RSET
RAJAGIRI SCHOOL OF
ENGINEERING & TECHNOLOGY
(AUTONOMOUS)

Project Phase 2 Report On

Lung Disease Detection using Respiratory Sounds

Submitted in partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology

in

Computer Science and Engineering

By

Abu Jose (U2003008)
Azmina Iqbal (U2003054)
Cathrin Raju (U2003057)
Dona Francis (U2003072)

Under the guidance of

Ms. Dincy Paul

Department of Computer Science and Engineering
Rajagiri School of Engineering & Technology (Autonomous)
(Parent University: APJ Abdul Kalam Technological University)
Rajagiri Valley, Kakkanad, Kochi, 682039
April 2024

CERTIFICATE

This is to certify that the project report entitled ***"Lung Disease Detection using Respiratory Sounds"*** *is a bonafide record of the work done by* ***Abu Jose (U2003008), Azmina Iqbal (U2003054), Cathrin Raju (U2003057), Dona Francis (U2003072)****, submitted to the Rajagiri School of Engineering & Technology (RSET) (Autonomous) in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology (B. Tech.) in Computer Science and Engineering during the academic year 2023-2024.*

Ms. Dincy Paul
Project Guide
Assistant Professor
Dept. of CSE
RSET

Dr. Sminu Izudheen
Project Coordinator
Professor
Dept. of CSE
RSET

Dr. Preetha K G
Head of Department
Professor
Dept. of CSE
RSET

ACKNOWLEDGMENT

We wish to express our sincere gratitude towards **Prof.(Dr). P. S. Sreejith**, Principal of RSET, and **Dr. Preetha K G**, Head of the Department of ”Department of Computer Science” for providing us with the opportunity to undertake our project, ”Lung Disease Detection using Respiratory Sounds”.

We are highly indebted to our project coordinator, **Dr. Sminu Izudheen**, Professor, Department of Computer Science, for her valuable support.

It is indeed our pleasure and a moment of satisfaction for us to express our sincere gratitude to our project guide **Ms. Dincy Paul** for her patience and all the priceless advice and wisdom she has shared with us.

Last but not the least, we would like to express our sincere gratitude towards all other teachers and friends for their continuous support and constructive ideas.

Abu Jose
Azmina Iqbal
Cathrin Raju
Dona Francis

Abstract

Pulmonary diseases represent a significant global health challenge, with conditions such as pneumonia, chronic obstructive pulmonary disease (COPD), and asthma imposing substantial morbidity and mortality. Timely detection and intervention are critical for effective management. This project presents an innovative approach for pulmonary disease detection, leveraging deep learning techniques to analyze respiratory sounds acquired from electronic stethoscope recordings.

Our research harnesses the power of deep learning, specifically Convolutional neural network (CNN) and Gated Recurrent Unit (GRU) , to analyze and classify respiratory sounds captured by electronic stethoscopes. We've assembled a diverse and extensive dataset comprising respiratory sounds from both healthy individuals and patients with various pulmonary conditions. Signal augmentation and pre-processing techniques such as noise addition, time stretching and pitch shifting have been employed to prevent overfitting. Feature extraction using methods such as mean values of Mel-frequency cepstral coefficients (MFCC), zero-crossing rate, chromagram, root mean square energy, and mel spectrogram for extraction of audio features from the original audio files and their augmented versions which have been tailored to the specific needs of deep learning models, ensuring the preservation of essential acoustic characteristics.

The aim of this project is to develop a robust and accurate CNN-GRU model for the classification of respiratory sounds into different disease categories.

This project represents a promising avenue for advancing the accessibility, early diagnosis and monitoring of pulmonary diseases, driven by the power of deep learning. It holds the potential to significantly reduce initial healthcare tool installation costs and enhance patient outcomes. By harnessing the capabilities of electronic stethoscopes and cutting-edge deep learning technology, this research serves as a crucial step toward creating a non-invasive, accessible tool for healthcare practitioners to revolutionize the detection

and management of pulmonary conditions.

Contents

List of Abbreviations

- ALSDNet - Automatic Lung Sound Diagnosis Network
- CNN – Convolutional Neural Networks
- COPD – Chronic Obstructive Pulmonary Disease
- CRUD - Create, Read, Update, Delete
- CT – Computed Tomography
- DAE – De Noising Autoencoder
- DCNN – Deep Convolutional Neural Network
- DCT – Discrete Cosine Transform
- EMD – Empirical Mode Decomposition
- FFT – Fast Fourier Transform
- GPU – Graphics Processing Unit
- GRU – Gated Recurrent Unit
- GUI - Graphical User Interface
- GTCC – Gamma Tone Cepstral Coefficients
- HIPAA – Health Insurance Portability and Accountability Act
- ICBHI – International Conference on Biomedical and Health Informatics
- IMF – Intrinsic Mode Functions
- IMFCC – Improved Mel Frequency Cepstral Coefficients
- KAUH – King Abdullah University Hospital

- LSTM – Long Short Term Memory
- MFCC – Mel Frequency Cepstral Coefficients
- ORM - Object-Relational Mapping
- RAM – Random Access Memory
- RELU – Rectified Linear Unit
- URTI – Upper Respiratory Tract Infection
- WHO – World Health Organization

List of Figures

Chapter 1

Introduction

1.1 Background

The primary objective of this project is to create an innovative lung disease detection tool leveraging advancements in technology. This tool aims to accurately identify various pulmonary diseases using audio signals obtained from patients through an electronic stethoscope. By harnessing the power of machine learning algorithms and signal processing techniques, the tool seeks to efficiently analyze these audio signals to detect respiratory ailments. The significance of this endeavor lies in its potential to revolutionize healthcare accessibility, particularly in rural or underserved areas where access to specialized medical resources like Computed Tomography (CT) Imaging or Spirometry Data Analysis is limited. This initiative not only addresses the challenges of geographical barriers but also aligns with the crucial need for timely intervention and early detection in managing lung-related health issues. The overarching goal is to democratize healthcare by providing an affordable and efficient means of diagnosing pulmonary diseases, ensuring that individuals in remote or financially constrained regions receive the necessary medical attention promptly. .

1.2 Problem Definition

This project aims to develop a model using deep learning for the detection of respiratory diseases like pneumonia, COPD (Chronic Obstructive Pulmonary Disease), bronchiolitis, URTI from respiratory sound obtained using an electronic stethoscope.

1.3 Scope and Motivation

This project delves into the realm of respiratory sound classification using advanced machine learning techniques, primarily focusing on Gated Recurrent Units (GRUs). The scope encompasses comprehensive data preprocessing, feature engineering, and model development to classify respiratory sounds into distinct categories, aiding in the diagnosis of respiratory conditions. Leveraging extensive audio processing libraries such as Librosa, the project aims to extract meaningful features from raw sound data, augmenting it to enhance model robustness. The model architecture, a CNN-GRU model, is designed to learn intricate patterns within respiratory sounds, aiming to accurately categorize diverse respiratory conditions.

The motivation behind this project lies in the critical need for accurate and efficient diagnostic tools for respiratory conditions. By harnessing the power of machine learning and signal processing, the project seeks to contribute to the field of healthcare by offering a non-invasive means of diagnosing respiratory disorders through sound analysis. Accurate classification of respiratory sounds can significantly aid healthcare professionals in timely and precise diagnoses, potentially leading to improved patient care and treatment strategies. Additionally, addressing class imbalances in the dataset and employing rigorous model evaluation techniques aim to ensure the reliability and practical applicability of the developed solution in real-world healthcare settings.

1.4 Objectives

- Data Preparation: Clean and organize raw respiratory sound data.
- Feature Extraction: Extract key audio features and augment data for diversity.
- Model Build: Develop a specialized GRU model(uses CNN-GRU model) for classification.
- Balance Classes: Address imbalances in disease representation.
- Evaluate Performance: Assess model accuracy and robustness.

- Clinical Impact: Aim for practical diagnostic applicability in healthcare.

1.5 Challenges

The project confronts few challenges that could impact its success. Firstly, the quality and quantity of respiratory sound data pose potential hurdles, as incomplete or noisy data might hinder accurate feature extraction and subsequent model training. Additionally, the class imbalance within the dataset could bias the model, affecting its ability to effectively classify various respiratory conditions. Extracting informative features from audio data while avoiding noise introduction is a complex task that demands careful handling. Furthermore, the risk of over-fitting looms large, particularly with complex models like GRU and CNN-GRU models, emphasizing the need for effective regularization techniques. Finally, ensuring the model's predictions align with clinical interpretations and diagnoses is vital for its practical applicability in healthcare, necessitating thorough validation aligned with medical expertise.

1.6 Assumptions

1. Feature Relevance: The project assumes that the selected audio features (e.g., MFCC, chroma features) extracted from respiratory sound data are informative and crucial for distinguishing between different respiratory conditions.

2. Data Quality: There's an assumption that the respiratory sound dataset is representative and of sufficient quality, containing accurate recordings without substantial artifacts or inconsistencies that could negatively impact the model's performance.

3. Model Suitability: The chosen CNN-GRU architecture is assumed to be suitable for capturing patterns within respiratory sounds and capable of learning and generalizing well across various respiratory conditions.

4. Augmentation Impact: The augmentation techniques applied to the data (such as noise addition, time-stretching, pitch-shift) are assumed to enhance the diversity and richness of the dataset.

1.7 Societal / Industrial Relevance

This project holds significant industrial and societal relevance by revolutionizing respiratory health diagnostics. Its accurate classification of respiratory conditions from sound data promises immense value in clinical settings, potentially streamlining diagnoses and offering a supplementary tool for healthcare practitioners. Moreover, in the realm of telemedicine, this innovation could greatly aid remote diagnosis, expanding access to quality healthcare, especially in underserved areas. Industries dealing with hazardous respiratory environments stand to benefit, as continuous monitoring of workers' respiratory health could enhance workplace safety. On a societal level, the project's ability to swiftly identify respiratory issues could bolster epidemiological surveillance, aiding in early outbreak detection and response. This advancement could optimize healthcare resources, potentially reducing costs and improving resource allocation. Finally, its success could fuel further research and collaboration between technology and healthcare sectors, fostering innovation in AI-driven diagnostic tools and medical technology.

1.8 Organization of the Report

The project roadmap encompasses several key stages. It initiates with an Introduction, outlining the project's scope and the significance of automated respiratory condition classification. Following this, the Data Collection and Preprocessing section delve into details about the dataset, exploration insights, and preprocessing steps, including data cleaning and augmentation techniques. Feature Engineering elucidates the extracted audio features and their relevance, accompanied by an explanation of augmentation methods used for data enrichment. The subsequent section, Model Development, articulates the selected CNN-GRU architecture, training strategies, and regularization techniques employed. Results and Evaluation display performance metrics, like accuracy and precision, through classification reports and confusion matrices for model assessment. The Discussion interprets model outcomes, delving into their clinical implications and addressing encountered limitations. Finally, the Conclusion and Future Directions summarize findings, propose potential research avenues, while the Industrial and Societal Implications section analyzes the project's broader impacts on healthcare, industries, and public health.

This chapter sets the stage by articulating the project's overarching aim: the development of an automated system for classifying respiratory conditions from sound data. The scope delineates the project's boundaries, outlining the dataset, methodologies, and intended outcomes. Motivation underscores the critical need for accurate and swift diagnosis in respiratory health, highlighting gaps in current diagnostic approaches that this project aims to address. The objectives succinctly lay out the project's goals, emphasizing the creation of a reliable classification model. Challenges are acknowledged, acknowledging hurdles like data variability, model complexity, and potential limitations in available resources. Lastly, the practical relevance underscores the project's significance in healthcare advancement, its potential societal impact, and the industrial implications, setting the tone for an exploration that intertwines technology, health, and societal welfare. Overall, this chapter serves as a compass, orienting the project towards its objectives while illuminating the significance and challenges inherent in this pursuit of automating respiratory condition classification.

The report is so organized such that after this introduction chapter detailing about the background of taking up this topic, the problem definition, the project's scope and motivation are discussed which is followed by its objectives, challenges, assumptions and the social and industrial relevance. The next chapter is composed of the literature review having a detailed overview of five papers or articles and then identifying the gaps in them to find the optimal model for the problem definition at hand. The subsequent chapter states the hardware, software and functional requirements needed for designing the solution for the problem definition that has been identified after the literature review. The next chapter outlines the system architecture starting with system overview and then giving details about the architectural design, module division which gives details on the data acquisition, data augmentation, data preprocessing, the CNN model, the GRU model, dense layers and model compilation, training and inference. This chapter also describes the work schedule using a gantt chart. The next chapter is results and discussions where the results are discussed and the quantitative results are specified.The last chapter of this report gives a detailed overview on the conclusions and the future scope. This is followed by references and the appendix.

Chapter 2

Literature Survey

2.1 Improvise approach for respiratory pathologies classification with multilayer convolutional neural networks

2.1.1 Introduction

The research aims to tackle the challenges in respiratory disease diagnosis stemming from a global shortage of physicians, with nearly 45% of WHO member states having less than one physician per 1000 population. Recognizing the urgency of timely and accurate diagnoses, particularly in regions with overburdened healthcare systems, the proposal suggests leveraging computer-based methods for automatic disease diagnosis. The dataset under consideration contains audio samples representing seven respiratory pathologies and healthy individuals, with a specific focus on distinguishing between diseases that manifest similar symptoms. Notably, smoking is identified as a significant contributor to these respiratory pathologies, alongside genetic and environmental factors. The paper underscores the complexities of distinguishing diseases with overlapping symptoms, such as Chronic Obstructive Pulmonary Disease (COPD) and asthma. Providing detailed insights into specific respiratory diseases, including COPD, Upper Respiratory Tract Infection (URTI), and bronchiectasis, the research underscores the critical importance of early detection, particularly for conditions like pneumonia, which exhibit variable recovery times and, in severe cases, can lead to fatality. Acknowledging recent technological advancements, the study highlights the positive impact of employing Deep Learning algorithms, particularly Convolutional Neural Networks (CNNs), in the realm of medical data analysis. The authors advocate for the integration of these advancements into the diagnostic process, drawing parallels with successful applications in other medical domains, such as brain tumor classification and prostate issues. The research introduces a novel approach to preprocessing audio files through a CNN architecture, specifically employing a matrix of Mel-Frequency Cepstral Coefficients (MFCC) features. The choice of a 1D CNN is rationalized for its ability to capture essential features, utilizing a kernel size of 1 to prevent overlooking crucial diagnostic information. The study references a prior work that utilized a CNN for audio sample classification, achieving an accuracy rate of 83%. In presenting a new CNN architecture and preprocessing methodology, the research aims to empirically compare its results with the previous approach, illustrating a commitment to iterative enhancements in diagnostic methodologies through the integration of advanced deep learning techniques.

