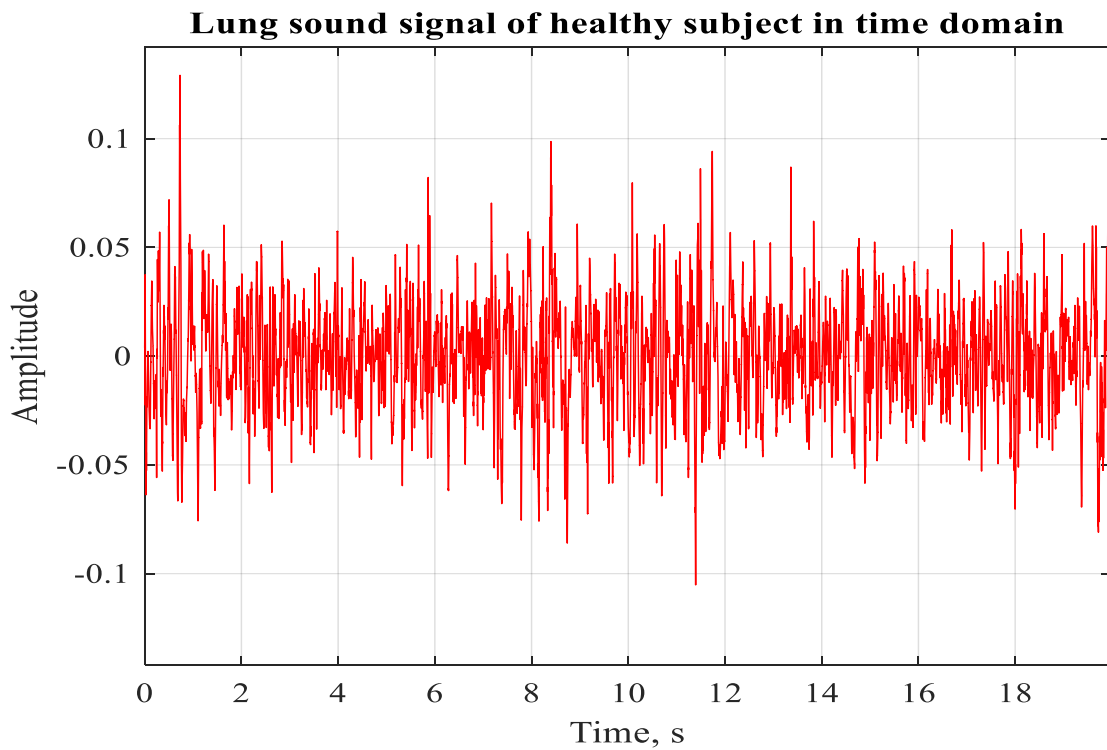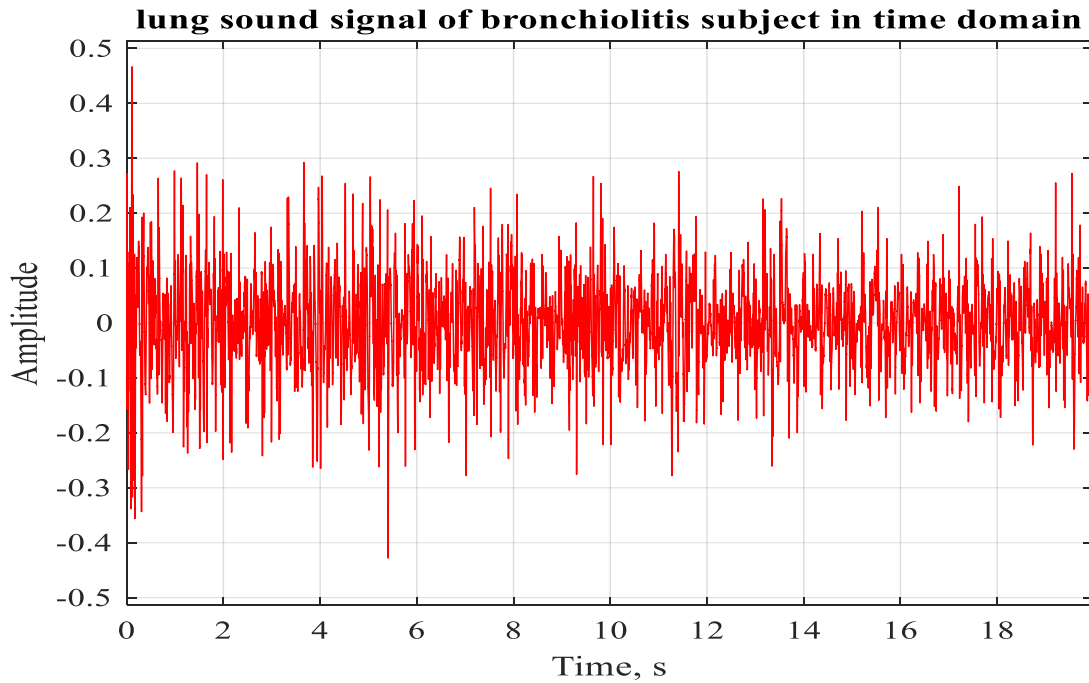
Figure 2.9: Time-domain waveform representation of lung sound signals; (top), bronchiolitis (bottom) healthy subjects.
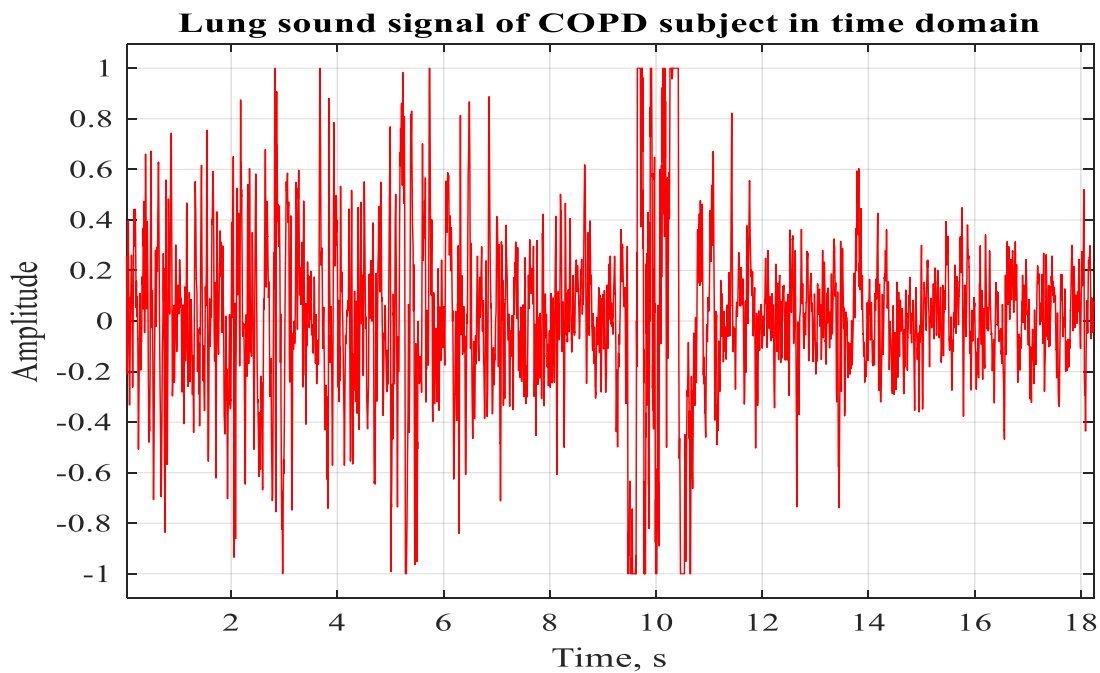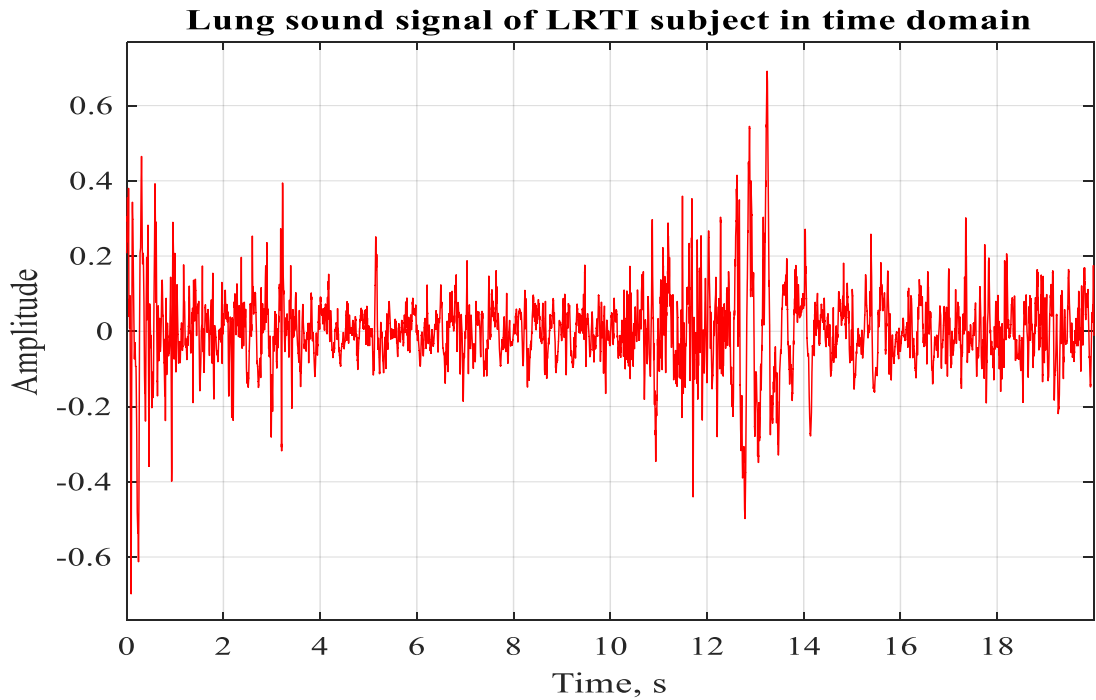
Figure 2.10: Time-domain waveform representation of lung sound signals; (top), LRTI (bottom) COPD subjects.

Lung sound signal of URTI subject in time domain

Lung sound signal of Pneumonia subject in time domain
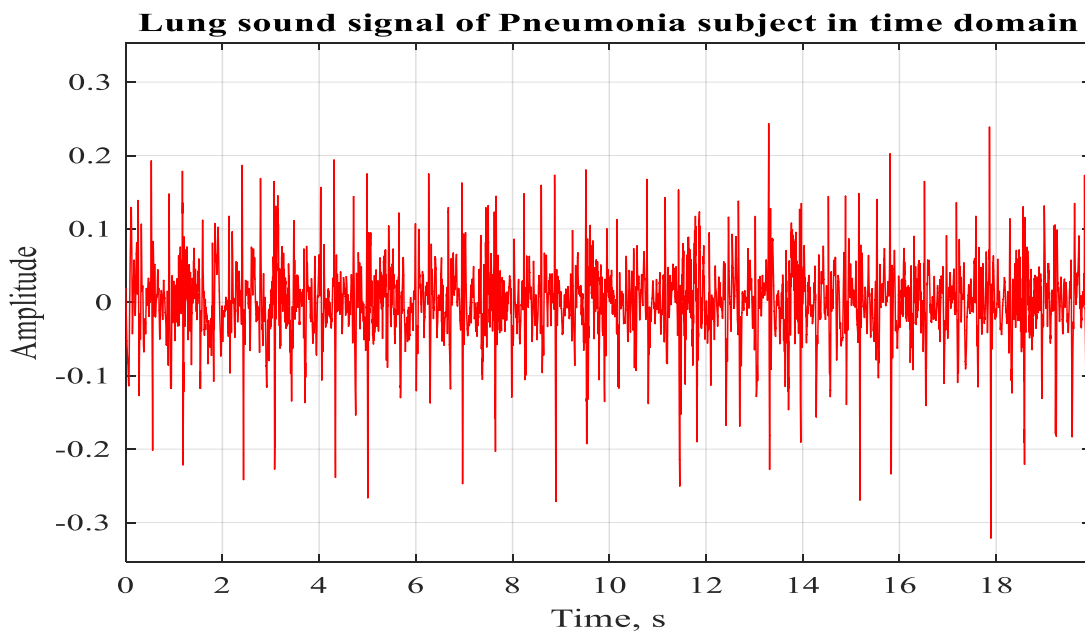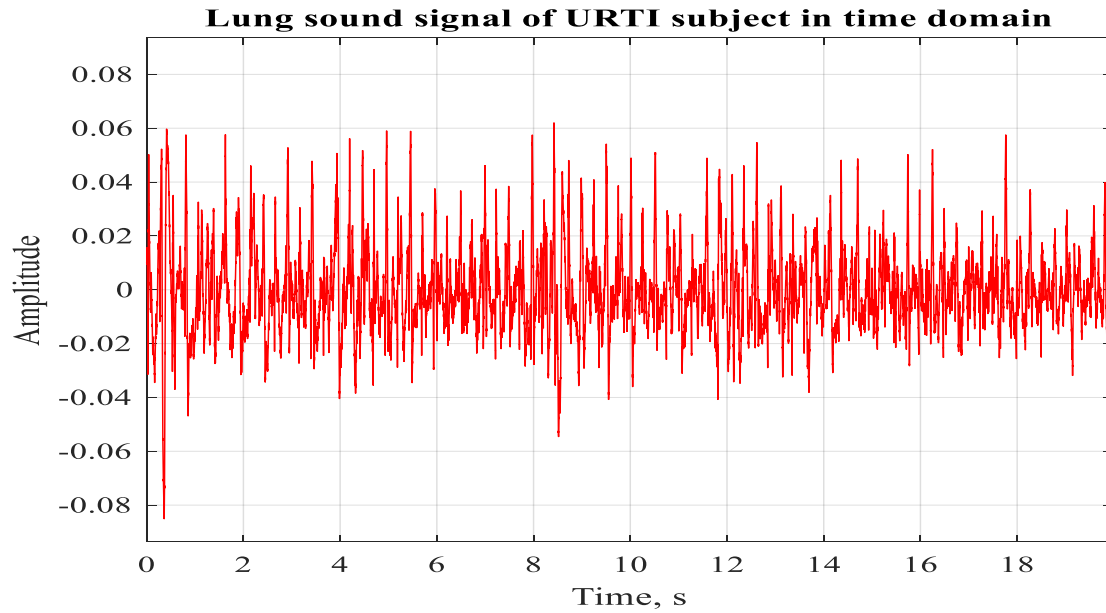
Figure 2.11: Time-domain waveform representation of lung sound signals; (top), URTI (bottom) pneumonia subjects.

The time domain graph presented using Figures 2.8 - 2.11 showed us the changes in a lung sound signal over a range of time while the frequency domain graph displays how much of the signals exist within a given frequency band concerning a span of frequencies.

2.3.2. Frequency-domain Analysis

Both time domain analysis and frequency domain analysis are widely used in different fields such as acoustics (sound processing), telecommunications, electronics and many other areas [45]. However, time-domain analysis provides only temporal information of signals. Time-domain analysis didn't provide frequency or spectral information of signals whereas frequency-domain analysis can provide it. Frequency domain analysis is the analysis of mathematical functions, physical signals, biological or electrical signals in reference to frequency. It is a method used to analyze data over range of frequencies. It is widely used in different fields such as electronics, statistics, and control systems engineering [45]. Frequency domain analysis is most widely used to analyze signals or functions which are periodic over time, but this doesn't mean that frequency domain analysis cannot be used to analyze signals or functions which are not periodic. Transformation is the most important concept used in the frequency domain analysis. It is used to convert a time domain function to a frequency domain function and vice versa. Fourier transformation is the most common transformation used in the frequency domain analysis. It is used to convert a signal of any shape into a sum of infinite number of sinusoidal waves [46]. Analyzing sinusoidal functions is much easier than analyzing a general shaped function.

Frequency Domain (Fourier)
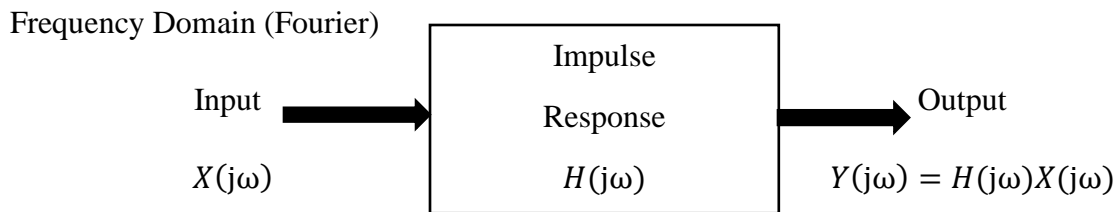
Figure 2.12: Frequency-domain approach [46].

The Fourier analysis representation of any energy function x(t) of a finite duration is given by the sum of sinusoids $e^{j\omega}$, shown in Equation (2.5) [47].

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(\omega)\, e^{j\omega t} d\omega \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.(2.5)$$

The amplitude x(ω) for each of the sinusoidal wave function $e^{j\omega t}$ is equal to correlation of $e^{j\omega t}$ with x, which is also called as Fourier transform, given in Equation (2.6) [47].

$$x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt \; \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. (2.6)$$

The more regular the function, $x(t)$, the faster will be the decay of the amplitude $|x(\omega)|$, when the frequency ω is increased. The above pair of equation (Equation (2.5) and Equation (2.6)) is called a Fourier transform pair. FFT is the most commonly used frequency-domain analysis technique used for sound processing. Let k be currently sampling point and n for all N sampling points within a period, then FFT converts time domain signals into the frequency domain to facilitate energy distribution in the frequency spectrum, shown in Equation (2.7) [47].

$$X[k] = \sum_{n=0}^{N-1} x^{\sim}[n] e^{-\frac{j2\pi nk}{N}} \; \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2.7)$$

The frequency-domain waveform representations of the lung sound signals of the COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma, and healthy subjects have been found using fast Fourier transform (FFT), shown using Figures 2.13, 2.14, and 2.15.

Lung sound signal of bronchiectasis subject in frequency domain
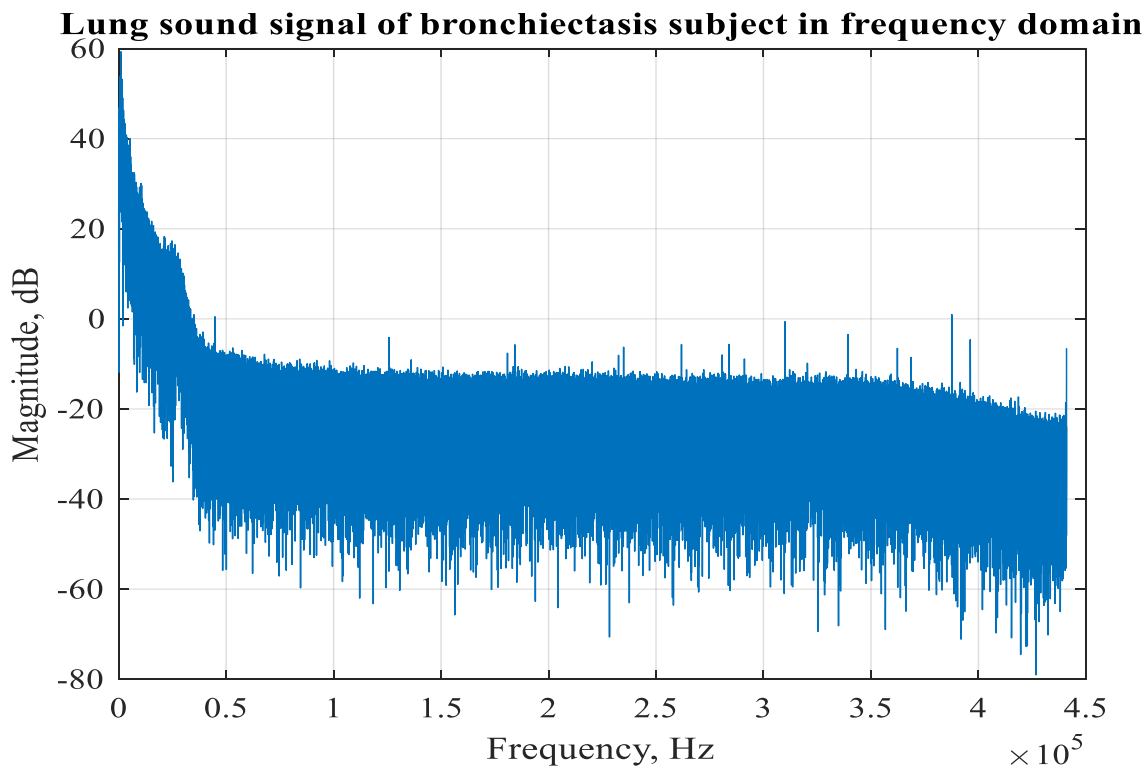
(a)
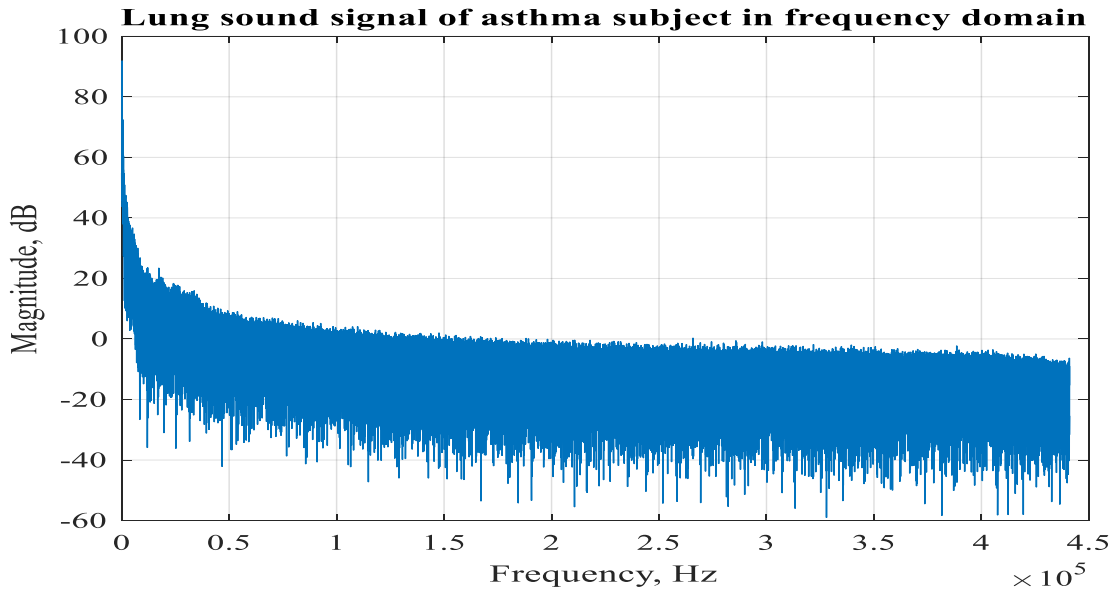
(b)
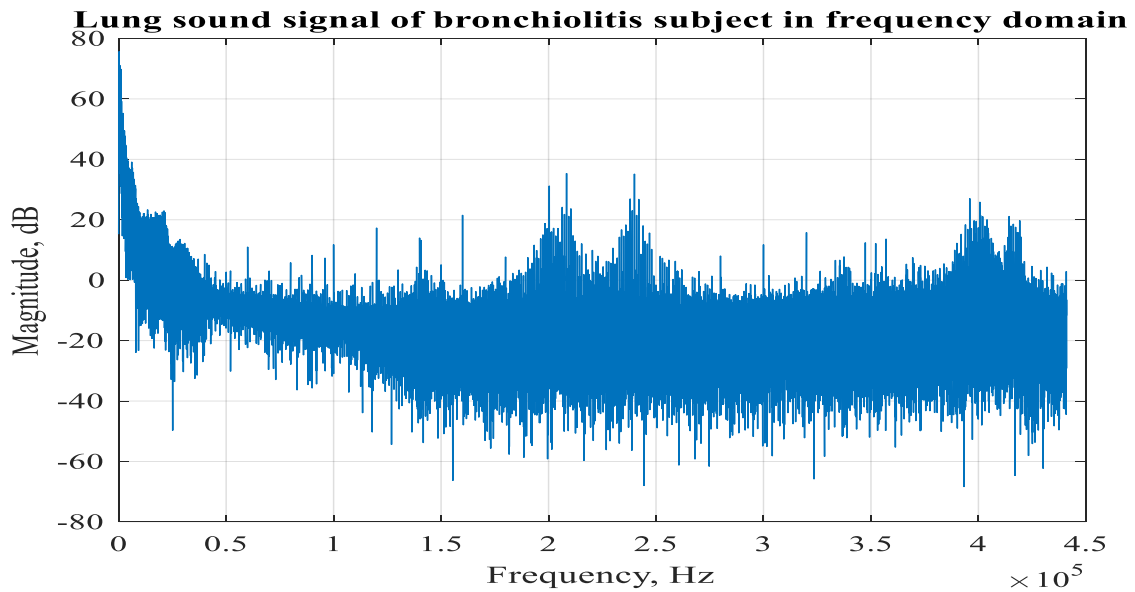
(c)

Figure 2.13: Frequency-domain waveform representation of lung sound signals; (a), bronchiectasis (b) asthma and (c) bronchiolitis subjects.

Lung sound signal of healthy subject in frequency domain

(d)

Lung sound signal of LRTI subject in frequency domain
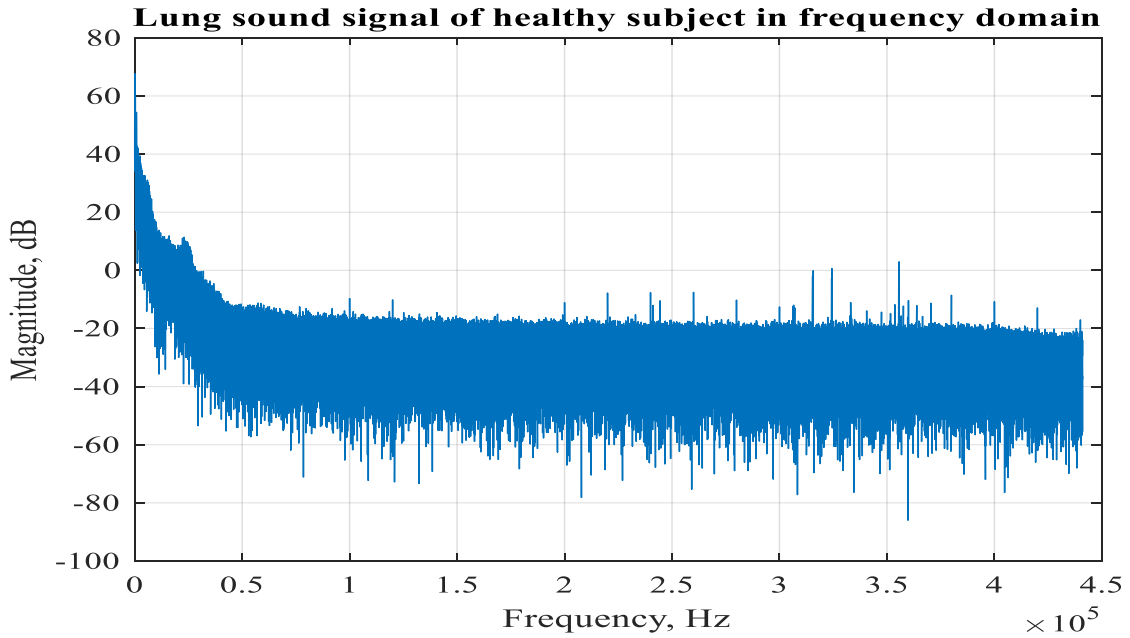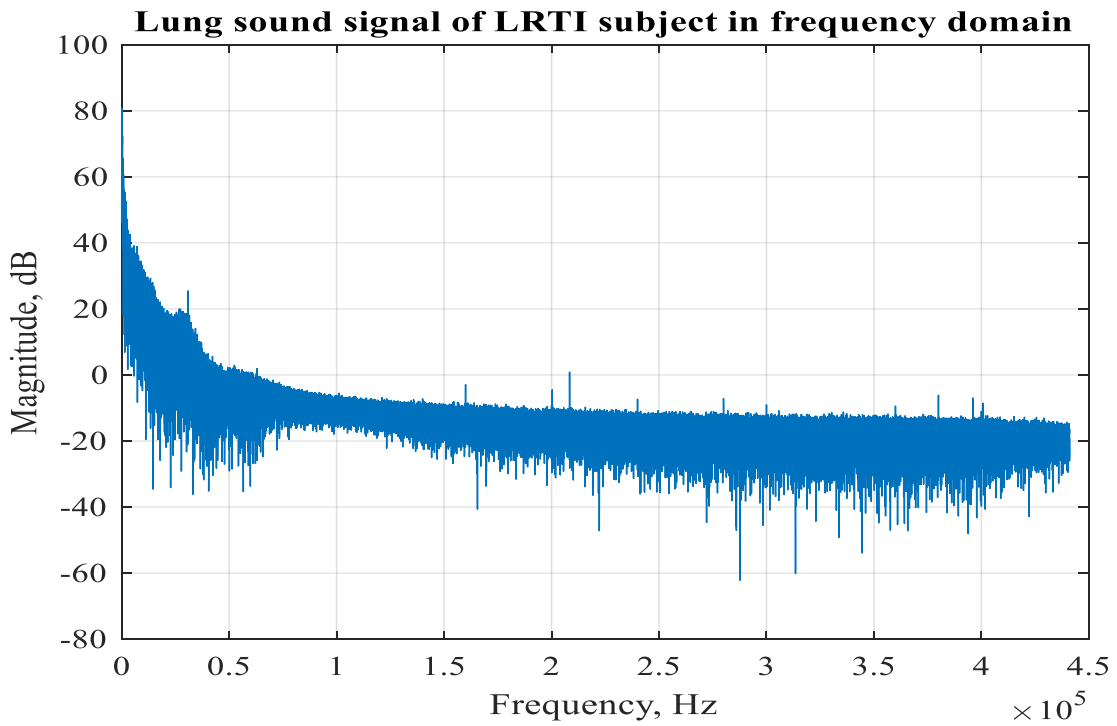
(e)
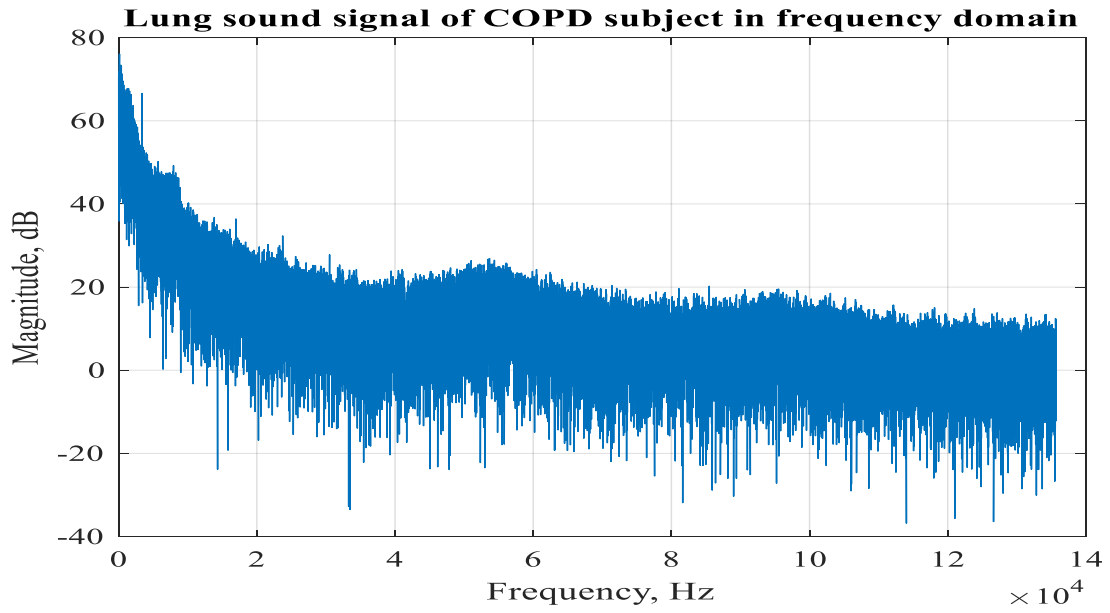
(f)

Figure 2.14: Frequency-domain waveform representation of lung sound signals; (e), healthy (d) LRTI and (f) COPD subjects.
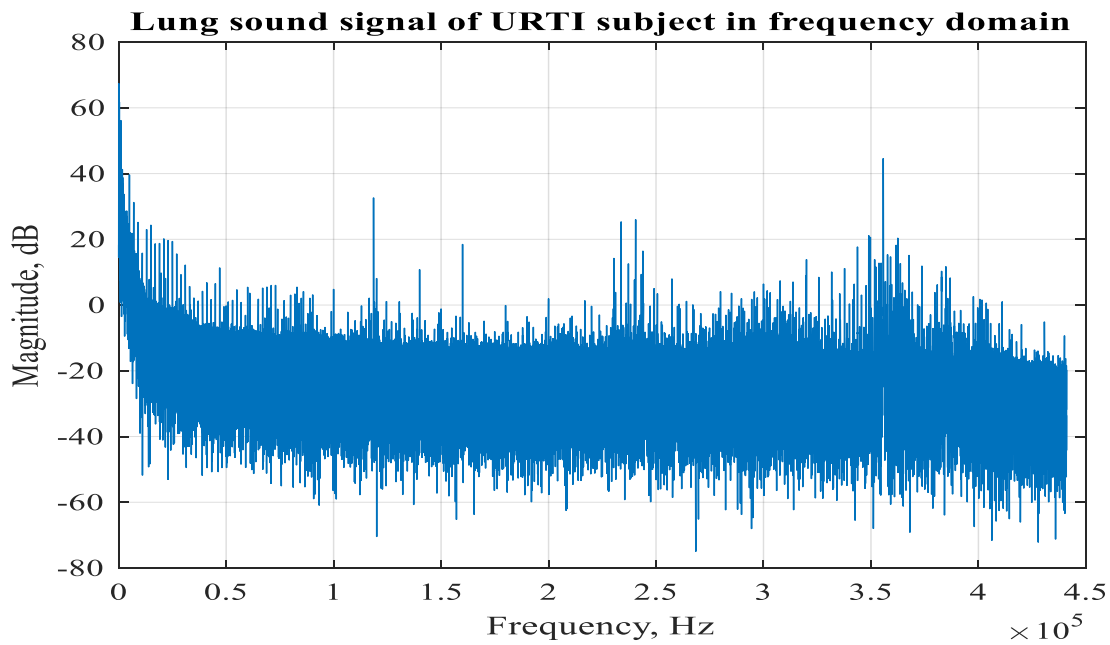
(g)
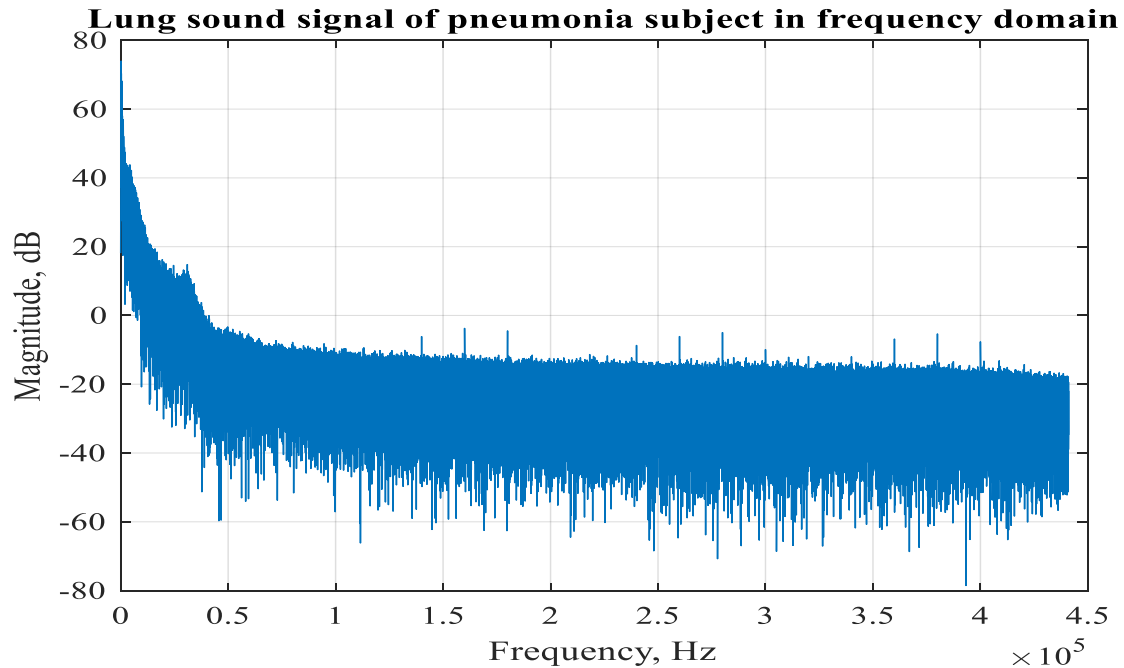
(h)

Figure 2.15: Frequency-domain waveform representation of lung sound signals; (g), URTI (h) pneumonia subjects.