2.1.2 Model Used

The proposed research represents a groundbreaking advancement in the realm of audio-based respiratory health diagnosis. The core innovation lies in a comprehensive approach that integrates Mel-frequency cepstral coefficients (MFCC), Melspectrogram, and Chroma CENS with Convolutional Neural Networks (CNNs) to enrich the representation of audio features. This amalgamation is designed to overcome limitations of prior research, offering a more robust and accurate analysis of respiratory sounds.[3]

To push the boundaries further, the research incorporates state-of-the-art deep learning feature extraction methods like d-vectors and i-vectors. This addition not only embraces contemporary techniques but also aims to capture nuanced patterns in audio signals that may hold crucial information for respiratory health diagnostics.[4]

The study's experimental foundation rests on the analysis of the ICBHI dataset, a rich repository encompassing chronic, non-chronic, and normal audio classes. This dataset serves as a stringent testing ground, enabling a meticulous comparison of three distinct features – MFCC, Melspectrogram, and Chroma CENS. This comparative analysis across different health conditions promises to shed light on the discriminative capabilities of each feature type.

Moving into the preprocessing phase, the research employs a novel data augmentation strategy. It involves the application of a delay function, introducing four unique versions of each audio sample with delays of 250ms, 500ms, 750ms, and 1000ms. This augmentation technique not only expands the dataset but also introduces temporal variations, enhancing the model's ability to generalize across different instances.

The choice of features is equally critical. The research opts for 39 features, comprising the three aforementioned types along with double delta features within the MFCCs. This nuanced selection aims to capture both spectral and temporal characteristics, providing a richer representation of respiratory audio signals.

Post-feature transformation, the individual features are stacked into a single multidimensional feature. This amalgamated feature set is then harnessed for the classification task, forming a crucial bridge between feature extraction and model training.

The final stage of the proposed methodology involves the application of a newly conceptualized CNN architecture tailored for the classification of audio into three classes: chronic, non-chronic, and healthy. This architecture is a key element contributing to the overall innovation of the research. Its design is presumably optimized to discern intricate patterns within the stacked feature set, enhancing the model's ability to accurately classify respiratory audio samples.[4]

In conclusion, the proposed research encapsulates a sophisticated and holistic approach to respiratory health diagnostics through audio analysis. Its integration of diverse features, contemporary deep learning techniques, data augmentation, and an innovative CNN architecture positions it as a potentially transformative advancement in the field, promising heightened accuracy and accessibility for respiratory health diagnosis.

2.1.3 Mel-frequency cepstral coefficients (MFCCs), Melspectrogram & Chroma energy features

The classification of pathologies in audio samples involves the extraction and transformation of various features, among which Mel-frequency cepstral coefficients (MFCC) have demonstrated effectiveness in audio classification tasks. The MFCC features are grounded in the human auditory system's perception, specifically designed to mimic the critical bandwidth variations of the ear.

The process of deriving MFCC involves seven computational steps:

1. Pre-emphasis: The audio sample undergoes filtering to increase the emphasis of higher frequencies.

2. Framing: The audio is segmented into frames with a duration typically ranging from

20 to 40 milliseconds during the analog-to-digital conversion.

3. Windowing: The Hamming window is applied to the frames for subsequent processing.

4. Fast Fourier Transform (FFT): The frames are converted into the frequency domain using FFT.

5. Mel Filter Bank Processing: The Mel filter bank, based on the Mel scale, is applied to the FFT results. Triangular filters are used to calculate the weighted sum of filter spectral components, aligning the output with the Mel scale.

6. Discrete Cosine Transform (DCT): The log Mel spectrum is converted into the time domain using DCT, resulting in coefficients.

7. Delta Energy and Delta Spectrum: To account for cepstral feature changes over time, additional features like double delta or delta (velocity) features are introduced.

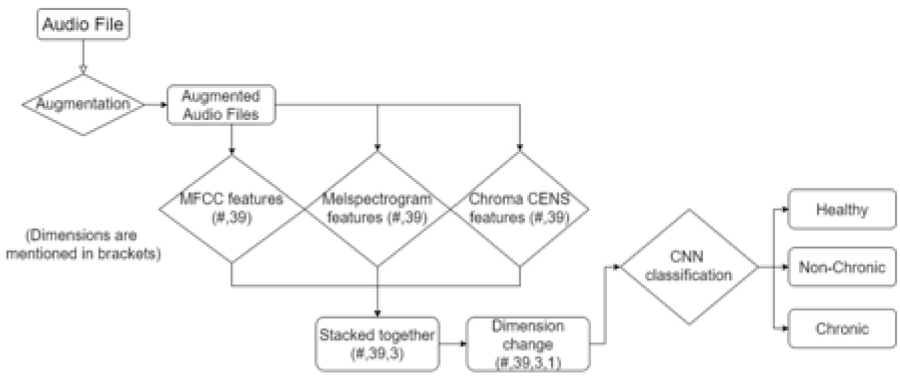


Figure 2.1: Proposed execution flow with flowchart

To address variable audio length, the mean of every 39 MFCC features across all frames is calculated, resulting in a final vector of 39 features for a single audio file. Notably, MFCC features exhibit greater decorrelation compared to Melspectrogram features, making them beneficial for linear models such as Gaussian mixture models.

The conversion from frequency to the Mel scale is expressed by the formula:
$M(f)=1125\ln(1+f/700)$

This formula encapsulates the transformation that aligns frequencies with the perceptual characteristics of the human auditory system, emphasizing the importance of the

Mel scale in audio processing. Overall, the detailed process underscores the nuanced and comprehensive nature of MFCC extraction, contributing to its effectiveness in audio classification tasks.

In contrast, Melspectrogram calculation involves the use of the librosa package and includes the following steps:

1. Separate windows: Input is sampled with a window size of 2048, with the window shifted by 512 samples for each subsequent calculation.

2. Compute FFT: Fast Fourier Transform is applied to each window, converting the time domain to the frequency domain.

3. Generating a Mel scale: The frequency spectrum is separated into 39 evenly spaced frequencies, forming the basis for the Melspectrogram.

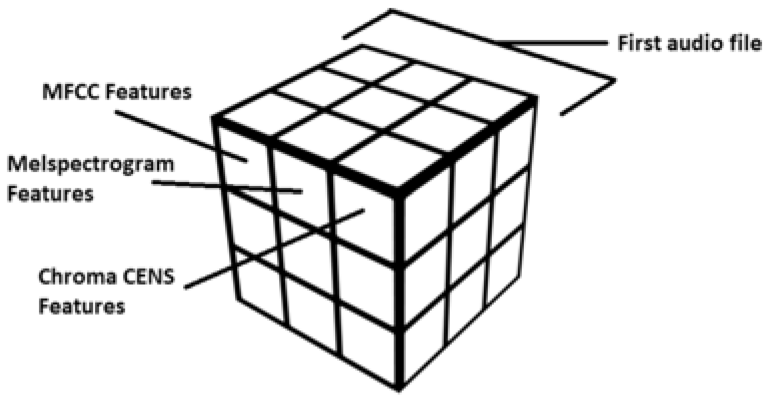


Figure 2.2: Feature Visualization

Chroma Energy Normalized Statistics (CENS) features are designed to enhance audio feature representations by smoothing local deviations in musical elements. The process involves normalizing chroma using L1-norm for relative energy distribution, quantizing based on a threshold to simplify patterns, smoothing to reduce local variations, and downsampling for efficiency. CENS features are valuable for tasks like chord recognition and musical genre classification.

2.2 Deep Learning Models For Detecting Respiratory Pathologies From Raw Lung Auscultation Sounds

The research methodology involves several key phase such as data acquisition and preparation, feature extraction, training, classification, and performance evaluation.[5]

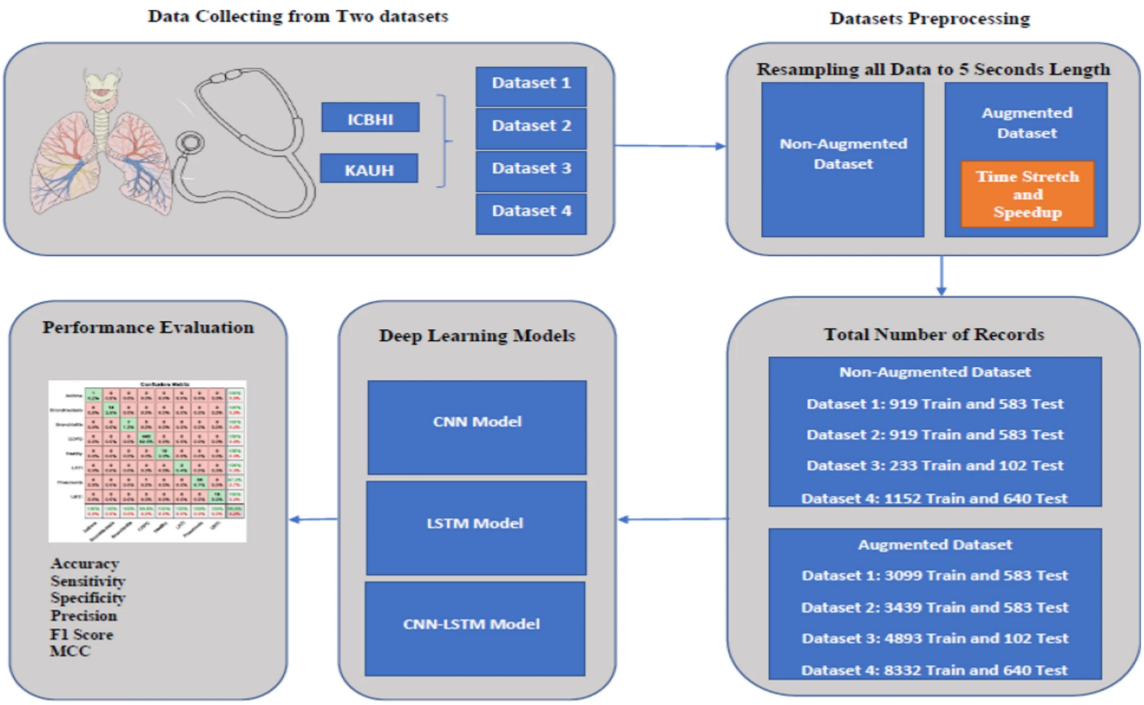


Figure 2.3: Model With CNN, LSTM, CNN-LSTM.

2.2.1 Data Acquisition

Two datasets, ICBHI 2017 Challenge and KAUH, were used, merged to create four datasets with varying classes. The KAUH dataset included 70 individuals with respiratory diseases and 35 healthy controls, providing 308 lung sound recordings. The ICBHI 2017 dataset contained 920 audio samples of 126 participants, annotated for various diseases and anomalies.

2.2.2 Data Preparation

Data augmentation techniques were applied to artificially expand the training dataset, including time stretch, time shift, add noise, and control volume.

- Time Stretch: randomly slow down or speedup the sound.

- Time Shift: shift audio to the left or the right by a random amount.
- Add Noise: add some random values to the sound.
- Control Volume: randomly increasing or decreasing the volume of the audio.

2.2.3 Deep Learning Models

Deep learning, a cutting-edge technology within artificial intelligence, has become pivotal in response to the abundance of massive datasets. [6] Its significance lies in a unique architecture inspired by the intricate structure of the human brain, featuring multiple sequential layers for intricate input data processing. This architecture facilitates the extraction and abstraction of deep features across various levels and perspectives. A myriad of deep learning algorithms has been introduced, with Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) standing out prominently. CNNs excel in image-related tasks, leveraging convolutional layers for hierarchical pattern learning, while LSTMs specialize in handling sequential data, making them ideal for tasks like natural language processing. The text pledges an in-depth exploration of developed CNN and LSTM models, both separately and in hybrid forms, underscoring their versatility in extracting deep features. CNNs prove adept in tasks involving grid-like data, such as image classification and object detection, while LSTMs shine in capturing long-term dependencies in sequential data. Hybrid models, combining CNN and LSTM strengths, address complex tasks, particularly those requiring simultaneous consideration of spatial and temporal information. This comprehensive approach underscores the power of deep learning across diverse domains.

Convolutional Neural Network

Convolutional Neural Networks (CNNs) are characterized by a sophisticated architecture with numerous hidden layers that leverage convolution and subsampling techniques for extracting deep features from input data.The architecture of a CNN comprises various types of layers, including input, convolution, RELU (rectified linear unit), fully connected, classification, and output layers, each serving a distinct purpose. These layers are intricately combined to form a CNN model capable of accomplishing specific tasks. In particular, CNNs have demonstrated remarkable success in diverse scientific fields, notably in the medical domain. The strength of CNNs lies in their ability to extract deep, representative, and discriminative features through the intricate arrangement of layers. Specifically, in the context of CNN layers, preceding layers undertake tasks such as downsampling and feature selection, ultimately leading to the generation of data categorization. This model excels in capturing intricate patterns and features, making it a valuable tool in various scientific applications, especially those demanding sophisticated analysis, such as medical image processing.

Long Short Term Memory

The Long Short-Term Memory (LSTM) architecture, initially proposed by Hochreiter and Schmidhuber in 1997. The fundamental structure of the LSTM typically comprises a memory cell, input gate, output gate, and forget gate. Considering an iteration at time t, denoting the input, cell, and hidden states as x_t, c_t, and h_t respectively, the LSTM produces the cell state c_t and hidden state h_{t-1} for the current input x_t, the previous cell state c_{t-1}, and its corresponding previous hidden state h_{t-1} . This architecture addresses the vanishing gradient problem in traditional recurrent neural networks, allowing LSTMs to capture and retain long-term dependencies in sequential data. The inclusion of gates enables the model to selectively update and forget information, contributing to its ability to effectively handle tasks involving sequential data, such as natural language processing and time series prediction. [7] The continuous refinement and expansion of LSTM variants underscore their versatility and enduring significance in the realm of deep learning.

CNN–LSTM model

In this hybrid model designed for lung sound analysis, the critical task of deep feature extraction and selection is efficiently managed by the CNN blocks. Specifically, the 1D convolutional layer and the max pooling layer within these CNN blocks play a pivotal role in extracting meaningful features from lung sound data. Simultaneously, the LSTM layer is incorporated to process these extracted characteristics as time-dependent features, enabling the model to learn and capture contextual time information from the dynamic nature of lung sounds. The hybrid architecture of 1D CNN–LSTM proves advantageous in comparison to methods solely based on CNN or LSTM.[8] This hybrid approach not only outperforms its counterparts in deep feature extraction and sound classification but also allows for the construction of considerably shallower models than pure CNN architectures. The synergistic combination of CNN and LSTM elements in this hybrid model underscores its efficacy in extracting intricate features and contextual information, particularly in the context of lung sound analysis.