The frequency-domain analysis technique such as Fourier transform will give very good results when used to analyze signals which are uniformly regular or linearly time-invariant. Fourier transform is the most popular transformation which provides better result in the analysis of stationary signals whose frequency content is unchanged over time [48]. However, it is not suitable for analysis of non-stationary signals whose statistical characteristics change with time. Because Fourier transform only provides the signal's frequency information, it loses the signal's time information. Although Fourier transform is a suitable frequency domain analysis method used for best analysis of time-invariant signals, it is not suitable for the analysis of non-stationary signals. For signals such as lung sounds, which are non-stationary, it is difficult to use the Fourier analysis alone. Therefore, to get important diagnostic information from lung sounds, the mutual inclusive time and frequency analysis of lung sounds is highly needed. Other signal analysis techniques such as STFT, and wavelet analysis can be used for time-frequency analysis of signals.

2.3.3. Time-Frequency domain Analysis

2.3.3.1. Short time Fourier transforms (STFT)

STFT is a sequence of Fourier transforms of a windowed signal. The standard Fourier transform provides the frequency information of a signal averaged over the entire signal time interval; whereas STFT provides the time-localized frequency information of a signal whose frequency components change over time. Fourier transform (FT) decomposes a given signal into sinusoidal components invariant over time. Considering a signal (t), then its Fourier transform is given in Equation (2.7) [49].

$$FT_X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.7)$$

The amplitude of the complex value of $FT_X(f)$ represents strength of the oscillatory components at the frequency (f) contained in the signal $x(t)$. However, no information is given on the time localization of such component of the signal. The STFT introduces a temporal dependence of a signal by just applying the FT to the portion of it but not to the entire time interval of the signal. Hence, the short time Fourier transform of a signal is given in Equation (2.8) [49].

$$STFT_{X,W}(\tau, f) = \int_{-\infty}^{\infty} x(t)w * (t - \tau)e^{-j2\pi ft}dt \dots \dots \dots \dots \dots \dots \dots \dots (2.8)$$

At each instant time (τ), we can get a spectral decomposition obtained through applying the FT to the portion of the given signal $x(t)$ viewed through the window, $w * (t - \tau)$ centered at (τ). Therefore, STFT is made up of those spectral components relative to the signal around the instant time (τ). The mathematical expression of STFT given in Equation 2.8, can be interpreted as a convolution and then as output of a filter. Meaning, we can consider STFT as frequency shifting of the signal $x(t)$ by $(-f)$ followed by a low pass filtering given by convolution with the function $w(-t)$ as shown in Equation (2.9) [49].

$$STFT_{X,W}(\tau, f) = \int_{-\infty}^{\infty} [x(t)e^{-j2\pi ft}dt]w(t - \tau)dt \dots \dots \dots \dots \dots \dots \dots \dots (2.9)$$

The window $w(t - \tau)$ should be normalized to unitary energy to preserve the energy and to get the energy distribution of the signal $x(t)$ in the time-frequency plane as shown in Figure 2.16 [49].
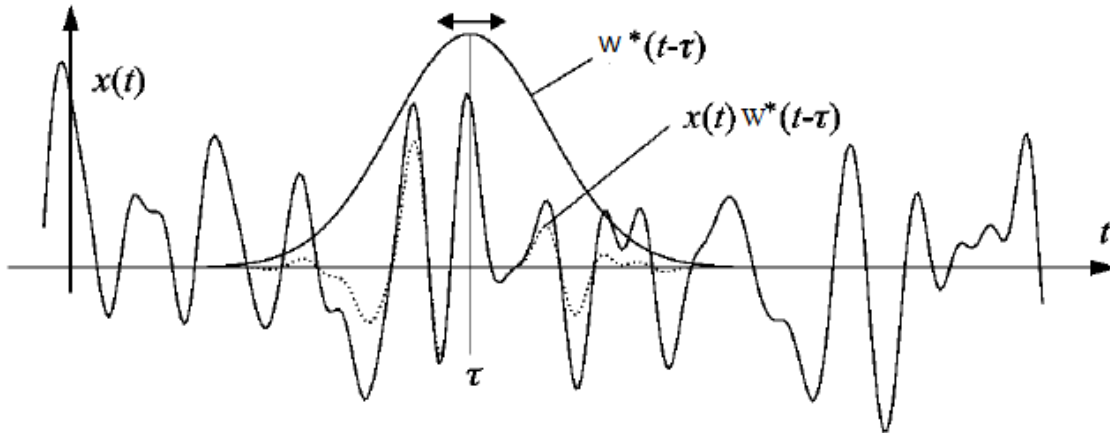
Figure 2.16: Short Time Fourier Transform (STFT) [49].

Generally, STFT is one of the most frequently used tools in audio and speech signals analysis and processing. The basic procedure for computing STFT of a given signal is to divide a longer time signal into shorter segments of equal length, then compute the FT separately over each shorter segment. This reveals the Fourier spectrum on each shorter segment. Though the Fourier spectrum is a vector, the output of STFT is a matrix; i.e., complex-valued. As a result, we can't directly visualize the complex-valued output of STFT. Instead, the outputs of STFT are usually visualized using their two-dimensional log-spectra called spectrogram. A spectrogram is a visual way of representing the signal strength or loudness (for audio signals) of a signal over time at different frequencies present in a particular waveform [50]. From the spectrogram of a signal, one can see whether there is more or less energy at different frequencies, and can also see how energy levels change over time.

2.3.3.2. Wavelet Analysis

The term wavelets mean small waves. The sinusoids used in Fourier analysis are big waves while a wavelet is small wave, an oscillation that decays quickly. Wavelet analysis is a new branch of mathematics which is widely used in signal processing and analysis, image processing, and numerical analysis [51].

It is a new method used for solving different difficult problems in mathematics, physics, and engineering. Wavelet analysis has been applied in a wide range of applications such as data compression, signal processing, image processing, pattern recognition, computer graphics, the detection of artifact and submarines and another medical image technology. Wavelet transform decomposes a signal into a set of basic functions called wavelets. wavelets allow to decomposed complex information such as music, sound, speech, images and patterns into elementary forms at different positions and scales and can subsequently reconstruct them with high precision [51]. Wavelet transform (WT) can be manipulated in two ways such as translation and scaling. Translation is shifting the central portion of wavelet along the time axis, meaning in WT the operation of frequency shifting characteristic of STFT is replaced by a change of scale over time. Translation is done to extract the time information of a signal. In scaling the amplitude and time duration of the wavelet functions are changed to obtain frequency information of a signal. As a result, due to translation and scaling, the wavelet is localized in time-frequency domain analysis of a given signal. Mathematically, given the function $h(t)$ in Equation (2.8) [52], which is a mother wavelet then the basic elements of the transform can be obtained through change of scale.

$$h_{\tau,\alpha}(t) = \frac{1}{\sqrt{|\alpha|}} \, h\left[\frac{t-\tau}{\alpha}\right] \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2.8)$$

where $\alpha \; \varepsilon \; R$ is a scale factor. The term $\frac{1}{\sqrt{|\alpha|}}$, is introduced to normalize the energy of the different wavelets. Therefore, the continuous wavelet transform (CWT) is defined as follows using Equation (2.9) [52].

$$CWT_X(\tau,\alpha) = \int_{-\infty}^{\infty} x(t) \, h_{\tau,\alpha}*(t)dt \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2.9)$$

Based on Equation (2.9), the WT at a given time instant τ and at a certain scale α is the inner product between the signal $x(t)$ and the wavelet $h_{\tau,\alpha}(t)$ with scale α centered at the time τ. So, the CWT indicates that how the signal $x(t)$ close to the function $h_{\tau,\alpha}(t)$ and as the function $h_{\tau,\alpha}(t)$ will be more or less concentrated in time around τ. This similarity criterion will take into account the behavior of the signal $x(t)$ in a neighborhood of the time τ.

Since, $h_{\tau,\alpha}(t) = \dfrac{1}{\sqrt{|\alpha|}} h\left[\dfrac{t-\tau}{\alpha}\right]$, then CWT can also be written as follows as shown in Equation (2.10).

$$CWT_X(\tau, \alpha) = \frac{1}{\sqrt{|\alpha|}} \int_{-\infty}^{\infty} x(t)\ h*\left[\frac{t-\tau}{\alpha}\right] dt \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.10)$$

With increasing the scale factor (growing α), the function h becomes wider over time and WT takes account of the slow behavior of the signal $x(t)$. Then, the CWT behaves as a filter bank with impulse response h, whose extension increases with a scale factor α.

On the other hand, the following expression of CWT, shown in Equation (2.11) emphasizes that, as the scale factor α grows, the signal $x(t)$ is compressed then a wider interval is covered by the function h, which now remains constant in width. In this way, the more compressed signal $x(t)$ can be filtered by the same filter with fixed impulse response h. Large scale (large α) factor provides an overview while the small-scale factor (small α) allows highlighting of details.

$$CWT_X(\tau, \alpha) = \frac{1}{\sqrt{|\alpha|}} \int_{-\infty}^{\infty} x(\alpha t)\ h*[t - \frac{\tau}{\alpha}] dt \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.11)$$

Though the continuous wavelet transforms (CWT) expressed in terms of scale factor (α) shown in the above equations, it can also be defined in terms of frequency by choosing the following expression for the parameter α.

$$\alpha = \frac{f_o}{f} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.12)$$

where, f_o is the center frequency of the Fourier spectrum of h(t) with $f > 0$. Then by substituting Equation (2.12) into Equation (2.11), we can obtain the CWT defined in terms of time and frequency shown in Equation (2.13).

$$CWT_x^h(\tau, f) = \sqrt{|\frac{f}{f_o}|} \int_{-\infty}^{\infty} x(t)\ h*[\frac{f}{f_0}(t - \tau] dt \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.13)$$

A comparison between STFT and WT is shown using Figure 2.17 [52]. The difference between them is the former has constant width; hence their amplitude spectrum is only shifted in frequency, the latter are scaled versions of each other, so with increasing frequency the time width decreases and the spectrum widens (meaning, with the increase of frequency, the wavelet becomes narrower in time but wider in frequency). In short, a constant window region is used in STFT analysis while the variable window region is used in WT analysis. Figure 2.17 also showed that the wavelet is compressed if the scale is low and stretched if the scale is high.
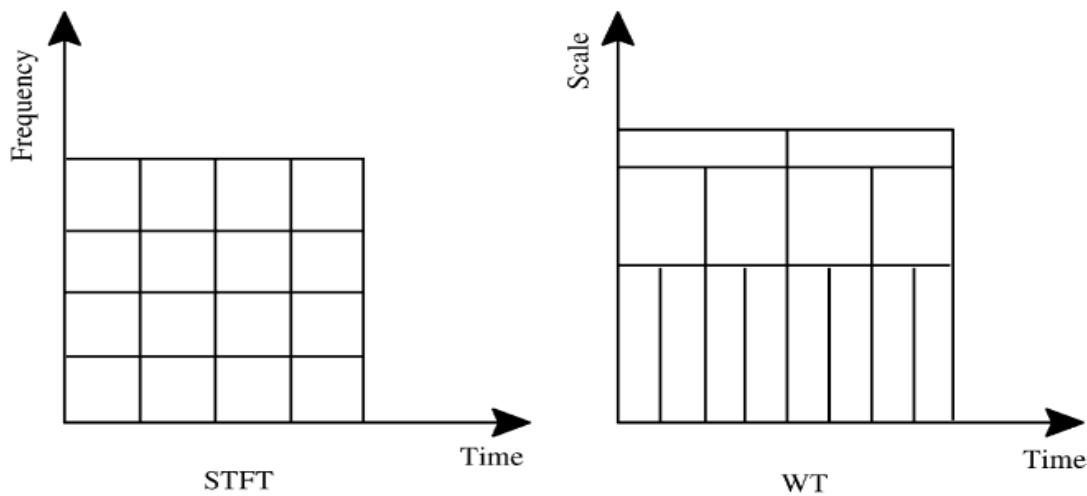
Figure 2.17: Window regions of STFT (left) and WT (right) analyses [52].

Generally, CWT measures how the signal $x(t)$ close to the mother wavelet, i.e., the function $h_{\tau,\alpha}(t)$. Meaning, it measures the similarity between the signal and the wavelet by continuously translating and scaling the mother wavelet. To represent the wavelet coefficients, an infinite number of wavelets are required which increases the redundancy. CWT is extremely redundant because a two-dimensional representation for one dimensional entity through the use of a continuous translation and a continuous scale is extremely redundant. Similarly, STFT overcome the drawback of Fourier transform by mapping a signal into a two-dimensional space of time and frequency. It mainly shows the presence of some specific frequency components of a signal in a single range of location. STFT can't show the existence of frequency components of a signal at any specific location. The main drawback of STFT is that it uses a single and fixed window which will make it unable to extract and provide different features of a signal. Therefore, analysis of the lung sound signals in a multiresolution perspective is highly important.

In this research, we employed wavelet multiresolution analysis technique to analyze the lung sound signals by decomposing them using discrete wavelet transforms (DWT). DWT is the best-suited wavelet signal processing method for multiresolution analysis of signals. Because DWT can decompose low-frequency components of a signal with fine frequency resolution and coarse time resolution and can also decompose high-frequency components of a signal with a fine time resolution and coarse frequency resolution. Moreover, DWT-based multiresolution analysis helps to a better understanding of a signal and is helpful in feature extraction applications. Generally, as compared to the Fourier transform and STFT, Wavelet signal processing using DWT will enable multiresolution analysis of the non-stationary lung sound signals.

2.4. Related Works

Following the development of modern digital signal analysis techniques, different scholars made lots of contribution for improving the current method of lung auscultation by applying different analysis techniques to the recorded lung sounds. In this section, we reviewed some of the earlier related works conducted on classification of lung diseases via processing of lung sounds using different techniques.

Gogus, et al. [53], performed a comparative study for identification of pulmonary disorders using different spectral analysis methods such as fast Fourier transform (FFT), autoregressive (AR) and autoregressive moving average (ARMA). The highest classification accuracy was found by analyzing the lung sounds in AR-Burg than FFT-Welch and ARMA spectral analysis methods. The highest classification accuracy of 93% was achieved by using artificial neural network (ANN) classifier trained with back-propagation algorithm. Similarly, Islam, et al. [54], used Welch spectral method for estimation of power spectral density (PSD) of lung sounds. Features were extracted from the PSD for classification of lung diseases as normal, asthma, and COPD. ANN and SVM classifiers were used for classification, and claimed the highest classification accuracy (93.3%) using SVM classifier.

Kurt, et al. [55], did detection of lung abnormalities using Mel-frequency cepstral coefficient (MFCC) features of the lung sounds. An accuracy of 86%, 75%, 80%, and 62% respectively were found using SVM classifier for the four classification categories such as healthy versus pathological classification (1), rale, rhonchus, and normal sound classification (2), Singular respiratory sound classification (3), and audio type classification with all sound types (4).

Likewise, Saha, et al. [56], also evaluated cepstral features for recognition of lung sounds. ANN was used for the classification of lung sounds as normal, wheeze, and crackle, and claimed an overall accuracy of 97.83%. Choudhry, et al. [57], applied empirical mode decomposition (EMD) to analyze lung sounds for further for classification of pulmonary conditions as normal, bronchitis, and asthma. Further investigation has been done by using the same features on different kernels of KNN and demonstrated the highest accuracy using Fine KNN. Mondal, et al. [58], also used EMD and a first order differentiation algorithm for further decomposition and denoising of lung sound signals. Detection of the lungs status has been done by analyzing the morphological complexities of the lung sound signals. Hilbert transform (HT) algorithm was used for the extraction of the morphological or structural features of lung sounds. Extreme learning machine (ELM) classifier was employed for final classification of data (normal or abnormal class) and claimed an accuracy of 92.86%, specificity of 86.90% and sensitivity of 86.30%. Haider, et al. [59], also conducted respiratory sound-based classification of pulmonary diseases as normal or COPD. A maximum classification accuracy of 83.6% was achieved by the SVM classifier using the most relevant lung sound parameters such as median frequency and linear predictive coefficients. Vora, et al. [60], performed a comparative study between the SVM and K-nearest neighbor (KNN) classifiers to classify COPD severity into four stages. An accuracy of 96.7% and 92.30% were claimed using SVM and KNN respectively.

Moreover, Hidayat, et al. [61], analyzed the lung sound signals complexity using fractal dimension for improving accuracy of lung sound classification. Fractal dimension was calculated in multiscale scheme using coarse-grained procedure. The highest classification accuracy of 99% was found using SVM with the scale of 1-5. Flietstra, et al. [62], also classified crackles found in patients with pneumonia, congestive heart failure, and interstitial pulmonary fibrosis. The crackle features were analyzed using a 16-channel lung sound analyzer and classified using SVM and back propagation neural network. An overall classification accuracy of 88% and 78% were found using SVM and ANN respectively. Don [63], performed random subset feature selection and sequential forward selection algorithms to choose the best subset of features for classification of lung sounds. The random subset feature selection (RSFS) algorithm provides significant improvement for classification of lung sounds as normal or abnormal, and claimed highest accuracy (94.1%) using KNN classifier. An accuracy of 89.5% was obtained for both SVM and Naive Bayes classifiers.

Rizal, et al. [64], performed pulmonary crackle detection using multiscale tsallis entropy. Tsallis entropy measurements were carried out through several orders and the pulmonary crackle signals were decomposed using coarse grained procedure. Multilayer perceptron neural network was used for classification of pulmonary conditions as normal or crackle. The highest classification accuracy of 97.67% was obtained at the second order tsallis entropy on the scale of 1-15. Similarly, Nugroho, et al. [65], compared multiscale entropy techniques for lung sound classification. The results of the comparison showed that the multiscale permutation entropy (MPE) provides the highest accuracy of 97.98% using Multilayer perceptron for five lung sound classes (asthma, friction rub, normal bronchial, and stridor).

Moreover, Lmtiaz, et al. [66], reviewed a paper which used features derived from continuous wavelet transform (CWT) for classification of lung sounds. ANN classifier trained using resilient back propagation algorithm was implemented to classify the lung sounds and an overall accuracy of 94.02% was claimed. Aykanat, et al. [67], used CNN to perform lung abnormalities detection involving four classification categories such as healthy versus pathological classification (1), rale, rhonchus, and normal sound classification (2), singular respiratory sound classification (3), and audio type classification with all sound types (4). By using spectrogram images of the lung sound signals obtained through STFT, they obtained an accuracy of 86%, 76%, 80%, and 62% respectively for the above-mentioned four classification categories. Likewise, Quan, et al. [68], performed classification of asthma severity using DNN trained with back-propagation algorithm to classify the data obtained from NIS (National Inpatient Sample) and MIMIC-III ('Medical Information Mart for Intensive Care'). An accuracy of 86% and 91.1% were found to classify the data obtained from NIS and MIMIC-III databases respectively, using the demographic information found in the databases.

Hosseini, et al. [69], applied deep convolutional neural networks (DCNNs) to diagnose pulmonary diseases using auditory and demographic information and a 78% classification accuracy was claimed using only the auditory information and 83% accuracy using a combination of auditory information with demographic information. Similarly, Demir, et al. [70], employed CNN based approach for efficient classification of lung diseases (crackles, crackles combined with wheezes, wheezes, and normal). An accuracy of 65.5% and 63.09% was claimed using SVM classifier and DCNNs via spectrogram images, respectively.

Fraiwan, et al. [71], used a deep learning network architecture consisted of two stages such as CNN and bidirectional long short-term memory (BDLSTM) units for recognition of pulmonary diseases from lung sounds. The highest average accuracy of 99.62% was obtained using CNN and BDLSTM. Table 2.2 shows a summary of the above-mentioned previous research works done by various scholars in different times. In Table 2.2, the methods which the researchers used, the accuracy they achieved, and the existing gaps associated with each of the researches have been summarized.

Table 2.2: A summary of previous related researches.

Authors	Applied methods	Results (Accuracy %)	Research gap
Islam, et al., 2018. [54]	FFT-Welch and SVM	93.3	The study didn't consider temporal auditory information of signals. Limited spectral features extracted from the power spectral density (PSD) of lung sounds for classification.
Kurt, et al., 2017. [55]	MFCC and SVM	Max accuracy of 86% for category 1	Only spectral features of lung sounds (i.e., Mel-frequency cepstral coefficients of lung sounds) were used for classification.
Haider, et al., 2019. [59]	Median frequency, Linear predictive coefficients & SVM	83.6	Limited spectral features were used for binary classification (normal or COPD).
Vora, et al., 2019. [60]	SVM KNN	96.7 92.3	Limited for classification of COPD using KNN and SVM algorithms.
Hidayat, et al., 2018. [61]	Fractal dimension and SVM	99	The study didn't consider spectral auditory information, only used temporal features of lung sounds obtained through the coarse-grained procedure for classification.

_____

Mondal, et al., 2014. [58]	EMD, HT and ELM	92.86	Limited spectral features characterizing the morphology of lung sounds were considered for binary classification (normal or abnormal).
Don, 2020. [63]	RSFS and KNN	94.1	Computationally costly due to random subset feature selection method for binary classification (normal or abnormal).
Gogus, et al., 2016. [53]	ANN with AR-Burg	93	The study didn't consider temporal auditory information of lung sounds. Limited spectral features were considered for classification.
Saha, et al., 2016. [56]	Cepstral features and ANN	97.83	Only spectral features (i.e., cepstral coefficients of lung sounds) were employed for classification. The study didn't consider temporal auditory information of lung sounds.
Rizal, et al., 2018. [64]	MSE and Multilayer perceptron neural network	97.67	The study didn't consider spectral auditory information, only temporal features of lung sounds obtained through multiscale tsallis technique for binary classification (normal or crackle).
Nugroho, et al., 2018. [65]	MSE and Multilayer perceptron neural network	97.98	Only temporal features of lung sounds obtained through multiscale tsallis technique for classification, i.e., there is loss of spectral information.
Lmtiaz, et al., 2019. [66]	CWT and ANN	94.02	Extremely redundant due to continuous translation and scaling, and is computationally costly.
Aykanat, et al., 2017. [67]	CNN, STFT	Max accuracy of 86% for category 1	Unable to extract different features from the lung sound's spectrograms due to fixed window region of STFT.
Quan, et al., 2017. [68]	Demographic info, DNN with	91.1	The study didn't consider auditory information of lung sounds, and the study is

	MIMIC-III		limited for classification of asthma.
Hosseini, et al., 2020. [69]	Auditory & Demographic info with DCNN	83	Unable to extract different features from the lung sound's spectrograms due to fixed window region of STFT.
Demir, et al., 2019. [70]	STFT and SVM	65.5	Unable to extract different features from the lung sound's spectrograms due to fixed window region of STFT.

2.4.1. Summary

Generally, there are various signal analysis techniques which can be used for analyzing lung sounds. Most researchers proposed spectral analysis techniques such as FT, FFT-Welch, AR, AR-Burg, ARMA, MFCC, and others to analyze the acquired lung sound signals for further detection of the lung pathologies [54,55,59,60,63,53,56]. Time-domain analysis techniques such as statistical methods, empirical mode decomposition (EMD), or fractal analysis methods were used by some researchers for classifying abnormal lung sounds [61,62,57,58]. Multiscale entropy (MSE) proposed by Costa et al. with a coarse-grained procedure is one of the popular biological signal analysis methods. Meanwhile, MSE has been proposed in literature to analyze lung sounds for further detection of pulmonary abnormalities [64,65]. However, frequency-domain as well as time-domain analysis techniques alone couldn't provide enough information when used on non-stationary signals like lung sounds. Therefore, to get important diagnostic information from the lung sounds, time-frequency analysis of lung sounds is crucial. Some researchers were employed time-frequency analysis techniques such as short time Fourier transform (STFT) and wavelet transform (WT) for analysis of lung sounds [66,67,70]. STFT overcome the drawback of frequency-domain analysis techniques by mapping a signal into a two-dimensional space of time and frequency but with a single fixed window, unable to extract different features. Studying the lung sounds in a multiresolution perspective is very important. Naturally, signals including the lung sounds usually contain both the low and high frequency components which will vary with time. The low-frequency components of such signals vary slowly with time, which need fine frequency resolution and coarse time resolution.

While the high frequency components of a signal vary quickly with time, which need fine time resolution and coarse frequency resolution. Multiresolution analysis (MRA) method is used to analyze such non-stationary signals having both high and low frequency components and whose frequency vary along with time [72]. To do this, wavelet signal processing is naturally a multiresolution analysis method because of the window dilation process. Lung sound signal features derived from continuous wavelet transform were used in literature for classification of lung sounds using ANN [66]. CWT is extremely redundant because a two-dimensional representation for one dimensional entity through the use of a continuous translation and a continuous scale is extremely redundant.

Discrete wavelet transforms (DWT) are best-suited processing method for multiresolution analysis of signals [72]. Because, DWT can decompose high-frequency components of a signal with a fine time resolution and coarse frequency resolution and can decompose low-frequency components of a signal with fine frequency resolution and coarse time resolution. So, DWT enable multiresolution analysis of the non-stationary lung sound signals by decomposing the signal into detail (high frequency) coefficients and approximation (low frequency) coefficients using successive high pass and low pass filtering operations. Therefore, a method involving DWT based analysis of lung sounds for further multiple classification of pulmonary diseases has been proposed in this study. After acquisition, pre-processing, and extraction of features from the lung sound signals, important features were selected and given as input to different machine learning classifiers. At last, after comparing classification accuracy of the different machine learning classifiers, SVM algorithm particularly fine gaussian SVM has been chosen for final classification of lung diseases as COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma or healthy.

CHAPTER THREE

METHODS and MATERIALS

3.1. Overview

The general procedure used in this study for classification of pulmonary diseases from the lung sounds is shown in Figure 3.1. An electronic stethoscope has been constructed by modifying the existing traditional acoustic stethoscope and is used for lung sounds acquisitions. Discrete wavelet transform (DWT) was used for multiresolution decomposition of the non-stationary lung sound signals. Moreover, a DWT-based denoising of lung sounds which proceeds three steps such as decomposition, detail coefficients thresholding, and reconstruction have been done. Following signal decomposition and reconstruction, different features have been extracted for classification of the lung sound signals of healthy, COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, and asthma diseased subjects. Feature ranking using filter selection method has been done to select the most relevant features. Additionally, feature normalization has been done to normalize the extracted features before using them for model training and data classification. Finally, we trained a supervised machine learning classifier for final classification of lung diseases by feeding the selected and normalized features as inputs. The activity of a task using machine learning requires the accomplishment of tasks such as: import data and structure it, data pre-processing, feature extraction, feature selection or ranking, train, and optimize classification models. Therefore, the general procedure presented using Figure 3.1 consists of series of steps including construction of electronic stethoscope, signal acquisition, pre-processing, feature extraction, feature selection, feature visualization and normalization, data splitting, model training and evaluation, model optimization, and data classification.
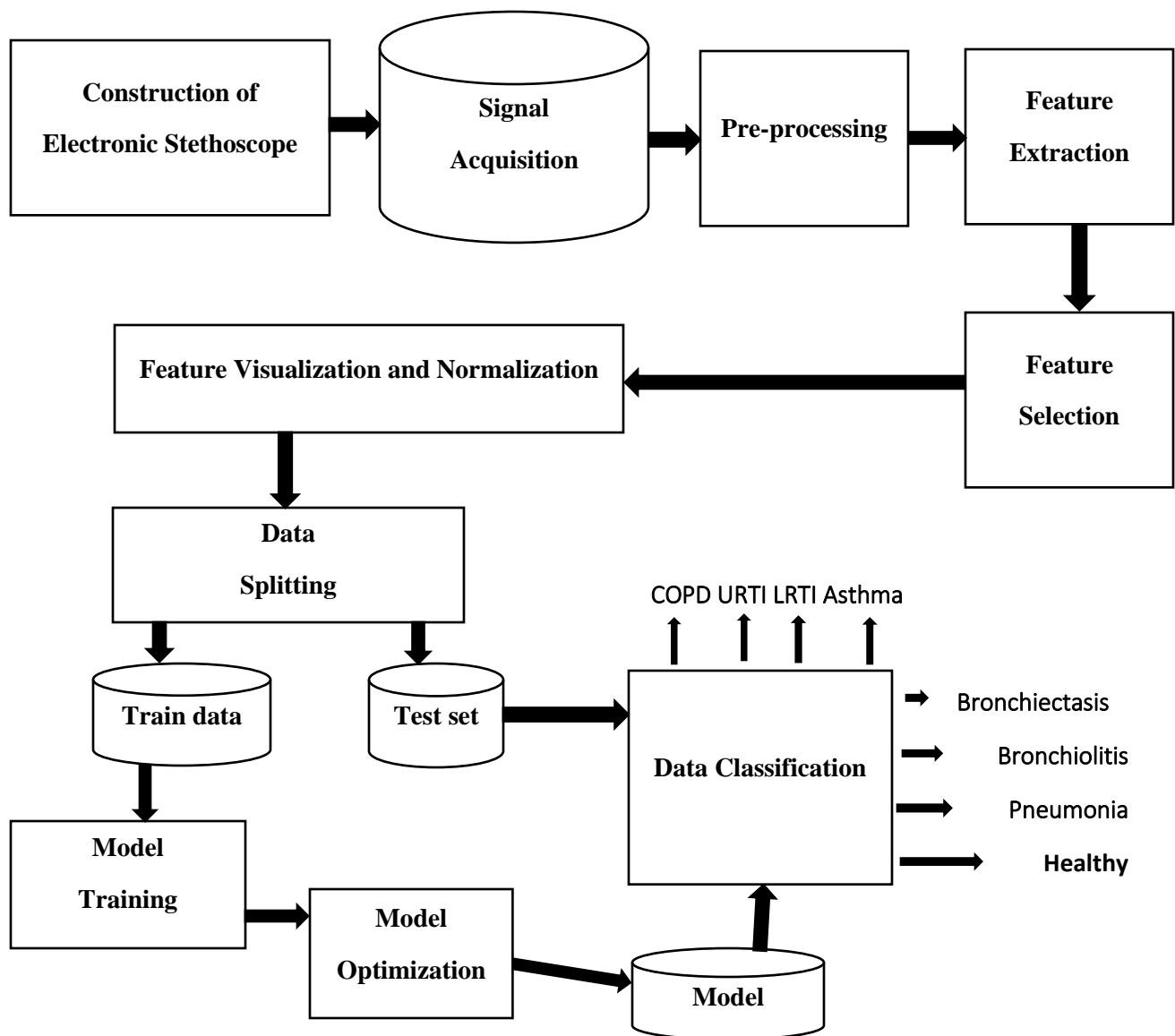
Figure 3.1: A general procedure for classification of lung diseases.

3.2. Construction of Electronic Stethoscope

Lung sounds are best recorded using stethoscope. Stethoscope is a medical device used for listening the internal sounds of human body such as Korotkoff sounds, heart sounds, lung sounds, and other internal body sounds [73,74]. The main parts of the stethoscope are ear tips, ear tube, tubing, headset, stem, chest-piece, diaphragm, and bell.

A proper combination of these parts creates a functioning stethoscope, which allows doctors, nurses, and other medical professionals to listen to a wide range of sounds so that they can quickly determine and treat potential medical issues. The air tips of the stethoscope are the parts which goes into the user's ear and used to receive the sounds that come from the chest-piece. The ear tubes are the metal or steel parts of the stethoscope connecting the ear tips and the synthetic or PVC tubing, which connects to the stem of the chest-piece. Tubing is the soft flexible line of the stethoscope and its purpose is to maintain and transfer the sound level that is acquired by the diaphragm or bell to the ear tubes. The headset is the part of the stethoscope holding the combined components of the upper half of the stethoscope including ear tubes, tension springs and ear tips. Stem is another metal or steel part of the stethoscope which basically connects the tubing to the chest-piece. The chest-piece is the head of the stethoscope which is composed of the connected stem, diaphragm and /or bell. The diaphragm is the larger circular end of the chest-piece used to listen higher-frequency sounds while bell is the smaller circular end of the chest piece used to listen lower-frequency sounds. Figure 3.2 shows labeled image of a stethoscope with its main components discussed above.
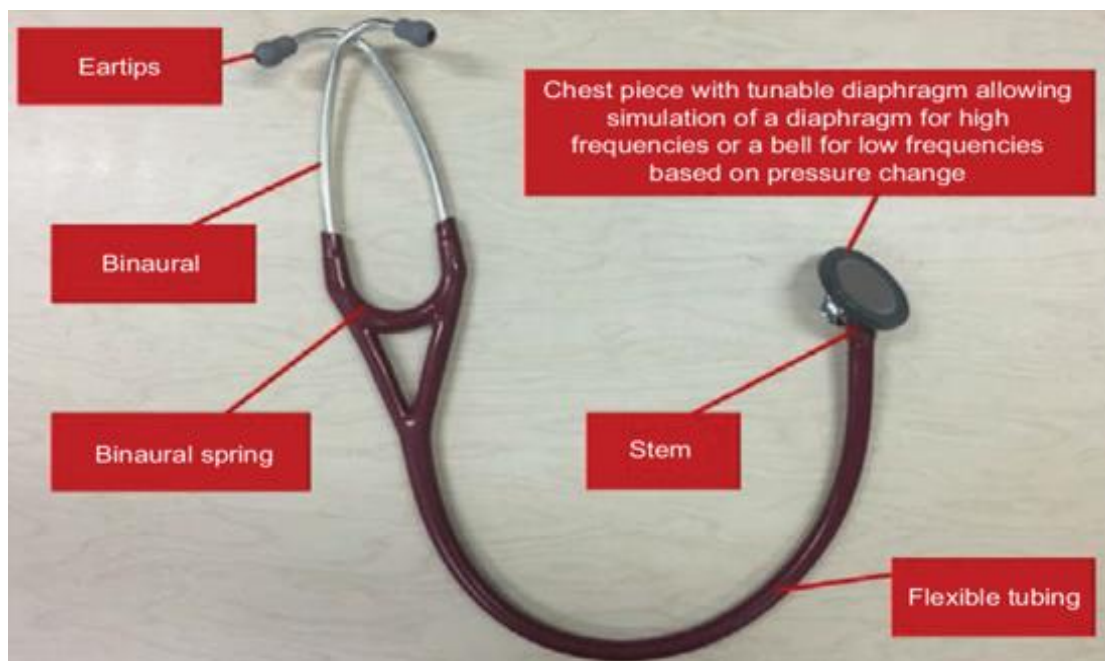
Figure 3.2: Labeled image of a stethoscope [75].

Generally, stethoscope is a useful medical device used by doctors, nurses, and other medical professionals to perform checks on the heart and lung functions [75]. However, the function of traditional stethoscopes particularly an acoustic stethoscope types which are available in many health care settings of the developing nations, are limited to analog filtering and amplification of bio-sounds for further interpretations by trained health professionals [76]. To avoid such limitations and subjectivity issues, developing an electronic stethoscope by doing certain modifications over the traditional acoustic stethoscope is highly important. The main function of an electronic stethoscope is to convert analog sound signals into digital ones.

One of the goals of this research was to develop a cost-effective electronic stethoscope which can be used for lung sounds acquisitions and interfaced it with computer for further analysis of lung sounds. To create an electronic stethoscope via modifying the traditional acoustic stethoscope, a transduction mechanism is required which will be used for converting acoustic waves into an equivalent electrical signal. After looking for different types of sensors which will be used for transducing acoustic signals into electrical signals, we used a condenser microphone due to its flat frequency response, high SNR, better sensitivity and can be affordable at low-cost [76]. Moreover, selection of the sensor (microphone) has been done following the context of computerized respiratory sound analysis (CORSA) specification recommendations [76]. CORSA offers standard considerations for proper selection of a condenser microphone which will be used for respiratory recordings, described in Table 3.1.

Table 3.1: Recommended specifications of sensor that detects human lung sounds [76].

Parameter	Specification
Frequency response	✓ 50-5000 Hz.
Dynamic range	✓ 60 dB
Sensitivity	✓ It is the voltage being generated in response to an input sound pressure of 0.1 Pa. ✓ Its typical value is 1mv/Pa.

_____

	✓ It must be stable across all frequencies, sound directions, and varying noise.
SNR	✓ It is the ratio of the output voltage to the noise when no input signal is applied. ✓ According to CORSA recommendation, it is recommended to be > 60 dB
Coupling	✓ Piezoelectric: contact. ✓ Condenser microphones: conical, diameter 10 to 25 mm.
Fixing	✓ Adhesive ring for piezoelectric sensors. ✓ Adhesive ring or elastic belt for condenser microphones.
Noise and interference	✓ Shielded microphone for protection from mechanical vibrations or EM shielded twisted pair or coaxial cable.

The above table 3.1 describes the recommended specifications provided by computerized respiratory sound analysis (CORSA) for sensors that detect human respiratory sounds. Therefore, based on the standard specification set by CORSA, we used an electret microphone, shown in Figure 3.7. An electret condenser microphone is the most common type of transducer used to detect sound signals and produces an equivalent electrical signal. It has two plates inside it just like a capacitor. The distance between the plates is directly proportional to the sound present in the environment. The capacitor is rated for 20 Hz to 16,000 Hz, so any sound waves in that range will be picked up by this microphone. The frequency range of lung sounds found in this range; hence it can be used to detect the lung sounds acquired by the stethoscope head and convert it into an equivalent electrical signal. Therefore, in this research we used an electret microphone having an operating frequency range 20 Hz to 16,000 Hz, high SNR (more than 62 dB), and had high sensitivity (High SNR).

3.3. Signal Acquisition

In this research, both online and local lung sound records were used. We used some annotated lung sound recordings from the international conference on biomedical and health informatics (ICBHI) respiratory sound dataset. This dataset was created by two independent research teams based in Greece and Portugal [77]. The lung sound records found in this dataset were acquired from 126 participants annotated with eight types of respiratory conditions such as COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma, and healthy. The duration of each record ranges from 10 to 90 seconds and were recorded from different chest locations such as left and right anterior, left and right posterior, and left and right lateral regions of the chest, shown in Figure 3.3.
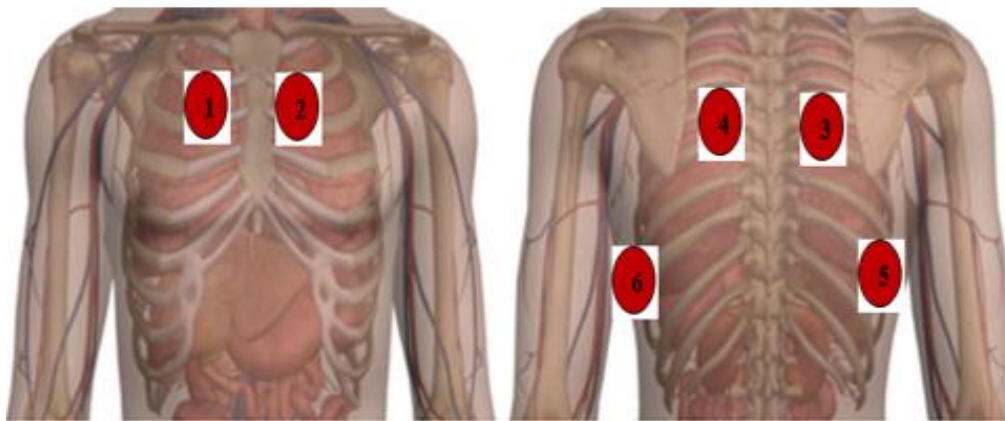
Figure 3.3: Lung auscultation points for the recordings of lung sounds marked in red [77].

The lung sound records found in the ICBHI respiratory sound dataset were acquired through auscultation of the stethoscope at different points such as, left and right anterior (marked 1 and 2), left and right posterior (marked 3 and 4), and left and right lateral (marked 5 and 6) regions of the chest wall shown in Figure 3.3. The ICBHI respiratory sound dataset is highly imbalanced with around 86% of the data belonging to COPD. Therefore, in this study, besides to the online lung sound data, we collected some local lung sound recordings using our electronic stethoscope. The collection of local lung sound data was held at Jimma University Institute of Health, Jimma University Medical Center (JUMC), Department of Internal Medicine, Pulmonology Unit. A total of 287local lung sound records where;60 asthma, 49bronchiectasis, 23 pneumonia, 15 URTI, 47bronchiolitis, 48LRTI, 30 healthy, and 15COPD records were collected. Each of the records were taken from the clinically admitted subjects.

_____

9 subjects with asthma condition, 6 subjects with bronchiectasis condition, 4 subjects with pneumonia condition, 3 subjects with URTI condition, 6 subjects with bronchiolitis condition, 7 subjects with LRTI condition, 6 subjects with COPD condition, and 6 healthy subjects were selected to record the local lung sound records. Once we acquired the online and local lung sounds, we did multiresolution analysis of the signals using DWT.

3.4. Multiresolution Decomposition of Lung Sound Signals using DWT

The discrete wavelet transforms (DWT) have been widely used in biomedical signal processing of different signals including the non-stationary signals. In this study, DWT has been used for multiresolution decomposition of the non-stationary lung sound signals. Because DWT can decompose high frequency components of a signal with a fine time resolution and coarse frequency resolution and can also decompose low frequency components of a signal with fine frequency resolution and coarse time resolution. Therefore, DWT can be efficiently realized by decomposing the given signal into detail (high frequency) coefficients and approximation (low frequency) coefficients. The central frequency and the frequency bandwidth of the detail coefficients decrease by half when the decomposition level increases by one. For example, the central frequency and the frequency bandwidth of the second detail coefficient (D2) are half that of the first detail coefficient (D1). The approximation at a certain resolution contains all of the information about the signal at any coarser resolutions. For instance, the frequency band of the second approximation (A2) covers the frequency bands of the third approximation and detail coefficients (A3 and D3).

Generally, DWT is used to analyze the lung sound signals by decomposing the signal into its coarse approximation and its detailed information by using successive high pass and low pass filtering operations. The procedure and the stages involved during multiresolution decomposition of a signal $x[n]$ using DWT is schematically shown using Figure 3.4. Each stage involves two digital filters (low pass and high pass filters) and two down samplers by 2. Each filter divides the input signal into two half-band frequencies. The first filter $h[n]$ is the discrete mother wavelet, which is high pass in nature and is used to extract the upper half of the frequencies of the signal which is referred to as its detail information. The second filter $g[n]$ is the mirror version of $h[n]$, which is low pass in nature and is used to extract the lower half of the frequencies of the signal which is referred to as its approximation information.

The scheme presented using Figure 3.4 also showed that the down sampled outputs of the first high pass and low pass filters provide the first detail information D1 and the approximation information A1 respectively. As shown in Figure 3.4, every decomposition level has a filtering and down sampling blocks in which the down sampling is done by a factor of 2. The first approximation information (A1) of the signal is further decomposed and the process will be continued until the end of the required level of decomposition. The approximation obtained at the first level i.e., A1 is the summation of the approximation and the detail at the second level. Similarly, the approximation obtained at the second level i.e., A2 is the summation of the approximation and the detail at the third level. Hence, as the level of decomposition increases, the approximation (low frequency components of a signal) will be obtained in such a way. Figure 3.4 shows a three-level wavelet decomposition of a signal. In this research, we implemented a six-level wavelet decomposition for better analysis of lung sound signals by following the same procedure shown in Figure 3.4.
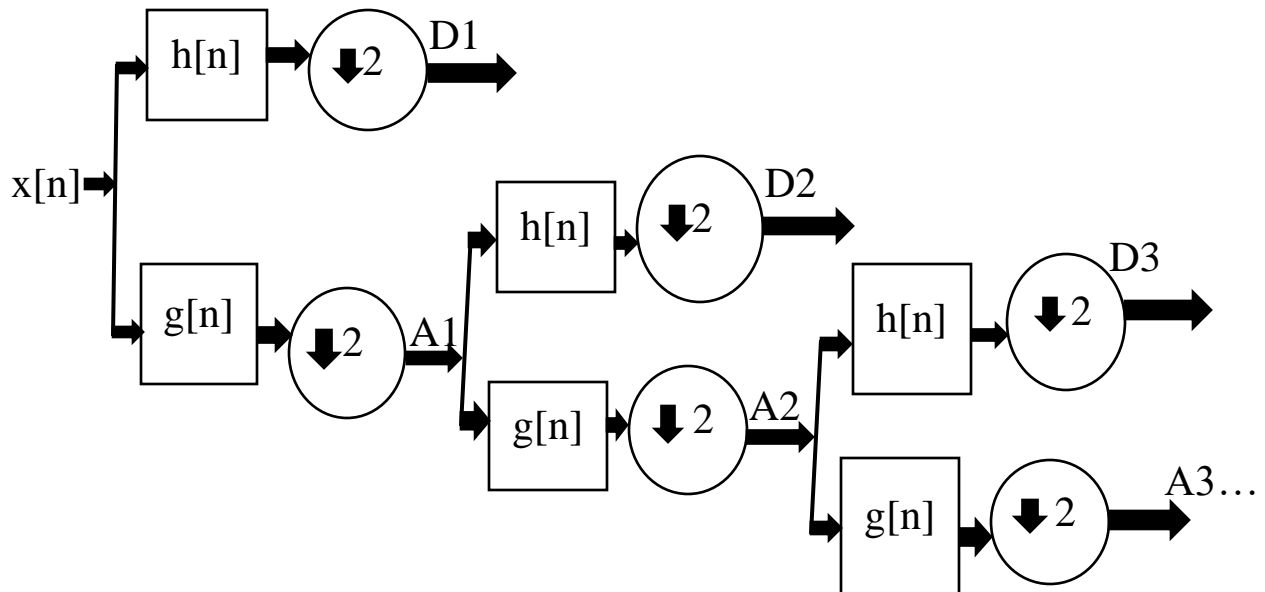
Figure 3.4: A procedure for a three-level wavelet decomposition of a signal.

3.5. Pre-processing

Signal pre-processing, particularly signal denoising is the critical step in signal processing which involves the removal of artifacts that can corrupt the acquired signals. The stethoscope recording is also vulnerable to different noises. Therefore, to denoise the noises without altering important features of the lung sounds, wavelet-based denoising technique is used in this research.

Since wavelet transform is a time-frequency signal analysis method and has a good localization in both time and frequency domains, it is highly helpful to remove unwanted signals altering the lung sounds both in time and frequency content.

3.5.1. Pre-processing of Lung Sound Signals using Wavelet Denoising Technique

A discrete wavelet transform (DWT)-based denoising of lung sounds proceeds in three steps such as decomposition, detail coefficients thresholding, and reconstruction. The decomposition step involves selection of an appropriate wavelet function or the mother wavelet and determination of the number of wavelet decomposition level. So, this is done by choosing a suitable mother wavelet and conduct the wavelet decomposition of the original signal S at level N. In the second step, for each wavelet decomposition level from 1 to N, we applied soft thresholding to the detail (high frequency) coefficients of the signal S. The final step is the reconstruction. The reconstruction is done upon the original approximation coefficients of level N and the modified detail coefficients of levels from 1 to N. In this research, we followed the following procedure shown in Figure 3.5 to implement the DWT-based denoising of the lung sound signals.
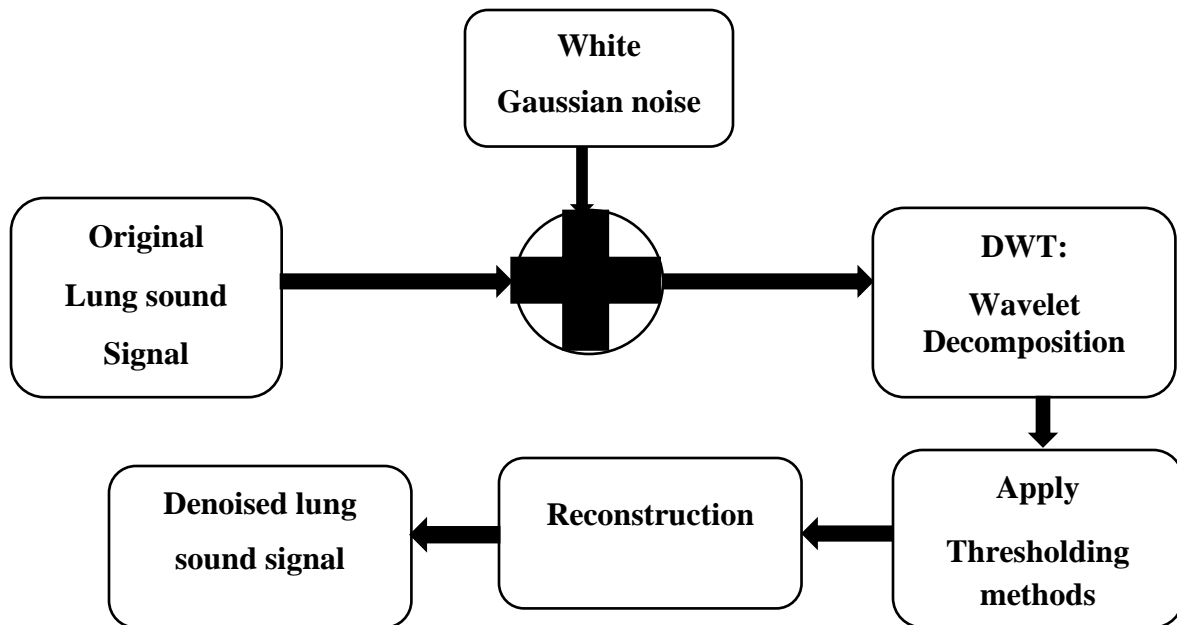
Figure 3.5: A procedure for wavelet-based denoising of lung sound signals.

The objective of this wavelet-based denoising technique, shown in Figure 3.5 is to recover a signal with less noise than the original signal by using an appropriate wavelet function and a suitable thresholding method. As shown in Figure 3.5, decomposition is the first step in DWT based denoising procedure. DWT is basically used to analyze the signal by splitting the given signal into its detail (high frequency) coefficients and approximation (low frequency) coefficients. Such splitting process is called signal decomposition. Therefore, DWT is used to decompose the lung sound signals into two parts such as high frequency and low frequency components by employing successive high pass and low pass filtering operations. During DWT-based signal decomposition, the signal is passed through a series of high pass filters to analyze the high frequency components of the signal and low pass filters to analyze the low frequency components of the signal. After looking the results of different wavelet functions, we conducted a 6th level decomposition of lung sound signals using *Sym13* wavelet function. Following the decomposition, the next step is to apply the thresholding method to remove the noise from the wanted signal. For thresholding, two most commonly used thresholding methods such as soft and hard thresholding methods were used in this study. Soft thresholding method is based on the process of setting zero coefficients whose absolute values are lower than the settled threshold value (λ), otherwise the coefficient value is modified, as shown in Equation (3.1). The thresholding is equal to the value of a sign function which multiplies the difference value between a coefficient (W_j) and the threshold (λ). The mathematical equation for the soft thresholding function representing the input-output characteristics of the process is given in Equation (3.1) [78].

$$f_S = \begin{cases} sgn(W_j)(|W_j| - \lambda), & for \ |W_j| \geq \lambda \\ 0, othrwise \end{cases} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. (3.1)$$

Where, (W_j) is the wavelet coefficient, (λ) is the threshold value and f_S is the soft thresholding function.

Hard thresholding is the process of establishing to zero to the coefficients whose absolute values are lower than the threshold vale (λ), otherwise the coefficient value is not modified as shown in Equation (3.2). The mathematical equation for the hard thresholding function representing the input-output characteristics of the process is given in Equation (3.2) [78].

$$f_H = \begin{cases} W_j, & for \ |W_j| \geq \lambda \\ 0, othrwise \end{cases} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.2)$$

Where, (W_j) is the wavelet coefficient, (λ) is the threshold value and f_H is the hard thresholding function.

Hard thresholding method had a problem of discontinuity, it leads to a certain fluctuation in the reconstruction of the original signal. conversely, soft thresholding method is more stable than hard thresholding method. Moreover, in this study, the performances of these thresholding methods for the denoising of lung sounds were investigated using the four well-known standard threshold selection rules such as *heursure, rigrsure, minimaxi and sqtwolog*.

Sqtwolog rule: This rule was proposed by Donoho and Johnstone. Based on this criterion, the threshold values (λ) are calculated by using universal threshold (square root log) method given by Equation (3.3) [79].

$$\lambda_j = \sigma_j\sqrt{2log(N_j)}\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(3.3)$$

σ_j is given as, $\sigma_j = \frac{MAD_j}{0.6745} = \frac{median(|W|)}{0.6745} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.4)$

Where, N_j is the length of the noisy signal at j^{th} scale, σ_j is Median Absolute Deviation (MAD) at j^{th} scale, and W represents wavelet coefficients at scale j.

Rigrsure rule: This is an adaptive thresholding method, where the threshold value (λ) is selected based on the principle of the Stein's unbiased risk estimate (SURE) quadrature loss function. It is a soft threshold estimator, starting with an estimate of risk for a particular threshold value (λ), the algorithm minimizes the risks in λ to yield an appropriate threshold value. In *rigrsure* criterion, the threshold can be calculated using Equation (3.5) [79].

$$\lambda_j = \sigma_j\sqrt{w_b} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.5)$$

Where, w_b is the bth squared wavelet coefficient (coefficient at minimal risk) chosen from the vector W containing the square of wavelet coefficients from small to large as $W = [w1, w2, .. wN]$ and σ is the standard deviation of the noisy signal.

Heursure rule: It is a mixed threshold method. In this criterion, the threshold is selected using a combination of *Sqtwolog* and *Rigrsure* methods. If the SNR (signal to noise ratio) is very small, the SURE method's estimation is poor. In such cases, the *Sqtwolog* method gives better threshold estimation [80].

Minimaxi rule: The principle of *minimaxi* method is to minimize the maximum risk of estimation. Based on the *minimaxi* criterion, a fixed threshold is chosen to yield minimax performance for mean square error against an ideal procedure [80].

In this research, we applied soft thresholding method using *Sqtwolog* threshold selection rule for the denoising of the lung sound signals. After decomposition and thresholding, the final step is the reconstruction of the original signal. This is the step where the noise free lung sound signal or the denoised lung sound signal is generated by performing inverse discrete wavelet transform (IDWT). During DWT-based decomposition, successive high pass and low pass filters were used. The outputs of the high pass and low pass filters are called the DWT coefficients. These coefficients in turn enable the reconstruction of the original signal, such process is called IDWT. Finally, after the denoising process, a performance measure called signal to noise ratio (SNR) was used to measure the performance of the denoising algorithms. In DWT-based denoising process, after decomposition, thresholding, and reconstruction, looking for the performance of the denoising algorithm is paramount. Otherwise, it is difficult to say the algorithm is best without showing its denoising performance. In case of signal denoising, computing SNR is the common way to show the performance of the denoising algorithm. To do this, addition of white Gaussian noise to the lung sound signals is the most suitable way to see the effect of noise added to the lung sound signals [81]. Next, for the lung sound signals containing the desired level of white Gaussian noise, the performance of the denoising algorithms can be measured by calculating SNR. We computed the SNR values using the following mathematical equation shown in Equation (3.6) [82].

$$SNR_{dB} = 10\log_{10}\left(\frac{X_{signal}}{X_{noise}}\right)^2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.6)$$

Where, X_{signal} is the the root mean square amplitude of the signal and X_{noise} is the root mean square amplitude of the noise.

3.6. Feature Extraction

Audio signal processing algorithms involves analysis of signals in different domains, extracting its features, predicting its behaviors, recognizing if any pattern is present in the signal, and determining how a particular signal is correlated to another similar signals [83]. Any parameter which has a potential to discriminate between different classes is termed as a feature. So, feature extraction (i.e., extraction of discriminative features of signals) is one of the most vital parts of machine learning process. Moreover, feature extraction or featurization is the process of mapping an input from X to a vector in R^d, shown in Figure 3.6.

Raw input   Feature Extraction   Feature vector
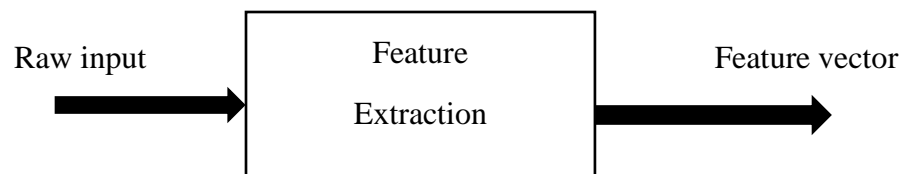
Figure 3.6: Feature extraction [84].

Generally, feature extraction is the process of transforming the input data into set of discriminatory features. The extraction of discriminatory features in the signal enhances the reduction of the length of the input data vector by eradicating redundancy in the signal and compressing the relevant information into a feature vector of significantly lower dimension. Once the lung sound signals are acquired, analyzed, and preprocessed, we need to determine features by employing feature extraction techniques. There are different feature extraction techniques. The temporal analysis techniques can be used for extraction of statistical distribution features by analyzing the time domain waveform representation of a signal while the spectral analysis techniques utilize the spectral representation of the audio signal for extraction of spectral features. In this research, different features have been extracted using discrete wavelet transform (DWT) feature extraction technique to discriminate between the lung sound signals of healthy, COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, and asthma diseased subjects. The DWT feature extraction technique is highly useful for extraction of wavelet features from the non-stationary lung sound signals. Thus, for such non-stationary lung sound signals an appropriate method for analysis and extraction of features should be used. Therefore, DWT feature extraction technique has been chosen for this purpose.

Basically, the DWT-based feature extraction procedure involves two steps. First, decompose the given signal into N levels using filtering and decimation to get the approximation and detailed coefficients. This can be done by employing successive high pass and low pass filtering operations. The Second step is followed by the reconstruction of the signal. This can be done by performing inverse discrete wavelet transform (IDWT). The features extracted from the IDWT of signals are considered highly useful features for input of the classifiers due to their effective time-frequency representation of the non-stationary signals. In this research, we applied the following procedure shown in Figure 3.7 to implement the DWT-based feature extraction algorithm for the extraction of features of the lung sound signals. As shown in Figure 3.7, the DWT feature extraction algorithm involves series of steps. The various steps involved during the DWT feature extraction algorithm are discussed as follows:

Step 1: The lung sound signals decomposed into six detail sub bands using DWT at level $N = 6$. The sub bands are the details (high-frequency band coefficients) and the approximation (low-frequency band coefficients).

Step 2: The approximation coefficients are further decomposed to extract localized information from the sub bands of detail coefficients. In this study, a six-level decomposition was done using Sym13 wavelet.

Step 3: All the six level detail band coefficients were taken for further analysis and processes.

Step 4: For six detail sub bands, the frequency vector (in radians/sample) is extracted using periodogram function in MATLAB.

Step 5: Finally, after decomposition the signals are reconstructed using IDWT.

Step 6: Then the features were computed from the reconstructed signal by using the syntax or by implementing the formula.

Step 7: At last, the extracted features for all the lung sound classes were tabulated in feature table for classification.
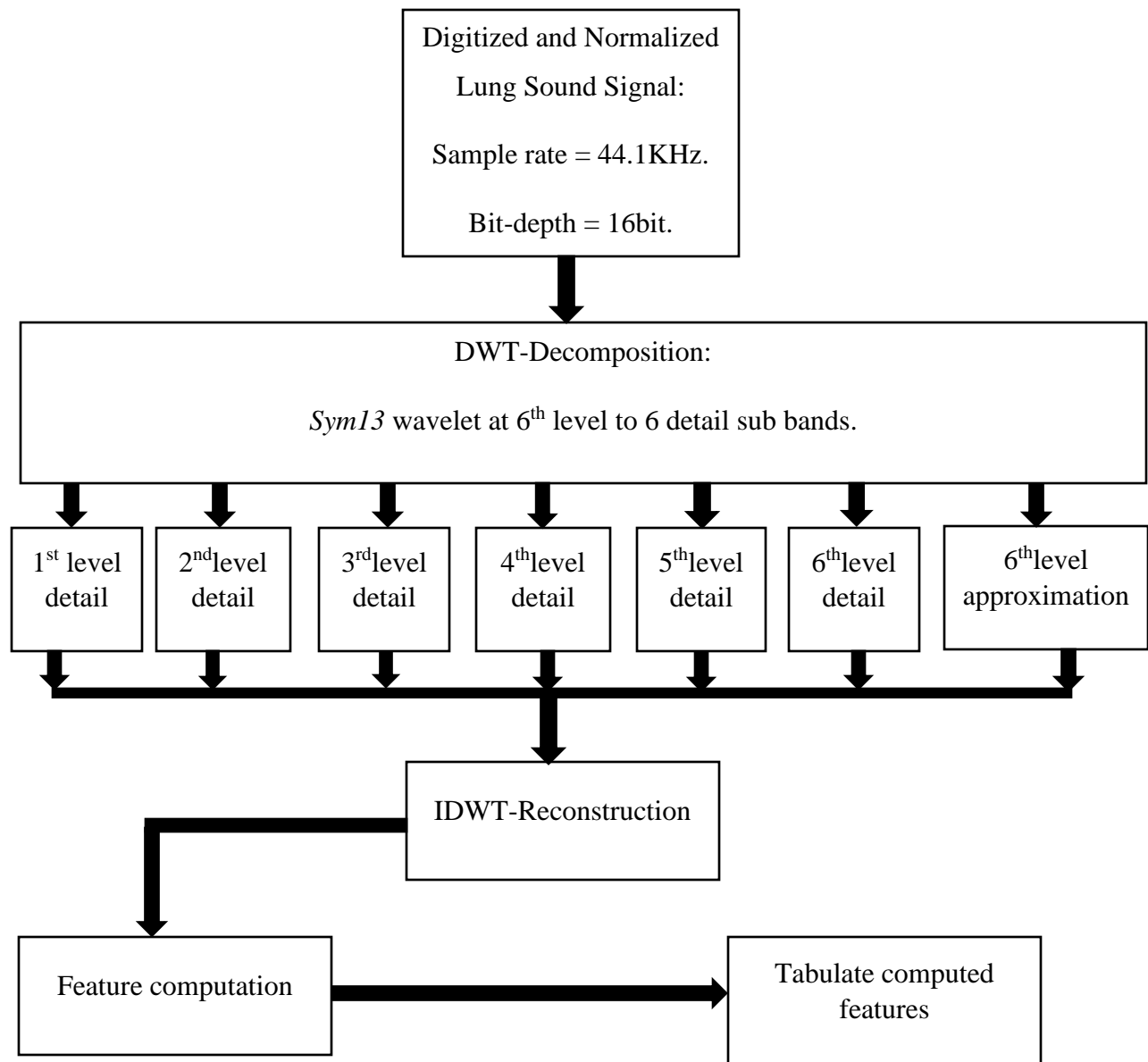
Figure 3.7: A procedure for DWT-based feature extraction.

Following the procedure presented in Figure 3.7, different features were extracted from the lung sound signals. These features extracted for the classification of lung diseases are clearly discussed as follows.

Mean: Mean is one of the basic statistical time domain features of a signal. Mean value is the average value of a signal. It can be calculated by summing all the samples and dividing by the number of samples (N). Mathematically, the arithmetic mean value of a signal can be calculated using Equation (3.7).

$$X_{mean} = \frac{\sum_{i=1}^{N} X_i}{N} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.7)$$

Standard deviation ($\boldsymbol{\sigma}$): Standard deviation is the amount of variation in the set of data values from the mean value. It is a statistic which measures the dispersion of a dataset relative to its mean value. Mathematically, σ can be calculated as the square of the variance, given in Equation (3.8). If the data points are further from the mean, there is the higher deviation within the dataset; hence, the more spread out the data, the higher the σ.

$$\sigma = \frac{\sqrt{(\sum(X_i - X_{mean})^2)}}{N} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.8)$$

Root mean square (RMS): RMS value of a signal can be calculated by squaring each value and finding the arithmetic mean of those squared values, then taking the square root of the result. Simply, RMS represents the average power of a signal, mathematically given in Equation (3.9).

$$X_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} X_i^2} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.9)$$

Shape factor: Shape factor is another important basic statistical time domain features of a signal. It is dependent on the shape of the signal while being independent of the signal dimensions. Mathematically, shape factor can be calculated by dividing RMS by the mean of the absolute value of a signal, given in Equation (3.10).

$$X_{SF} = \frac{X_{RMS}}{\frac{1}{N}\sum_{i=1}^{N} |X_i|} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.10)$$

Kurtosis: Kurtosis is one of the higher order statistical time domain features of a signal. It is the length of a signal distribution, or equivalently it is to mean how outlier prone the signal is. The positive value of kurtosis shows that the signal is peaked and the negative value of kurtosis shows the signal is flat. Mathematically, kurtosis can be determined using Equation (3.11).

$$X_{kurt} = \frac{\frac{1}{N}\sum_{i=1}^{N}(X_i - X_{mean})^4}{(\frac{1}{N}\sum_{i=1}^{N}(X_i - X_{mean})^2)^2} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (3.11)$$

Skewness: Skewness is another important higher order statistical time domain feature of the signal. It is the asymmetry of a signal distribution. Meaning, it is the measure of the asymmetry of the data. Mathematically, skewness can be calculated using Equation (3.12).

$$X_{skew} = \frac{\frac{1}{N}\sum_{i=1}^{N}(X_i - X_{mean})^3}{(\frac{1}{N}\sum_{i=1}^{N}(X_i - X_{mean})^2)^{3/2}} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . . (3.12)$$

Peak value: Peak value is among the extracted impulsive metrics statistical feature. It is the maximum absolute value of the input signal given as $X_p = \max |X_i|$.

Crest factor: Crest factor is another extracted impulsive metrics feature. It is the ratio of the maximum absolute value to the RMS value of the input signal. Therefore, mathematically crest factor can be calculated by dividing the peak value by the signal RMS, given in Equation (3.13).

$$X_{crest} = \frac{X_p}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}X_i^2}} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (3.13)$$

Impulse factor: Impulse factor is the ratio of the maximum absolute value to the mean of the absolute value of the input signal. Meaning, it is a measure of comparison between the height of a peak to the mean level of the signal. Mathematically, impulse factor can be calculated using Equation (3.14).

$$X_{IF} = \frac{X_P}{\frac{1}{N}\sum_{i=1}^{N}|X_i|} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.14)$$

Clearance factor: It is the ratio of the maximum absolute value to the squared mean of the absolute value's square root. Therefore, clearance factor can be calculated as the peak value divided by the squared mean value of the square roots of the absolute value amplitudes, given in Equation (3.15).

$$X_{clear} = \frac{X_p}{\left(\frac{1}{N}\Sigma_1^N \sqrt{|X_i|}\right)^2} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.15)$$

Signal to noise ratio (SNR): SNR is the ratio of signal power to noise power, and its unit of measurement is typically decibels (dB). It is a measurement parameter that compares the level of the desired signal to the level of background noise.

Total harmonic distortion (THD): THD is the measure of the harmonic distortion present in the signal. It is the ratio of total harmonic component power to fundamental power. Meaning, it is the ratio of the sum of all the powers of harmonic components to the power of fundamental frequency, given in Equation (3.16).

$$average\ frequency\ (f_{avg}) = \frac{\Sigma_{i-1}^n f_i X p_i}{\Sigma_{i=1}^n p_i} \dots \dots \dots \dots \dots \dots \dots \dots . (3.16)$$

Where p is power spectral density and f is frequency vector.