2.3 GTCC-based BiLSTM deep-learning framework for respiratory sound classification using empirical mode decomposition

The human respiratory system is a complex and intricate network of organs that plays a crucial role in gas exchange, vital for sustaining life. Respiratory sounds, the audible vibrations generated during the inhalation and exhalation of air, offer valuable insights into the health of the lungs. While normal lung sounds are typically described as smooth and clear, various abnormalities can manifest as distinct adventitious sounds, such as crackles, wheezes, and rhonchi. These adventitious sounds are often associated with respiratory diseases, including asthma, chronic obstructive pulmonary disease (COPD), and even pneumonia. Early detection and accurate diagnosis of these diseases are crucial for effective management and improved patient outcomes.This is where the field of automatic lung sound classification comes into play. By leveraging advancements in deep learning and signal processing techniques, researchers are developing powerful tools to analyze and classify lung sounds. This paper delves into a novel method for lung sound classification that utilizes the synergistic combination of empirical mode decomposition (EMD) and a stacked BiLSTM architecture. [9]

EMD serves as the foundation for the proposed method. This signal processing technique decomposes the complex lung sound signal into a set of intrinsic mode functions (IMFs). Each IMF represents a distinct frequency component within the original signal, providing a deeper understanding of the underlying spectral characteristics. This decomposition process enables the identification of specific frequency bands associated with different types of adventitious sounds, paving the way for more effective feature extraction.Traditionally, Mel-frequency cepstral coefficients (MFCCs) have been widely used for feature extraction in audio classification tasks. However, the proposed method introduces a novel approach using Gammatone cepstral coefficients (GTCC). Inspired by the human auditory system, GTCCs capture the spectral information of the signal in a manner that closely aligns with how we perceive sound. This biologically inspired approach has been shown to outperform MFCCs in various audio classification tasks, including lung sound analysis. [10]

The extracted GTCC features, carrying valuable information about the spectral characteristics of the IMFs, are then fed into a stacked BiLSTM network. Bidirectional Long Short-Term Memory (BiLSTM) networks are a specific type of recurrent neural network (RNN) known for their ability to learn long-range dependencies within data sequences. The stacked architecture involves multiple BiLSTM layers, allowing the network to progressively learn more complex patterns within the feature representations. This deep learning architecture facilitates robust and accurate classification of lung sounds into different categories.The proposed method has been rigorously evaluated on a dataset of lung sounds from patients with normal and abnormal respiratory conditions. The experimental results demonstrate the effectiveness of the approach, highlighting several key strengths.

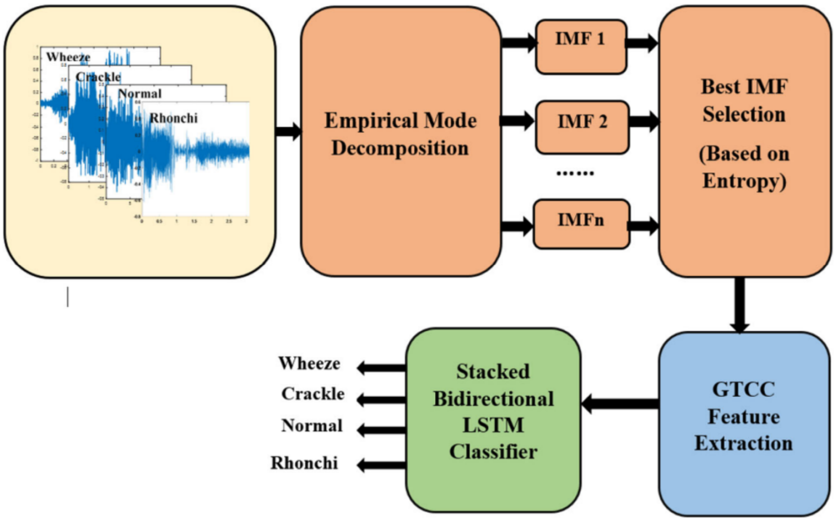


Figure 2.4: Lung Disease Prediction from Lung Sounds using GTCC and Empirical Mode Decomposition

Among the decomposed IMFs, IMF 3 consistently provides the most informative features for classification. This IMF captures the frequency range associated with several critical adventitious sounds, making it a valuable source of information for the BiLSTM network.The extracted GTCC features, capturing the spectral information with greater fidelity, consistently outperform MFCCs in terms of classification accuracy. This emphasizes the advantage of utilizing features inspired by the human auditory system for lung sound analysis.The stacked BiLSTM network achieves high classification accuracy for all types of lung sounds, demonstrating its ability to effectively learn and classify complex patterns in the extracted GTCC features.

By combining EMD, GTCC, and a stacked BiLSTM architecture, the proposed method offers a powerful tool for automatic lung sound classification. This approach demonstrates the potential to improve the early detection and diagnosis of various respiratory diseases, ultimately contributing to improved patient care and management. Further research and development could explore the integration of this method with other diagnostic tools and clinical workflows, paving the way for a future where AI-powered lung sound analysis plays a crucial role in respiratory healthcare.

2.4 Automatic diagnosis of COVID-19 disease using deep convolutional neural network with multi-feature channel from respiratory sound data: Cough, voice, and breath(2022) [1]

- De Noising Autoencoder
- GFCC
- IMFCC
- Deep CNN model

2.4.1 De Noising Autoencoder

- DAE is used as one of the feature extraction channels for the input function of the deep convolutional network to obtain deep features of respiratory sounds by removing background noise .
- DAE is designed to extract sound features by removing noise from background sound signals, thereby enhancing the discriminative features of the respiratory sounds .
- The DAE algorithm is employed to denoise the input respiratory sound signals, ensuring that the extracted features are robust and representative of the underlying patterns in the data .

2.4.2 GFCC

- GFCC is a feature extraction technique based on the Gamma-tone filter-bank, which is used to capture the spectral envelope of the input sound signal .
- The GFCC method involves applying the Gamma-tone filter-bank to the input sound signal and computing the logarithmic power spectrum of the output signals. The resulting power spectrum is then transformed using the Discrete Cosine Transform (DCT) to obtain the GFCC coefficients .
- GFCC provides transient respiratory sound features and is utilized to extract discriminative features from the input respiratory sound data

2.4.3 Improved Multi-frequency Cepstral Coefficients (IMFCC)

- IMFCC is employed to provide rich features from respiratory sounds, contributing to the extraction of deep features in the proposed approach .
- IMFCC is utilized to capture the multi-frequency characteristics of the respiratory sound signals, enhancing the representation of the input data for subsequent classification by the DCNN model .

2.4.4 Deep Convolutional Neural Network (DCNN) Model

- The DCNN model is designed to learn the underlying patterns in the input respiratory sound data and classify the signals into different categories, such as COVID-19 positive or negative .
- The DCNN provides better distinguishable features and comparable feature representations for accurate classification of sound signals by integrating various signal processing techniques that obtain various kinds of information .
- The proposed DCNN model, along with multi-feature channels, has shown promising results in achieving high accuracy for diagnosing COVID-19 disease with human respiratory sounds

In summary, the DAE, GFCC, IMFCC, and DCNN model work together to extract discriminative features from the input respiratory sound data and classify the signals into different categories, ultimately contributing to the accurate diagnosis of COVID-19 disease.

2.5 ALSD-Net: Automatic lung sounds diagnosis network from pulmonary signals[2]

This work aims to diagnose major lung diseases using artificial intelligence by analyzing lung sounds. Specific sounds like rhonchi/wheeze, fine crepitation, and coarse crept are associated with conditions like asthma, COPD, interstitial lung disease, pulmonary edema, and bronchiectasis.

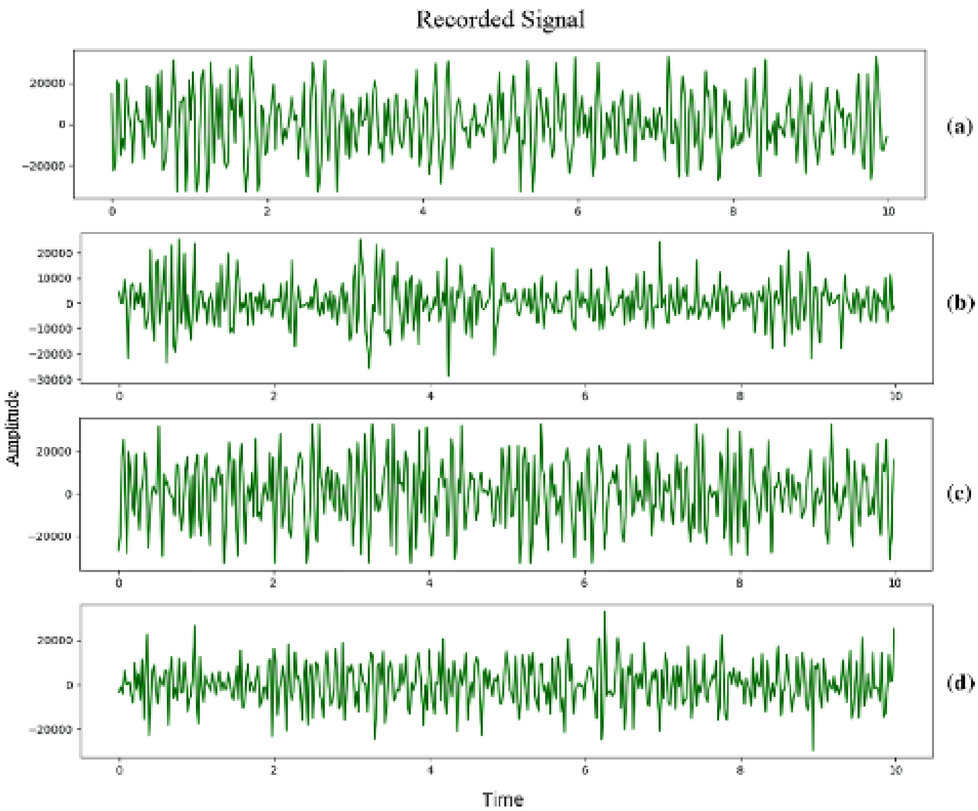


Figure 2.5: The plot of pulmonary signals amplitude concerning time domain with zoom view. a Crepitation signal, b Normal signal, c Rhonchi signal, d Wheezing signal

To address miss classifications, a Convolutional Neural Network (CNN) model, called ALSD-Net, is introduced. This model, coupled with a unique pulmonary signal augmentation method, outperforms existing methods, achieving a remarkable accuracy of

94.24%.[2] The approach enhances robustness in noisy environments and effectively tackles data scarcity issues by employing audio augmentation theory. Additionally, a signal processing model is introduced to remove noise from recorded pulmonary signals, ensuring high fidelity. The CNN model exhibits low error rates, high accuracy in multi-classification of diseases, and efficiency in real-time applications with low computational demands.[11] This work significantly improves automatic lung sound diagnosis, offering a promising solution for accurate and efficient disease classification.

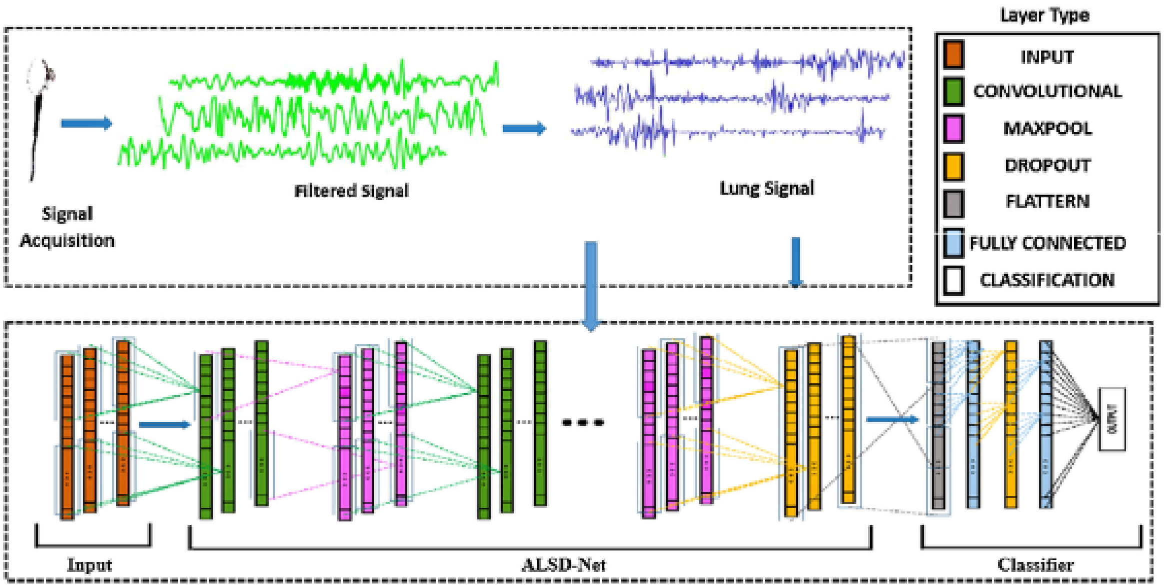


Figure 2.6: The structure of ALSD-Net for the multi-classification of lung sounds

This paper also outlines the representation and augmentation of acquired pulmonary signals, emphasizing their discrete-time nature composed of amplitude points represented by real numbers. The signal, denoted as P, consists of amplitude values pi at sequential points, forming a signal dimension represented by 'n'. Background deformations, labeled as signal 'b,' share the same dimensions as the original signal. An augmented signal 'a' is derived by adding the background deformations to the original signal using a formula that incorporates a deformation control parameter 'k,' set at 1000 Hz for converting deformation values into three-digit parameters.[12]

$$P = \coprod_{i=0}^{n-1} p_i$$

$$\beta = \coprod_{i=0}^{n-1} \beta_i$$

$$\alpha = P + \beta * \lambda$$

For training and validation, augmented signals are generated by combining original signals with background noise using the deformation technique.[13] This augmentation process ensures the creation of additional samples, maintaining the same quantity as the original dataset. Specifically, randomly generated noise samples within the [0, 1000] range, when added to the original signals, introduce different background noises.[14] Table 2 details the dataset composition, specifying the number of pulmonary signals and their respective labels utilized for the proposed CNN-based multi-class classification of pulmonary disorders. This augmentation method broadens the dataset by introducing diverse background noises to aid in training CNN models for improved multi-classification of pulmonary disorders.[2]

2.6 Summary and Gaps Identified

In recent years, there has been significant research interest in leveraging deep learning techniques for the automatic classification and diagnosis of respiratory diseases through the analysis of lung sounds. One study, based on a literature survey, focuses on addressing the global shortage of physicians by proposing a computer-based method for automatic respiratory disease diagnosis. This approach utilizes deep learning algorithms, particularly Convolutional Neural Networks (CNNs), to analyze audio samples representing various respiratory pathologies and healthy individuals. The study emphasizes the importance of early detection, especially for diseases like pneumonia, and introduces a novel approach involving Mel-frequency cepstral coefficients (MFCC), Melspectrogram, and Chroma CENS features with CNNs and Long Short-Term Memory (LSTM) networks[15].