Signal to noise and distortion ratio (SINAD): SINAD is the ratio of total signal power to total noise plus the distortion power. In addition, other frequency domain/spectral features were also extracted from the lung sound signals of the different subjects for the classification of most common lung diseases.

Peak amplitude: The peak amplitude of a sinusoidal waveform signal is the maximum or negative deviation of a waveform from its zero-reference level. For a symmetrical sinusoidal waveform, the positive peak value is the same as the negative peak value. Generally, peak amplitude is the spectral feature generated based on the amplitude of the peaks. It is the maximum amplitude value of the signal.

Peak frequency: Peak frequency (f_p) is the maximum frequency value of the signal within the selected spectral region. f_p is commonly estimated from the spectra found from sampled data. It is the spectral peak generated based on the frequency of the peaks.

Band power: Band power is defined as the area under the spectrum curve within the selected band limits. It represents the power of the signal in the chosen frequency band. After pertinent feature extractions, the next step was feature selection to select the most discriminative features used for classification of data.

3.7. Feature Selection

Feature selection is the process of reducing input variables or features used for training the given machine learning model. It is the process of selecting and retaining only the most important features used for classification. Feature selection is highly important because it reduces the computational cost and makes the model time efficient. It reduces the training time, avoids overfitting, and improves accuracy of the model. The two main types of feature selection methods used in machine learning are supervised and unsupervised methods. Since we used the supervised machine learning models for classification of lung diseases, we focused on the supervised feature selection methods. There are various supervised feature selection methods which can be grouped into three main categories such as wrapper, filter, and embedded methods [85]. The wrapper methods include forward selection, backward selection, and stepwise selection algorithms. These methods used the model performance to select the best features for classification. The two main disadvantages of wrapper methods are increasing the overfitting risk and are computationally expensive. Wrapper methods are computationally more expensive compared to filter and embedded methods due to the repeated learning steps. For example, in case of sequential forward selection algorithm, it is a greedy search algorithm to get the optimal feature subset by interactively selecting the features based on the performance of the model. In doing so, we have to train the model for each feature subset combination until the optimal features are found. Therefore, this approach is much more computationally expensive than filter and embedded approaches. In filter methods, the features are selected or filtered based on the general characteristics of dataset like the correlation with the dependent variable to predict.

These methods include ANOVA, Pearson correlation, Chi-square test, information gain, and variance thresholding. Filter methods are faster, better approach, and avoids overfitting during classification. Embedded methods have been recently proposed by combing the advantages of wrapper and filter methods. These methods perform the feature selection in the learning or the model building phase. It includes the regularization methods such as LASSO, Ridge regression, Elastic net, and decision trees. Embedded methods are less computationally expensive than wrapper methods and less prone to overfitting. Table 3.2 shows a short comparison between the above-mentioned feature selection methods.

Table 3.2: Comparison among the three feature selection methods [85].

Criteria	Feature selection methods		
	Wrapper method	Filter method	Embedded method
Selection mechanism	It uses predictive model.	It uses poxy measure.	Feature selection is embedded in the model building phase.
Speed	Slower.	Computationally faster.	Medium.
Overfitting	Prone to overfitting.	Avoids overfitting.	Less prone to overfitting.

From Table 3.2, we can see that filter methods are fast, and effective methods to select relevant features. These methods are also helpful to avoid overfitting that will happen during final data classification. In this study, we select and applied one of the statistical filter feature selection methods i.e., one-way ANOVA. We used Figure 3.8 for the selection of ANOVA method among the other filter methods. Figure 3.8 is an overview sketch used for selection of filter methods in machine learning.
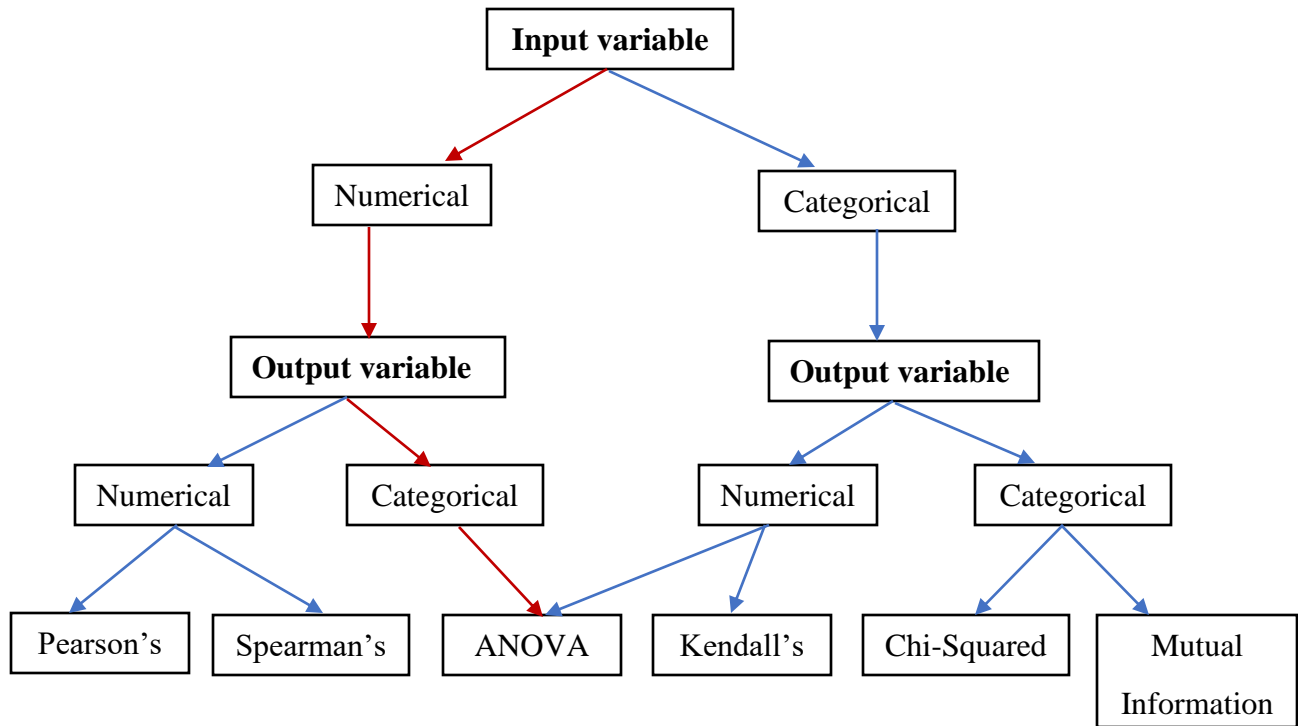
Figure 3.8: An overview sketch for selection of filter methods [86].

From Figure 3.8, we can see that for numerical input and categorical output classification issues, ANOVA is the preferred method [86]. Therefore, we used one-way ANOVA for selecting the most relevant features for final classification of lung diseases. Feature selection using one-way ANOVA is the process of ordering or ranking the extracted features by the value of some scoring function, $S(f_i)$. It usually measures the relevance of features [86]. The scoring function ($S(f_i)$) is computed from the training data, measuring some criteria of feature f_i. Selection of relevant features using feature or variable ranking is the simplest method. It is simply to select the k highest ranked feature according to S. Generally, one-way ANOVA feature selection method is simple, computationally faster, and avoids overfitting during model training and final data classification. Therefore, in this research, we implemented the one-way ANOVA feature ranking method using $minmax$ normalization scheme, given in Equation (3.17). The selected features have the discriminative power to discriminate the lung sound signals of COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma and healthy subjects.

$$F_{new} = \frac{F - \min(F)}{\max(F) - \min(F)} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.17).$$

3.8. Feature Visualization and Normalization

After ranking the features and selecting the most relevant features, we noticed that the features were not on the same scale. This will bias the classifier during data training and classification [87]. It is always better to make the features at the same scale. Therefore, after visualizing the features, we normalized the features to have the same scale before training the classifier. Next data splitting was done.

3.9. Data Splitting

Data splitting is the step to prepare the data for training and testing. In data science, the dataset is usually split into three parts such as training, validation and testing sets. The training set is the dataset in which the model is trained on while the validation dataset is used to validate the model. The testing set is the one used to see or evaluate how well the model performs on unseen data. There are two most commonly used validation methods in machine learning used for data splitting. These are hold-out validation and cross-validation methods. The hold-out validation method split the dataset into train and test sets. It is recommended for large datasets. The hold-out validation method is dependent on just one train-test split. This makes the hold-out validation method score dependent on how the dataset is split into train and test sets. On the other hand, cross-validation method is dependent on multiple train-test splits. It is usually the preferred method, because it gives our model the opportunity to train on multiple train-test splits. This gives a better indication of how well our model will perform on unseen or new data. Additionally, the cross-validation method protects against overfitting by partitioning the data into folds and estimating accuracy on each fold. Moreover, cross-validation or $'k'$ fold cross-validation is when the dataset is randomly split into multiple $'k'$ groups. Hence, one of the groups is used as the test set and the remaining are used as the training set. Then the model is trained on the training set and scored on the test set, the process is repeated until each unique group used as the test set. In this research, first the dataset has been split using the hold-out method. Then we applied an 8-fold cross-validation, shown in Figure 3.9.

Split 1	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 1
Split 2	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 2
Split 3	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 3
Split 4	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 4
Split 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 5
Split 6	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 6
Split 7	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 7
Split 8	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Metric 8

Training data　　**Test data**

Figure 3.9: 8-fold cross-validation.

The $'k'$ fold cross-validation (8-fold cross-validation in our case), shown in Figure 3.9 is a statistical method of evaluating and comparing different learning models by dividing data into two segments which are the training and validation segments. Then the training and validation sets should cross-over in successive rounds so that each data point has a chance of being validated against 8 separate times.

3.10. Model Training and Evaluation

After a successful completion of all the above tasks (data acquisition, data structuring, data pre-processing, feature extraction, feature ranking or selection, feature visualization and normalization, and data splitting), the final step is model training for data classification i.e., classification of most common lung diseases using features extracted from the lung sounds. This is a very important stage to interactively train, validate and tune different machine learning classification models. In this study, we select and trained four important machine learning models for final classification of most common lung diseases. Selection of the best classifier types depends on the data being processed. For selection of our models, we took into account the characteristics of different supervised machine learning algorithms, shown in Table 3.3.

Table 3.3: Characteristics of different supervised machine learning algorithms [88].

Classifier	Prediction speed	Memory usage	Interpretability
Decision Trees	Fast	Small	Easy
Discriminant Analysis	Fast	Small for linear and large for quadratic	Easy
Logistic Regression	Fast	Medium	Easy
Support Vector Machines	Medium for linear and slow for others	Medium for linear, medium for multiclass and large for binary.	Easy for linear SVM and hard for all other kernel types
Nearest Neighbor Classifiers	Slow for cubic and medium for others	Medium	Hard
Ensemble Classifiers	Fast to medium depending on the choice of algorithm	Low to high depending on the choice of the algorithm	Hard
Naïve Bayes Classifiers	Medium for simple distributions and slow for kernel distributions or high dimensional data	Small for simple distributions and medium for kernel distributions or high dimensional data	Easy

Table 3.3 shows the typical characteristics of different supervised machine learning algorithms and we used it as a guide for selection of our models. After choosing the best classification models, we trained them to classify the lung diseases as COPD, URTI, LRTI pneumonia, bronchiectasis, bronchiolitis, asthma, or as healthy. In this research, due to their best characteristics presented in Table 3.3, we selected and trained four important machine learning models including Naive Bayes, K-nearest neighbor (KNN), Ensemble and support vector machine (SVM) models for classification of most common lung diseases. Naive Bayes models or classifiers are a group of classification algorithms based on the Bayes' theorem. Meaning, it is not a single algorithm but a family of algorithms which considers that each attribute in unclassified tuple X is conditionally independent [89].

Naive Bayes algorithms classify the data in two steps. In the first step, using the training data, the method estimates probability distribution of the parameters by assuming each predictor are conditionally independent given the class. Finally, in the prediction step for any unseen data, the method computes the posterior probability distribution of that sample belonging to each class.

$$P(C_1|X) = \prod_1^n P(X_i \,|\, C_1)P(C_1) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.18)$$

Where: $P(C_1|X)$ is the probability of tuple X belongs to class 1, $P(C_1)$ is the probability of class 1 that exists in the training set and $\prod_1^n P(X_i \,|\, C_1)$ is the production of each attribute in tuple X for each labeled class, and the tuple will be classified to the class with the maximum probability [89].

KNN classifiers classify the object based on the distance between the new object and the defined objects. The new object is assigned to the class K that has the shortest distance to class K which defined as the nearest neighbor [89]. Ensemble classifiers mix results from many weak learners into one high-quality ensemble model. SVM algorithm represents the training data as points in a flat separated space by an apparent gap. Then the new objects are mapped into space with the forecast category based on which side of the gap they fall [89]. It classifies the data by finding the best hyperplane which separates data points of one class from those of the other classes. Finally, after we trained different machine learning models, we choose the best model with higher classification accuracy. In the first run result, we found SVM particularly Fine Gaussian SVM is the best model with higher classification accuracy.

Therefore, due to its higher classification accuracy, a Fine Gaussian SVM algorithm has been selected for final classification of data. Next, we assessed the classification performance of the selected algorithm (Fine Gaussian SVM algorithm) using confusion matrices. Confusion matrix is a table which helps us to visualize the performance of a supervised machine learning algorithm. It helps us to identify the areas where the classification algorithm has performed poorly. It holds the number of observations for each class, true positive rates (TPR), false negative rates (FNR), positive predictive values (PPV), and false discovery rates (FDR). In each of the confusion matrix presented in the results section, each of the rows show the instances in the true class while each of the columns show the instances in the predicted class. A brief summary about the SVM model is presented below.

3.10.1. Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm used for solving classification and regression issues. Today, it is widely used in different research zones such as speech recognition, face recognition, medicinal conclusion, and so on. SVM algorithm represents the training data as points in a flat separated space by an apparent gap or a separating hyperplane. Then the new objects are mapped into space with the forecast category based on which side of the gap they fall. However, as shown in Figure 3.10 (a), there could be an infinite number of separating hyperplanes. SVM solves this problem using the concept of the smallest distance between any of the data samples and the decision boundary. Meaning, as shown in Figure 3.10 (b), the best hyperplane which corresponds to the one giving the largest margin between the classes is selected. Therefore, SVM classifies the data by finding the best hyperplane which separates data points of one class from those of the other class [90]. The best hyperplane for an SVM means the one with the largest margin between the two classes. Margin is the maximal width of the slab parallel to the hyperplane that has no interior support vectors. Support vectors are the data points which are closest to the separating hyperplane. These points are found on the boundary of the slab, shown in Figure 3.10. Figure 3.10 illustrates an overview sketch of SVM algorithm.
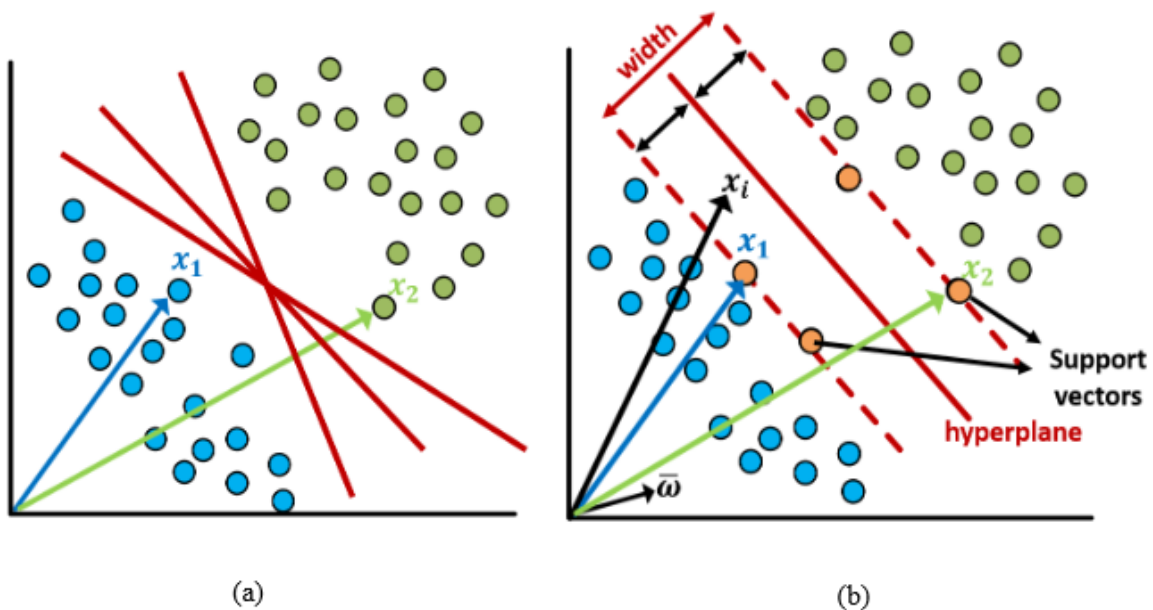
Figure 3.10: An overview sketch of SVM algorithm [90].

Support vector machines (SVM) algorithm was originally designed for binary case classification issues. However, real world problems need classification of more than two categories. As a result, multiclass SVM become widely used in a wide range of applications such as speech recognition, optical character recognition, and in the area of bioinformatics. Multiclass SVM aims to designate labels to instances by using support vectors, where the labels are drawn from a finite set of several elements. The classification learner uses $fitcsvm$ function to train the classifier for binary case classification, where there are exactly two classes. When there are more than two classes (multiclass classification case), the classification learner uses $fitcecoc$ function to reduce the multiclass classification problem to a set of binary classification subproblems with one SVM learner for each subproblem. Figure 3.11 shows a multi-class SVM method which can be used for classification of more than two categories.
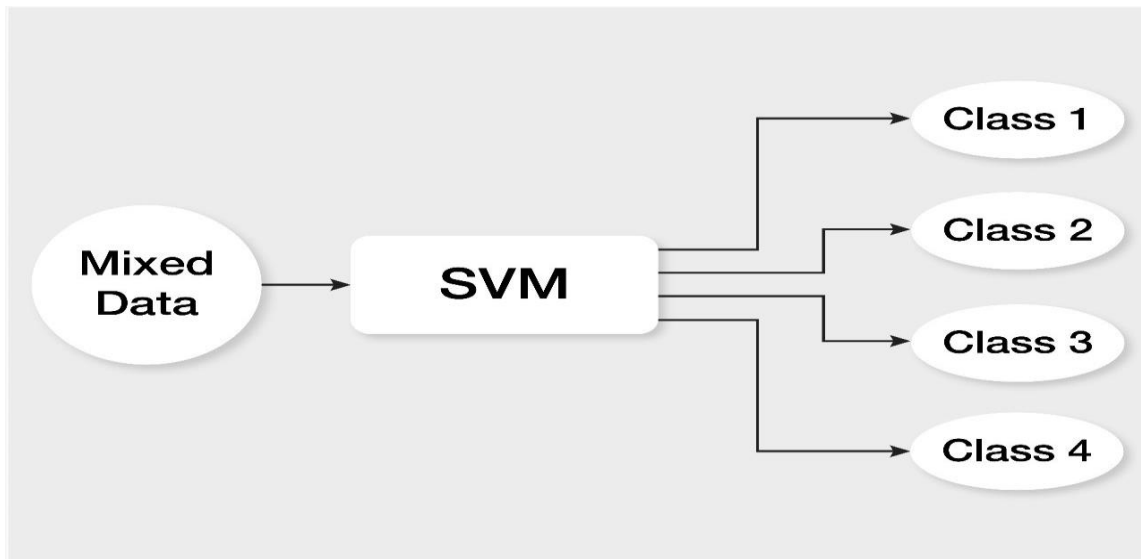
Figure 3.11: Multi-class SVM method.

Generally, the basic approach for using binary classification algorithms for multi-classification issues is to split the multi-class classification dataset into multiple binary classification datasets and fit a binary classification model on each. There are two important heuristic approaches such as One-vs-All (OVA) and One-vs-One (OVO) strategies, used to split a multiclass classification problem into multiple binary classification datasets and train a binary classification model. The One-vs-All strategy also called One-vs-Rest is a heuristic method used for using binary classification algorithms for multi-class classification.

It splits the multi-class classification into one binary classification problem per class. Meaning, One-vs-All method splits a multi-class dataset into multiple binary classification problems. Then a binary classifier is trained on each binary classification problem so that predictions are made using the most confident model. One-vs-One is another heuristic approach used for using binary classification algorithms for multi-class classification. It splits the multi-class classification into one binary classification problem per each pair of classes. Meaning, like OVA, OVO splits a multi-class classification dataset into binary classification problems. Unlike OVA which splits a multi-classification dataset into one binary for each class, OVO approach splits a multi-classification dataset into one dataset for each class versus every other class. So, it trains one learner for each pair of classes to distinguish one class from every other class. Figure 3.12 illustrates the SVM principle and of the One-vs-One (OVO) multiclass classification SVM for three classes. In this research, One-vs-One (OVO) multiclass classification SVM was applied to classify most common lung diseases into eight separate classes by following the same procedure shown in Figure 3.12.
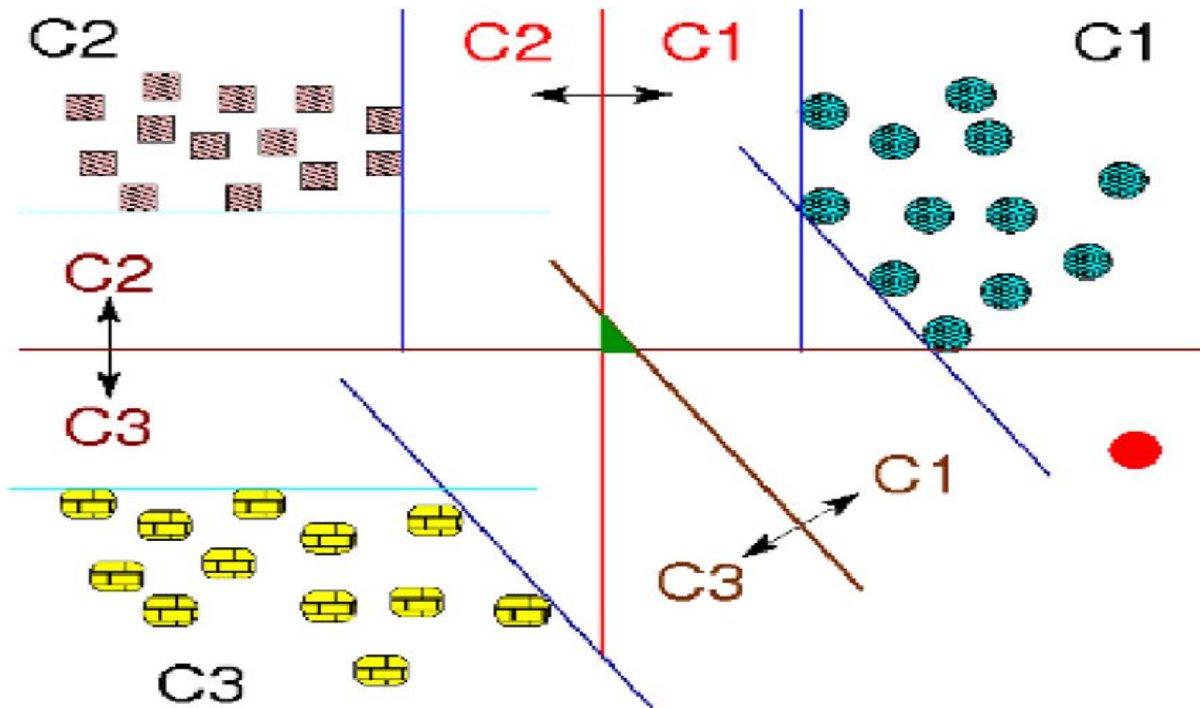
Figure 3.12: Illustration of One-vs-One (OVO) multiclass SVM for three classes.

3.11. Model optimization

The fundamental goal of machine learning is to create a model that performs well and gives higher classification accuracy in predicting a particular set of cases. To achieve a more accurate model, we need to apply machine learning optimization techniques. Optimization by itself is the problem of finding a set of inputs to an objective function that results in a maximum or minimum function evaluation. It is the process of adjusting the hyperparameters in order to minimize the cost function by using one of the optimization techniques [91]. Hyperparameters can have a direct impact on training of machine learning models. Hence, to achieve a maximal performance, it is better to know how to optimize them.

Some of the common hyperparameter tunning strategies used in machine learning includes grid search, random search, and Bayesian optimization [91]. Grid search method is the most basic hyperparameter tunning method. In this method, a model will be built for each possible combination of the hyperparameter values provided. Then evaluating each model, and the architecture which produces the best results will be selected. Random search builds up a grid of hyperparameter values and selects random combinations to train the model. This method allows us to explicitly control the number of parameter combinations. However, tunning by means of using these common hyperparameter tunning techniques can be a time-consuming challenge especially with large parameters spaces.

Hyperparameters can also be tunned manually. Traditionally, hyperparameters were optimized or tuned manually by means of trial-and-error method. However, there is a continual improvement for faster, better, and more automatic methods to optimize hyperparameters. Bayesian optimization is another most important optimization technique used to optimize hyperparameters. There are different ways to perform Bayesian optimization technique including *fitcauto* and *fitrauto*, classification learner and regression learner apps, fit function, and *bayespot*. After we choose a particular type of model to train, for example a decision tree or a support vector machine (SVM), we can tune the model by selecting different advanced options. For instance, we can change the maximum number of splits for a decision tree or the box constraint of an SVM. Instead of manually selecting these options, we can use hyperparameter optimization within the classification learner app to automate the selection of hyperparameter values.

_____

For a given model type, the app tries different combinations of hyperparameter values using an optimization scheme that seeks to minimize the model classification error, and returns a model with the optimized hyperparameters. In this study, we used classification learner app to automatically build optimized model using Bayesian optimization scheme. Therefore, in this study, we used Bayesian optimization technique to optimize our model for improving accuracy of the selected model. At last, several performance metrics such as TPR, FNR, TNR, and FPR were measured for determining accuracy, sensitivity, and specificity of the final model. Therefore, by using the performance metrics obtained from the confusion matrices, the accuracy, sensitivity, and specificity of the final model were calculated using Equations 3.19, 3.20, and 3.21.

Sensitivity is the performance metric which evaluates a model's ability to predict true positives of each available category. The equation for calculating sensitivity of a model is given in Equation (3.19).

$$Sensitivity = \frac{TP}{TP + FN} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.19)$$

Specificity is the performance metric which evaluates a model's ability to predict true negatives of each available category. It determines a model's ability to predict if an observation does not belong to a specific category. The equation for calculating specificity of a model is given in Equation (3.20).

$$Specificity = \frac{TN}{TN + FP} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.20)$$

Accuracy is another performance metric that evaluates a model's ability to correctly differentiate observations from different category. It is defined as the percentage of correct predictions for the test data. Accuracy can be calculated by dividing the number of correct predictions by the number of total predictions, mathematically given in Equation (3.21).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3.21)$$

Where, true positives (*TP*) are when a model predicts an observation belongs to a class and it actually belong to that class.

True negatives (*TN*) are when a model predicts an observation doesn't belong to a class and it actually doesn't belong to that class. False positives (*FP*) are when a model predicts an observation belongs to a class when in reality it doesn't belong to that class. False negatives (*FN*) are when a model predicts an observation doesn't belong to a class when in reality it does belong to that class.

3.12. Materials

For a successful completion of this research, different software and hardware materials have been used. The software materials were used for solving, and performing mathematical calculations, managing codes, files, and data, processing and analysis of signals. Two important software materials such as MATLAB (an abbreviation of "matrix laboratory") and audacity software were used in this study. We used MATLAB R2019b software to perform the above-mentioned tasks. Additionally, we used the latest version of audacity software (audacity 2.4.2). Audacity is a free and open-source digital multi-track audio recorder and editor application software used to display, edit, and amplify lung sound records. It is also used to annotate the lung sound records or label sound tracks to identify specific events. Moreover, the hardware materials (traditional stethoscope head, medical grade tubing, condenser microphone, audio cable and computer installed with audacity software) were used for the construction of our electronic stethoscope.

CHAPTER FOUR

RESULTS and DISCUSSION

4.1. Construction of Electronic Stethoscope

First of all, we made an electronic stethoscope for lung sound acquisitions by modifying the existing conventional or traditional acoustic stethoscope which is readily available in our most health care institutions. As shown in Figure 4.1, the materials we used for the construction of our electronic stethoscope were head of traditional stethoscope, medical grade tubing, condenser microphone, audio cable and computer installed with audacity software.

Figure 4.1: Hardware materials used for the construction of electronic stethoscope.

The head of the normal stethoscope was placed over the patient chest wall as usual to pick up the lung sounds through its diaphragm and bell end. We used head of yuwell traditional stethoscope, which is locally and readily available in most health care organizations of Ethiopia, shown in Figure 4.2.

_____

Incorporating the head of traditional stethoscope in the design gave significant advantages such as patient comfort and most importantly to get standard and valuable sound. Figure 4.3 shows head of yuwell traditional stethoscope (diaphragm and bell ends) which we used for the construction of our electronic stethoscope.

Figure 4.2: Yuwell traditional acoustic stethoscope.

Figure 4.3: Head of yuwell traditional stethoscope; (a), the diaphragm end (b) bell end.

The medical grade tubing was used to create connection between head of the traditional stethoscope and the condenser microphone. In this research, a moisture proof PVC tube which can replace the stethoscope tubing was used to create connection between the stethoscope head and the microphone. Figure 4.4 shows one end of the medical grade tubing connected to the head of the stethoscope.

Figure 4.4: Medical grade tubing connected to head of stethoscope.

For the transduction purpose a reliable, cost-efficient, small and lightweight condenser microphone having reproducible frequency response has been selected and mounted onto the head of the stethoscope. A condenser microphone (SG electret microphone, shown in Figure 4.5) was selected and mounted onto the head of the stethoscope to convert the acoustic sound picked up by the stethoscope head into an equivalent electrical signal.
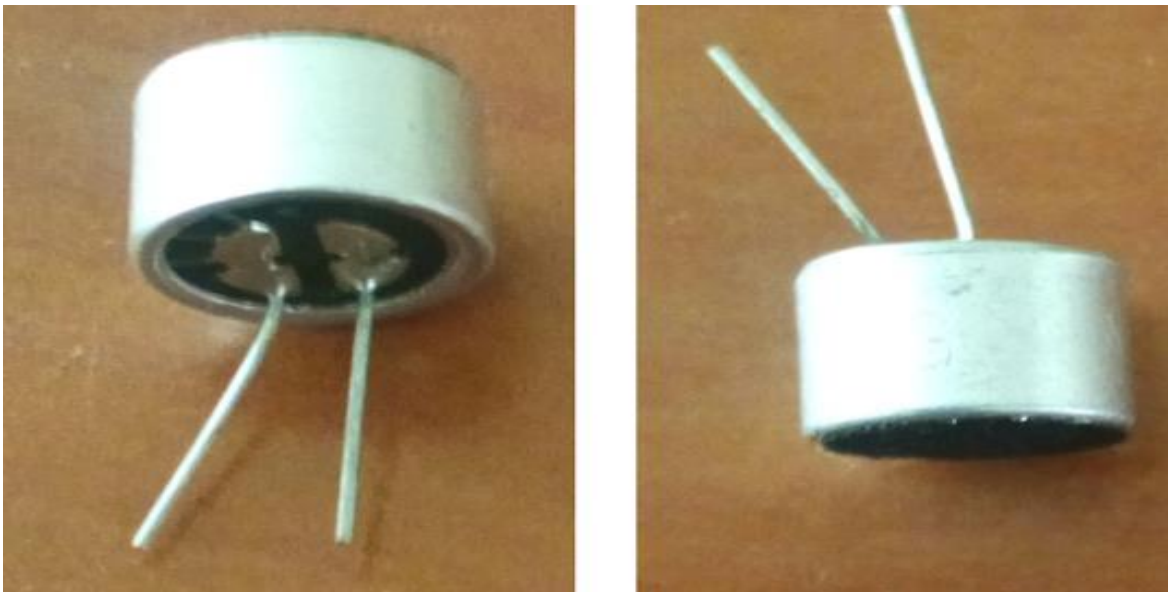
Figure 4.5: SG electret microphone.

_____

One end of the condenser microphone is mounted on the head of the stethoscope through the medical grade tubing and the other end is connected to computer through audio cable; so that it can easily interfaced with computer for further recording and processing of lung sounds. We used an audio cable shown in Figure 4.6 to connect one end of the condenser microphone to computer. Figure 4.7 shows a condenser microphone sensor connected to an audio cable.

Figure 4.6: Audio cable.
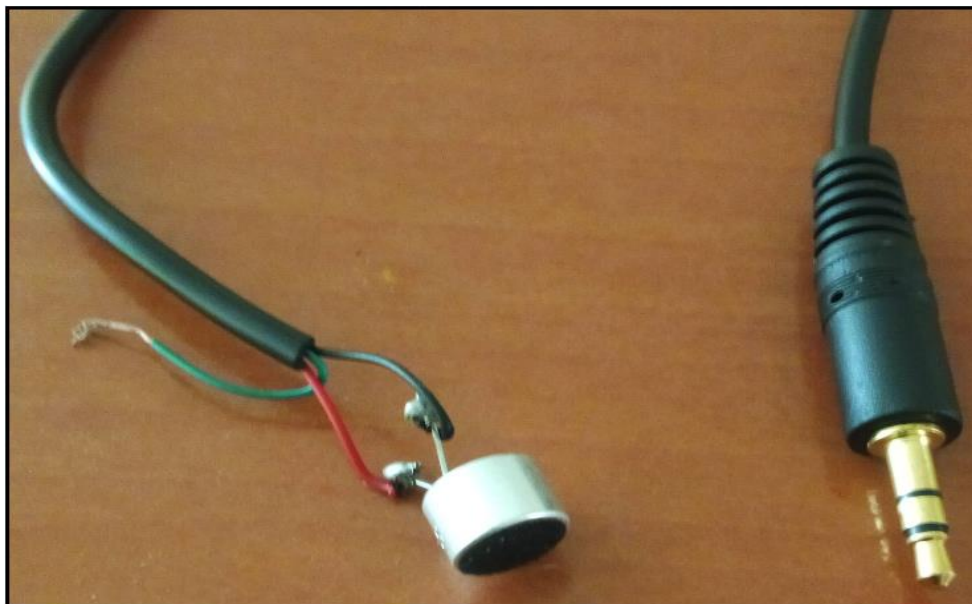
Figure 4.7: SG electret microphone connected to an audio cable.