Another proposed method involves a GTCC-based BiLSTM deep-learning framework, as presented in a specific study, which integrates empirical mode decomposition (EMD) and Gammatone cepstral coefficients (GTCC) features with a stacked BiLSTM network for accurate lung sound classification. This approach aims to capture nuanced patterns in audio signals associated with respiratory diseases, ultimately improving diagnostic accuracy. Similarly, an automatic diagnosis method for COVID-19 utilizes a deep convolutional neural network (CNN) with multi-feature channels, including De Noising Autoencoder (DAE), GFCC, and IMFCC, as detailed in another study, to classify respiratory sounds and diagnose COVID-19 disease. By combining various signal processing techniques with deep learning, this method achieves high accuracy in identifying COVID-19 positive cases from respiratory sound data.

Additionally, the ALSD-Net, described in a separate study, introduces a Convolutional Neural Network (CNN) model coupled with a unique pulmonary signal augmentation method for diagnosing major lung diseases. This model outperforms existing methods by employing audio augmentation theory to enhance robustness in noisy environments and tackle data scarcity issues. The proposed CNN model exhibits low error rates and high accuracy in multi-classification of diseases, demonstrating its potential for real-time applications in respiratory healthcare.

Overall, these studies collectively showcase the significant strides made in leveraging artificial intelligence for the analysis and classification of lung sounds in the medical domain. By integrating deep learning techniques with innovative signal processing methods, researchers aim to improve early detection, accurate diagnosis, and ultimately patient care in respiratory healthcare. These advancements hold promise for the future integration of AI-powered tools with diagnostic workflows, contributing to enhanced respiratory disease management.

Gaps identified include: Data Generalization and Validation: Although the described methods demonstrate high accuracy in classifying respiratory diseases, there may be challenges in generalizing these results to diverse populations or clinical settings. Additionally, the passage does not extensively discuss the validation of these methods on external datasets or real-world clinical scenarios, which is crucial for assessing their robustness and reliability.

Interpretability of Models: Deep learning models, such as CNNs and LSTMs, are often regarded as "black boxes" due to their complex architectures and high-dimensional feature representations. While these models achieve impressive results in disease classification, understanding the underlying features driving their predictions remains a challenge. Improved interpretability could enhance trust and adoption of these AI-powered diagnostic tools in clinical practice.

Addressing Data Imbalance and Bias: The passage does not explicitly address potential issues related to data imbalance or bias in the training datasets used for developing the deep learning models. Imbalanced datasets, where certain classes are underrepresented, can lead to biased model predictions and reduced generalizability. Future research should explore methods for mitigating data imbalance and ensuring equitable representation of diverse patient populations.

Chapter 3

Requirements

3.1 Hardware and Software Requirements

1. Processing Power:

   - Minimum: A dual-core processor (e.g., Intel Core i3) is the minimum requirement for basic functionality.
   - Recommended: For improved performance in training large neural networks, a quad-core processor or higher is recommended (e.g., Intel Core i5 or equivalent)

2. Memory (RAM):

   - Minimum: 8 GB RAM is necessary to handle the computational load.
   - Recommended: For handling larger datasets and complex neural network models, a minimum of 16 GB RAM is recommended

3. Storage:

   - Minimum: A 256 GB SSD provides sufficient storage for the operating system and essential software.
   - Recommended: A 512 GB SSD or higher is recommended for efficient storage of datasets, deep learning models, and intermediate results.

4. Graphics Processing Unit (GPU):

   - Minimum: An integrated GPU is sufficient for basic tasks and functionalities.
   - Recommended: For accelerated deep learning tasks, a dedicated GPU with CUDA support is recommended (e.g., NVIDIA GeForce GTX series).

5. Network Connectivity:

- Minimum: Standard Ethernet or Wi-Fi connectivity is required.
- Recommended: A high-speed internet connection is recommended for efficient dataset downloads, updates, and potential usage of cloud-based services.

6. Operating System:
   - Supported: The system supports Windows 10, macOS, or Linux (Ubuntu).
   - Recommended: Linux is preferred for better compatibility with deep learning frameworks and tools.
7. Python Environment:
   - Version: Python 3.7 or later is required..
   - Packages: Essential Python libraries such as NumPy, pandas, scikit-learn, TensorFlow, Keras, librosa, and Matplotlib must be installed
8. Deep Learning Framework:
   - Framework: TensorFlow 2.x is the primary deep learning framework.
   - Dependencies: The installation of CUDA Toolkit and cuDNN is necessary for GPU acceleration.
9. Development Environment:
   - IDE: Jupyter Notebook, VSCode, or any preferred Python IDE can be used for development.
10. Additional Tools:
    - Audio Processing: ffmpeg or pydub libraries are used for audio file manipulation.

3.2 Functional Requirements

1. User Interface for Audio Submission: Users should have an intuitive interface for submitting audio recordings, providing relevant patient information such as age, gender, and medical history.

2. Data Preprocessing: The system preprocesses audio data by extracting relevant features (MFCC, Zero Crossing Rate) using the librosa library. This step ensures the data is suitable for model input.

3. Convolutional Neural Network - Gated Recurrent Unit (CNN-GRU) Model: The system employs a CNN-GRU model designed using TensorFlow and Keras. This model is crucial for learning patterns and making predictions based on the provided audio features.

4. Training and Evaluation: The system undergoes a training phase using labeled datasets, where the CNN-GRU model learns from the provided audio features. The model's performance is then evaluated on a separate test set.

5. Disease Prediction: The trained model predicts respiratory diseases based on the input audio features. Predictions can include specific diseases such as asthma, bronchitis, or pneumonia.

6. Progress Display: During training and processing, the system displays progress updates to the user. These updates provide insights into the ongoing tasks without overwhelming the user with unnecessary details.

7. Results Presentation: Once the predictions are obtained, the system generates a comprehensive disease detection report. This report includes details such as predicted diseases, confidence scores, and potentially suggested courses of action.

8. Data Security: The system ensures secure handling of patient data, implementing encryption and access controls to maintain privacy and comply with relevant healthcare regulations (e.g., HIPAA).

9. Database Management: The system manages patient information, diagnosis details, and audio file metadata efficiently. It supports functionalities like data retrieval, storage, and updating.

10. Scalability:The system is designed to be scalable, accommodating varying dataset sizes and model complexities. This ensures adaptability and readiness for future enhancements or increased data volumes.

Chapter 4

System Architecture

In this chapter, we look into comprehensive overview that includes its architecture, design, module division, and visual representation of the planned timeline through Gantt Chart. The System Overview section provides a bird's-eye view, outlining the core components and their interactions. Following this, the Architectural Design section delves into the structure and organization of the system, elucidating the principles guiding its construction. The Module Division segment breaks down the system into distinct modules, elucidating their specific functions and interconnections. To provide a tangible sense of project Progression, the chapter concludes with a Work Schedule presented in the form of a Gantt Chart. This chart serves as a road map, detailing the planned timeline for the various project activities, ensuring a clear understanding of the anticipated milestones and their inter-dependencies.

4.1 System Overview

The project constitutes a comprehensive system for the classification of respiratory sounds using 1D Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs). The system architecture begins with the acquisition of respiratory audio files, sourced from the Respiratory Sound Database. These audio files are then subjected to a multi-step data processing pipeline. Patient IDs and corresponding diagnoses are extracted, and labels for each audio file are determined. Subsequently, data augmentation functions are applied to introduce variability into the training dataset, encompassing techniques like adding noise, time-stretching, and random-shifting.

Following data augmentation, feature extraction functions are invoked to capture the intrinsic characteristics of the respiratory sounds. These features include mean values of

MFCC, Zero Crossing Rate, Chromagram, Root Mean Square Energy, and Mel Spectrogram. Each of these features contributes to the holistic representation of the audio data, facilitating the subsequent learning process of the CNN-GRU model.

To ensure the model's training effectiveness, the dataset undergoes additional curation by excluding classes associated with rare diseases. This step streamlines the classification task to focus on more prevalent respiratory conditions. The processed features are then flattened and labels are encoded, preparing the data for input to the 1D CNN model.

The heart of the system lies in the meticulously designed architecture of the CNN-GRU model. Comprising convolutional layers for feature extraction, max-pooling layers for spatial down-sampling, batch normalization layers, GRU layers for learning the temporal relationship between the features also the patterns in them and dense layers for classification, the model is engineered to discern intricate patterns within the audio data. The use of the Adam optimizer and categorical cross-entropy loss during training ensures efficient convergence towards an optimal solution.

The evaluation phase is conducted rigorously, with a dedicated test set utilized to assess the model's performance. Standard metrics, including loss and accuracy, provide quantitative insights into the model's predictive capabilities. The confusion matrix further enhances the interpretability of results, breaking down the classification outcomes for each respiratory condition. Visualization tools, such as loss and accuracy plots, track the model's learning trajectory over training epochs, aiding in the identification of potential overfitting or underfitting.

4.2 Architectural Design

The architecture is of a Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) hybrid model designed for respiratory sound classification. It comprises convolutional layers for feature extraction, max-pooling layers for down-sampling, batch normalization layers to normalize the activation,and Leaky ReLU activation functions. Multiple

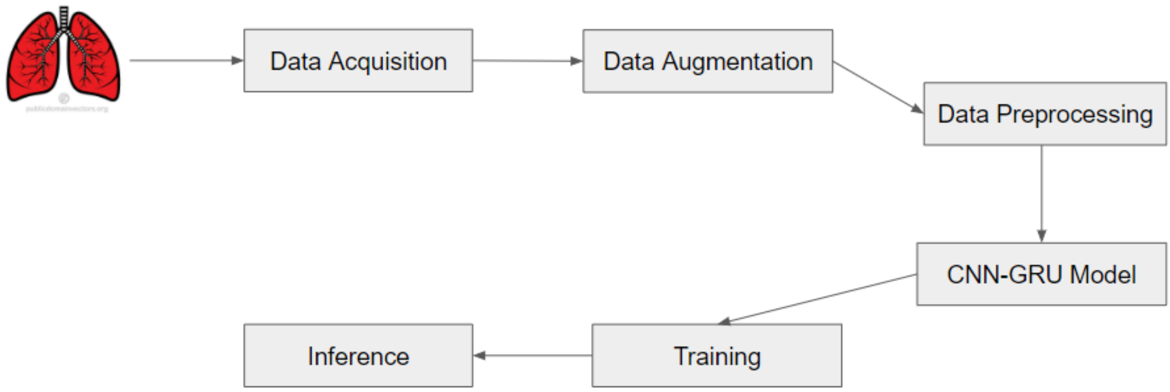


Figure 4.1: Architecture Design

GRU layers are stacked after the CNN layers such that each GRU layer passes the output sequence to the next layer. The model processes mean values of various audio features, and the final dense layer with softmax activation produces multi-class probabilities. The architecture is trained using the Adam optimizer and categorical cross-entropy loss. Dropout layers mitigate overfitting, and an early stopping mechanism halts training when performance plateaus. Model checkpointing saves the model's best performing weights during training, allowing it to resume or evaluate later. The design facilitates effective learning of temporal patterns in respiratory sounds, resulting in robust classification capabilities.

4.3 Module Division

- Data Acquisition and Augmentation
- Data Pre-processing
- CNN-GRU model
- Training and Classification
- Inference

4.3.1 Data Acquisition

The data acquisition process in this project involves a meticulous exploration of the Respiratory Sound Database. Patient IDs are associated with each audio file, establishing a crucial linkage between the data and the corresponding patients. This step ensures a comprehensive understanding of the dataset's structure and allows for the precise labeling of each audio file with its respective respiratory condition. Diagnoses and file paths are extracted, providing the necessary context for subsequent data processing. The comprehensive dataset, comprising patient information, diagnoses, and file paths, forms the foundational structure for the subsequent stages of the respiratory sound classification system.

4.3.2 Data Augmentation

Following data acquisition, a thoughtful data augmentation strategy is employed to enhance the diversity and generalization capabilities of the model. Three distinct techniques are applied to the audio data. Firstly, the addition of noise introduces variability, mimicking real-world conditions where respiratory sounds are often accompanied by environmental noise. Secondly, time-stretching modifies the temporal dynamics of the sounds, presenting the model with instances of varying duration. Lastly, time-shifting alters the timing of the audio signals, addressing timing variations commonly encountered in diverse respiratory conditions. These augmentation techniques are systematically integrated into the feature extraction process, contributing to the creation of an augmented dataset that encompasses varied instances of noise, time-stretching, and time-shifting. This augmentation strategy enriches the training data, enabling the model to learn more robust features and enhancing its adaptability to the diverse and dynamic nature of real-world respiratory sounds.

4.3.3 Data Preprocessing

Data preprocessing in this project is a critical step aimed at optimizing the respiratory sound data for effective model training. The initial transformation involves extracting relevant features from each audio file, including mean values of Mel-Frequency Cepstral Coefficients (MFCC), Zero Crossing Rate, Chromagram, Root Mean Square Energy, and Mel Spectrogram. These features capture essential characteristics of respiratory sounds. Additionally, it implements a comprehensive data augmentation strategy, introducing variability through techniques such as adding noise, time-stretching, and time-shifting. This augmentation diversifies the training dataset, enhancing the model's ability to generalize across different conditions.

Furthermore, the system undertakes the task of filtering out audio files associated with very rare diseases, such as Asthma and LRTI, to ensure a balanced and representative dataset. Class counts are visualized through a bar plot, providing insights into the distribution of respiratory conditions in the final dataset. The data is then split into training and testing sets, and standard scaling is applied to normalize the features. The final data preprocessing step involves expanding the dimensions of the feature sets to accommodate the input requirements of the 1D CNN and GRU hybrid architecture. This meticulous data preprocessing pipeline ensures that the model is trained on a well-structured, diverse, and standardized dataset, optimizing its performance for accurate respiratory sound classification.