Jimma University, Jimma Institute of Technology, School of BME, Bio-Instrumentation Stream

Besides to the above discussed components, an audacity software was used in the construction of our electronic stethoscope. Audacity is a free and open-source digital multi-track audio recorder and editor application software available for windows, macOS, GNU/Linux and other operating systems. Once we captured the lung sounds, audacity software gives us the power to display, edit, amplify, filter, changing playback rate, and show spectrograms of signals. Additionally, it is used to annotate the lung sound records or used to label sound tracks to identify specific events. Generally, as compared to the traditional stethoscope, our electronic stethoscope allows amplification of the acoustic signals, electronic removal of noise artifacts, and recording for playback onto computer for further processing of lung sound signals. Once, the lung sounds are acquired through this constructed electronic stethoscope, the next computerized processing of lung sounds held using different modern digital signal processing techniques.

Figure 4.8: SG electret microphone connected to head of traditional stethoscope and an audio cable.

Finally, we made our electronic stethoscope by doing certain modifications over the available traditional stethoscope; so that it will be best solution for limited resource health care settings in terms of simplicity, patient comfort, and cost effectiveness by solving the existing limitations of current traditional stethoscopes.

Audio cable   **System**   **Electret microphone**   **Stethoscope head**
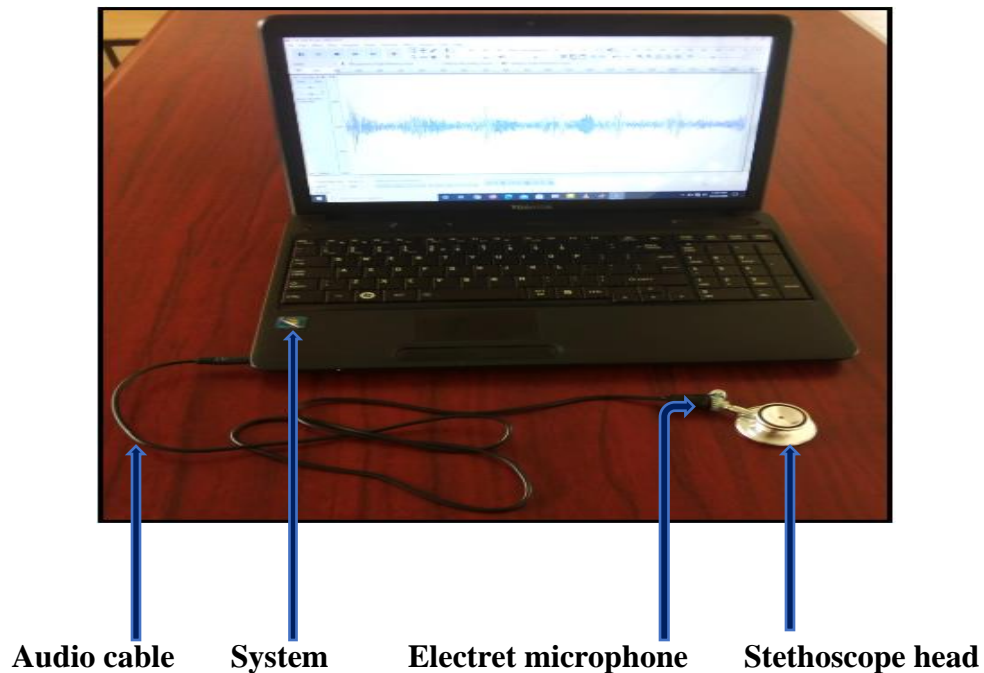
Figure 4.9: An electronic stethoscope for acquisition of lung sounds.

4.2. Signal Acquisition

The local lung sound collection was held at Jimma University Institute of Health, Jimma University Medical Center (JUMC), Department of Internal Medicine, Pulmonology Unit. The lung sounds were recorded sequentially from each of the subjects at sitting position, shown in Figure 4.10.

Figure 4.10: Lung sound recording from patients at JUMC, Department of Internal Medicine, Pulmonology unit; (top), COPD (bottom) bronchiectasis subjects.

Table 4.1 summarizes a number of lung sound records for each class used in this study. According to the data presented using Table 4.1, a total of 500 lung sound records where; 75 healthy, 70 asthma, 60 pneumonia, 60 URTI, 50 LRTI, 60 COPD, 65 bronchiectasis, and 60 bronchiolitis records were used in this study.

Table 4.1: Number of lung sound records for each class used in the study.

Pulmonary conditions	Number of lung sound records		Total
	Local	Online	
Asthma	60	10	70
Bronchiectasis	49	16	65
Pneumonia	23	37	60
URTI	15	45	60
Bronchiolitis	47	13	60
LRTI	48	2	50
Healthy	30	45	75
COPD	15	45	60
Total	**287**	**213**	**500**

4.3. Pre-processing

A discrete wavelet transform (DWT)-based denoising of the lung sound signals which involves the three important steps such as decomposition, detail coefficients thresholding, and reconstruction have been done. Signal decomposition is the first step in DWT-based denoising process. Therefore, the wavelet decomposition of the lung sound signals of healthy, COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, and asthma subjects are demonstrated in Figures 4.11 – 4.14.
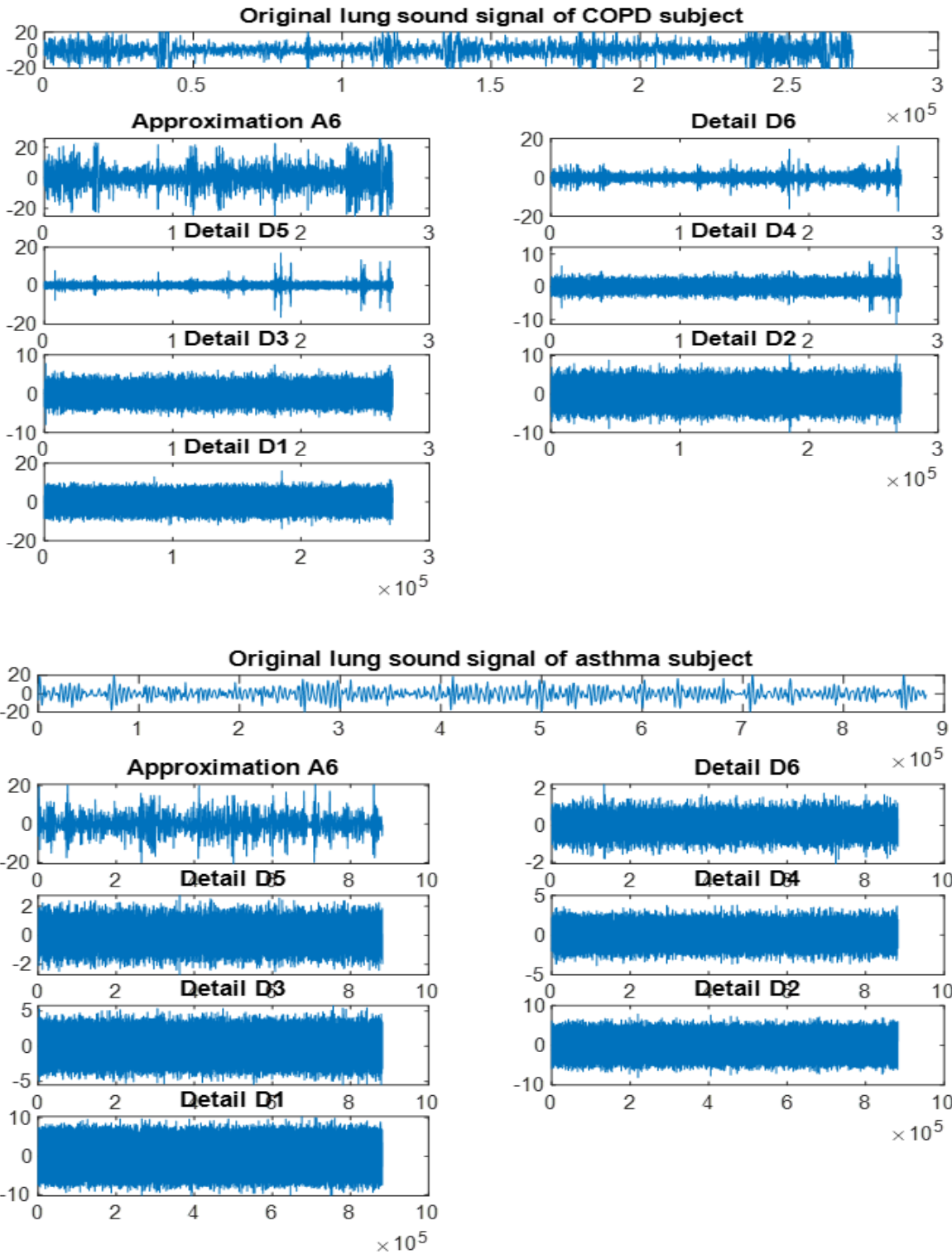
Figure 4.11: Wavelet decomposition of lung sound signals using *Sym13* at 6th level; (top), COPD (bottom) asthma subjects.
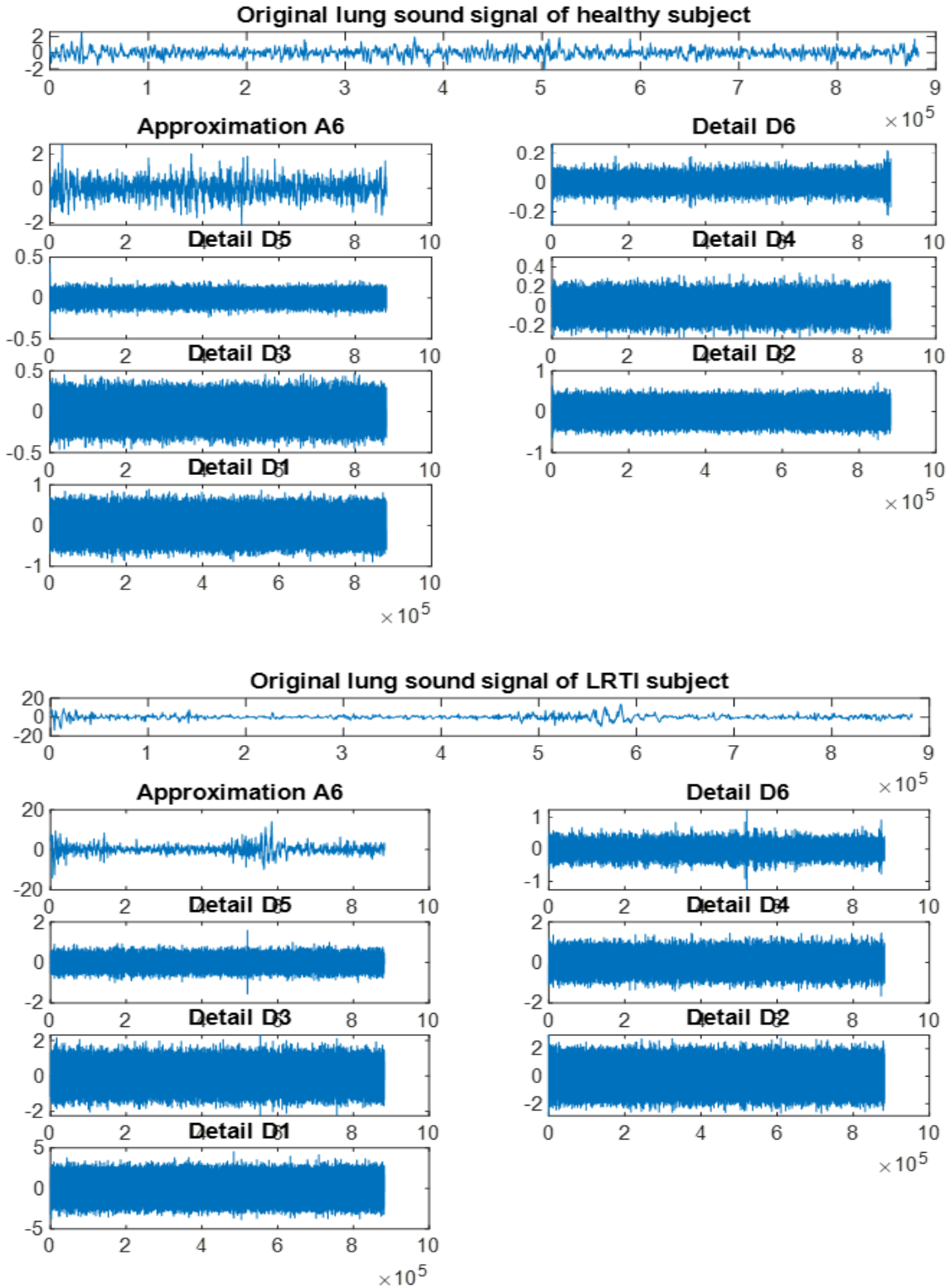
Figure 4.12: Wavelet decomposition of lung sound signals using *Sym13* at 6th level; (top), healthy (bottom) LRTI subjects.
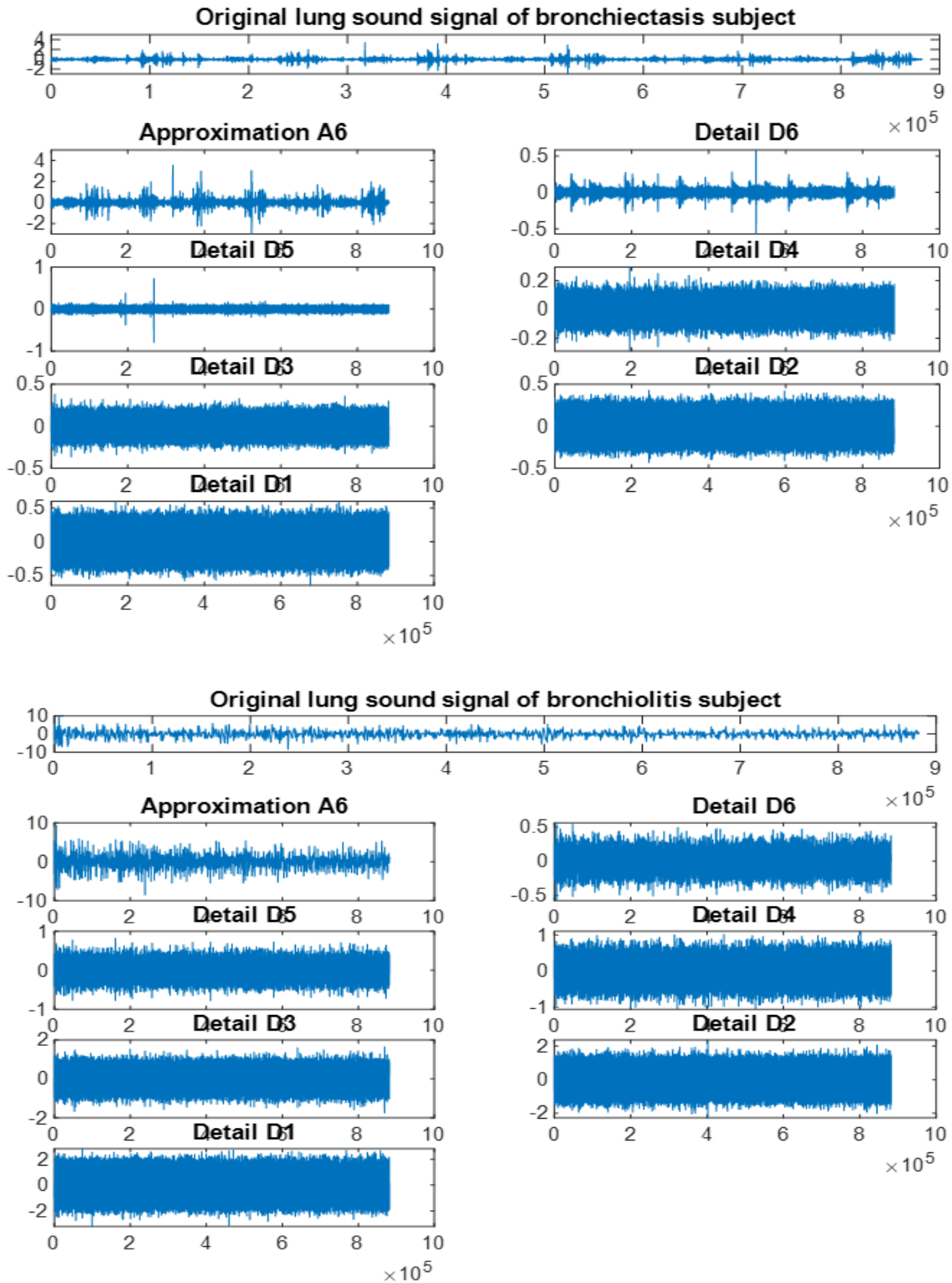
Figure 4.13: Wavelet decomposition of lung sound signals using *Sym13* at 6th level; (top), bronchiectasis (bottom) bronchiolitis subjects.
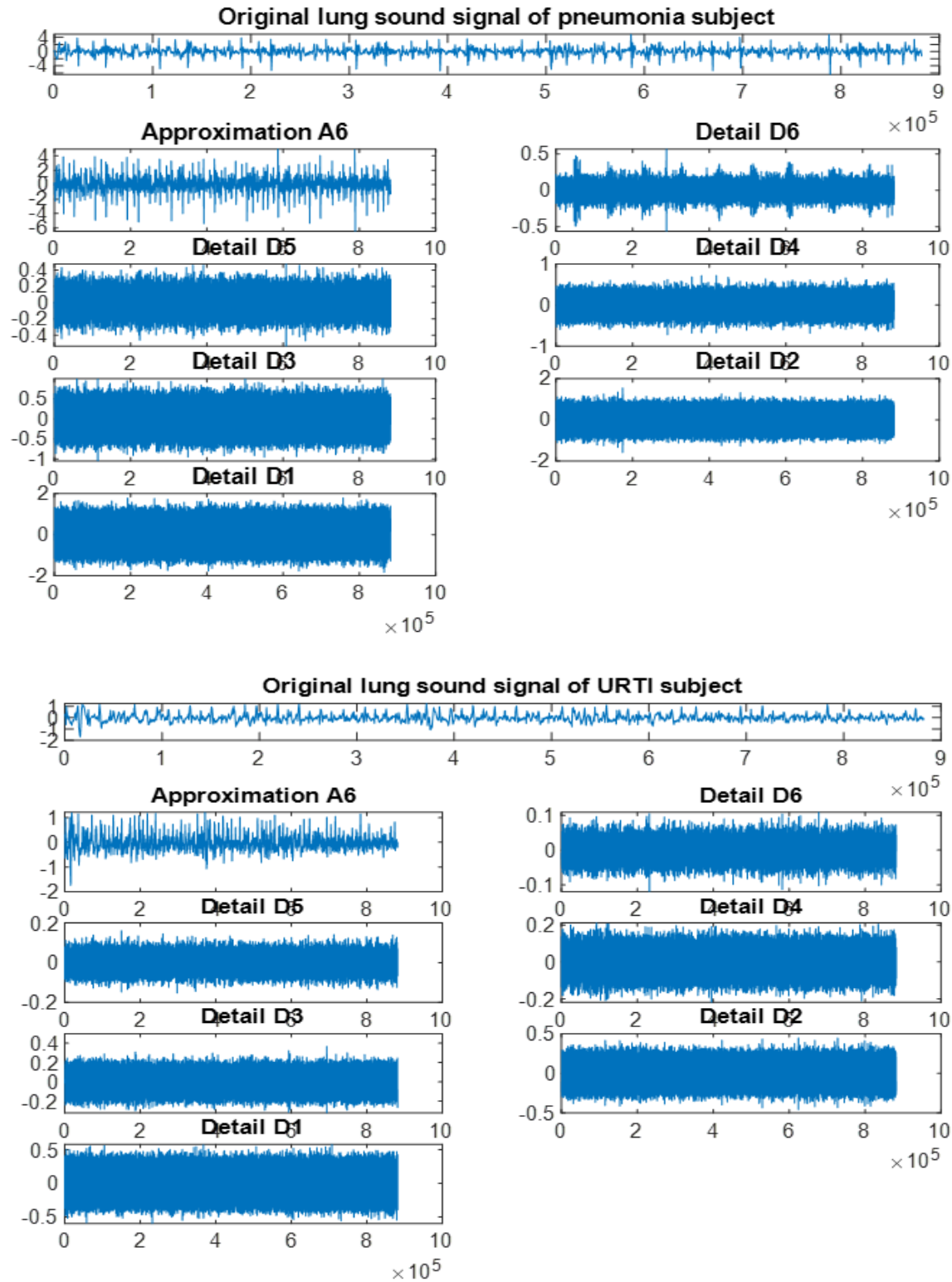
Figure 4.14: Wavelet decomposition of lung sound signals using *Sym13* at 6th level; (top), pneumonia (bottom) URTI subjects.

_____

Following the decomposition, the denoising was performed by applying *Sym13* wavelet function with soft thresholding method and *Sqtwolog*, threshold selection rule. Figures 4.15, 4.16, and 4.17 shows the effect of *Sym13* wavelet function on denoising of the lung sound signal of healthy and some diseased subjects such as COPD, bronchiectasis, bronchiolitis, pneumonia and URTI.
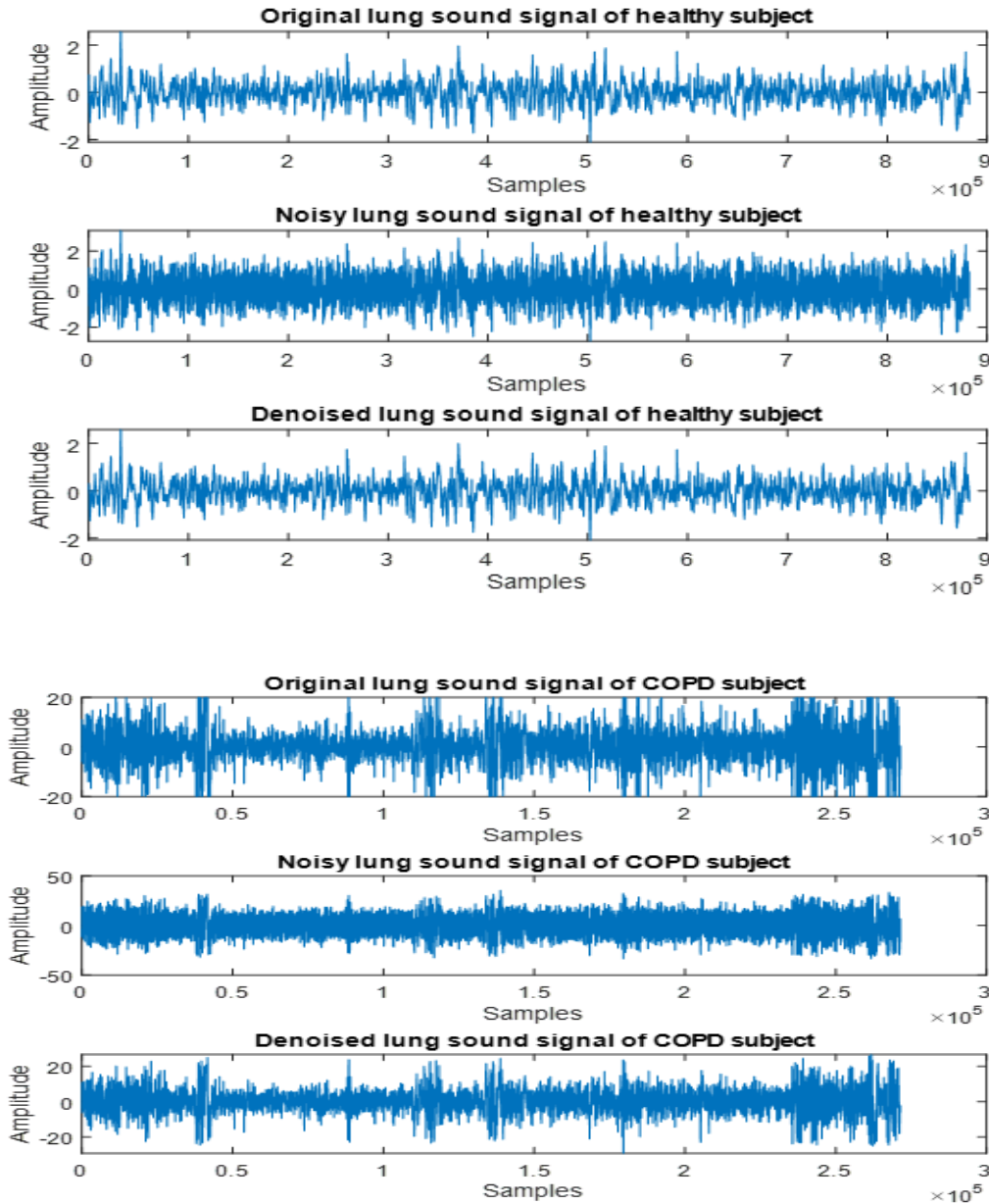
Figure 4.15: Denoising of lung sound signals using *Sym13* wavelet function at 6th level with soft thresholding method; (top), healthy (bottom) COPD subjects.
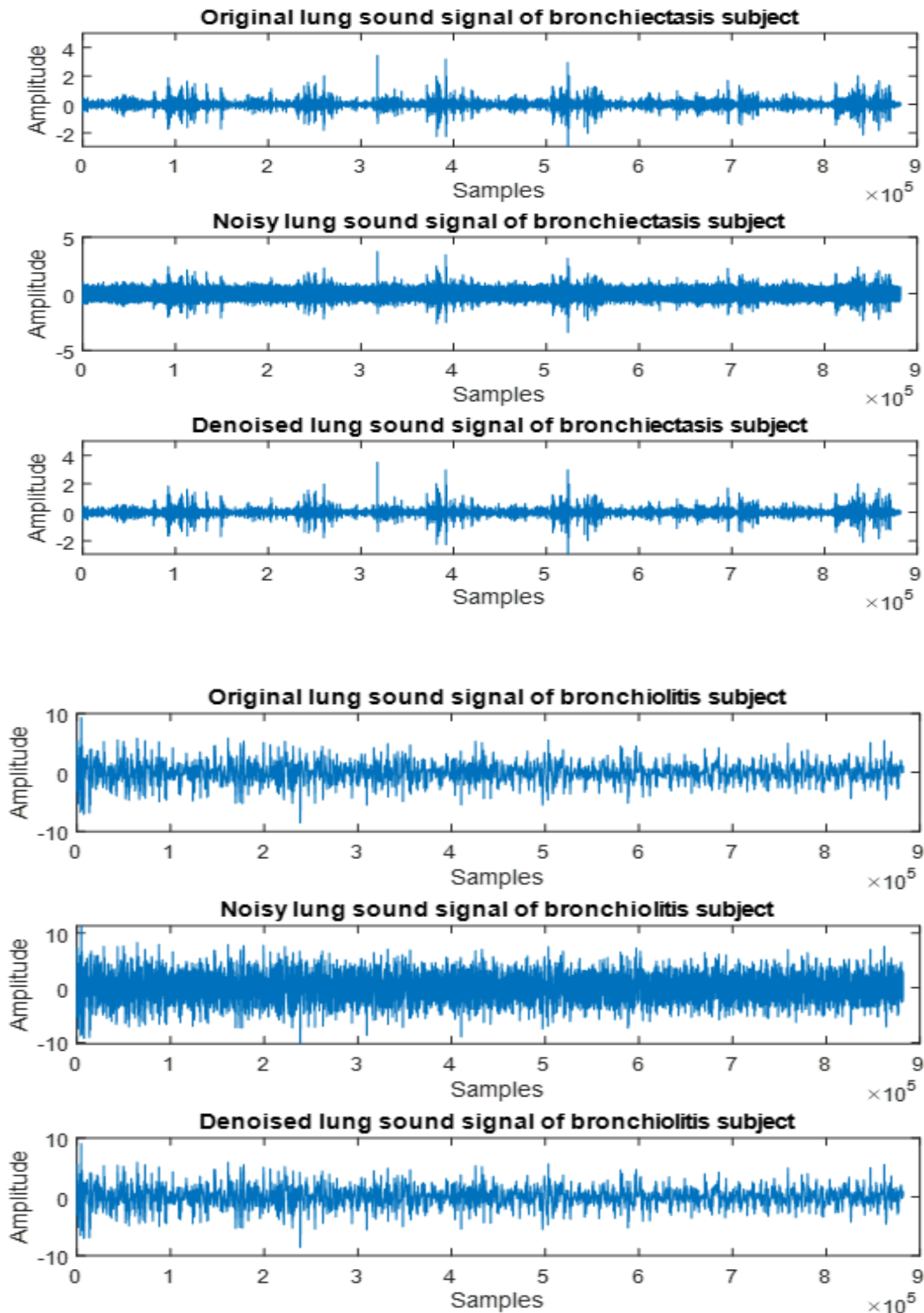
Figure 4.16: Denoising of lung sound signals using *Sym13* wavelet function at 6th level with soft thresholding method; (top), bronchiectasis (bottom) bronchiolitis subjects.
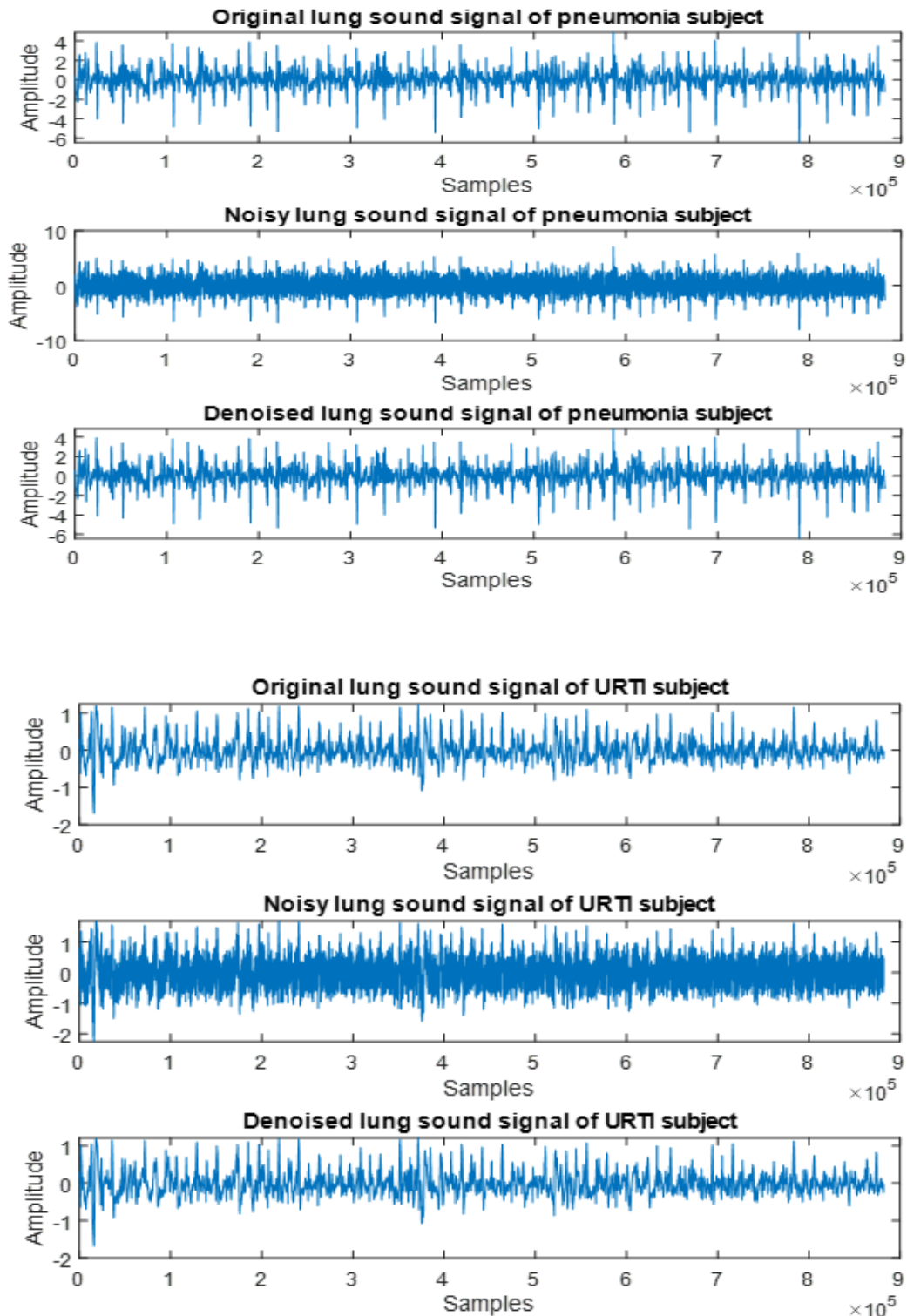
Figure 4.17: Denoising of lung sound signals using *Sym13* wavelet function at 6th level with soft thresholding method; (top), pneumonia (bottom) URTI subjects.

_____

Moreover, the effect of the two thresholding methods (soft and hard thresholding) were studied by computing the SNR results when denoising the lung sound signals using different wavelet functions. Table 4.2 shows the SNR results obtained when denoising the lung sound signals using *Db4, Db10, Sym5,* and *Sym13* wavelet functions at different level of decomposition with the two thresholding methods.

Table 4.2: SNR results obtained when denoising the lung sound signals using different wavelet functions at different level of decomposition with soft and hard thresholding methods (*Sqtwolog* threshold selection rule was applied).

Wavelet function	SNR (dB)									
	Level 3		Level 4		Level 5		Level 6		Level 7	
	Soft	Hard	Soft	Hard	Soft	Hard	Soft	Hard	Soft	Hard
Db4	16.7345	15.9934	18.3210	17.8921	20.2870	20.2946	20.3753	20.3253	20.014	19.8754
Db10	17.3240	16.6587	19.1032	18.9821	21.4465	21.4312	**22.3861**	**22.3758**	21.3201	20.986
Sym5	16.8341	16.0241	19.6565	18.7694	20.3309	20.4181	21.3416	21.3472	20.9788	19.8320
Sym13	17.8934	17.2301	19.6587	19.2789	21.5857	21.4880	**22.5011**	**22.4848**	21.6542	21.3057

According to Table 4.2, the SNR results for the different wavelet functions were calculated at different level of decomposition (3rd to 7th level of decomposition). Moreover, the results were found by applying *Sqtwolog* threshold selection rule with the two thresholding methods such as soft and hard thresholding methods. A total of 240 lung sound signals (30 for each class) for the initial round and 400 lung sound signals (50 for each class) for final test were used and the better SNR results were found still at six level of decomposition. During the selection of the wavelet function, the level of decomposition and the thresholding method are the important parameters to be taken into account.

This is due to the fact that the level of decomposition and the thresholding method could affect the performance or the SNR values of the wavelet function, shown in Table 4.2. Based on the result shown in Table 4.2, the highest SNR values were observed for the *Db10* and *Sym13* wavelet functions at 6th level of decomposition. The SNR values for the Db10 wavelet function were 22.3861 dB and 22.3758 dB with soft and hard thresholding methods respectively. For *Sym13* wavelet function, the SNR results were 22.5011dB and 22.4848 dB with soft and hard thresholding methods respectively. Furthermore, we used the following column chart to visually compare these SNR values of the different wavelet functions. Figure 4.18 shows a column chart used for comparing the SNR values obtained when using *Db4, Db10, Sym5,* and *Sym13* wavelet functions with soft and hard thresholding methods.
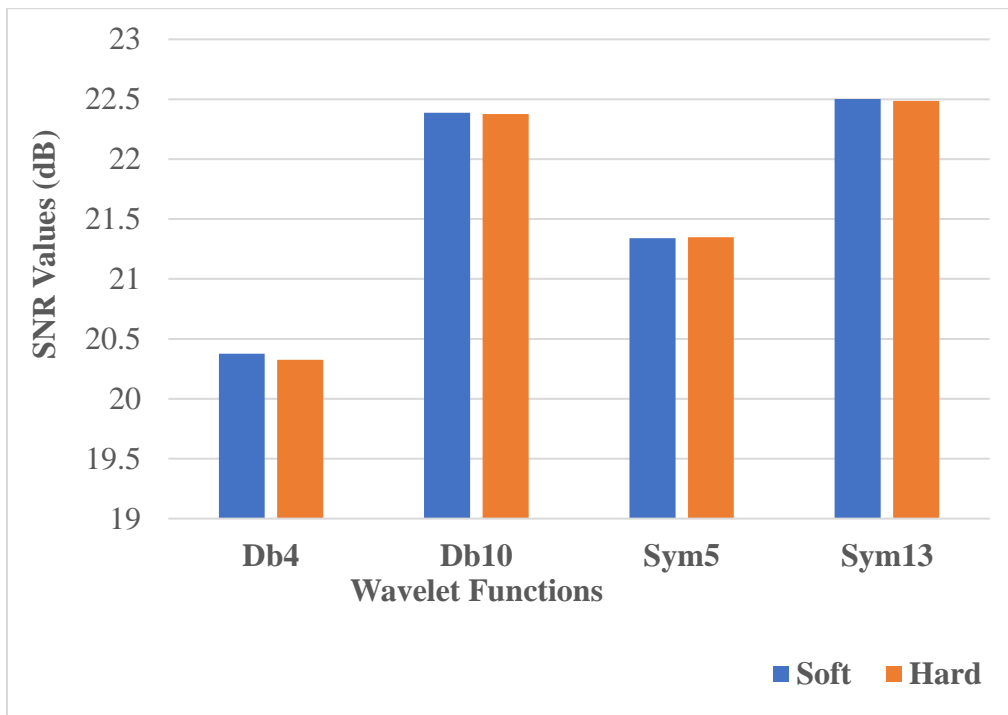
Figure 4.18: A comparison of SNR values obtained using different wavelet functions with soft and hard thresholding methods at 6th level of decomposition.

According to the result shown in Figure 4.18, the wavelet function particularly *Sym13* function gives the highest SNR value with soft thresholding method. The SNR value of *Sym13* at 6th level of decomposition with soft thresholding method was 22.5011 dB. This is the highest SNR value as compared to the other wavelet functions SNR values.

Therefore, in this research, Sym13 wavelet function at 6th level decomposition with soft thresholding method has been selected for decomposition and denoising of the lung sound signals. At last, the effect of the four threshold selection rules such as *Rigrsure*, *Sqtwolog*, *Heursure*, and *Minmax* were also studied by computing the SNR results when denoising the lung sound signals using the selected wavelet function. Table 4.3 shows the performance of the selected wavelet function (*Sym13*) in terms of SNR for denoising of lung sound signals of healthy, COPD, URTI, LRTI, bronchiectasis, bronchiolitis, pneumonia, and asthma subjects.

Table 4.3: SNR values of the four threshold selection rules for denoising some lung sound signals using *Sym13* wavelet function at 6 level decomposition with soft thresholding method.

Thresholding method	Soft			
Wavelet function	Sym13 at 6th level decomposition			
Threshold selection rules	**Heursure**	**Rigrsure**	**Minimax**	**Sqtwolog**
Performance measure	SNR (dB)	SNR (dB)	SNR (dB)	SNR (dB)
Asthma	22.5274	7.9084	20.8074	22.5437
Bronchiectasis	18.3032	18.3383	18.3459	18.3573
Bronchiolitis	22.7513	22.7919	22.7030	22.7142
COPD	11.8043	6.7031	11.6804	11.8150
Healthy	22.5302	22.4980	22.4411	22.4537
Pneumonia	21.8823	21.9152	21.9101	21.8736
LRTI	22.5367	22.6272	22.4426	22.5723
URTI	22.3549	22.2753	22.3364	22.3901

Table 4.3 shows the performance of *Sym13* wavelet function in terms of SNR values for denoising of different lung sound signals.

From this table, we can observe that the performance (SNR values) of *Sym13* wavelet function with soft thresholding method was different for the four threshold selection rules. According to the result shown in Table 4.3, the highest SNR values of Sym13 wavelet function were observed in using *Sqtwolog* threshold selection rule. *Sqtwolog* threshold selection rule gave the maximum performance of the selected wavelet function for denoising of lung sound signals of healthy, COPD, URTI, LRTI, bronchiectasis, bronchiolitis, pneumonia, and asthma subjects, shown using Table 4.3. Therefore, in this research, *Sqtwolog* was found the best threshold selection rule for the best denoising of the lung sound signals using *Sym13* wavelet function with soft thresholding method. Finally, the spectrograms for the input or original, noisy, and denoised lung sound signals are shown using Figure 4.19.
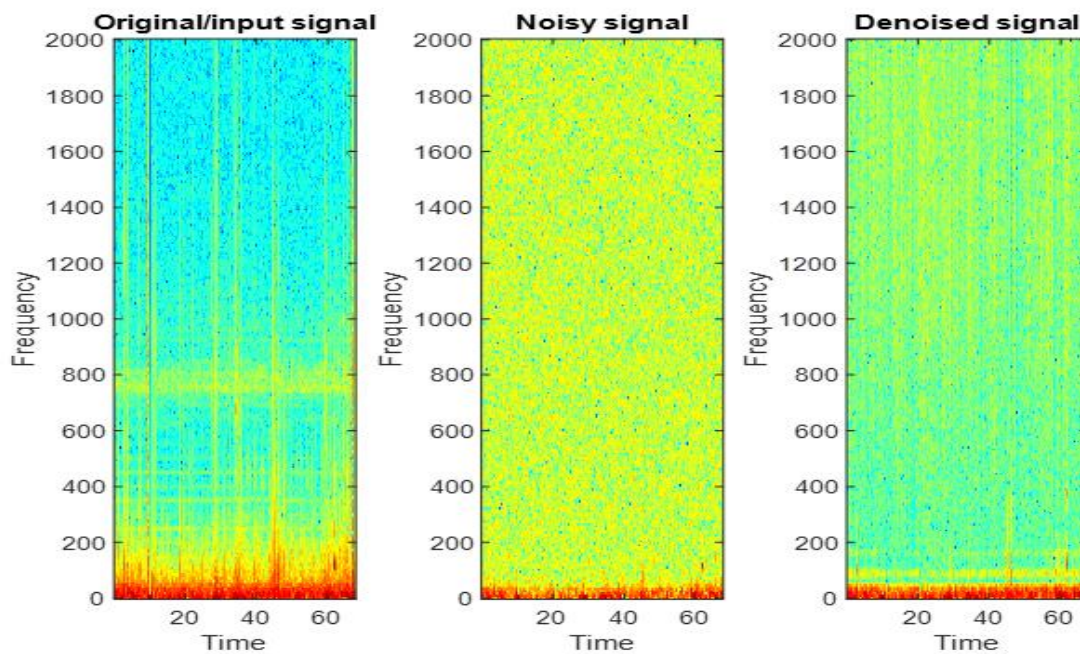
Figure 4.19: Spectrograms of lung sound signals; original/input lung sound signal (left), noisy lung sound signal (middle), and denoised lung sound signal (right).

The spectrograms shown using Figure 4.19 are used to show the clarity of lung sound segments obtained after applying the selected algorithm. After applying *Sym13* wavelet function at 6th level with soft thresholding method using *Sqtwolog* threshold selection rule, we can observe that the lung sounds are clear in the denoised lung sound signal spectrogram, shown in Figure 4.19.

4.4. Feature Extraction

Important features were extracted from the lung sound signals. The extracted features were used for final classification of data. Some of the extracted features include mean, standard deviation, root mean square, kurtosis, skewness, peak frequency, band power, total harmonic distortion, and others. Therefore, we create a feature table to view the values of the selected features. Table 4.4 shows sample of the extracted features stored in the feature table.

Table 4.4: Feature table holding sample of extracted features.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Clearance	CrestFact	ImpulseFa	Kurtosis	Mean	PeakValu	RMS	SINAD	SNR	ShapeFac	Skewness	Std	THD	PeakAmp	PeakFreq	BandPow	Label
2	4.3883	2.3767	3.3689	3.8487	0.0018	1	0.4208	-24.2055	-24.2055	1.4175	-0.0187	0.4208	-24.2579	62.0905	0.0987	1.1958	Asthma
3	1.6835	1.3045	1.4871	1.5341	0.005	1	0.7666	-23.8833	-23.8831	1.14	-0.0115	0.7666	-18.02	161.689	0.0764	3.6781	Asthma
4	2.3404	1.5838	1.9493	2.0601	0.001	1	0.6314	-24.9632	-24.9573	1.2307	3.27E-04	0.6314	-2.8167	71.629	0.0885	2.5305	Asthma
5	2.3696	1.5642	1.9432	2.0613	-3.14E-04	1	0.6393	-22.0281	-22.028	1.2423	0.0017	0.6393	-22.4962	73.9254	0.0972	2.574	Asthma
6	4.8792	3.1452	4.0374	3.5117	5.31E-05	1	0.3179	-25.8311	-25.8301	1.2836	0.1293	0.3179	-9.8336	39.0964	0.0376	0.653	Asthma
7	4.3982	2.393	3.3859	3.8666	0.0016	1	0.4179	-22.3177	-22.3177	1.4149	-0.0187	0.4179	-27.2819	52.6413	0.0946	1.1069	Asthma
8	4.3913	2.3909	3.3811	3.8556	0.0016	1	0.4182	-21.8857	-21.8856	1.4141	-0.0191	0.4182	-27.6799	52.0115	0.0313	0.9929	Asthma
9	2.3189	1.5746	1.9338	2.0419	0.0013	1	0.6351	-25.0626	-25.0586	1.2281	-4.30E-04	0.6351	-4.7128	61.7554	0.0885	2.5242	Asthma
10	2.3056	1.5696	1.925	2.0317	0.0011	1	0.6371	-25.0751	-25.0692	1.2264	-3.93E-05	0.6371	-2.4275	66.7234	0.0169	2.5798	Asthma
11	2.3333	1.5498	1.918	2.031	-3.69E-04	1	0.6453	-24.5863	-24.5862	1.2376	0.0016	0.6453	-22.1426	73.617	0.0861	2.5586	Asthma
12	2.244	1.7285	1.9952	1.797	-3.31E-04	1	0.5785	-33.0024	-33.0018	1.1543	0.0028	0.5785	-4.7319	37.7981	0.0067	2.0991	Asthma
13	2.3374	1.5523	1.9218	2.0359	-2.49E-04	1	0.6442	-24.6177	-24.6176	1.238	0.0012	0.6442	-21.1406	74.5058	0.0861	2.5708	Asthma
14	2.2441	1.7289	1.9955	1.7969	-5.35E-04	1	0.5784	-32.7891	-32.7874	1.1542	0.0034	0.5784	-0.9844	39.4214	0.0067	2.1001	Asthma
15	1.6722	1.3004	1.4798	1.5259	0.0049	1	0.769	-22.6702	-22.6701	1.138	-0.0114	0.769	-19.9387	174.225	0.0764	3.6679	Asthma
16	1.6738	1.3008	1.4807	1.5269	0.0051	1	0.7687	-22.9101	-22.91	1.1383	-0.0118	0.7687	-20.3852	176.457	0.0837	3.6808	Asthma
17	4.3311	2.3602	3.3334	3.8004	0.0018	1	0.4237	-23.8698	-23.8698	1.4124	-0.0186	0.4237	-24.6017	61.1598	0.0987	1.2151	Asthma
18	4.3617	2.3688	3.3521	3.8259	0.0018	1	0.4221	-24.0214	-24.0213	1.4151	-0.0187	0.4221	-24.5802	58.2606	0.0946	1.1849	Asthma
19	4.3474	2.3647	3.3434	3.8143	0.0018	1	0.4229	-23.9583	-23.9582	1.4139	-0.0186	0.4229	-24.4738	57.415	0.0946	1.1878	Asthma
20	4.3238	2.3579	3.3287	3.7937	0.0018	1	0.4241	-23.8558	-23.8557	1.4117	-0.0186	0.4241	-24.6575	56.7247	0.0946	1.1897	Asthma
21	4.3541	2.3668	3.3476	3.82	0.0018	1	0.4225	-23.9996	-23.9996	1.4144	-0.0186	0.4225	-24.4611	61.8416	0.0987	1.2118	Asthma

Now, the extracted features are computed and stored in feature table as shown in Table 4.4. The next step is to rank features to see which features contributes significant roles to distinguish the lung sound signals of the different subjects. For example, when we create and look for feature histogram plot of the lung sound signals of the different subjects, there are lots of overlapping between the features as shown in Figure 4.20.
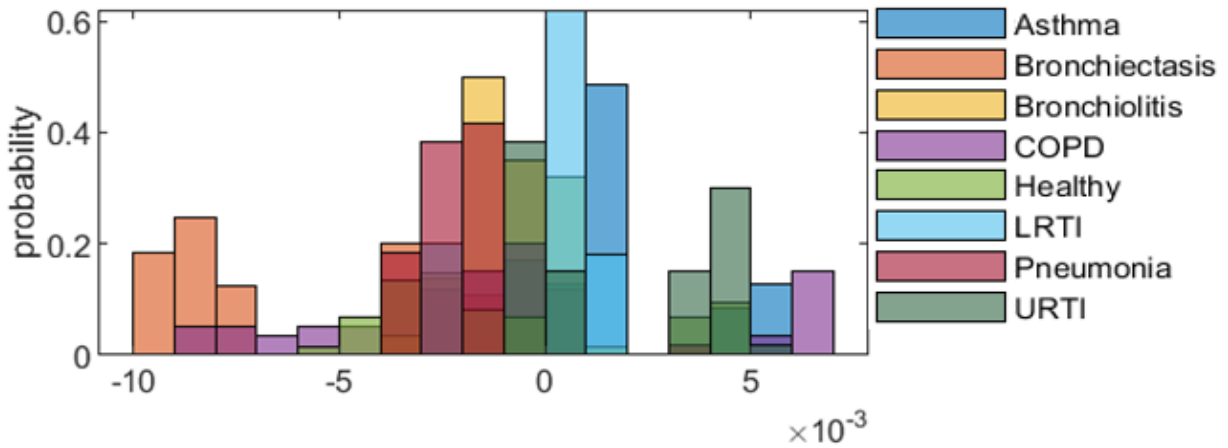
(a)
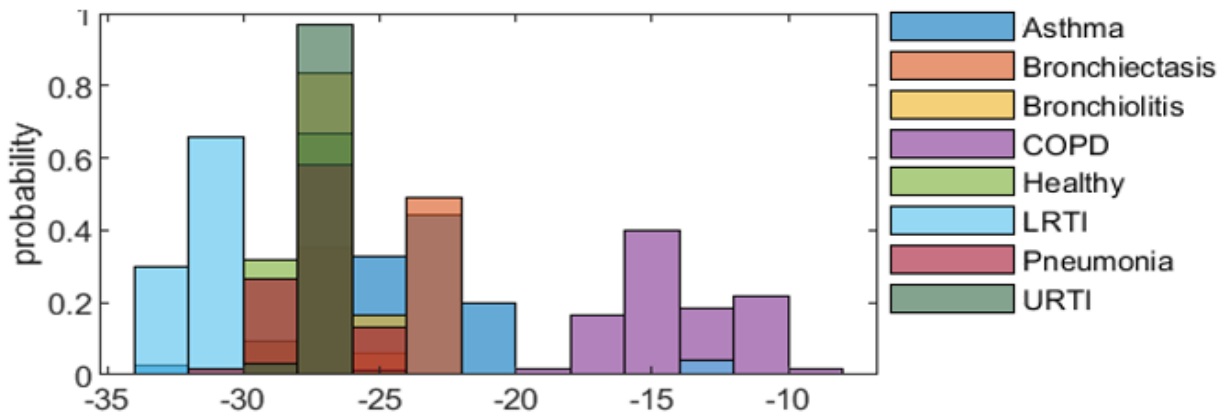
(b)
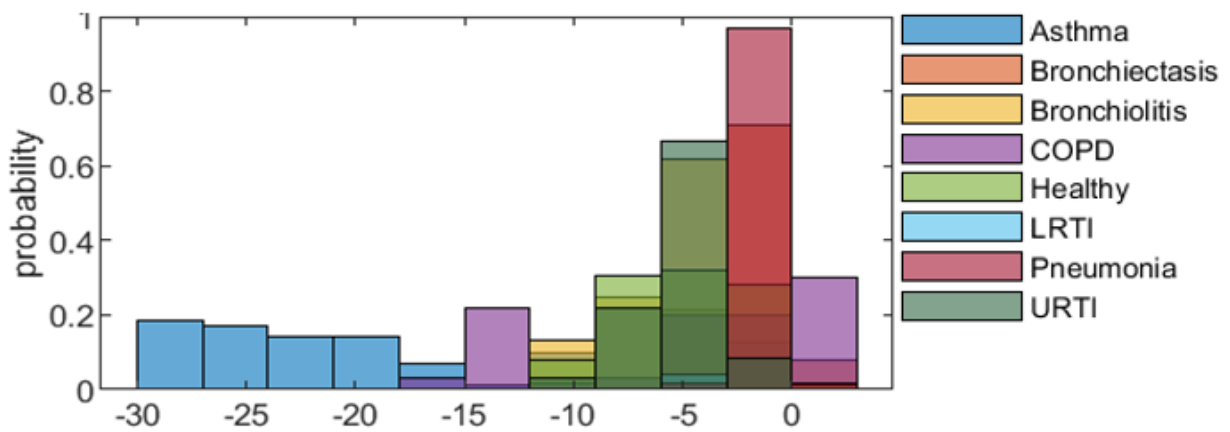
(c)

Figure 4.20: A sample of feature histogram plot of some of the extracted features; (a), mean (b) SNR and (c) THD.

There is a lot of overlapping between the feature's histogram plot of the different subject's lung sound signals, shown in Figure 4.20. Due to this overlapping, it is really hard to tell the most useful discriminative feature by looking the histogram plots. As a result, feature ranking is highly important. It helps us to get the best predictors or features used to discriminate different classes.

4.5. Feature Selection

In this research, One-way ANOVA feature ranking method was employed for selection of most discriminative features. One-way ANOVA is a common technique used to characterize features by doing one-way analysis of variance. The result of ranking features using one-way ANOVA is shown using Table 4.5 and Figure 4.21. Table 4.5 shows the features sorted by importance using one-way ANOVA feature ranking method.

Table 4.5: Features sorted by importance using one-way ANOVA feature ranking method. (Std=standard deviation, RMS= root mean square, SNR= signal-to-noise ratio, SINAD= signal-to-noise and distortion ratio.)

Feature	One-way ANOVA score values
Shape factor	875.078
Std	550.53
RMS	550.312
Crest factor	450.617
Clearance factor	432.095
Impulse factor	419.692
SNR	419.064
SINAD	416.626
Band power	231.186
Peak amplitude	95.6772
Mean	65.935
Kurtosis	61.0681
Peak frequency	29.0683
Skewness	19.7747
Peak value	2.3584
THD	0

Moreover, we used the following bar chart to see the importance of extracted features for classifying the data using one-way ANOVA, shown using Figure 4.21.
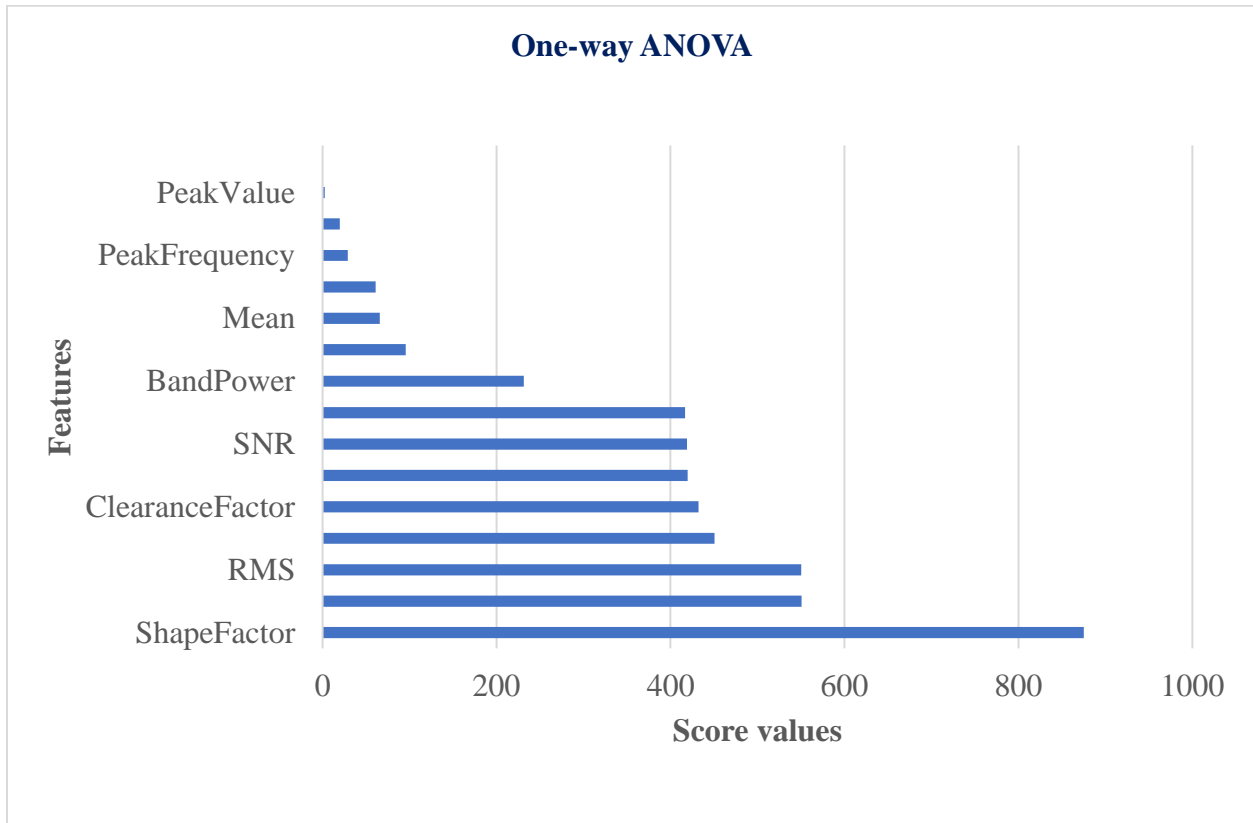
Figure 4. 21: One-way ANOVA feature ranking result.

Generally, when we are using machine learning models for classification, we choose features with high ANOVA score and leave out others with much smaller ANOVA score. Meaning, a high score is an indicative for a valuable or relevant feature. From the result shown using Table 4.5 and Figure 4.21, we can notice that the last three time-domain features such as skewness, peak value, and total harmonic distortion (THD) were not playing important role in distinguishing lung sound signals of different subjects. Therefore, we only removed these features while all the other features were kept in the feature table for training. Generally, ranking features using one-way ANOVA helps us to select the top relevant features. This in turn help us to get higher accuracy value during training and classification of final data. Table 4.6 shows the feature table holding only the selected features and their corresponding values for different classes.

Table 4.6: Feature table holding a sample of selected features after feature ranking.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Clearance	CrestFact	ImpulseFa	Kurtosis	Mean	RMS	SINAD	SNR	ShapeFac	Std	PeakAmpl	PeakFreqı	BandPowı	Label
2	4.3883	2.3767	3.3689	3.8487	0.0018	0.4208	-24.206	-24.206	1.4175	0.4208	62.0905	0.0987	1.1958	Asthma
3	1.6835	1.3045	1.4871	1.5341	0.005	0.7666	-23.883	-23.883	1.14	0.7666	161.689	0.0764	3.6781	Asthma
4	2.3404	1.5838	1.9493	2.0601	0.001	0.6314	-24.963	-24.957	1.2307	0.6314	71.629	0.0885	2.5305	Asthma
5	2.3696	1.5642	1.9432	2.0613	-3.14E-04	0.6393	-22.028	-22.028	1.2423	0.6393	73.9254	0.0972	2.574	Asthma
6	4.8792	3.1452	4.0374	3.5117	5.31E-05	0.3179	-25.831	-25.83	1.2836	0.3179	39.0964	0.0376	0.653	Asthma
7	4.3982	2.393	3.3859	3.8666	0.0016	0.4179	-22.318	-22.318	1.4149	0.4179	52.6413	0.0946	1.1069	Asthma
8	4.3913	2.3909	3.3811	3.8556	0.0016	0.4182	-21.886	-21.886	1.4141	0.4182	52.0115	0.0313	0.9929	Asthma
9	2.3189	1.5746	1.9338	2.0419	0.0013	0.6351	-25.063	-25.059	1.2281	0.6351	61.7554	0.0885	2.5242	Asthma
10	2.3056	1.5696	1.925	2.0317	0.0011	0.6371	-25.075	-25.069	1.2264	0.6371	66.7234	0.0169	2.5798	Asthma
11	2.3333	1.5498	1.918	2.031	-3.69E-04	0.6453	-24.586	-24.586	1.2376	0.6453	73.617	0.0861	2.5586	Asthma
12	2.244	1.7285	1.9952	1.797	-3.31E-04	0.5785	-33.002	-33.002	1.1543	0.5785	37.7981	0.0067	2.0991	Asthma
13	2.3374	1.5523	1.9218	2.0359	-2.49E-04	0.6442	-24.618	-24.618	1.238	0.6442	74.5058	0.0861	2.5708	Asthma
14	2.2441	1.7289	1.9955	1.7969	-5.35E-04	0.5784	-32.789	-32.787	1.1542	0.5784	39.4214	0.0067	2.1001	Asthma
15	1.6722	1.3004	1.4798	1.5259	0.0049	0.769	-22.67	-22.67	1.138	0.769	174.225	0.0764	3.6679	Asthma
16	1.6738	1.3008	1.4807	1.5269	0.0051	0.7687	-22.91	-22.91	1.1383	0.7687	176.457	0.0837	3.6808	Asthma
17	4.3311	2.3602	3.3334	3.8004	0.0018	0.4237	-23.87	-23.87	1.4124	0.4237	61.1598	0.0987	1.2151	Asthma
18	4.3617	2.3688	3.3521	3.8259	0.0018	0.4221	-24.021	-24.021	1.4151	0.4221	58.2606	0.0946	1.1849	Asthma
19	4.3474	2.3647	3.3434	3.8143	0.0018	0.4229	-23.958	-23.958	1.4139	0.4229	57.415	0.0946	1.1878	Asthma
20	4.3238	2.3579	3.3287	3.7937	0.0018	0.4241	-23.856	-23.856	1.4117	0.4241	56.7247	0.0946	1.1897	Asthma
21	4.3541	2.3668	3.3476	3.82	0.0018	0.4225	-24	-24	1.4144	0.4225	61.8416	0.0987	1.2118	Asthma

4.6. Feature Visualization and Normalization

After extraction of different features from the lung sound signals, we ranked and selected the most useful features for final training and classification of most common lung diseases. However, we saw the selected features were not on the same scale. This will bias the classifier and affect the accuracy during data training and classification. Therefore, it is always encouraged to first visualize the data features before training the classifiers. Sometimes, feature visualization will give some insight on how to increase the accuracy. We visualized the extracted data features as shown in Table 4.7, which shows sample of feature values before normalization.

Table 4.7: Sample of feature values before normalization.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Clearance	CrestFact	ImpulseFa	Kurtosis	Mean	RMS	SINAD	SNR	ShapeFac	Std	PeakAmpl	PeakFreqı	BandPowı	Label
2	4.3883	2.3767	3.3689	3.8487	0.0018	0.4208	-24.206	-24.206	1.4175	0.4208	62.0905	0.0987	1.1958	Asthma
3	1.6835	1.3045	1.4871	1.5341	0.005	0.7666	-23.883	-23.883	1.14	0.7666	161.689	0.0764	3.6781	Asthma
4	2.3404	1.5838	1.9493	2.0601	0.001	0.6314	-24.963	-24.957	1.2307	0.6314	71.629	0.0885	2.5305	Asthma
5	2.3696	1.5642	1.9432	2.0613	-3.14E-04	0.6393	-22.028	-22.028	1.2423	0.6393	73.9254	0.0972	2.574	Asthma
6	4.8792	3.1452	4.0374	3.5117	5.31E-05	0.3179	-25.831	-25.83	1.2836	0.3179	39.0964	0.0376	0.653	Asthma
7	4.3982	2.393	3.3859	3.8666	0.0016	0.4179	-22.318	-22.318	1.4149	0.4179	52.6413	0.0946	1.1069	Asthma
8	4.3913	2.3909	3.3811	3.8556	0.0016	0.4182	-21.886	-21.886	1.4141	0.4182	52.0115	0.0313	0.9929	Asthma
9	2.3189	1.5746	1.9338	2.0419	0.0013	0.6351	-25.063	-25.059	1.2281	0.6351	61.7554	0.0885	2.5242	Asthma

From the result shown in Table 4.7, we can notice that the features are not on the same scale. This will bias the classifier. So, normalizing the features is needed before training the classifier. In this research, normalization of features is done by subtracting the mean and dividing the standard deviation of each column. The result of feature normalization is shown in Table 4.8. The features are now almost on the same scale.

Table 4.8: Sample of feature values after normalization.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Clearance	CrestFactor	ImpulseFact	Kurtosis	Mean	RMS	SINAD	SNR	ShapeFact	Std	PeakAmp	PeakFreq	BandPow	Label
2	0.33317	0.043648	0.045166	0.015824	0.72865	0.52327	0.39233	0.39217	0.33317	0.52394	0.35089	0.03139	0.32296	Asthma
3	0.00244	0.0001641	0.0001739	0.015824	0.93113	0.99673	0.40594	0.40578	0.00244	0.99674	0.91618	0.02429	0.99815	Asthma
4	0.11058	0.011494	0.011225	0.00364	0.68193	0.81162	0.36034	0.36044	0.11058	0.8119	0.40503	0.02814	0.68601	Asthma
5	0.12433	0.010698	0.011078	0.003647	5.98E-01	0.82247	0.48427	0.48407	0.12433	0.82273	0.41806	0.0309	0.69783	Asthma
6	0.17365	0.074816	0.061149	0.013528	6.21E-01	0.38252	0.3237	0.3236	0.17365	0.38339	0.22038	0.01192	0.17534	Asthma
7	0.33011	0.044311	0.045574	0.015946	0.71869	0.51934	0.47204	0.47184	0.33011	0.52002	0.29726	0.03008	0.29879	Asthma
8	0.3292	0.044226	0.045459	0.015871	0.71817	0.51984	0.49028	0.49008	0.3292	0.52052	0.29368	0.00991	0.2678	Asthma
9	0.01206	0.011119	0.010853	0.003515	0.69667	0.8167	0.35614	0.35616	0.10746	0.81697	0.34899	0.02814	0.68428	Asthma

4.7. Data Splitting

After feature normalization (making features to have the same scale) shown in Table 4.9, the next step is to prepare the data for training and testing programmatically. At last, the result of data splitting into sample data features used for training and testing is shown using Table 4.9 and 4.10 respectively. Therefore, the data features presented using Table 4.9 were used for training and validating our model while the data features given in Table 4.10 were used to test our model.