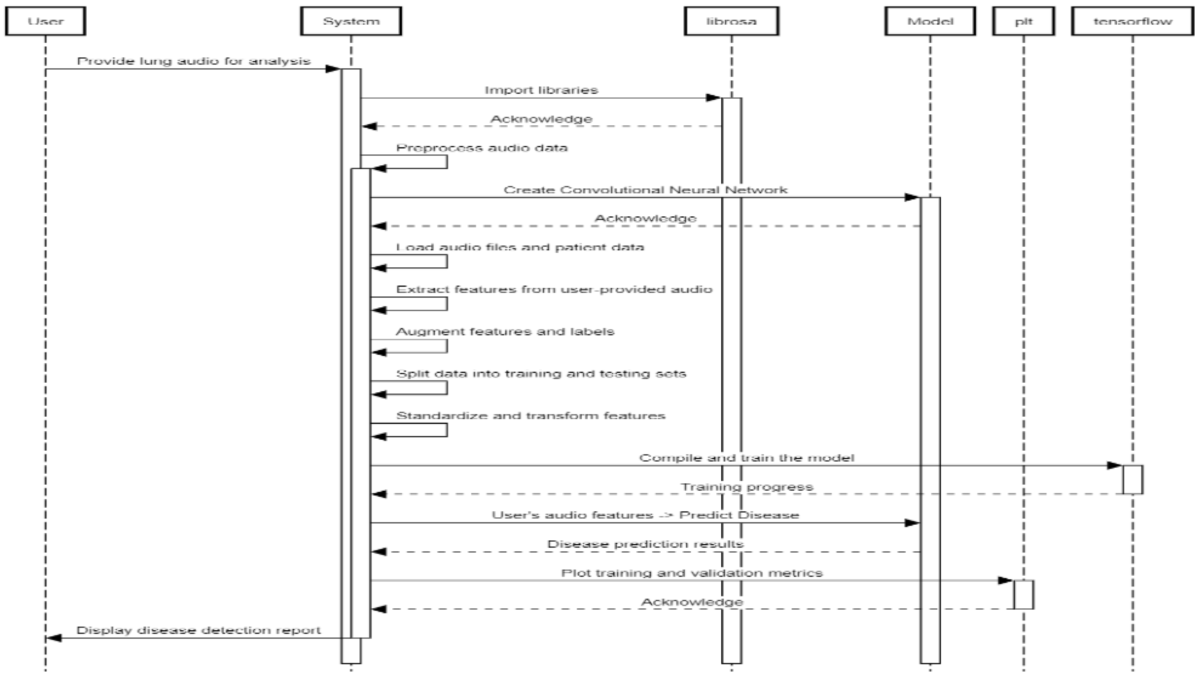


Figure 4.2: Sequence Diagram

4.3.4 1D CNN Model

The Convolutional Neural Network (CNN) model is designed for the classification of respiratory sounds based on their features. Comprising a series of convolutional layers, pooling layers, and batch normalization layers, this 1D CNN architecture is tailored to effectively capture spatial patterns inherent in audio data. The initial convolutional layers, each followed by max-pooling layers, serve as feature extractors. These layers are instrumental in recognizing hierarchical features in the time domain, crucial for understanding the nuanced characteristics of respiratory sounds.

The model architecture includes multiple convolutional layers with varying filter sizes and strides, enabling the extraction of both local and global features from the input data. The max-pooling layers are used to downsample the extracted features, reducing the computational complexity while retaining the most salient information and the resulting feature vectors. The batch normalization layers normalize that is scales the values to mean of zero and standard deviation equaling one. These are then fed into the Gated Recurrent Unit.

4.3.5 Gated Recurrent Unit Model

Gated Recurrent Unit is designed to capture the dependencies in the sequential data such that the output of each of the GRU layer is not a single output thus ensuring the flow of information to the succeeding layers. The input sequences of the the GRU model have a hierarchical representation which is achieved by stacking multiple GRU layers successively. The model is used to learn patterns in the data and its capacity and complexity to do this is determined by the number of units. After the processing of the input sequence using the different GRU layers the outputs of some of these layers are combined that is an add function is used allowing the model to enhance its representation competence as this aggregates information from different parts of the input sequence.

4.3.6 Dense Layers And Model Compilation

In this model that combines Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) layers, Dense layers play a crucial role in the final stages of the model. After the sequence data undergoes processing by the convolutional and GRU layers to extract and learn features, the output needs to be transformed into a format suitable for classification tasks. GRU layers processes the sequence data, and the output is then passed through two dense layers with Leaky Rectified Linear Unit (ReLU) activation functions.Leaky ReLU introduces non linearity to the function so as to better understand the input to output mappings. These dense layers introduce non-linearity and enhance the model's capacity to learn complex patterns. The final dense layer employs a softmax activation function to produce class probabilities, making it ideal for multi-class classification tasks.

Following the definition of the architecture, the model is compiled to prepare for training. This compilation step involves specifying additional parameters such as the loss function, optimizer, and evaluation metrics. The model is compiled using the Adam optimizer with a learning rate of 0.001, an adaptive learning rate optimization algorithm combining the benefits of AdaGrad and RMSProp. The categorical cross-entropy loss function is selected as it effectively quantifies the difference between the predicted class probabilities and the true labels, making it suitable for multi-class classification problems. Additionally, accuracy is chosen as the evaluation metric to monitor the model's performance during training, providing a straightforward and intuitive measure of classification accuracy. Together, these steps define the architecture and training settings of the CNN-GRU model, laying the groundwork for effective learning and prediction on sequence data.

4.3.7 Training

The CNN-GRU model undergoes training using a combination of labeled respiratory sound data and the stochastic gradient descent optimization algorithm. The training process occurs over multiple epochs, where each epoch consists of iterations through the entire training dataset. During each iteration, the model computes predictions, compares them to the actual labels, and adjusts its internal parameters to minimize the discrepancy between predicted and actual outcomes. The loss function, in this case, categorical cross-entropy, quantifies the model's performance, and the Adam optimizer efficiently updates weights to enhance accuracy. The training performance is monitored through metrics like accuracy and loss, and an early stopping mechanism is employed to prevent overfitting by terminating training when the model's performance on a validation set ceases to improve. This meticulous training regimen ensures that the CNN-GRU model learns intricate patterns in respiratory sounds, enabling it to generalize well to diverse conditions and make accurate predictions on new instances.

4.3.8 Inference

In the final inference , the trained Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) hybrid model undergoes evaluation on a previously unseen test set, aiming to assess its generalization capabilities and accuracy in classifying respiratory sound samples. The evaluation involves feeding the preprocessed test data into the model, which then generates predictions for the respiratory conditions associated with each sound recording. These predictions are compared against the ground truth labels to compute metrics such as loss and accuracy, providing a comprehensive understanding of the model's performance on novel data. Additionally, a confusion matrix is employed to visually represent the distribution of predicted classes and their alignment with the true classes, offering detailed insights into the model's classification accuracy for individual respiratory conditions. This final inference step ensures that the CNN-GRU model is reliable and effective in accurately identifying various respiratory disorders, thus showcasing its potential for real-world applications in automated respiratory condition diagnosis.

4.4 Work Schedule - Gantt Chart

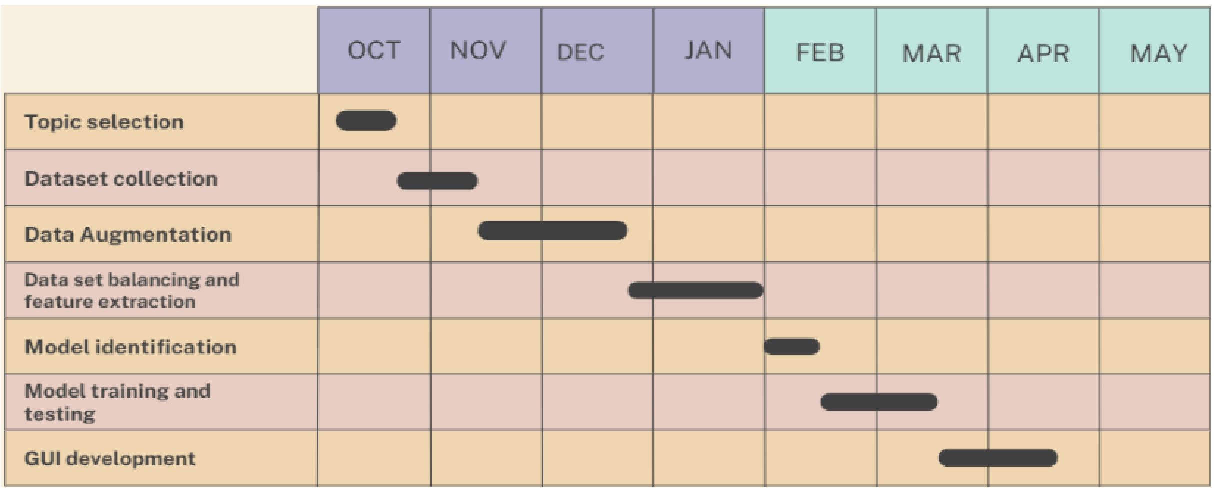


Figure 4.3: Gantt Chart

To sum up, this chapter thoroughly examined the important facets of our project, starting with a thorough System Overview that offered understanding of the overall design and operation.The section on Architectural Design examined the system's architecture and clarified the fundamental design ideas that influenced its creation. By breaking the system down into smaller, more manageable parts, Module Division clarified the functions and relationships between them. The Gantt Chart ensures transparency and a clear execution road map by graphically outlining our tasks and deadlines.

Chapter 5

System Implementation

In this chapter we discuss about the overall system implementation and details on the dataset identified, proposed methodology and the algorithm, user interface design, database design and the description of implementation strategies.

Here detailed review of the dataset used, the number of files , the number of patients and the file time length etc are specified.The libraries imported, the functions used, etc are specified and their uses are explained. The structure of the User Interface and the languages and its methodology of implementation is also specified.The details of the database design is then explained along with its schema and the reason for using a Django model.The last part before the conclusion of this chapter is giving a description of the implementation strategies by specifying the python libraries associated for performing the functions in the system, the implementation of CNN and GRU models as per the needs of the system and the different evaluation methods to understand the prediction scores and the training results of the model.

5.1 Datasets Identified

The Respiratory Sound database was originally compiled to support the scientific challenge organized at Int. Conf. on Biomedical Health Informatics - ICBHI 2017.

The database consists of a total of 5.5 hours of recordings containing 6898 respiratory cycles, of which 1864 contain crackles, 886 contain wheezes, and 506 contain both crackles and wheezes, in 920 annotated audio samples from 126 subjects.

The cycles were annotated by respiratory experts as including crackles, wheezes, a combination of them, or no adventitious respiratory sounds. The recordings were collected using heterogeneous equipment and their duration ranged from 10s to 90s.

5.2 Proposed Methodology/Algorithms

Detailed Respiratory Sound Classification Algorithm:

1.Import Required Libraries: Import necessary libraries such as NumPy for numerical operations, librosa for audio processing, scikit-learn for machine learning tasks, Matplotlib for data visualization, Seaborn for enhanced plotting, and TensorFlow for building neural network models.

2.Define Audio Data Processing Functions: Implement functions to preprocess audio data: add noise: Add white Gaussian noise to simulate real-world conditions and improve model robustness.

shift: Shift audio signals in the time domain to mimic variations in sound recordings due to patient movement or device placement.

stretch: Stretch audio signals to simulate changes in speaking rate or breathing patterns.

pitch shift: Perform pitch shifting to simulate variations in pitch due to different respiratory conditions or individual characteristics.

Create a function to visualize audio waveforms and their corresponding spectrograms or features using Matplotlib and librosa.

3.Feature Extraction and Augmentation: Load respiratory sound data from the specified directory containing audio files. Extract Mel-frequency cepstral coefficients (MFCCs) features from the audio files using librosa, as MFCCs are commonly used for audio classification tasks due to their effectiveness in capturing relevant audio characteristics.Augment the dataset to increase its size and diversity, which can help improve the model's generalization performance. Augmentation techniques may include adding noise, shifting, stretching, and pitch shifting.Organize the augmented data into feature arrays and corresponding labels for further processing.

4.Model Building: Split the dataset into training, validation, and testing subsets to assess the model's performance accurately.Design a neural network model architecture using TensorFlow's Keras API.Utilize Gated Recurrent Unit (GRU) layers for sequential data processing, as GRUs are well-suited for capturing temporal dependencies in time-series data like audio.Introduce dense layers with LeakyReLU activation functions to introduce non-linearity into the model and enable it to learn complex patterns.Apply BatchNormalization to normalize the activations of each layer, which can accelerate training and improve model stability. Configure the output layer with softmax activation to output class probabilities for multi-class classification.Define the model architecture by specifying the input shape and connecting the layers in a sequential manner.

5.Model Training: Compile the neural network model with appropriate loss function, optimizer, and evaluation metric to train it effectively.Train the compiled model using the training dataset while validating it on the validation set to prevent overfitting.Monitor the training process using metrics such as accuracy to evaluate the model's performance on both training and validation data.Utilize early stopping to halt training if the validation loss fails to improve after a specified number of epochs, preventing overfitting.

6.Model Evaluation: Assess the trained model's performance using the testing subset to evaluate its generalization ability on unseen data.Compute performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix to quantify the model's classification performance.Visualize the confusion matrix using Matplotlib and Seaborn to gain insights into the model's strengths and weaknesses in classifying different respiratory conditions. Interpret the confusion matrix to identify any patterns or trends in misclassifications and assess the model's overall effectiveness.

5.3 User Interface Design

Home Page Wireframe: The home page wireframe serves as the main entry point for users into the pulmonary disease detection system using lung sound. Lung Sound Analysis Page Wireframe: The lung sound analysis page wireframe illustrates the interface for recording, analyzing, and interpreting lung sounds collected from patients. It features audio recording controls, visualization of sound waveforms, spectrograms, and diagnostic markers indicating potential pulmonary abnormalities.

5.4 Database Design

The database been designed using Django database model.This database model can store the data into distinct tables and thus makes it easy for managing and querying the data. This model of database can be used to design and create relationships between different entities as per the requirements using foreign keys and links to different entities and relationships.This database is easily modifiable and thus flexible and thus can include, remove and modify the data according the needs of the user using the GUI.This changes do not reflect significantly on the overall structure of the database making it scalable according to the requirements of the application and the user.The Django's high level API can be used to perform CRUD and it's built-in ORM and thus is relatively straightforward and easy to understand which in turn simplifies maintenance and the future in the development of the database and the system, The database has to stores relevant information such as the lung sound files, patient id, and after the diagnosis the disease and the accuracy of prediction.The database can be modified to store additional information such as name of the patient, their date of birth, etc as per the user requirements.

5.5 Description of Implementation Strategies

Audio Data Capture and Preprocessing

Python Library: librosa

Methods: Use librosa.load() to load audio files into Python, which provides an easy-to-use interface for reading audio files and extracting various features such as MFCCs (Mel-frequency cepstral coefficients). Utilize librosa.feature.mfcc() to extract MFCC features from audio signals, a commonly used representation for audio classification tasks. Apply preprocessing techniques such as adding noise, shifting, stretching, and pitch shifting using custom functions or libraries like noisereduce or audiomentations to augment the dataset for improved model generalization.

Model Design

Language: Python with TensorFlow and Keras

Methods: Design a CNN architecture using Keras, a high-level neural networks API running on top of TensorFlow. Stack convolutional layers followed by max-pooling layers to extract hierarchical features from audio spectrograms or MFCCs. Then a GRU model is designed to relate the dependencied in the patterns and the relationship between the features extracted that are to be used for the dense layer to be classified into different classes by the softmax activation to output class probabilities.