Table 4. 9: Sample of data features used for training.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Clearance	CrestFact	ImpulseFa	Kurtosis	Mean	RMS	SINAD	SNR	ShapeFac	Std	PeakAmpl	PeakFreq	BandPow	Label
2	0.05066	0.04365	0.04517	0.01582	0.728651	0.52327	0.39233	0.39217	0.33317	0.52394	0.35089	0.03139	0.32296	'Asthma'
3	0.00021	0.00016	0.00017	5.59E-05	0.931133	0.99673	0.40594	0.40578	0.00244	0.99674	0.91618	0.02429	0.99815	'Asthma'
4	0.01246	0.01149	0.01122	0.00364	0.681933	0.81162	0.36034	0.36044	0.11058	0.8119	0.40503	0.02814	0.68601	'Asthma'
5	0.01301	0.0107	0.01108	0.00365	5.98E-01	0.82247	0.48427	0.48407	0.12433	0.82273	0.41806	0.0309	0.69783	'Asthma'
6	0.05981	0.07482	0.06115	0.01353	6.21E-01	0.38252	0.3237	0.3236	0.17365	0.38339	0.22038	0.01192	0.17534	'Asthma'
7	0.05071	0.04423	0.04546	0.01587	0.718175	0.51984	0.49028	0.49008	0.3292	0.52052	0.29368	0.00991	0.2678	'Asthma'
8	0.01206	0.01112	0.01085	0.00352	0.696674	0.8167	0.35614	0.35616	0.10746	0.81697	0.34899	0.02814	0.68428	'Asthma'
9	0.01181	0.01092	0.01064	0.00345	0.684378	0.81944	0.35562	0.35572	0.10539	0.81971	0.37718	0.00534	0.69942	'Asthma'
10	0.01233	0.01011	0.01048	0.00344	0.594278	0.83061	0.37626	0.3761	0.11874	0.83086	0.41631	0.02736	0.69364	'Asthma'
11	0.01241	0.01022	0.01057	0.00347	6.02E-01	0.82916	0.37493	0.37478	0.11926	0.82941	0.42135	0.02736	0.69695	'Asthma'
12	0.01067	0.01738	0.01233	0.00185	5.84E-01	0.73908	0.02992	0.02998	0.01935	0.73946	0.22222	0.0021	0.56893	'Asthma'
13	0	0	0	0	9.26E-01	1	0.45716	0.45697	0	1	0.98734	0.02429	0.99538	'Asthma'
14	0.04959	0.04298	0.04432	0.01549	7.29E-01	0.52731	0.40651	0.40634	0.32708	0.52797	0.34561	0.03139	0.32822	'Asthma'
15	0.05016	0.04333	0.04477	0.01567	0.728669	0.52518	0.40011	0.39994	0.33034	0.52585	0.32915	0.03007	0.32001	'Asthma'
16	0.0499	0.04316	0.04456	0.01559	0.728616	0.52619	0.40277	0.40261	0.3289	0.52686	0.32435	0.03007	0.32078	'Asthma'
17	0.04945	0.04289	0.04421	0.01545	0.72897	0.52786	0.4071	0.40693	0.32629	0.52852	0.32043	0.03008	0.32131	'Asthma'
18	0.05002	0.04325	0.04466	0.01563	0.727535	0.52569	0.40103	0.40086	0.32955	0.52636	0.34948	0.03139	0.32734	'Asthma'
19	0.04975	0.04307	0.04444	0.01554	0.728377	0.52675	0.40445	0.40428	0.32802	0.52741	0.32274	0.03007	0.32105	'Asthma'
20	0.0503	0.04342	0.04488	0.01571	0.72837	0.52464	0.39821	0.39804	0.33113	0.52531	0.34998	0.03139	0.32623	'Asthma'
21	0.00018	0.00015	0.00015	5.05E-05	0.931914	0.99696	0.40887	0.40871	0.002	0.99696	0.97665	0.02429	0.99928	'Asthma'

Table 4.10: Sample of data features used for testing.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Clearance	CrestFactor	ImpulseFa	Kurtosis	Mean	RMS	SINAD	SNR	ShapeFac	Std	PeakAmpl	PeakFreq	BandPow	Label
2	0.05084	0.04431073	0.04557	0.01595	0.718694	0.51934	0.47204	0.47184	0.33011	0.52002	0.29726	0.03008	0.29879	'Asthma'
3	0.01066	0.01736151	0.01232	1.85E-03	0.596699	0.73927	0.02092	0.02093	0.01946	0.73965	0.21301	0.0021	0.56867	'Asthma'
4	2.93E-05	1.65E-05	2.13E-05	6.75E-06	0.937747	0.99967	0.44703	0.44685	0.00039	0.99967	1	0.0266	0.99889	'Asthma'
5	0.01237	0.01146127	0.01116	0.00362	6.82E-01	0.81206	0.35929	0.3594	0.1094	0.81234	0.36251	0.00564	0.68277	'Asthma'
6	0.01267	0.01168202	0.01141	0.0037	6.81E-01	0.8091	0.36438	0.36446	0.1122	0.80938	0.36658	0.00564	0.68043	'Asthma'
7	0.01284	0.01051773	0.01092	0.00359	0.600273	0.82496	0.49413	0.49393	0.12334	0.82522	0.41354	0.02783	0.6966	'Asthma'
8	0.05879	0.07397729	0.06028	0.01333	0.619768	0.38541	0.32803	0.32793	0.16992	0.38627	0.39442	0.00593	0.1741	'Asthma'
9	0.05881	0.0740266	0.0603	0.01331	0.619937	0.38524	0.32944	0.32934	0.16969	0.3861	0.44825	0.00593	0.17695	'Asthma'
10	0.0588	0.07399704	0.06029	0.01331	0.620541	0.38534	0.33017	0.33008	0.16977	0.3862	0.39564	0.00593	0.17454	'Asthma'
11	0.0129	0.01055355	0.01096	0.00361	5.96E-01	0.82446	0.49252	0.49232	0.1239	0.82472	0.41343	0.02783	0.69676	'Asthma'
12	0.01295	0.01061476	0.01101	0.00363	5.97E-01	0.82361	0.48298	0.48278	0.12415	0.82388	0.41128	0.02783	0.697	'Asthma'
13	0.05104	0.04444201	0.04574	0.01601	7.20E-01	0.51857	0.47002	0.46982	0.33125	0.51925	0.3003	0.03008	0.30152	'Asthma'
14	0.00016	0.00013936	0.00014	4.51E-05	9.22E-01	0.99723	0.46501	0.46483	0.00174	0.99723	0.97827	0.02429	0.99462	'Asthma'
15	0.05047	0.04399804	0.04523	0.01581	0.718123	0.52119	0.47889	0.47869	0.32847	0.52186	0.32831	0.00991	0.29061	'Asthma'
16	0.01056	0.01721288	0.01221	0.00184	0.398666	0.74095	0.2581	0.25806	0.01925	0.74132	0.20707	9.35E-05	0.57214	'Bronchi'
17	0.01058	0.01729523	0.01225	0.00183	0.418175	0.74002	0.24367	0.24369	0.0188	0.74039	0.18265	9.35E-05	0.57229	'Bronchi'
18	0.01058	0.01730816	0.01226	0.00183	0.475124	0.73987	0.23384	0.23394	0.01866	0.74025	0.1827	9.35E-05	0.57059	'Bronchi'
19	0.01059	0.01730829	0.01226	0.00184	0.453149	0.73987	0.24238	0.24239	0.01883	0.74024	0.18467	9.35E-05	0.57143	'Bronchi'
20	0.01056	0.01730858	0.01225	0.00183	0.488974	0.73987	0.23339	0.23347	0.01832	0.74024	0.17957	9.35E-05	0.57009	'Bronchi'
21	0.01056	0.0172872	0.01224	1.82E-03	0.51808	0.74011	0.23037	0.23039	0.01856	0.74049	0.18443	1	0.57085	'Bronchi'

4.8. Model Training and Evaluation

This was the final step to interactively different classification models. In this research, we trained important machine learning models such as Naive Bayes, K-nearest neighbor (KNN), Ensemble and support vector machine (SVM) models for classification of most common lung diseases using the features extracted from the lung sound signals. The results of all these machine learning models for final classification of data are shown using Figure 4.22 and Table 4.11.
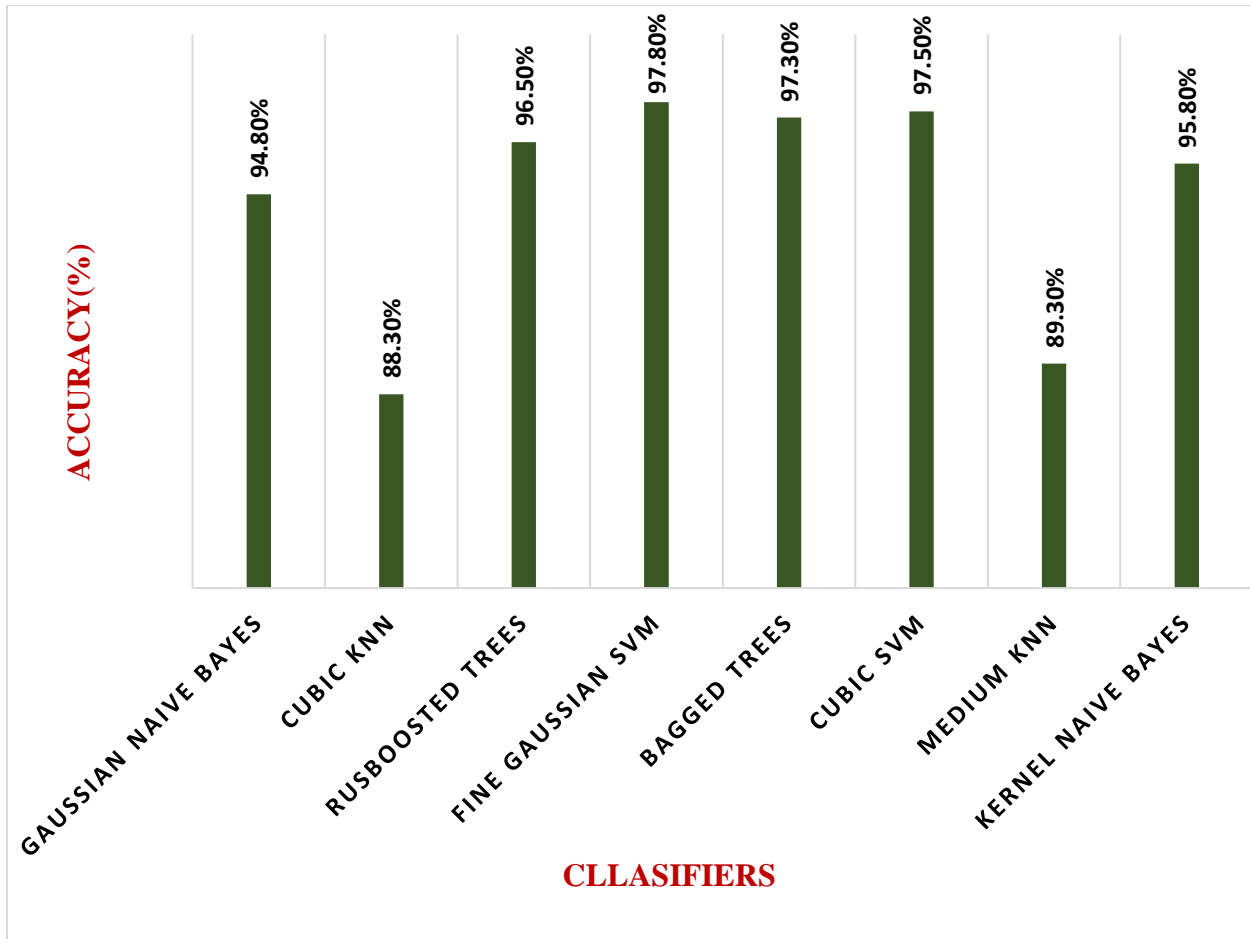
Figure 4.22: Accuracy achieved on different machine learning classifiers.

A column chart presented using Figure 4.22 shows the comparison in accuracy achieved by different machine learning classifiers for classification of lung diseases. The highest classification accuracy was achieved using SVM model particularly Fine Gaussian SVM algorithm. Table 4.11 also shows the accuracy of each classifier in ascending order.

Table 4.11: Classification accuracy results in ascending order.

Classifiers	Algorithms	Accuracy (%)
KNN	Cubic KNN	88.3%
	Medium KNN	89.3%
Naive Bayes	Gaussian Naive Bayes	94.8%
	Kernel Naive Bayes	95.8%
Ensemble	RUSBoosted trees	96.5%
	Bagged trees	97.3%
SVM	Cubic SVM	97.5%
	Fine Gaussian SVM	**97.8%**

After training multiple models, we compare their classification accuracy side-by-side as shown in Table 4.11. The accuracy of all the classifiers is ranged between 88.3% and 97.8%. According to the results shown using Table 4.11, in the first run result we found that SVM model particularly Fine Gaussian SVM is the best model with its higher classification accuracy of 97.8%. Next, we assessed the classification performance of the selected classifier (Fine Gaussian SVM) using confusion matrices.

Table 4.12: A number of correctly and misclassified observations for each class.

True class	Asthma	Bronchiectasis	Bronchiolitis	COPD	Healthy	LRTI	Pneumonia	URTI
Asthma	55				1			
Bronchiectasis	1	51						
Bronchiolitis			48					
COPD				48				
Healthy					60			
LRTI	3					37		
Pneumonia					2		46	
URTI					2			46

Predicted class

Table 4.12 shows a number of correctly and misclassified observations of each class obtained in training Fine Gaussian SVM. From this table we have noticed that 56 lung sound records acquired from asthma patients were used for training of which 55 records were correctly classified and 1 record is misclassified. Similarly, 52 lung sound records acquired from bronchiectasis patients were used for training of which 51 records were correctly classified and 1 record is misclassified. For bronchiolitis patients, 48 lung sound records were used for training and all of them were correctly classified. Likewise, 48 lung sound records acquired from COPD patients were used for training and all of them were correctly classified. For healthy subjects, 60 lung sound records were used for training and all the records were correctly classified. Among 40 lung sound records of LRTI patients used for training, 37 records were correctly classified and the remaining 3 records were misclassified. Similarly, 48 lung sound records acquired from pneumonia patients were used for training of which 46 records were correctly classified and 2 records were misclassified. At last, 48 lung sound records acquired from URTI patients were used for training of which 46 records were correctly classified and 2 records were misclassified.

We also visualized the performance of a classifier performed per true class in terms of true positive rates (TPR) and false negative rates (FNR). TPR is the proportion of correctly classified observations per true class while FNR is the proportion of incorrectly classified observations per true class. The confusion matrix presented using Table 4.13 shows the performance of a classifier performed per true class. The last two columns on the right side of Table 4.13 show TPR and FNR per true class, whereas the first column is the TPR column and the second column is the FNR column.

Table 4.13: Performance of the first selected classifier performed per true class.

True class	Asthma	Bronchiectasis	Bronchiolitis	COPD	Healthy	LRTI	Pneumonia	URTI		TPR	FNR
Asthma	98%				2%					98%	2%
Bronchiectasis	2%	98%								98%	2%
Bronchiolitis			100%							100%	
COPD				100%						100%	
Healthy					100%					100%	
LRTI	8%					93%				93%	7%
Pneumonia					4%		96%			96%	4%
URTI					4%			96%		96%	4%

Predicted class

For example, from Table 4.13, the first row from the top shows all lung sound records with the true class asthma. The columns show the predicted classes. 98% of the lung sound records from asthma are correctly classified as from asthma, so 98% is the TPR for correctly classified points in the class, shown in the green cell in the TPR column. The other 2% lung sound records in the asthma row are incorrectly classified. Hence, the FNR for incorrectly classified points in the class is 2%, shown in the pink cell in the FNR column. Similarly, the second row from the top shows all lung sound records with the true class bronchiectasis. 98% of the lung sound records from bronchiectasis are correctly classified as from bronchiectasis, so 98% is the TPR for correctly classified points in the class, shown in the green cell in the TPR column. The other 2% lung sound records in the bronchiectasis row are incorrectly classified.

Hence, the FNR for incorrectly classified points in the class is 2%, shown in the purple cell in the FNR column. Generally, Table 4.13 can be interpreted in the same fashion for each row. We also assessed the performance of a classifier performed per predicted class instead of true class to investigate false discovery rates (FDR). In addition to FDR, we also investigate positive predictive values (PPV). FDR is the proportion of incorrectly classified observations per predicted class while PPV is the proportion of correctly classified observations per predicted class. The confusion matrix presented using Table 4.14 shows the performance of a classifier performed per predicted class. The last two rows found at the bottom of Table 4.14 show PPV and FDR per predicted class, where the first row is the PPV row and the second row is the FDR row.

Table 4.14: Performance of the first selected classifier performed per predicted class.

True class	Asthma	Bronchiectasis	Bronchiolitis	COPD	Healthy	LRTI	Pneumonia	URTI
Asthma	93%				2%			
Bronchiectasis	2%	100%						
Bronchiolitis			100%					
COPD				100%				
Healthy					92%			
LRTI	5%					100%		
Pneumonia					3%		100%	
URTI					3%			100%
Positive Predictive Value	93%	100%	100%	100%	92%	100%	100%	100%
False Discovery Rate	7%				8%			

Predicted class

Table 4.14 shows the performance of a classifier performed per predicted class in terms of PPV and FDR. Positive predicted values (PPV) are shown in green for the correctly predicted points in each class.

The false discovery rates (FDR) are shown in pink for the incorrectly predicted points in each class. Now, the performance of the first selected and trained model has been evaluated using confusion matrix. Next, we tried to optimize it to improve its' accuracy.

4.9. Model optimization

In this research, we applied Bayesian optimization technique to automatically optimize our previous trained and selected model. After, playing with the parameters/hyperparameters using Bayesian optimization technique, the accuracy of our model become 98.8%. After optimization, we evaluate the performance of an optimized model using confusion matrices as we did previously for the first trained and selected model.

Table 4.15: A number of correctly and misclassified observations for each class after optimization.

From Table 4.15, we have noticed that among 56 lung sound records acquired from asthma patients, 55 records were correctly classified and 1 record is misclassified. A total of 52 lung sound records acquired from bronchiectasis patients were used for training and all of them were correctly classified. For bronchiolitis and COPD patients, 48 lung sound records for each were used for training and all of them were correctly classified.

Likewise, among 60 lung sound records of healthy subjects used for training, 59 lung sound recordings were classified correctly and the remaining 1 record was misclassified. For LRTI patients, 40 lung sound records were used for training of which 39 records were correctly classified and the remaining 1 record was misclassified. A total of 48 lung sound records acquired from pneumonia patients were used for training of which 47 records were correctly classified and 1 record is misclassified. At last, 48 lung sound records acquired from URTI patients were used for training and all of the records were correctly classified. We also visualized the performance of an optimized classifier performed per true class in terms of true positive rates (TPR) and false negative rates (FNR). The confusion matrix presented using Table 4.16 shows the performance of an optimized classifier performed per true class. The last two columns on the right side of Table 4.16 showed TPR and FNR per true class, where the first column is the TPR column and the second column is the FNR column.

Table 4.16: Performance of an optimized classifier performed per true class.

True class \ Predicted class	Asthma	Bronchiectasis	Bronchiolitis	COPD	Healthy	LRTI	Pneumonia	URTI	TPR	FNR
Asthma	98%	2%							98%	2%
Bronchiectasis	2%	98%							98%	2%
Bronchiolitis			100%						100%	
COPD				100%					100%	
Healthy					98%		2%		98%	2%
LRTI	3%					98%			98%	3%
Pneumonia					2%		98%		98%	2%
URTI								100%	100%	

For example, from Table 4.16, the last row (the eighth row) shows all lung sound records with the true class URTI. 100% of the lung sound records from URTI were correctly classified as from URTI, so 100% is the TPR for correctly classified points in the class, shown in the green cell in the TPR column.

Previously, for the first trained unoptimized model, the TPR and FNR were 96% and 4% respectively. This shows that there is improvement in the model to correctly classify the true class URTI. Similarly, in Table 4.16, the seventh row shows all lung sound records with the true class pneumonia. 98% of the lung sound records from pneumonia were correctly classified as from pneumonia, so 98% is the TPR for correctly classified points in the class, shown in the green cell in the TPR column. Before optimizing the model, for the first trained unoptimized model the TPR and FNR were 96% and 4% respectively. This also shows that there is improvement in the model to correctly classify the true class pneumonia. Generally, Table 4.16 can be interpreted in the same fashion for each row. We also assessed the performance of an optimized classifier performed per predicted class instead of true class to investigate false discovery rates (FDR) and positive predictive values (PPV). The confusion matrix presented using Table 4.17 shows the performance of an optimized classifier performed per predicted class. The last two rows show PPV and FDR per predicted class, where the first row is the PPV row and the second row is the FDR row.

Table 4. 17: Performance of an optimized classifier performed per predicted class.

Table 4.17 shows the performance of an optimized classifier performed per predicted class in terms of PPV and FDR. Positive predicted values (PPV) are shown in green for the correctly predicted points in each class, and false discovery rates (FDR) are shown in pink for the incorrectly predicted points in each class. From Table 4.17, we have noticed that there are improvements in the optimized model to correctly predict classes such as healthy, LRTI, pneumonia, and URTI predicted classes. Furthermore, the minimum classification error plot for the optimized model was shown using Figure 4.23. The plot showed estimated minimum classification error, observed minimum classification error, best point hyperparameters, and minimum error hyperparameters. Each light blue point shown in the plot correspond to the estimated minimum classification error computed by the optimization process. The minimum classification error is estimated based on an upper confidence interval of the current classification error objective model, mentioned in the best point hyperparameters description. Each dark blue point in the plot correspond to the observed minimum classification error computed during the optimization process. For example, the dark blue point at the third iterations corresponds to the classification error observed in the first, second, and third iterations. Best point hyperparameters shown as red square in the plot indicates the iteration that correspond to the optimized hyperparameters. The optimized hyperparameters didn't always provide the observed minimum classification error. For example, when the hyperparameter optimization is done by using Bayesian optimization, it selects the set of hyperparameter values which minimizes an upper confidence interval of the classification error objective model, instead of the set which minimizes the classification error. The yellow point in the plot indicates that the iterations which correspond to the hyperparameters that yield the observed minimum classification error.
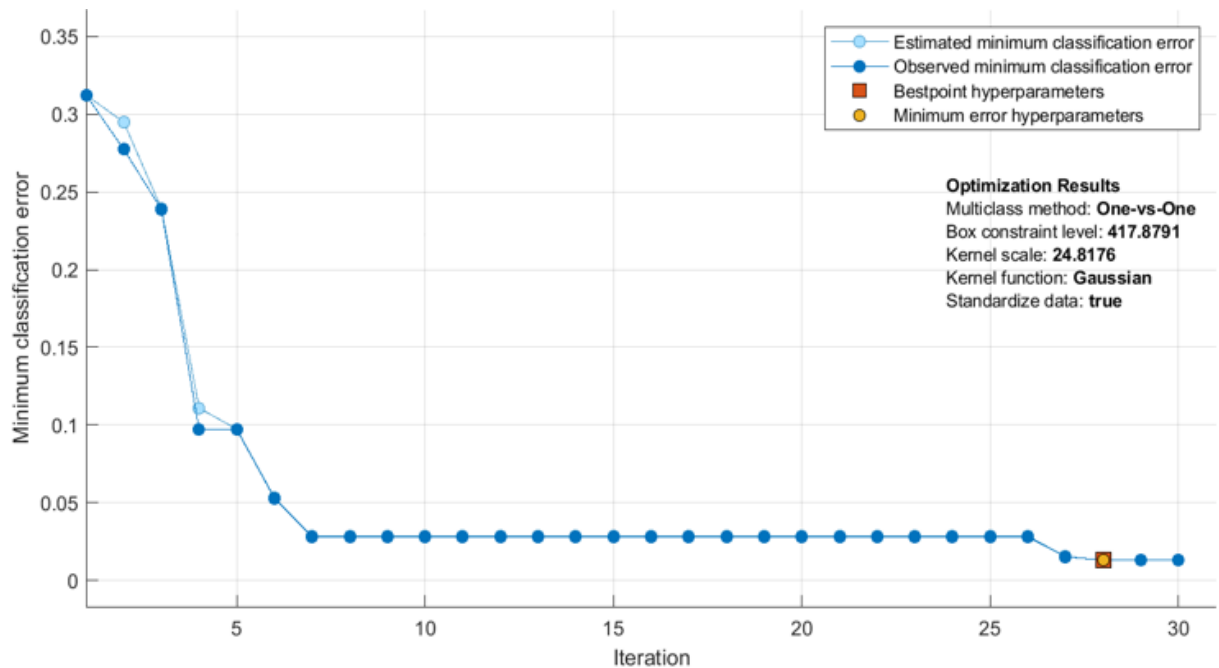
Figure 4.23: Minimum classification error plot for the optimized model.

From Figure 4.23, we have observed that the optimal model parameter results such as box constraint level and kernel scale were 417.8791 and 24.8176 respectively. Now, the performance of the optimized model has been evaluated using confusion matrices as well as the results of optimization are discussed above. After optimizing the model, we exported the final optimized model to make predictions using the new data. When we exported the final optimized model from the classification learner to MATLAB workspace, we get an optimized trained structure which we can use to make predictions using the un seen data. A total of 100 lung sound records (asthma = 14, bronchiectasis = 13, bronchiolitis = 12, COPD = 12, healthy = 14, LRTI = 10, pneumonia = 13, URTI = 12) were used for testing the final optimized model. The model correctly classified all new observations of the actual classes except observations from pneumonia class. Among the 13 observations or records from the true class pneumonia, 12 observations were correctly classified and only 1 record was misclassified. All the records from the true class asthma, bronchiectasis, bronchiolitis, COPD, healthy, LRTI, and URTI were correctly classified. At last, sensitivity, specificity, and accuracy values of the final optimized model were calculated. We have noticed that the overall accuracy of the final optimized model on the un seen data was 99% with sensitivity of 99.04% and specificity of 99.2%.

_____

4.10. Discussion

In this study, besides to the online lung sound records, we collected some local lung sound records using our electronic stethoscope. An electronic stethoscope has been constructed by modifying the existing acoustic stethoscope and we used it for local lung sound collection. The collection of local lung sounds was done following the legal procedures. Moreover, the collection of local lung sounds was done following the same fashion as the online lung sounds acquired. Therefore, the duration of each local lung sound records is similar to the online lung sound records which ranges from 10 to 90 seconds. Furthermore, the local lung sound records were normalized at a uniform sampling rate of 44.1KHz and bit depth of 16 bit-depth, as of the online lung sound records. In addition, the local lung sound records were acquired from different chest locations such as left and right anterior, left and right posterior, and left and right lateral regions of the chest wall. Since the ICBHI respiratory sound dataset is highly imbalanced, we took only some annotated lung sound records of each class. Besides to the ICBHI respiratory sound dataset, we also took some annotated lung sound records available online. Finally, we combined these online lung sound records with the local lung sound records collected at Jimma university medical center (JUMC). Generally including the online and local lung sound records, a total of 500 lung sound records where; 75 healthy, 70 asthma, 60 pneumonia, 60 URTI, 50 LRTI, 60 COPD, 65 bronchiectasis, and 60 bronchiolitis records were used in this study.

After collection of lung sound signals, wavelet-based denoising of the lung sound signals which involves the three important steps such as decomposition, detail coefficients thresholding, and reconstruction have been done. A 6th-level DWT multiresolution decomposition of lung sound signals have been carried out. Following the decomposition, the DWT-based denoising of lung sounds has been done by applying different wavelet thresholding to remove the noise from the lung sound signals. In this research, four wavelet functions (*Db4, Db10, Sym5, and Sym13*) and four different threshold selection rules (*Rigrsure, Sqtwolog, Heursure, and Minmax*) were used to analyze the required performance of denoising the lung sound signals. After applying a threshold, the effects of noise on the lung sound signals were removed and finally the denoised signals were reconstructed using IDWT.

The SNR results for the different wavelet functions were calculated at different level of decomposition (3rd to 7th level of decomposition). In doing so, the SNR values increased from 3rd to 6th level. But after 6th level of decomposition, the SNR values become decreased. This shows 6th level decomposition is enough to decompose the lung sound signals to get the required information (details and approximation). After this level, decomposing the signals will miss the required information. The algorithm was tested using the above-mentioned most widely used wavelet functions. The tested lung sound signals were noised by the white gaussian noise added at SNR = 5dB as an initial value to test the performance of the denoising algorithm for noise reduction. A total of 240 lung sound signals (30 for each class) were used for the initial test. Moreover, when we increased the tested signals to 400 (50 for each class), the better SNR result was found still at six level of decomposition. After evaluating the performance of the above-mentioned wavelet functions, *Sym13* wavelet function at 6th level decomposition with soft thresholding method and *Sqtwolog* threshold selection rule was found the best algorithm for denoising of the lung sound signals. Following signal acquisition and pre-processing, we extracted features which could be used for final classification of data.

In this research, we applied DWT feature extraction technique to extract important features from the lung sound signals. Fundamentally, the DWT-based feature extraction procedure involves two important steps. First, decompose the given signal into N ($N = 6$ in this study) levels using filtering and decimation. The second step is followed by the reconstruction of the signal using IDWT. The features extracted from the IDWT of signals are considered highly useful features for input of the classifiers due to their effective time-frequency representation of the non-stationary signals. A total of 16 features were extracted from the lung sound signals. Each of the features have been discussed in chapter three under feature extraction section. Subsequently, feature selection by a means of variable or feature ranking was done to select the most discriminative features used for classification. We implemented the one-way ANOVA feature ranking method using $minmax$ normalization scheme. After feature ranking, we got the three features such as skewness, peak value and total harmonic distortion (THD) had lesser score as compared to other features. Therefore, we removed these features and a total of 13 most discriminative features were selected to train different machine learning models.

After feature selection, feature visualization and normalization have been done to avoid bias of the classifier during model training and data classification. Normalization of features has been done by subtracting the mean and dividing the standard deviation of each column. After normalization, the features become on the same scale. The next step was data splitting to prepare the data for model training. In this research, first the dataset has been split using the hold-out method. A common split using the hold-out method is using 80% of data for training and the remaining 20% of data for testing. Hence, we split our dataset into training (80%) and testing (20%) using the function *cvpartition*. A total of 500 lung sound records were used in this study. Since we split the data as 80% of data for training and the remaining 20% of data for testing, the number of lung sound records for training and testing were 400 and 100 respectively. Once we create a cross-validation partition for data, next an 8-fold cross-validation has been applied to the training dataset. Then the training and validation sets should cross-over in successive rounds so that each data point has a chance of being validated against 8 separate times.

After data splitting, the next step was model training and evaluation. Before training and evaluating classification models, it was necessary to convert the label in training and testing datasets from categorical array to string array. Therefore, the label in training and testing data has been converted into string. The final step was to interactively train different classification models. We trained important machine learning models such as Naive Bayes, K-nearest neighbor (KNN), Ensemble and support vector machine (SVM) models for final classification of data. Cubic KNN and Medium KNN algorithms gives an accuracy of 88.3% and 89.3% respectively. KNN classifiers typically have best predictive accuracy in low dimensions but might not in high dimensions. Additionally, KNN classifiers have high memory usage and are not easy to interpret. Gaussian Naive Bayes and Kernel Naive Bayes algorithms gives an accuracy value of 94.8% and 95.8% respectively. Naive Bayes classifiers are easy to interpret and useful for multiclass classification. However, the Naive Bayes algorithms need a small amount of training data for estimating the vital parameters which made the algorithms extremely fast compared to more sophisticated methods. Moreover, ensemble algorithms such as RUSBoosted trees and Bagged trees were trained and provide an accuracy value of 96.5% and 97.3% respectively. Ensemble classifiers tend to be slow to fit because they often need many weak learners.

Moreover, SVM algorithms such as Cubic SVM and Fine Gaussian SVM gives higher classification accuracy of 97.5% and 97.8% respectively. After we trained different machine learning models, we choose the best model with higher classification accuracy (i.e., Fine Gaussian SVM algorithm) and we assessed its performance using confusion matrices. Next, we tried to improve optimized the accuracy of the selected model. There are different ways to improve the accuracy of a machine learning model. These includes reframe the problem, use meaningful features, apply cross-validation, comparing different classification algorithms, and model optimization i.e., hyperparameter tuning. In this research, after collecting enough data, we reframe our data in a suitable way. This is due to the fact that machine learning models are only good as good as our data. It is difficult to fit machine learning models on raw data directly due to some reasons such as implementations require data to be numeric, algorithms impose specific requirements, raw data contains errors, and columns may be redundant or irrelevant. After we prepared and reframe the data, we extracted meaningful features from the lung sound signals and we ranked the features to select the most discriminative features.

After extracting and ranking features, we visualize and normalized the features to have the same scale in order to avoid bias of the classifier. We also applied a cross-validation partition (8-fold cross-validation) to protect overfitting by partitioning the dataset into 8 folds and estimating accuracy on each fold. Moreover, we optimized the selected model using Bayesian optimization technique to improve its accuracy. After, playing with the parameters/hyperparameters using Bayesian optimization technique, the accuracy of our model has been improved to 98.8%. Furthermore, the minimum classification error plot for the optimized model was shown and the optimal model parameter values of box constraint level and kernel scale were found 417.8791 and 24.8176 respectively. Finally, we exported the final optimized model to make predictions for the new data and the overall accuracy of the final model has been reached to 99% with sensitivity of 99.04% and specificity of 99.2%. Table 4.18 shows a comparison between the proposed method and the previous conducted researches.

Table 4.18: A comparison between the proposed method and the previous methods.

S. No	Year	Authors	Applied methods	Results (Accuracy %)
1	2018	Islam, et al., [54].	FFT-Welch and SVM	93.3
2	2017	Kurt, et al., [55].	MFCC and SVM	75.75
3	2019	Haider, et al., [59].	(Median frequency, Linear predictive coefficients) and SVM	83.6
4	2014	Mondal, et al., [58].	(EMD, HT) and ELM	92.86
5	2020	Don [63].	RSFS and KNN	94.1
6	2016	Gogus, et al., [53].	AR-Burg and ANN	89
7	2016	Saha, et al., [56].	Cepstral features and ANN	97.83
8	2018	Rizal, et al., [64].	MSE and MPNN	97.67
9	2018	Nugroho, et al., [65].	MPE and MPNN	97.98
10	2019	Lmtiaz, et al., [66].	CWT and ANN	94.02
11	2017	Aykanat, et al., [67].	CNN and STFT	76
12	2019	Demir, et al., [70].	STFT and SVM STFT and DCNN	65.5 63.09
13	2020	Hosseini, et al., 2020. [69].	DCNN and (Auditory & Demographic info)	83
14	2021	**Proposed method.**	**MRA (i.e., DWT) and SVM**	**99**

From Table 4.18, we can see that different researchers applied different techniques for analysis of lung sounds and further classification of lung diseases.

Most of the researchers used spectral analysis techniques such as FT, FFT-Welch, AR, AR-Burg, ARMA, MFCC, and others [42,55,59,63,56]. These techniques can be used for spectral analysis of lung sound signals because the process of converting time domain to frequency domain can be easily achieved through such methods. However, frequency-domain analysis techniques couldn't provide temporal information of signals. Similarly, time domain analysis techniques such as empirical mode decomposition and multiscale entropy techniques have been used for analysis of lung sounds [58,64,65]. These techniques lost spectral information of signals. Time-frequency domain analysis techniques has been also used in literature for analysis of the non-stationary lung sound signals [66,67,70]. STFT-based analysis techniques provide both time and frequency information of signals but unable to extract multiple features due to single and fixed window region. CWT can also provide time-frequency analysis of signals but a two-dimensional representation for one dimensional entity through the use of a continuous translation and scaling is extremely redundant.

In this research, we applied MRA (i.e., DWT) analysis technique for better time-frequency analysis of the non-stationary lung sound signals. A 6th level decomposition of lung sound signals has been conducted using *Sym13* wavelet function. The performances of the two thresholding methods (soft and hard thresholding methods) have been evaluated using four important threshold selection rules (*Rigrsure, Sqtwolog, Heursure, and Minmax*). Through the use of DWT, we got important features of the signals which we used for final classification of lung diseases. In this study, we trained different important classical machine learning classifiers and we compared their classification performance. Finally, we choose the model with higher classification performance for further improvements. Furthermore, feature scaling for normalization of extracted features, 8-fold cross validation and model optimization using Bayesian optimization technique has been done for improving the accuracy of the final model. The overall accuracy of the model has been reached to 99% during prediction of the new data. Generally, we can conclude that our method is strong enough for the classification of lung diseases as COPD, URTI, LRTI, pneumonia, bronchiectasis, bronchiolitis, asthma, or healthy.

CHAPTER FIVE

CONCIUSION and RECOMMENDATION

5.1. Conclusion

Nowadays, improving a stethoscope-based auscultation diagnosis of pulmonary diseases by combining modern digital signal analysis techniques and AI applications is one of the main research hotspots. In this study, an attempt has been done to analyze the lung sound signals using wavelet multiresolution analysis for further classification of pulmonary diseases.

The study has showed the performance of different wavelet functions along with different thresholding methods and threshold selection rules for further better decomposition and denoising of the non-stationary lung sound signals. The performance of the denoising algorithm was evaluated by computing the SNR values at different level of decomposition. Moreover, DWT-based feature extraction was performed for extraction of different features from the lung sound signals. The study has also showed the features extracted from the IDWT of signals are highly useful features for the input of the classifiers due to their effective time-frequency representation of the non-stationary lung sound signals. At last, feature ranking was done to select the most discriminative features and a total of 13 features were used to train different machine learning models. During classification, we achieved 97.8% accuracy using Fine Gaussian SVM algorithm. Furthermore, we optimized the model by applying Bayesian optimization technique and the accuracy of the model was improved to 98.8%. Finally, the accuracy of the optimized model has been reached to 99% with sensitivity of 99.04% and specificity of 99.2% in making predictions for the new data. Generally, the proposed method delivered a considerable improved result for the classification of most common lung diseases on the base of lung sound analysis.

From this study we can conclude that, wavelet multiresolution analysis can be an ideal solution for better analysis of lung sound signals which can be used for further classification of pulmonary diseases. The proposed method has some major weakness. One of the major weaknesses of this study is the data acquisition system.

In this study, the lung sound signals were acquired using single channel data acquisition system where a single microphone sensor placed over a stethoscope head. Moreover, the study didn't apply the state-of-the-art deep learning techniques.

5.2. Recommendation

The proposed method can provide more reliable way of diagnosis of lung diseases by combing the model with the knowledge base. Therefore, it is recommended to design a user interface (UI) and connect the model proposed in this study with the knowledge base so as to develop an intelligent system to detect the type of lung disease.

Multichannel data acquisition system (which requires high cost) will be helpful to gather more adequate information about the lungs pathology. It can be used for high quality recordings of lung sounds, on the basis of which lung diseases and pathological lung conditions can be assessed more objectively. The accumulation of more information will help to improve better classification performance in combined channel scenario than in single channel scenario. Therefore, it is recommended to construct multichannel lung sound data acquisition system to collect huge data for classification with more classes.

In this research, limited numbers of wavelet functions were used for lung sound analysis. The performance of other wavelet families for best analysis of lung sounds needs to be investigated. The classification performance of other classical machine learning models can also be studied in future. Moreover, we recommend the accuracy of the presented algorithm can be further improved using other deep learning techniques for larger lung sound datasets. Deep learning techniques involve use of large amount of data and complex algorithms which require powerful computation hardware (i.e., computer).

Furthermore, the proposed method is an indication for classification of any abnormalities associated with other bio-sound signals such as Korotkoff sounds, heart sounds, and other internal human body sound signals. At last, we would like to recommend the researchers working in bio-sound signal processing to look for the opportunity of using lung sound signals for classification of Corona virus disease (COVID-19), the current world pandemic issue.

References

[1] Prasad, Bhairav. Chronic obstructive pulmonary disease (COPD). International Journal of Pharmacy Research & Technology. 2020; 10(1): 67-71.

[2] Vestbo J. COPD: definition and phenotypes. Clin Chest Med. 2014 Mar;35(1):1-6. doi: 10.1016/j.ccm.2013.10.010. Epub 2013 Dec 12. PMID: 24507832.

[3] Calverley.P, M.A., Georgopoulos, Dimitris. Chronic obstructive pulmonary disease: Symptoms and Signs. European Respiratory Monograph. 1998; 3(7). doi:10.1183/1025448x.00038002.

[4] Kirenga, B.J., Schwartz, J.I., De Jong, C., Van der Molen, T., & Okot-Nwang, M. (2016). Guidance on the diagnosis and management of asthma among adults in resource limited settings. African Health Sciences, 15(4), 1189. doi:10.4314/ahs.v15i4.18.

[5] Kim, H., Mazza, J. (2011). Asthma. Allergy, Asthma & Clinical Immunology, 7(Suppl 1), s2. doi:10.1186/1710-1492-7-S1-S2.

[6] Cukic, V., Lovre, V., Dragisic, D., Ustamujic, A. Asthma and Chronic obstructive pulmonary disease (COPD)-differences and similarities. Materia Socio Medica. 2012: 24(2): 100-105.

[7] Grief, S.N., & Loza, J.K. (2018). Guidelines for the evaluation and treatment of pneumonia. Primary care: Clinics in Office practice, 45(3), 485-503. doi:10.1016/j.pop.2018.04.001.

[8] Cotton, M., Innes, S., Jaspan, H., Madide, A., Rabie, H. Management of upper respiratory tract infections in children. S Afr Fam Pract. 2008; 50(2): 6-12.

[9] Rohilla, Ankur., Sharma, Vineet., Sonu, S.K. Upper respiratory tract infections: An overview. International Journal of Current Pharmaceutical Research. 2013; 5(3): 1-3.

[10] Mirkarimi, M., Alisamir, M., Saraf, S., Heidri, S., Barouti, S., Mohammad, S. Clinical and epidemiological determinants of lower respiratory tract infections in hospitalized pediatric patients. Int J Pediatrics. 2020 Nov 17;2020:884420. doi: 10.1155/2020/8844420. PMID: 33281906; PMCID: PMC7688345.

[11] Rademacher, J., & Welte, T. (2011). Bronchiectasis. Deutsches Aerzteblatt Online. doi:10.3238/arztebl.2011.0809.

[12] Rocha, B.M., Filos, D., Mendes, L., Vogiatzis, I., Perantoni, E., Kaimakamis, E., Natsiavas, P., Oliveira, A., Jacome, C., Marques, A., Paiva, R.P., Chouvarda, I., Carvalho, P., Maglaveras, N (2017). A respiratory sound database for the development of automated classification. IFMBE Proceedings, 33-37. doi:10.1007/978-981-10-7419-6_6.

[13] Fuster, Valentin. The stethoscope's prognosis. Journal of the American college of cardiology. 2016; 67(6): 118-119.

[14] Sengur, Abdulkadir., Bajaj, Varun., Demir, Fatih. Convolutional neural networks based efficient approach for classification of lung diseases. Health Information Science and Systems. 2019; 8(4).

[15] Andres, E., Gass, R., Charloux, A., Brandt, C., Hentzler, A. Respiratory sound analysis in the area of evidence-based medicine and the world of medicine 2.0. Journal of medicine and life. 2018; 11(2): 89-106.

[16] Hobbs, D.L., M.S, R.S., R, T. Chest radiography for radiologic technologists. Radiologic Technology. 2007; 78(6): 494-515.

[17] Islam, A.M., B, Irin., B, P. Classification of normal, asthma, and COPD subject using multichannel lung sound analysis. IEEE.2018: 0290-0294.

[18] Dukic, L., K, M.Lara., Dorotic, A., Barsic, I. Blood gas testing and related measurements: National recommendations on behalf of the Croatian society of medical biochemistry and laboratory medicine. Biochemia Medica. 2016; 26(3): 318-336.

[19] Altan, G., Kutlu, Y., Garbi, Y., Pekmezci, O.A., Nural, S. Multimedia respiratory database (RespiratoryDatabase@TR): Auscultation sounds and chest x-rays. Natural and Engineering Sciences. 2017; 2(3): 59-72.

[20] Rashid, Al-Hasib., Mazumder, N.A., Mohsenim, T., Hosseini, Morteza., Ren, Haoran., Prakash, B. Neural networks for pulmonary disease diagnosis using auditory and demographic information. arXiv:2011.13194v1 [cs.LG] 26 Nov 2020.

[21] Pramono.A, X.R., Lmtiaz, A.S., V-Rodriguez, Esther. Evaluation of features for classification of wheezes and normal respiratory sounds. PLoS ONE. 2019; 14(3): 1-3.

[22] Naqvi Hassan, Z.S., Choudhry, Ahmad.M (2020). An automated system for classification of chronic obstructive pulmonary disease and pneumonia patients using lung sound analysis. Sensors, 20(22), 6512. doi:10.3390/s20226512.

[23] Woldeamanuel, T.Berhanu., Gebreyesus, H.Legesse. Prevalence and risk factors for chest related symptoms of acute respiratory tract infections among under five children: Case of Ethiopia. Journal of trends in biomedical research (JTBR). 2019; 2(1): 1-7.

[24] Gogus, F.Z., Karhk, B., Harman, G. Identification of Pulmonary Disorders by Using Different Spectral Analysis Methods. International Journal of Computational Intelligence Systems. 2016; 9(4): 595-611.

[25] Islam, M.A., B, Irin., B, Parthasarathi., Saha, Goutam (2018). Classification of normal, asthma, COPD subjects using multichannel lung sound analysis. 2018 International conference on communication and signal processing (ICCSP). doi:10.1109/iccsp.2018.8524439.

[26] Kurt, Bahar., Saryal, Sevgi., et al (2017). Classification of lung sounds using convolutional neural networks. EURASIP Journal on Image and Video Processing. 65 (2017); 1-9.

[27] Haider, N.S., Singh, B.K., Periyasamy, R. et al. Respiratory sound-based classification of chronic obstructive pulmonary disease: A risk stratification approach in machine learning paradigm. Journal of medical systems 43, 255(2019).

[28] Jennings, Wesley G (2015). Experimental Research Design. The Encyclopedia of Crime and Punishment, 1-6. doi:10.1002/9781118519639.wbecpx113.

[29] Tharwat, Alaa (2018). Classification assessment methods. Applied computing and informatics, 17(1): 168-192. S2210832718301546–. doi:10.1016/j.aci.2018.08.003.

[30] Islam, M.S., Paul, G., Ong, H.X., Young, P.M., Gu, Y.T., & Saha, S.C. (2020). A review of respiratory anatomical development, air flow characterization and particle deposition. International journal of environmental research and public health, 17(2), 380. doi:10.3390/ijerph17020380.

[31] Betts, J.Gordon., Wise, James.A., Young, Kelly.A., Johnson, E., Poe, B., Kruse, D.H., Korol, Oksana, Johnson, J.E., Womble, M., DeSaix, Peter. Anatomy and Physiology. Textbook Equity Edition. Huston, Texas77005: OpenStax College; 2013.

[32] Tortora, G., Derrickson, B. Principles of anatomy and physiology. 14th ed. Hoboken, NJ: John Wiley & Sons; 2014.

[33] Lei, Yuan. Medical ventilator system basics: A clinical guide. 3rd ed. Walton Street, Oxford: Oxford University press; 2017.

[34] Kaufman, D.P., Kandle, P.F., Murray, I., Dhamoon, A.S. Physiology, Oxyhemoglobin Dissociation curve. 2020 Jul 26. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. PMID: 29762993.

[35] Andres, Emmanuel., Gass, Raymond., Brandt, Christian., Hentzler, Andre. Respiratory sounds analysis in the world of health 2.0 and medicine 2.0. EC Pulmonology and Respiratory Medicine. 2018; 564 – 585.

[36] Zimmerman, B., Williams, D. Lung sounds. [updated 2020 Sep 2]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021.

[37] Bohadana, A., Izbicki, G., Kraman, S.S. Fundamentals of lung auscultation. New England Journal of Medicine. 2014; 370(8): 744-751.

[38] L.H, Raymond., J.R, Murphy. In defense of the stethoscope. Respiratory Care. 2008; 53(3): 355-369.

[39] Melbye, Hasse. Auscultation of the lungs-still a useful examination? Tidsskrift for Den norske legeforening. 2001; 121(4): 451-452.

[40] Julius, O.Smith. Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications. 2nd ed. Stanford: W3K publishing; 2007.

[41] Welsby, P.D., Earis, J.E. Some high-pitched thoughts on chest examination. Postgrad Med J. 2001; 77: 617- 620.

[42] Sarkar, M., Madabhavi, I., Niranjan, N., Dogra, M. Auscultation of the respiratory system. Annals of thoracic medicine. 2015; 158 – 168.

[43] A.J, Bates., R.Cetto, H., H.Calmet, A.M., Gambaruto, N.S., Tolley, G. Houzeaux., R.C, Schroter., D.J, Doorly. Dynamics of airflow in a short inhalation. Journal of the royal society interface. 12(102): 20140880, 2015.

[44] Kandaswamy, A., Kumar, S.C., et al. Neural classification of lung sounds using wavelet coefficients. Computers in Biology and Medicine. 2004; 34(6): 523-537.

[45] Ambaye, G.A. Time and frequency domain analysis of signals: A review. International journal of engineering research & technology (IJRET). 2020; 9(12): 271-276.

[46] Prof. Brad, Osgood. Electrical engineering department, Stanford University, Stanford. The Fourier Transform and its Applications. Stanford Engineering Everywhere EE261. 2016; 3: 65-86.

[47] Chen, Hasing-Chin., Huang, Tzeng-Wen., Tan, Hasu-Tan., Chang, Chun-Cheng., Chang, Jen-Yuan. (2015). Using K-Nearest Neighbor classification to diagnose abnormal lung sounds. Sensors, 15(6), 13132-13158. doi:10.3390/s150613132.

[48] Aydin, Seda Guzel., Kaya, Turgay., Guler, Hasan. Wavelet-based study of valence-arousal model of emotions on EEG signals with LabVIEW. Brain informatics. 2016; 3: 109-117.

[49] Rioul, Olivier., Vetterli, Martin (1991). Wavelets and Signal Processing. IEEE signal processing magazine, 8(4), 14-38. doi:10.1109/79.91217.

[50] Wyse, Lonce. Audio spectrogram representations for processing with convolutional neural networks. Proceedings of the first international workshop on deep learning and music joint with IJCNN. 2017; 1(1): 37-41.

[51] Hariharan, G. Wavelet analysis -an overview. In: wavelet solutions for reaction-diffusion problems in science and engineering. Forum for interdisciplinary mathematics. Singapore: Springer, 2019: p. 15-31.

[52] Debnath, L. (Ed.). (2003). Wavelets and Signal Processing. Journal of the Indian Institute of Science. doi:10.1007/978-1-4612-0025-3.

[53] Gogus, F.Z., Karhk, B., Harman, G. Identification of Pulmonary Disorders by Using Different Spectral Analysis Methods. International Journal of Computational Intelligence Systems. 2016; 9(4): 595-611.

[54] Islam, M.A., B, Irin., B, Parthasarathi., Saha, Goutam (2018). Classification of normal, asthma, COPD subjects using multichannel lung sound analysis. 2018 International conference on communication and signal processing (ICCSP). doi:10.1109/iccsp.2018.8524439.

[55] Kurt, Bahar., Saryal, Sevgi., et al (2017). Classification of lung sounds using convolutional neural networks. EURASIP Journal on Image and Video Processing. 65 (2017); 1-9.

[56] Saha, G., S, Md., Sengupta, N. Lung sound classification using cepstral-based statistical features. Computers in Biology and Medicine. 2016; 75(1): 118-129.

[57] Choudhry, M.A., Khan, A.Z. et al. Intelligent system for classification of pulmonary diseases from lung sound. 13th International conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS). 2019; 1-6.

[58] Mondal, Ashok., Bhattacharya, P., Saha, Goutam. (2014). Detection of lungs status using morphological complexities of respiratory sounds. The scientific world journal, 2014, 1-9. doi:10.1155/2014/182938.

[59] Haider, N.S., Singh, B.K., Periyasamy, R. et al. Respiratory sound-based classification of chronic obstructive pulmonary disease: A risk stratification approach in machine learning paradigm. Journal of medical systems 43, 255(2019).

[60] Shah, Chintan., Vora, Siddhi. COPD classification using machine learning algorithms. International research journal of engineering and technology (IRJET). 2019; 6(4): 1-4.

[61] Hidayat, R., Nugroho, H.A., Rizal, A. Fractal dimension for lung sound classification in multiscale scheme. Journal of computer science. 2018; 14(8): 1081-1096.

[62] Flietstra, B., Markuzon, N., Vyshedskiy, A., Murphy, R. (2011). Automated analysis of crackles in patients with interstitial pulmonary fibrosis. Pulmonary medicine, 2011, 1-7. doi:10.1155/2011/590506.

[63] Don, S. Random subset feature selection and classification of lung sound. International conference on computational intelligence and data science (ICCIDS 2019). 2020; 167: 313-322.

[64] Rizal, A., Hidayat, R., Nugroho, H.A. Multiscale tsallis entropy for pulmonary crackle detection. International journal of advances intelligent informatics. 2018; 4(3): 192-201.

[65] Nugroho, H.A. et al. Comparison of multiscale entropy techniques for lung sound classification. Indonesian journal of electrical engineering and computer science. 2018; 12(3): 984-994.

[66] Pramono, RXA., Lmtiaz, S.A., Rodriguez-Villegas, E. Evaluation of features for classification of wheezes and normal respiratory sounds (2019). PLoS ONE 14(3): https://doi.org/10.1371/journal.pone.0213659.

[67] Aykanat, Murat., Kilic, Ozkan., Kurt, Bahar., Saryal, Sevgi. Classification of lung sounds using convolutional neural networks. EURASIP Journal on Image and Video Processing.2017; 1-9.

[68] Do, Quan., S.C, Tran., Chaudri, Jamil. Classifcation of asthma severity and medication using TensorFlow and multilevel databases. The 7th international conference on current and future trends of information and communication technologies in healthcare (ICTH 2017). 2017; 113(2017): 344-351.

[69] Hosseini, Morteza., Ren, Haoran., Rashid, Al-Hasib., Mazumder, N.A., Prakash, B., Mohsenim, T. Neural networks for pulmonary disease diagnosis using auditory and demographic information. arXiv:2011.13194 [cs.LG] 26 Nov 2020.

[70] Demir, Fatih., Sengur, Abdulkadir., Bajaj, Varun. Convolutional neural networks based efficient approach for classification of lung diseases. Health Information Science and Systems. 2019; 8(4).

[71] Fraiwan, M., Fraiwan, L., Alkhodari, M. et al. Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory units. J Ambient Intell Human comput (2021). https://doi.org/10.1007/s12652-021-03184-y.

[72] Guler, Hasan., Kaya, Turgay., Aydin, S.G. Wavelet-based study of valence-arousal model of emotions on EEG signals with LabVIEW. Brain Informatics. 2016; 3: 109-117.

[73] Permin, Henrik., Norn, Svend. Stethoscope-Over 200 years. Journal of Pulmonary and Respiratory Research. 2019; 3: 001-008. doi:10.29328/journal.jprr.1001010.

[74] Mangion, Kenneth. The stethoscope. Malta Medical Journal. 2007; 19(2): 41-44.

[75] Swarup, S., & Makaryus, A. (2018). Digital stethoscope: technology update. Medical Devices: Evidence and Research, volume 11, 29-36. doi:10.2147/mder.s135882.

[76] Rao, Adam. University of California, San Francisco. Acoustic methods of pulmonary diseases diagnosis. ProQuest Dissertations Publishing, 2019. 13859752.

[77] Rocha, B.M., Filos, D., Mendes, L., et al. A respiratory sound database for the development of automated classification. In proceedings of the international conference on Biomedical and Health Informatics, Thessaloniki, Greece, 2017. 33-37.

[78] Valencia, Daniel., Orejuela, David., Salazar, Jeferson., Valencia, Jose. (2016). Comparison analysis between rigrsure, sqtwolog, heursure and minimaxi techniques using hard and soft thresholding methods. 2016 XXI symposiums on signal processing, images and artificial vision (STSIVA). doi:10.1109/stsiva.2016.7743309.

[79] Verma, Neema., Verma, K.A. Performance analysis of wavelet thresholding methods in denoising of audio signals of some Indian musical instruments. International journal of engineering science and technology (IJEST). 2012; 4(05): 2027-2052.

[80] Devnath, Liton., Islam, Rafiqul. Selection of wavelet and thresholding rule for denoising the ECG signals. Annals of pure and applied mathematics. 2015; 10(1): 65-73.

[81] Chang, C.G., Cheng, P.Y. Investigation of noise effect on lung sound recognition. Proceedings of the seventh international conference on machine learning any cybernetics, Kunming. 2008; 1-14.

[82] Akkar, H.A.R., Al-Dosari, I.H.M., & H, W.A. H. A novel wavelet thresholding method for implementation of signal compressor design. Proceedings of the international conference on information and communication technology-ICICT'19. doi:10.1145/3321289.3321292.

---

[83] Sharma, Garima., Umapathy, Kartikeyan., Krishnan, Sridhar. (2020). Trends in audio signal feature extraction methods. Applied acoustics, 158, 107020. doi:10.1016/j.apacoust.2019.107020.

[84] Ghojogh, B., Samad, N.M., Mahhadi, A.S., Kapoor, T., Ali, W., Karray, C.M. Feature selection and feature extraction in pattern analysis: Literature review. arXiv:1905.02845 [cs.LG] 7 May 2019.

[85] Biswas, S.Kr., Bordoloi, M., & Purkayastha, B. (2016). Review on feature selection and classification using neuro-fuzzy approaches. International journal of applied evolutionary computation, 7(4): 28-44. doi:10.4018/IJAEC.2016100102.

[86] Omer Fadl Elssied, N., Ibrahim, O., Hamza Osman, A. (2014). A novel feature selection based on one-way ANOVA F-Test for E-mail Spam classification. Research Journal of Applied Sciences, Engineering and Technology, 7(3), 625-638. doi:10.19026/rjaset.7.299.

[87] D, Bollegala. Dynamic feature scaling for online learning of binary classifiers. arXiv:1407.7584 [cs.LG] 28 Jul 2014.

[88] Y, F.O., T, J.E.A., O, Awodele., O., J.H., O, Olakanmi., J, Akinjobi. Supervised machine learning algorithms: Classification and comparison. International journal of computer trends and technology (IJCTT). 2017: 2-12.

[89] Alsheref, K.F., Gomaa, H.W. Blood diseases detection using classical machine learning algorithms. (IJACSA) international journal of advanced computer science and applications. 2019; 7(10): 77-81.

[90] Awad, M., Khanna, R. (2015). Support vector machines for classification. In: Efficient learning machines. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4302-5990-9_3.

[91] Gambella, C., Ghaddar, B., & Naoum-Swaya, J. (2020). Optimization problems for machine learning: A survey. European journal of operational research. doi:10.1016/j.ejor.2020.08.045.

Annex

Sample MATLAB source codes written for some specific tasks

```matlab
## The following is part of the MATLAB script written for
structuring the data holding the lung sounds to categories in
folder.

clear
close all
clc
warning("off")

% structure the data into categories in folder
% import the dataset of each class
datafolder = "Respiratory_Sound_Database/audio";

% Labels Class/Categories
Data_Class =
categorical(["COPD","URTI","LRTI","Pneumonia","Bronchiectasis","
Bronchiolitis","Asthma","Healthy"]);

for i=1:1:length(Data_Class)

mkdir(strcat("Respiratory_Sound_Database\structure1\",string(Dat
a_Class(i))));
end
currentfolder = pwd;
cd(datafolder);
listdir=dir;
h=1;
L=1;
for i=4:1:length(listdir)
    cd(listdir(i).name)
    inside=dir;
for k=4:1:length(inside)
    index=strfind(inside(k).name,'-');
    Labels(h)=Data_Class(str2num(inside(k).name(index(1)+1))+1);
movefile(inside(k).name,strcat(currentfolder,"\Respiratory_Sound
_Database\structure1\",string(Labels(h))));
    h=h+1;
end
    cd(strcat(currentfolder,'\',datafolder));
end
```

The following is part of the MATLAB script written to visualize and listen the lung sounds data (Wav. file).

```matlab
% Visualize and Listen the Data (Wav. file)
figure()
datafolder = "Respiratory_Sound_Database/structure1";
currentfolder = pwd;
cd(datafolder);
listdir=dir;
for i=3:1:length(listdir)
    cd(listdir(i).name)
    inside=dir;
    subplot(3,4,i-2);
    [y,fs]=audioread(inside(4).name);
    plot(y(:,:));
    soundsc(y(:,:),fs);
    grid on;
    title(listdir(i).name)
    drawnow;
    pause(2)
    cd(strcat(currentfolder,'\',datafolder));
end
```

The following is part of the MATLAB script written for understanding the distribution of dataset, where only two commands (audioDatastore and countEachLabel) were used.

```matlab
% look the distribution of dataset
datafolder = "Respiratory_Sound_Database\structure1";

ads = audioDatastore(datafolder, ...
'IncludeSubfolders',true, ...
'FileExtensions','.wav', ...
'LabelSource','foldernames');

% count each label/class/category
G1=countEachLabel(ads);
```

The following is part of the MATLAB script written for checking how much percentage of data are on a specific sampling rate and bit-depth, and number of channels.

```matlab
% Check how much percentage of the data are on a specific
Sampling Rate
for i=1:1:length(ads.Files)
    [y,fs(i)]=audioread(ads.Files{i});
end
Overall
=table(categories(categorical(fs)),countcats(categorical(fs))');
Overall.Properties.VariableNames{1}='Frequency';
Overall.Properties.VariableNames{2}='Occurences';

% Check how much percentage of the data are recorded on a
specific Bit depth
for i=1:1:length(ads.Files)
    info = audioinfo(ads.Files{i});
if strcmp(info.CompressionMethod,'Uncompressed')
        Bits(i)=info.BitsPerSample;
else
        Bits(i)=0;
end
end
% Check how much percentage of the data are Mono / Stereo (i.e.,
number of channels)
mono=1;
stereo=1;
for i=1:1:length(ads.Files)
    y=audioread(ads.Files{i});
if size(y,2) == 1
        mono=mono+1;
else
        stereo=stereo+1;
end
end

## The following is part of the MATLAB script written for
normalizing our data in terms of sampling rate and bit-depth.

%To doup convert or down convert
%From the information in the previous steps choose which what
would you do?

for i=1:1:length(ads.Files)
    y =audioread(ads.Files{i});
if size(y,2) == 1
        result = y;
else
        result = (y(:,1)+y(:,2))/2;
end
```

```matlab
    audiowrite(ads.Files{i},result,44100,'BitsPerSample',16);
    clear yFsresult
end
% Let visualize and listen to the data again
figure()
datafolder = "Respiratory_Sound_Database/structure";
currentfolder = pwd;
cd(datafolder);
listdir=dir;
for i=3:1:length(listdir)
    cd(listdir(i).name)
    inside=dir;
    subplot(3,4,i-2);
    [y,fs]=audioread(inside(randi([44,100])).name);
    plot(y(:,:));
    soundsc(y(:,:),fs);
    grid on;
    title(listdir(i).name)
    drawnow;
    pause(2)
    cd(strcat(currentfolder,'\',datafolder));
end
```

The following is part of the MATLAB script written for denoising lung sound signals using discrete wavelet transform (DWT).

```matlab
clear, clc, close all

% Denoising lung sound Signal using discrete wavelet transform
(DWT)
%
% The denoising procedure proceeds in three steps:
%[1]
% Decomposition. select a wavelet (i.e. mother wavelet), and
choose a decomposition level (N)
% Then Compute the wavelet decomposition of the signal s at
level N.
%

%[2]
% Detail coefficients thresholding. For each level from 1 to N,
select
% a threshold and apply soft thresholding to the detail
coefficients.
%
```

```matlab
%[3]
% Reconstruction. Compute wavelet reconstruction based on the
original
% approximation coefficients of level N and the modified detail
coefficients
% of levels from 1 to N.

fprintf('--- Denoising lung sound signal using discrete wavelet
transform ---\n\n');

% load sample of lung sound
fprintf('-> Step 1/6: Load "track.wav" - ');
[originalsignal, fs] = audioread('track.wav');
amp = 20;
originalsignal = amp*originalsignal;
N = length(originalsignal);
fprintf('OK\n');

%---------------------------%
%  add white Gaussian noise  %
%---------------------------%
fprintf('-> Step 2/6: Add white Gaussian noise - ');
% the scalar SNR specifies the signal-to-noise ratio per sample,
in dB
sn = 5;
% add white Gaussian noise to a signal
originalsignalN = awgn(originalsignal,sn,'measured');
fprintf('OK\n');

%---------------------------%
%            DWT :           %
%    wavelet decomposition   %
%---------------------------%
fprintf('-> Step 3/6: Decompose the lung sound signal - ');
level = 6;
fprintf('\n   Input the number of specific wavelet: (1) db4, (2)
db10, (3) sym5 or (4) sym13');
wname = input('\n   wname = ');
if wname == 1
    wt = 'db4';
elseif wname == 2
    wt = 'db10';
elseif wname == 3
    wt = 'sym5';
elseif wname == 4
    wt = 'sym13';
end
```

```matlab
% computes four filters
[Lo_D,Hi_D,Lo_R,Hi_R] = wfilters(wt);
[C,L] = wavedec(originalsignalN,level,Lo_D,Hi_D);
cA6 = appcoef(C,L,wt,level);
% extract the levels 6, 5, 4, 3, 2, and 1 detail coefficients
from C
[cD1,cD2,cD3,cD4,cD5,cD6] = detcoef(C,L,[1,2,3,4,5,6]);
% reconstruct the level 6 approximation from C
A6 = wrcoef('a',C,L,Lo_R,Hi_R,level);
% reconstruct the details at levels 1, 2, 3, 4, 5, and 6, from C
D1 = wrcoef('d',C,L,Lo_R,Hi_R,1);
D2 = wrcoef('d',C,L,Lo_R,Hi_R,2);
D3 = wrcoef('d',C,L,Lo_R,Hi_R,3);
D4 = wrcoef('d',C,L,Lo_R,Hi_R,4);
D5 = wrcoef('d',C,L,Lo_R,Hi_R,5);
D6 = wrcoef('d',C,L,Lo_R,Hi_R,6);
% a = approximation
% d = detail
fprintf('OK\n');

%---------------------------%
%          thresholding        %
%---------------------------%
fprintf('-> Step 4/6: Thresholding - ');
% TPTR = 'rigrsure', adaptive threshold selection using
principle of Stein's
% Unbiased Risk Estimate
% TPTR = 'heursure', heuristic variant of the first option
% TPTR = 'sqtwolog', threshold is sqrt(2*log(length(X)))
% TPTR = 'minimaxi', minimax thresholding
fprintf('\n   Input the number of threshold selection rule : (1)
heursure, (2) rigrsure, (3) minimaxi or (4) sqtwolog');
tr = input('\n   threshold selection rule = ');
if tr == 1
    tptr = 'heursure';
elseif tr == 2
    tptr = 'rigrsure';
elseif tr == 3
    tptr = 'minimaxi';

elseif tr == 4
    tptr = 'sqtwolog' ;
end
thr_D1 = thselect(D1,tptr);
thr_D2 = thselect(D2,tptr);
thr_D3 = thselect(D3,tptr);
thr_D4 = thselect(D4,tptr);
```

```matlab
thr_D5 = thselect(D5,tptr);
thr_D6 = thselect(D6,tptr);
% Hard thresholding is the simplest method but soft thresholding
has nice
% mathematical properties. Hard threshold signal is x if x>thr,
and is 0 if
% x<=thr. And the soft threshold signal is sign(x)(x-thr) if
x>thr and is 0
% if x<=thr
fprintf('\n   Input the number of threshold type: (1) soft or
(2) hard');
sh = input('\n   threshold = ');
if sh == 1
    sorh = 's';
elseif sh == 2
    sorh = 'h';
end
% threshold coefficient of details
tD1 = wthresh(D1,sorh,thr_D1);
tD2 = wthresh(D2,sorh,thr_D2);
tD3 = wthresh(D3,sorh,thr_D3);
tD4 = wthresh(D4,sorh,thr_D4);
tD5 = wthresh(D5,sorh,thr_D5);
tD6 = wthresh(D6,sorh,thr_D6);
fprintf('OK\n');

%--------------------------%
%   compute Inverse DWT    %
%--------------------------%
fprintf('-> Step 5/6: Compute Inverse DWT - ');
denoised = A6 + tD1 + tD2 + tD3 + tD4 + tD5 + tD6;
err = max(abs(originalsignalN-denoised));
fprintf('OK\n');

%--------------------------%
%       compute SNR        %
%--------------------------%
fprintf('-> Step 6/6: Compute SNR - ');
% SNR - Signal to Noise Ratio
SNR = snr(originalsignal,originalsignalN);
NoisySNR = 20*log10(norm(originalsignal(:)) / norm
(originalsignal(:)-originalsignalN(:)) );
SNR = snr(originalsignal,denoised);
DenoisedSNR = 20*log10(norm(originalsignal(:)) / norm
(originalsignal(:)-denoised(:)) );
fprintf('OK\n');
```

```matlab
%---------------    Display Figures   -----------------------
------------

figure(1)
subplot(6,1,1); plot(originalsignal); title('Original lung sound
signal');
subplot(6,2,3); plot(A6); title('Approximation A6')
subplot(6,2,4); plot(D6); title('Detail D6')
subplot(6,2,5); plot(D5); title('Detail D5')
subplot(6,2,6); plot(D4); title('Detail D4')
subplot(6,2,7); plot(D3); title('Detail D3')
subplot(6,2,8); plot(D2); title('Detail D2')
subplot(6,2,9); plot(D1); title('Detail D1')
figure(2)
subplot(6,1,1); plot(originalsignal); title('Original lung sound
signal');
subplot(6,2,3); plot(A6); title('Approximation A6')
subplot(6,2,4); plot(tD6); title('Denoised Detail D6')
subplot(6,2,5); plot(tD5); title('Denoised Detail D5')
subplot(6,2,6); plot(tD4); title('Denoised Detail D4')
subplot(6,2,7); plot(tD3); title('Denoised Detail D3')
subplot(6,2,8); plot(tD2); title('Denoised Detail D2')
subplot(6,2,9); plot(tD1); title('Denoised Detail D1')
% display the comparison of original signal, noisy signal, and
denoised signal
figure(3)
subplot(3,1,1); plot(originalsignal); title('Original lung sound
signal');
xlabel('Samples'); ylabel('Amplitude');
subplot(3,1,2); plot(originalsignalN); title('Noisy lung sound
signal');
xlabel('Samples'); ylabel('Amplitude');
subplot(3,1,3); plot(denoised); title('Denoised lung sound
signal');
xlabel('Samples'); ylabel('Amplitude');
figure(4)
subplot(1,3,1); specgram(originalsignal,512,fs); title('Original
lung sound signal');
subplot(1,3,2); specgram(originalsignalN,512,fs); title('Noisy
lung sound signal');
subplot(1,3,3); specgram(denoised,512,fs); title('Denoised lung
sound signal');
```

Jimma University, Jimma Institute of Technology, School of BME, Bio-Instrumentation Stream

```matlab
## The following is part of the MATLAB script written for
visualizing and normalizing the extracted features.

clear, clc, close all
% the visualization will give me some insight on how to increase
the
% accuracy
tablefordiagnostic = FeatureTable1;
head(FeatureTable1);

% now, normalize the features by substracting the mean and
dividing the standard
% deviation of each column
FeatureTable1{:,1:13} =
normalize(FeatureTable1{:,1:13},'range');
head((FeatureTable1));

## The following is part of the MATLAB script written for
splitting the dataset into training, validation and test sets.

clear, clc, close all
% creat cross-validation partition for data.

rng(1)
cvp = cvpartition(FeatureTable1.Label,'HoldOut',0.2);

dataTrain_Validation = FeatureTable1(cvp.training,:);
dataTest = FeatureTable1(cvp.test,:);


## The following is part of the MATLAB script written for
converting the label in dataTrain-Validation and dataTest from
categorical to string array before training the model.
clear, clc, close all
% convert the lable in dataTrain-Validation and dataTest from
categorical to string.

dataTrain_Validation{:,15}=string(dataTrain_Validation{:,14});
dataTrain_Validation(:,14)=[];
dataTrain_Validation.Properties.VariableNames{14} = 'Label';

dataTest{:,15}=string(dataTest{:,14});
dataTest(:,14)=[];
dataTest.Properties.VariableNames{14} = 'Label';
```