Model Evaluation Methods

Methods: Use Keras's model.compile() method to compile the model with appropriate loss function, optimizer, and evaluation metric (e.g., categorical cross-entropy, Adam optimizer, accuracy). Train the model using model.fit() on the training dataset, validating it on a separate validation set to monitor performance and prevent overfitting. Evaluate the trained model's performance using model.evaluate() on the testing dataset, calculating metrics such as accuracy, precision, recall, and F1-score.

Chapter 6

Results & Discussions

6.1 Results and Discussion

The data preprocessing and data augmentation has been performed for the 30% of the project progress for the 1st phase of the project.Functions were created for augmenting the data by adding noise,stretching the audio and shifting the pitch.The spectograms were created using MFCC techniques in each of the audio file to extract the features needed to feed into the machine learning model. In the second phase of the project we designed a deep learning model, trained it to classify the pulmonary diseases into the categories such as COPD, Pneumonia, Bronchiolitis, URTI and as healthy. For the 60% evaluation we designed a simple GRU model and trained it but for the 100% evaluation we designed a CNN-GRU model and did the training and testing on this model as a CNN-GRU model has a better accuracy.We so compared the results obtained from a GRU model and a CNN-GRU model and analyzed the outputs of both to find that a CNN-GRU model is a more optimal deep learning model for our requirement.

6.2 Quantitative Results

```
[['Bronchiectasis' 'Bronchiolitis' 'COPD' 'Healthy' 'Pneumonia' 'URTI']
 ['16' '13' '793' '35' '37' '23']]
```

Figure 6.1: Dataset

The results show the count of each disease label after data augmentation. The bar plot provides a visual representation of the distribution of diseases in the dataset.

This is a graphical representation indicating the disease count in the sound files. It is to be noted that Asthma and LRTI have been omitted as there is insufficient data for it.

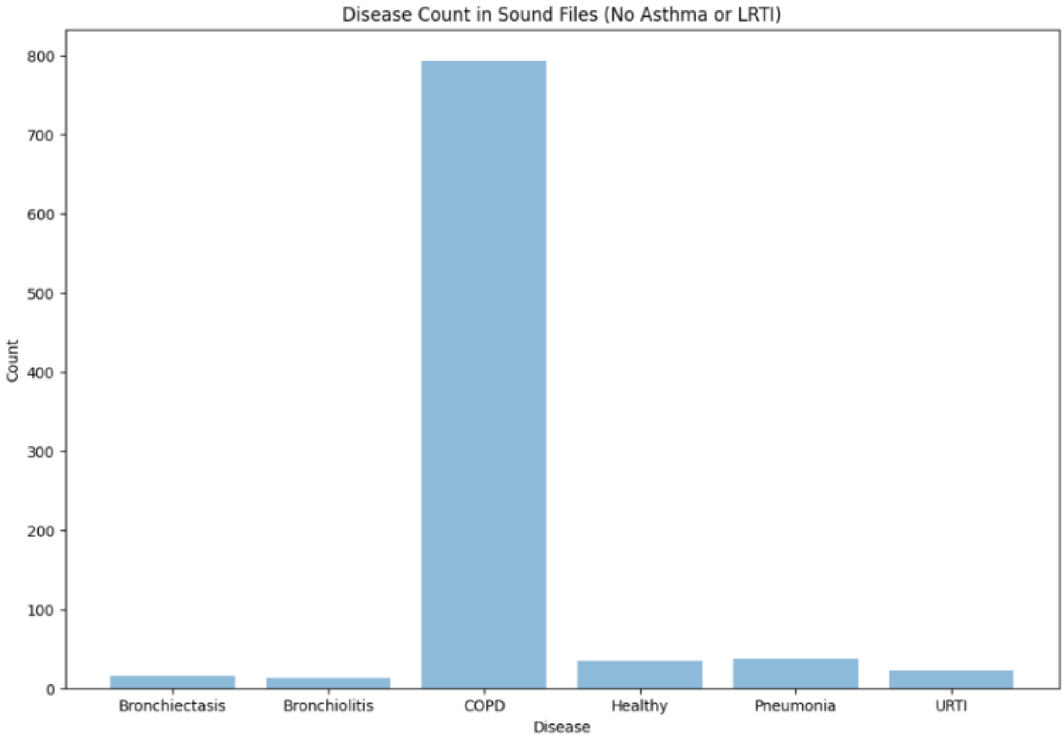


Figure 6.2: Disease Count in the Sound Files

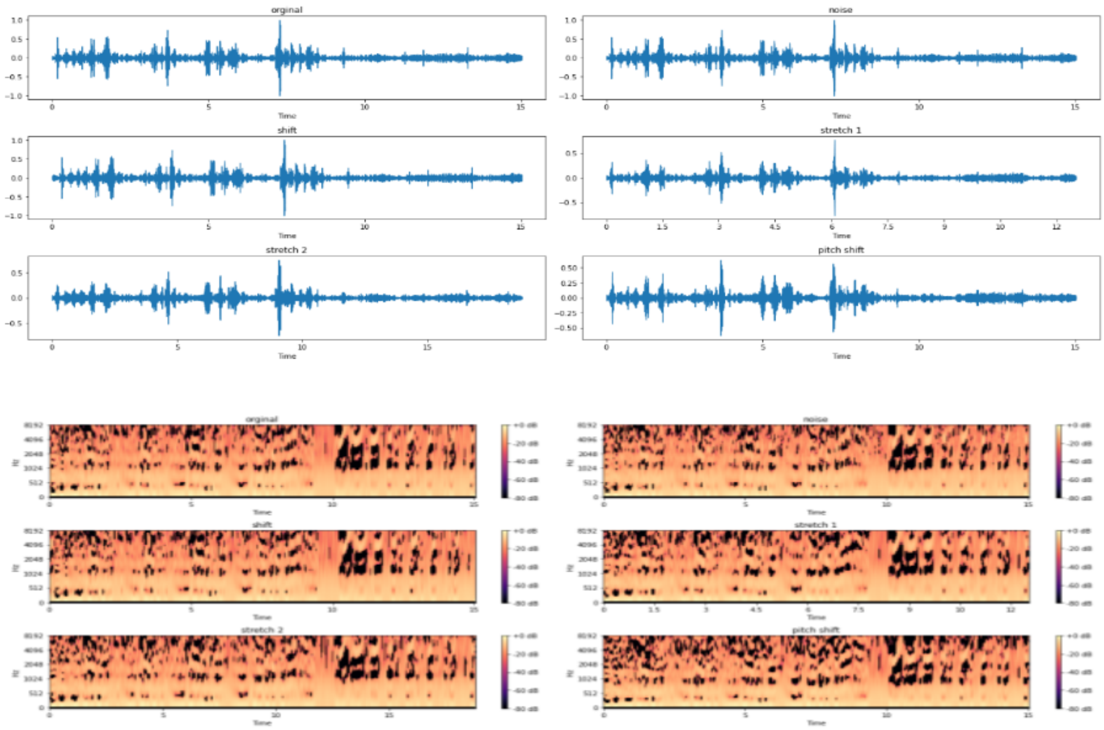


Figure 6.3: Spectograms and audio files

This figure shows the various audio files and their respective spectograms.The figure shows the different augmented forms of an audio and their spectograms.

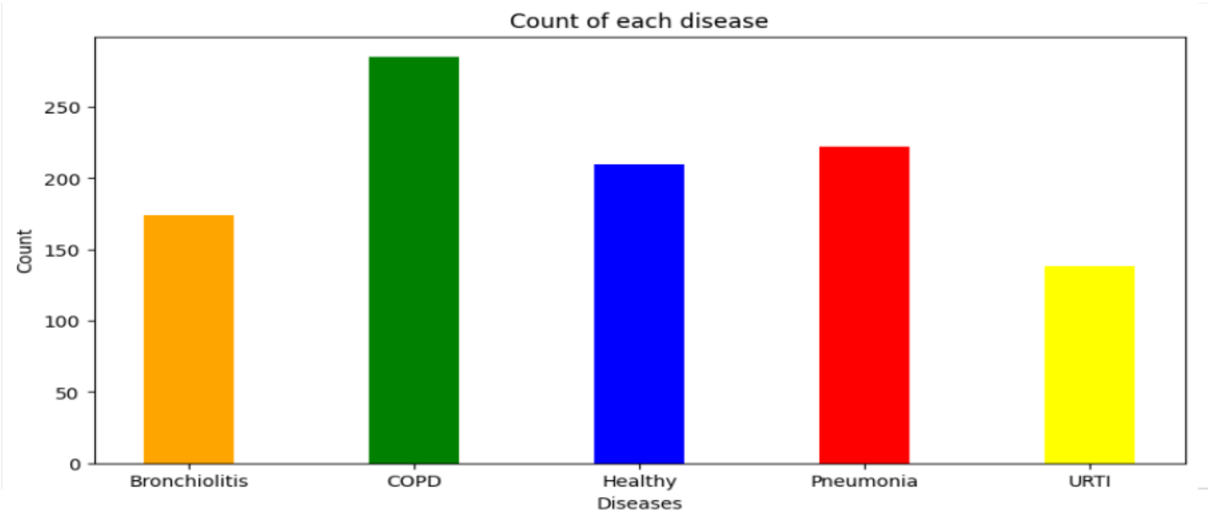


Figure 6.4: Augmented dataset

The dataset had a very prevalent imbalance in the number of classes of the diseases with COPD's cases being of a large number. The data of each class is augmented to reduce the imbalance in the initial dataset classes.

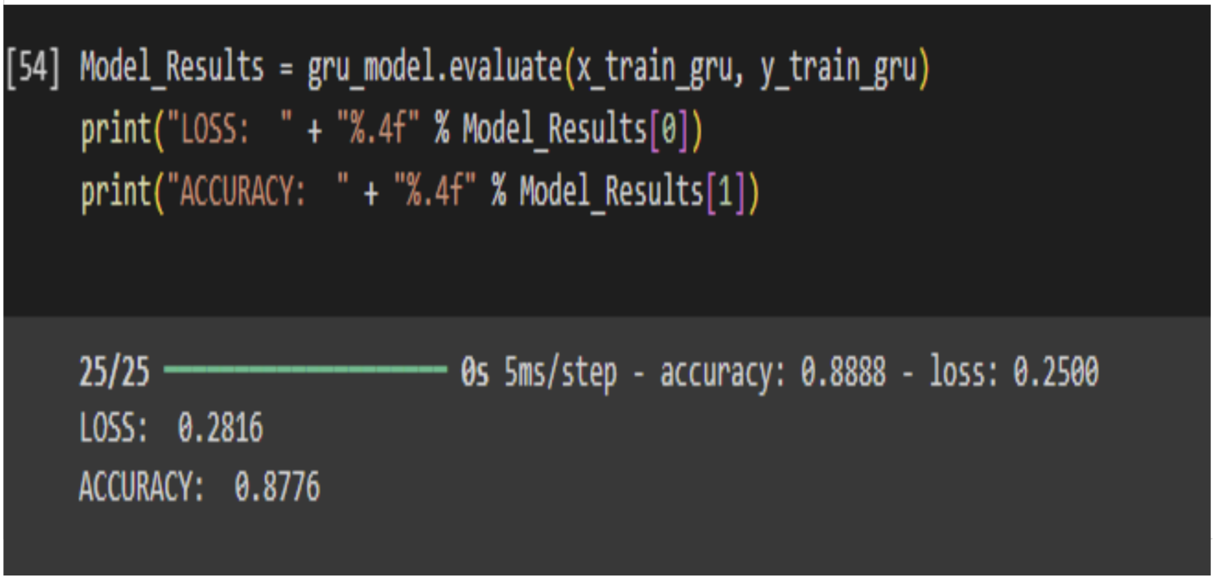


Figure 6.5: Model Training Results of the GRU model

This figure shows the model training results of a simple GRU model which include loss percentage of 28.16 and accuracy percentage of 87.76.

```
print(classification_report(y_testclass, classpreds, target_names=classes))
```

```
              precision    recall  f1-score   support

        COPD       1.00      0.94      0.97        17
 Bronchiolitis     0.92      0.92      0.92        13
   Pneumonia       0.71      1.00      0.83        10
        URTI       0.57      0.57      0.57         7
     Healthy       0.79      0.65      0.71        17

    accuracy                           0.83        64
   macro avg       0.80      0.82      0.80        64
weighted avg       0.84      0.83      0.83        64
```

Figure 6.6: Model Prediction Scores of the GRU model

The GRU model achieved performance on the test dataset with an overall accuracy of 87%, indicating its ability to make accurate positive predictions, capture relevant instances, and maintain a balance between precision and recall. However, we noted that the f1 score for URTI was 0.57.

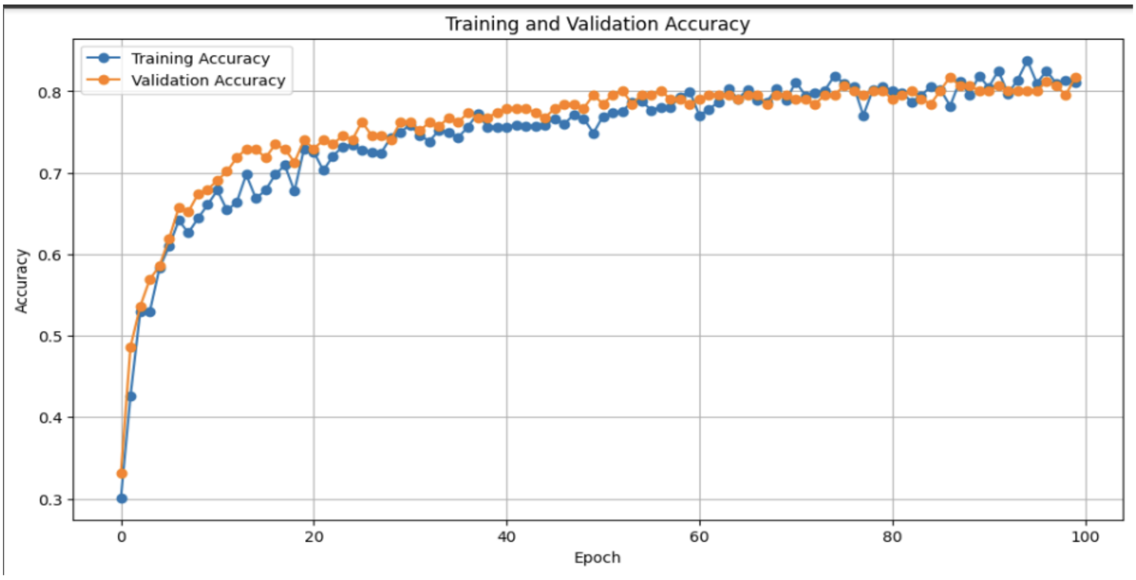


Figure 6.7: Training Validation Graph of the GRU model

The figure above shows the training and validation graph for our GRU model.The training process of the GRU model is depicted in the training and accuracy graph. As the number of epochs increases, the model's accuracy steadily improves, indicating effective learning. Initially, there might be fluctuations in the accuracy due to the model's adjustment to the training data. However, as training progresses, the accuracy stabilizes, eventually converging to the validation dataset. This graph demonstrates the effectiveness of the GRU model in learning from the training data and its ability to generalize well to unseen data.

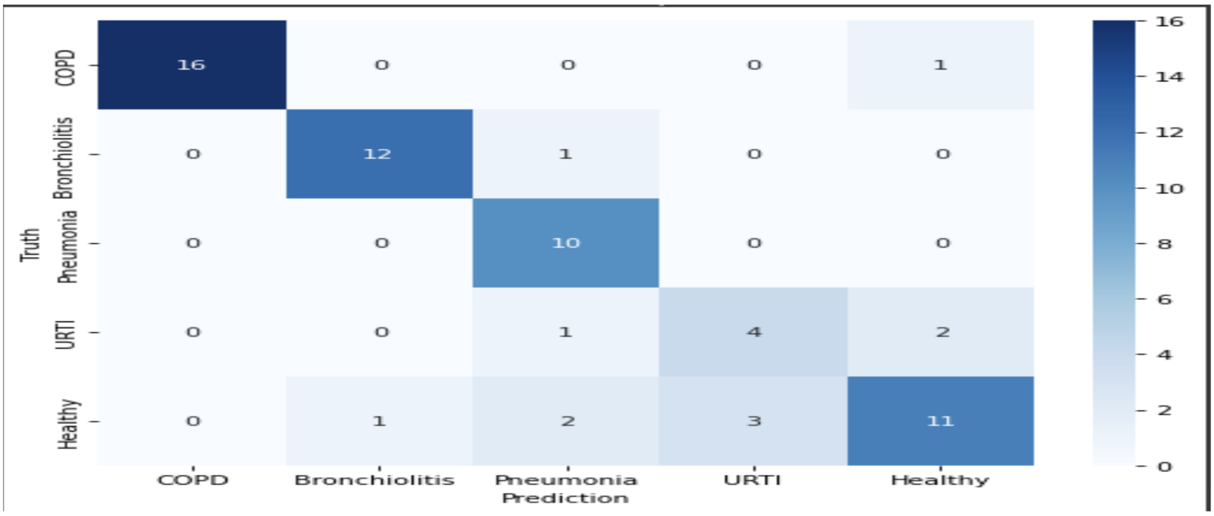


Figure 6.8: Confusion Matrix of the GRU model

This figure shows the confusion matrix of the simple GRU model with respect to the five different classes.

After the 60% evalution the augmentation process was again modified by randomizing the features selected from the class COPD for improving the balance of the data set.

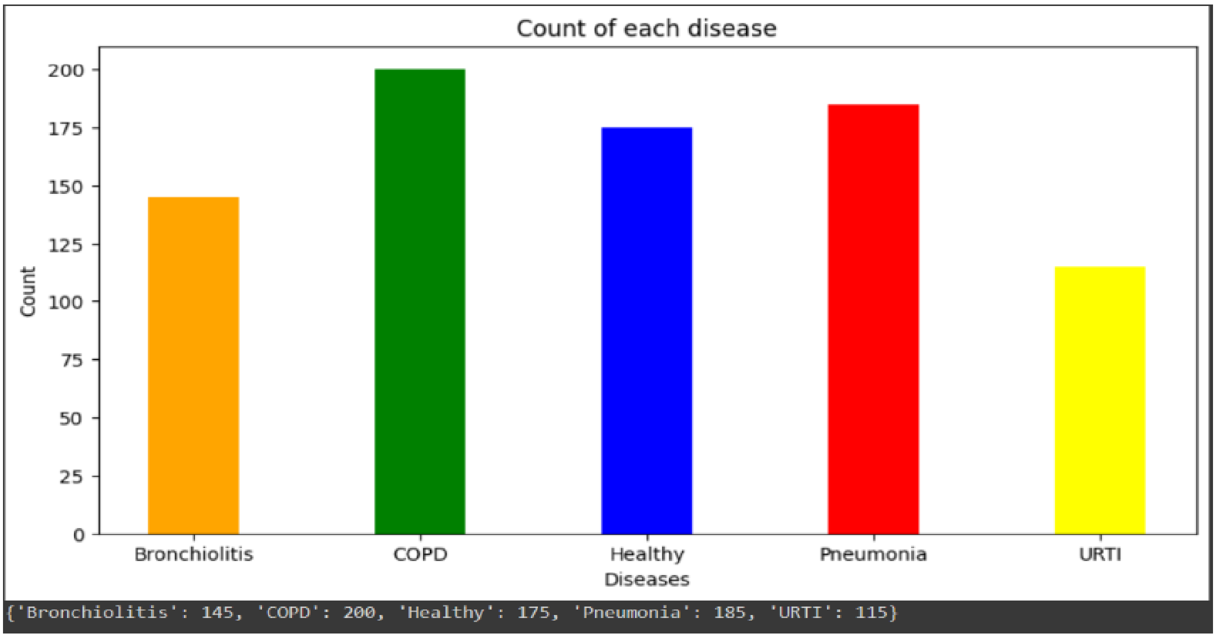


Figure 6.9: Modified Augmented Dataset

The CNN-GRU model has an accuracy of 96.32% and thus shows its peak ability to make more precise predictions.From this we can understand the accuracy of the trained data.

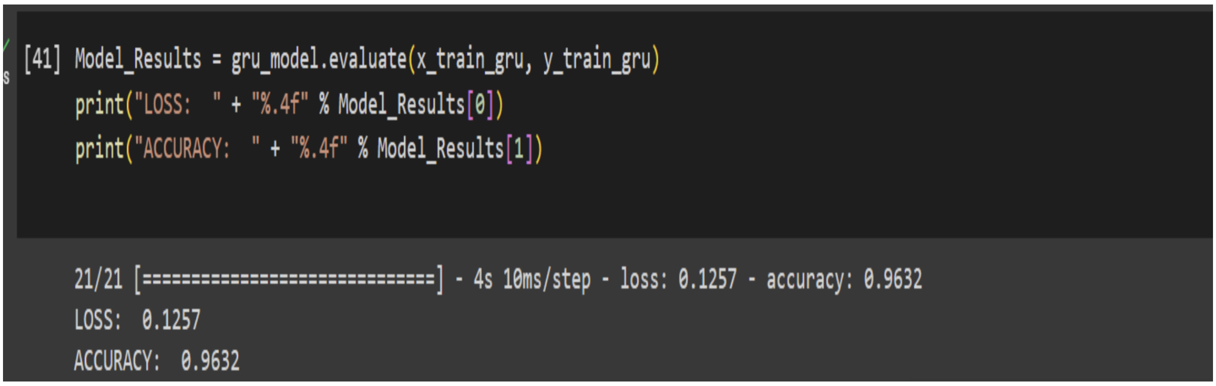


Figure 6.10: Model Training Results of CNN-GRU model

```
print(classification_report(y_testclass, classpreds, target_names=classes))

              precision    recall  f1-score   support

        COPD       1.00      0.95      0.97        20
Bronchiolitis      1.00      1.00      1.00        12
   Pneumonia       1.00      0.88      0.93         8
        URTI       1.00      1.00      1.00         3
     Healthy       0.80      1.00      0.89         8

    accuracy                           0.96        51
   macro avg       0.96      0.97      0.96        51
weighted avg       0.97      0.96      0.96        51
```

Figure 6.11: Model Prediction Scores of CNN-GRU model

The precision, recall, f1-score, support of the CNN-GRU model is calculated and its is seen to have produced a much better scores than the simple GRU model.

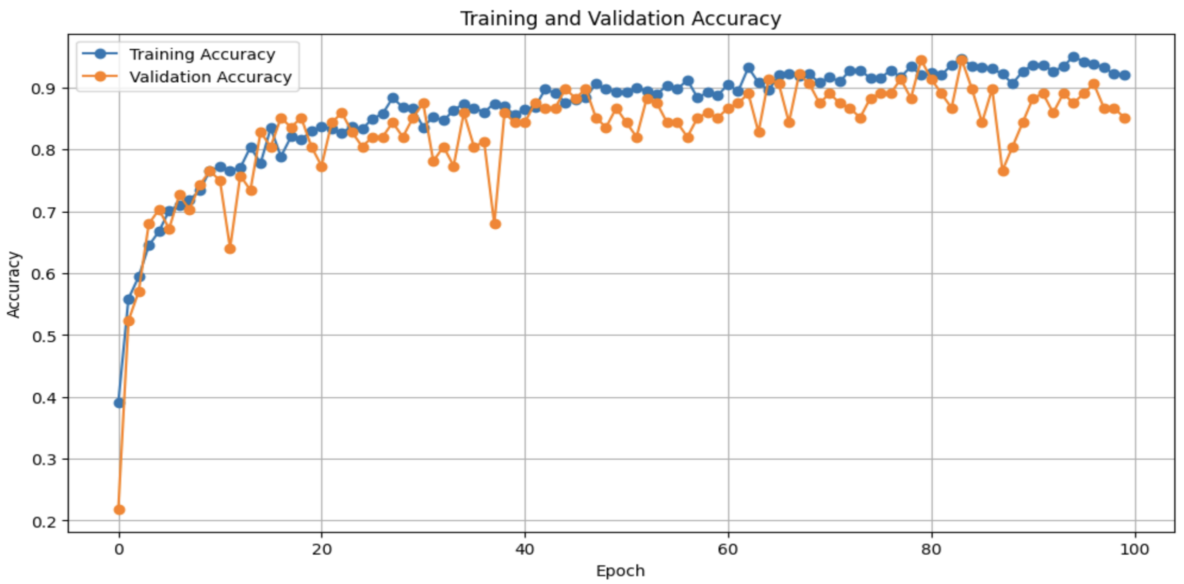


Figure 6.12: Training and Validatiom Graph of CNN-GRU model

A graph was plotted showing the training and validation accuracy calculated for each epoch in the training of the CNN-GRU model.It is observed that the values of the training accuracy and validation accuracy have most of their values near to each other.

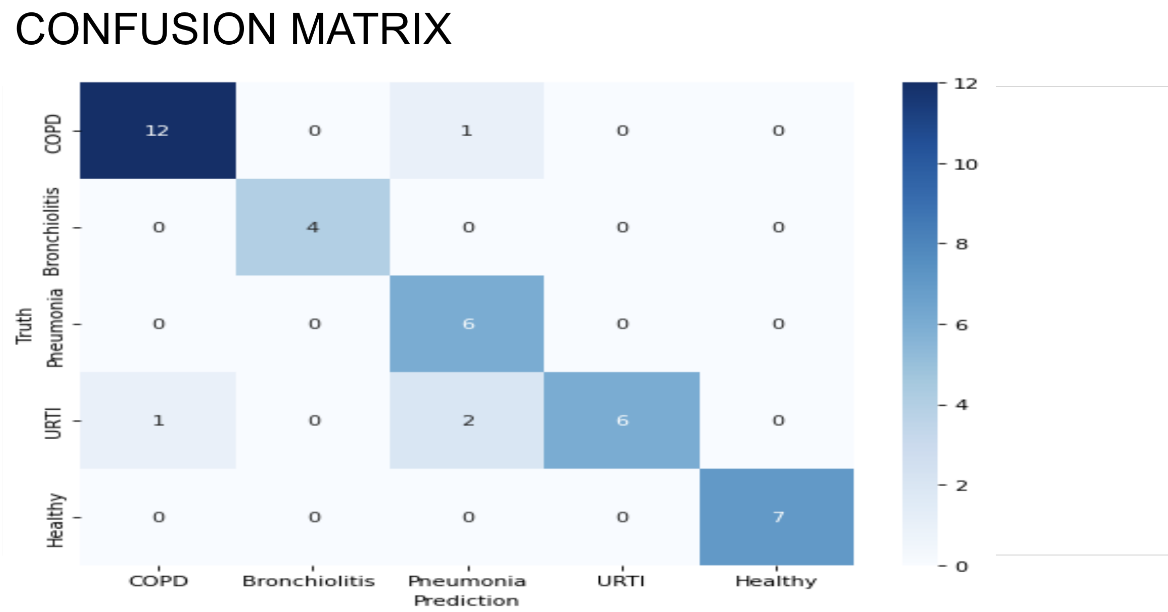


Figure 6.13: Model Confusion Matrix of CNN-GRU model

The confusion matrix provides a detailed breakdown of the GRU model's performance by presenting the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. It allows us to assess the model's performance across different classes. The diagonal elements represent the correctly classified instances, while off-diagonal elements indicate misclassifications. By analyzing the confusion matrix, we are able to evaluate the model's precision, recall, and overall accuracy, gaining valuable insights into its strengths and weaknesses

Chapter 7

Conclusions & Future Scope

In conclusion, this project code provides a robust and comprehensive framework for the analysis and classification of respiratory sound data, leveraging the capabilities of a Gated Recurrent Unit (GRU) and Convolutional Neural Network (CNN) hybrid model. The systematic stages of data acquisition, augmentation, and preprocessing lay a robust foundation, ensuring the creation of a balanced and enriched dataset for effective model training. The trained CNN-GRU model showcases commendable accuracy and validation performance, as underscored by the meticulous evaluation results on an unseen test set.

Looking ahead, the future scope of this work encompasses a spectrum of exciting possibilities and avenues for expansion. The continuous refinement of the model architecture and meticulous hyperparameter tuning holds the promise of elevating its performance to unprecedented levels. Exploring novel and sophisticated data augmentation techniques, as well as considering the integration of transfer learning from pre-trained models, could unlock new dimensions for enhancing the model's capabilities. The integration of real-time monitoring functionalities and the deployment of the model in clinical settings represent promising avenues for the development of an efficient and automated diagnostic tool with tangible real-world impact.

Furthermore, fostering collaboration with healthcare professionals and domain experts is pivotal for refining the model's interpretability and ensuring alignment with rigorous clinical standards. The ongoing adaptation of the framework to handle diverse datasets encompassing a wide array of respiratory conditions is vital for broadening its applicability and making meaningful contributions to the healthcare domain. As the fields of machine learning and medical technology continue to advance, the exploration of novel features and the integration of multi-modal data offer exciting prospects for gaining deeper insights into respiratory health, thereby enhancing the model's diagnostic capabilities.

In essence, while the current model exemplifies a powerful approach to respiratory sound analysis, its future trajectory holds the promise of continuous innovation and broader applicability. This work stands as a significant contribution to the evolving intersection of artificial intelligence and healthcare, with the potential to substantially improve patient outcomes and provide valuable support to healthcare professionals in the diagnosis of respiratory conditions.

References

[1] K. K. Lella and A. Pja, “Automatic diagnosis of covid-19 disease using deep convolutional neural network with multi-feature channel from respiratory sound data: cough, voice, and breath,” *Alexandria Engineering Journal*, vol. 61, no. 2, pp. 1319–1334, 2022.

[2] N. Baghel, V. Nangia, and M. K. Dutta, “Alsd-net: Automatic lung sounds diagnosis network from pulmonary signals,” *Neural Computing and Applications*, vol. 33, 12 2021.

[3] S. Borwankar, J. P. Verma, R. Jain, and A. Nayyar, “Improvise approach for respiratory pathologies classification with multilayer convolutional neural networks,” *Multimedia Tools and Applications*, vol. 81, no. 27, pp. 39 185–39 205, 2022.

[4] E. Messner, M. Fediuk, P. Swatek, S. Scheidl, F.-M. Smolle-Jüttner, H. Olschewski, and F. Pernkopf, “Multi-channel lung sound classification with convolutional recurrent neural networks,” *Computers in Biology and Medicine*, vol. 122, p. 103831, 05 2020.

[5] A. Alqudah, S. Qazan, and Y. Obeidat, “Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds,” *Soft Computing*, vol. 26, pp. 1–25, 09 2022.

[6] F. Jin, F. Sattar, and D. Goh, “New approaches for spectro-temporal feature extraction with applications to respiratory sound classification,” *Neurocomputing*, vol. 123, pp. 362–371, 01 2014.

[7] F. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with lstm,” *Neural computation*, vol. 12, pp. 2451–71, 10 2000.

[8] M. Fraiwan, L. Fraiwan, M. Alkhodari, and O. Hassanin, “Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term

memory," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, 04 2021.

[9] X. Valero and F. Alías-Pujol, "Gammatone cepstral coefficients: Biologically inspired features for non-speech audio classification," *Multimedia, IEEE Transactions on*, vol. 14, pp. 1684–1689, 12 2012.

[10] s. Jayalakshmy and G. Sudha, "Gtcc-based bilstm deep-learning framework for respiratory sound classification using empirical mode decomposition," *Neural Computing and Applications*, vol. 33, 12 2021.

[11] S. R. Dubey, S. Chakraborty, S. Roy, S. Mukherjee, S. K. Singh, and B. Chaudhuri, "diffgrad: An optimization method for convolutional neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, pp. 1–12, 12 2019.

[12] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional lstm," 12 2013, pp. 273–278.

[13] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," *30th International Conference on Machine Learning, ICML 2013*, pp. 1139–1147, 01 2013.

[14] D. Reynolds and R. Rose, "Robust text-independent speaker identification using gaussian mixture speaker models," *Speech and Audio Processing, IEEE Transactions on*, vol. 3, pp. 72 – 83, 02 1995.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735–80, 12 1997.

Appendix A: Presentation

FINAL PRESENTATION

Lung Disease Detection using Respiratory Sounds

By:
Abu Jose (U2003008)
Azmina Iqbal (U2003054)
Cathrin Raju (U2003057)
Dona Francis (U2003072)

Project Guide:
Ms. Dincy Paul

LIST OF CONTENTS

- Problem Definition
- Project Objective
- Novelty of Idea and Scope of Implementation
- Literature Review
- Proposed Method
- Architecture Diagram
- Modules in Detail
- Project Gantt Chart
- Hardware and Software Requirements
- Work done (30% Evaluation)
- Work done (60% Evaluation)
- Results (100% Evaluation)
- Future Scope
- Task Distribution
- Conclusion
- References
- Paper Publication

PROBLEM DEFINITION

- Traditional methods of pulmonary disease detection, such as auscultation with a traditional stethoscope may lead to misdiagnosis or other methods like chest X-Rays are expensive

- A model using deep learning for the detection of respiratory diseases like pneumonia, COPD (Chronic Obstructive Pulmonary Disease), bronchiolitis, URTI from respiratory sound obtained using an electronic stethoscope

PROJECT OBJECTIVES

The purpose of detecting pulmonary diseases using lung sound analysis is to improve the

- Early diagnosis
- Monitoring
- Management of respiratory conditions

Lung sound analysis involves the recording and analysis of sounds produced by the respiratory system during breathing. These sounds can provide valuable information about the health of the lungs and airways.

NOVELTY OF IDEA AND SCOPE OF IMPLEMENTATION

- This approach offers a non-intrusive means of diagnosis, minimizing patient discomfort and reducing the risk of complications associated with invasive techniques.
- Its scalability and adaptability allow for integration into hospitals, clinics, and homecare environments alike.
- With the potential for remote monitoring and telemedicine, our approach not only enhances patient care but also reduces the burden on healthcare systems.
- Its versatility ensures widespread adoption, facilitating timely interventions and ultimately improving patient outcomes on a global scale.

LITERATURE REVIEW

TITLE	AUTHOR	METHODOLOGY
GTCC based BiLSTM deep learning framework for respiratory sound classification	S.Jayalakshmy, Gnanou Florence Sudha	Classification using RNN - BiLSTM
Characteristics of breathe sounds and adventitious respiratory sounds	Charbonneau G, Vanderschoot J	Classification of lung sounds
Neural classification of lung sounds using wavelet coefficients	Kandaswamy A, Kumar C, Malmuruga n N	Comprehensive comparison between MFCCs and GTCCs
Classification of Lung sounds based on convolutional neural network	Aykanat M, Ozkan K, Bahar K, Sevgi S	Classification using Convolutional Neural Networks as comparison

TITLE	AUTHOR	METHODOLOGY
Predicting respiratory anomalies and diseases via recurrent neural networks	Perna D, Tagarelli A	Classification using RNN - BiLSTM
Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds	Ali Mohammad Alqudah, Shoroq Qazan, Yusra M	CNN used to extract deep features from the lung sound signals. LSTM used to classify the lung sound signals based on the extracted features.
Recognition of Pulmonary Disease from lung sounds using convolutional neural networks and long short term memory	Fraiwan M, Fraiwan L	Recognizing pulmonary disease from electronically recorded lung sounds using CNN and LSTM units
Lung sounds classification using convolutional neural networks	Dalal Bardou , Kun Zhang , Sayed Mohammad Ahmad	CNNs and SVMs, processed audio data with features such as MFCC and spectrograms.

TITLE	AUTHOR	METHODOLOGY
Detecting Respiratory Pathologies Using Convolutional Neural Networks and Variational Autoencoders for Unbalancing Data	María Teresa García-Ordás , José Alberto Benítez-Andrades	CNN and VAE for respiratory pathology classification. Mel Spectrograms are used to represent audio data.
Classification of lung sounds using convolutional neural networks	Murat Aykanat , Özkan Kılıç , Bahar Kurt , Sevgi Saryal	CNNs and SVMs, processed audio data with features such as MFCC and spectrograms
Automatic Diagnosis of COVID 19 disease using deep convolutional neural networks with multi feature channels from respiratory sound data:cough, voice and breath	Kranthi Kumar Lella, Alphonse Pja	Proposes a multichannel deep CNN model for detecting COVID 19 disease from human respiratory sounds like voice,cough and breath. Implements De noising Audio Encoder(DAE) and GFCC on the data to extract deep features

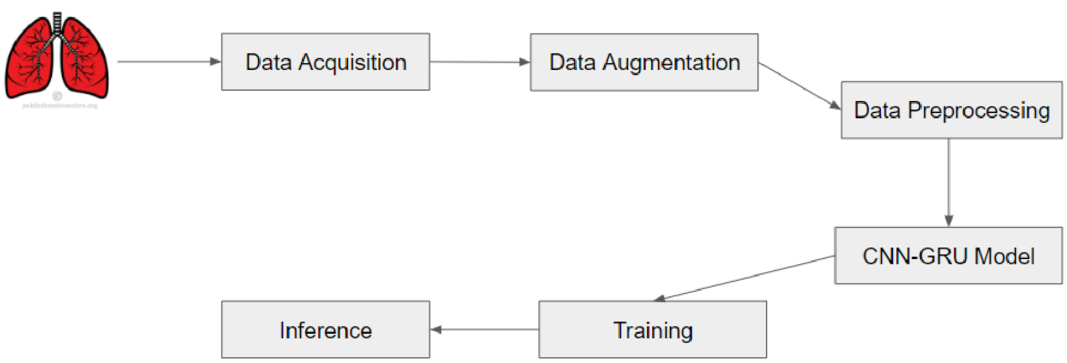
TITLE	AUTHOR	METHODOLOGY
Environment sound classification using Multiple Feature Channels and attention based Deep Convolutional neural network	J.Sharma	In this paper, we propose a model for the Environment Sound Classification Task (ESC) that consists of multiple feature channels given as input to a Deep Convolutional Neural Network (CNN) with Attention mechanism
Deep convolutional neural network and Data Augmentation for Environmental sound classification	J.Salamon , J.P Bello	Using deep CNN model to classify environmental sounds after extracting features from data and augmentation.
Feature extraction and classification of heart sounds using ID convolutional neural network	Fen Li, Ming Liu	proposed a one-dimensional convolutional neural network (CNN) model, which divides heart sound signals into normal and abnormal directly independent of ECG.
An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank	Malcolm Slaney	This report describes various filterbanks such as GFCC MFCC IMFCC and ERB and their implementation

TITLE	AUTHOR	METHODOLOGY
ALSD-Net: Automatic lung sounds diagnosis network from pulmonary signals	Neeraj Baghel, Vivek Nangia, Malay Kishore Dutta	Classification using CNN
Multi-channel lung sound classification with convolutional recurrent neural networks.	Messner E et al	Classification using C-RNN
Neural classification of lung sounds using wavelet coefficients	Kandaswamy A, Kumar C, Ramanathan R, Jayaraman S, Malmurugan N	Classification of lung sounds using ANN and wavelet transforms
Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes.	Bahoura M	Classification of lung sounds using pattern recognition
diffGrad: An Optimization Method for Convolutional Neural Networks	Dubey SR, Chakraborty S, Roy SK, Mukherjee S, Singh SK, Chaudhuri BB	Classification using CNN

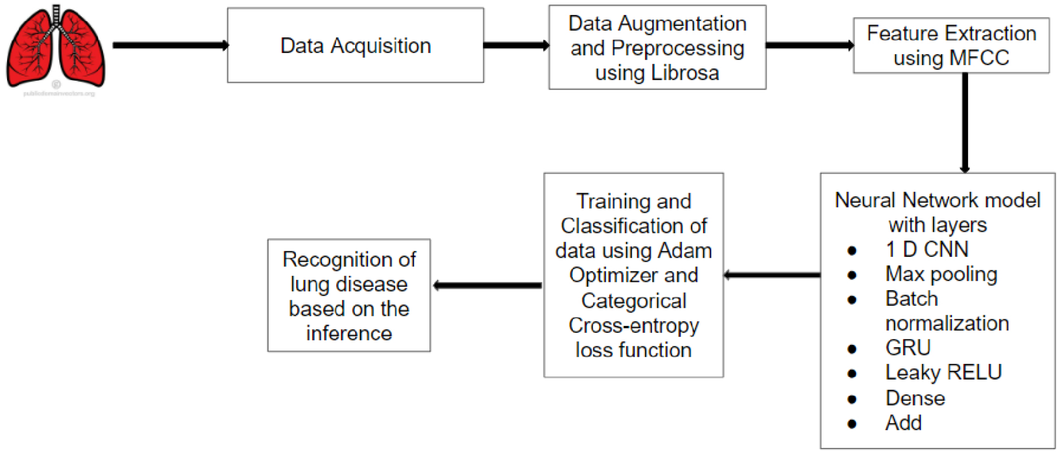
PROPOSED METHOD

- The proposed system in this project is a deep learning architecture specifically designed for **disease recognition using lung sound recordings**
- It will utilize advanced deep learning techniques such as convolutional neural networks (CNNs) to **extract relevant acoustic features from lung sound recordings**
- Signal Augmentation and Pre-processing techniques such as **noise addition, time stretching and pitch shifting (using Librosa)** will be applied to enhance the quantity of the recordings
- The system aims to accurately classify different pulmonary diseases, including **Pneumonia, COPD (Chronic Obstructive Pulmonary Disease), bronchiolitis, Bronchiectasis and URTI** based on extracted features
- Through training and testing, the system aims to **improve diagnostic accuracy and facilitate early detection** of pulmonary diseases using lung sound recordings.

ARCHITECTURE DIAGRAM



MODULES IN DETAIL



MODULES IN DETAIL

Data acquisition and augmentation:

- Setting up the dataset path and obtaining a list of filenames for audio files in the specified directory.
- Extracting patient IDs from filenames and obtaining corresponding labels from the patient diagnosis file.
- Defining functions for data augmentation, including adding noise, time-stretching, and time-shifting.
- Removing audio files associated with rare diseases ('Asthma' and 'LRTI') from the dataset to remove the imbalance.

Feature Extraction using MFCC:

In total **52 MFCC features** were extracted from each audio file they are:

- **Spectral Energy:** Represents the overall energy level of the audio signal across different frequency bands.
- **Spectral Centroid:** Indicates the center of mass of the spectrum, representing the "brightness" of the sound.
- **Spectral Bandwidth:** Represents the width of the frequency range where most of the signal energy is concentrated.
- **Spectral Contrast**: Captures the difference in amplitude between peaks and valleys in the spectrum.
- **Spectral Rolloff:** Indicates the frequency below which a certain percentage of the total spectral energy lies.
- **Zero-Crossing Rate:** Indicates the rate at which the audio signal changes its sign (from positive to negative or vice versa), often related to the frequency content and timbre of the signal.
- **MFCC Coefficients (1-13):** These coefficients capture the spectral envelope of the audio signal after applying a logarithmic transformation to the frequency spectrum.
- **Delta MFCC Coefficients (1-13):** Represent the rate of change of the corresponding MFCC coefficients over time.
- **Delta-Delta MFCC Coefficients (1-13):** Capture the acceleration or second-order rate of change of the MFCC coefficients.
- In total, there are 39 features derived from the MFCCs (13 MFCC coefficients, along with their first and second derivatives), plus additional features such as spectral energy, centroid, bandwidth, and zero-crossing rate, bringing the total to 52 features.

Neural Network Model:

- **Convolutional Blocks:**

  Two convolutional blocks are added, each consisting of a Conv1D layer with increasing filter sizes (256,512). Each Conv1D layer is followed by a MaxPooling1D layer with pool size 2 and 'same' padding. Both convolutional layers use ReLU activation.

- **GRU Layers:**

  Ten GRU Layers are added, each with different filter sizes (32,128,64,128,64,128,128,32,64,32). Following the first six layers an add layer is used to add those layers and then after the remaining four layers another add layer is used to add the output of the previous add layer and the following GRU layers

- **Dense Layers:**

  Two Dense layers are added with 32 units and ReLU activation in the first layer and the number of unique labels (6 in this case) with softmax activation in the output layer

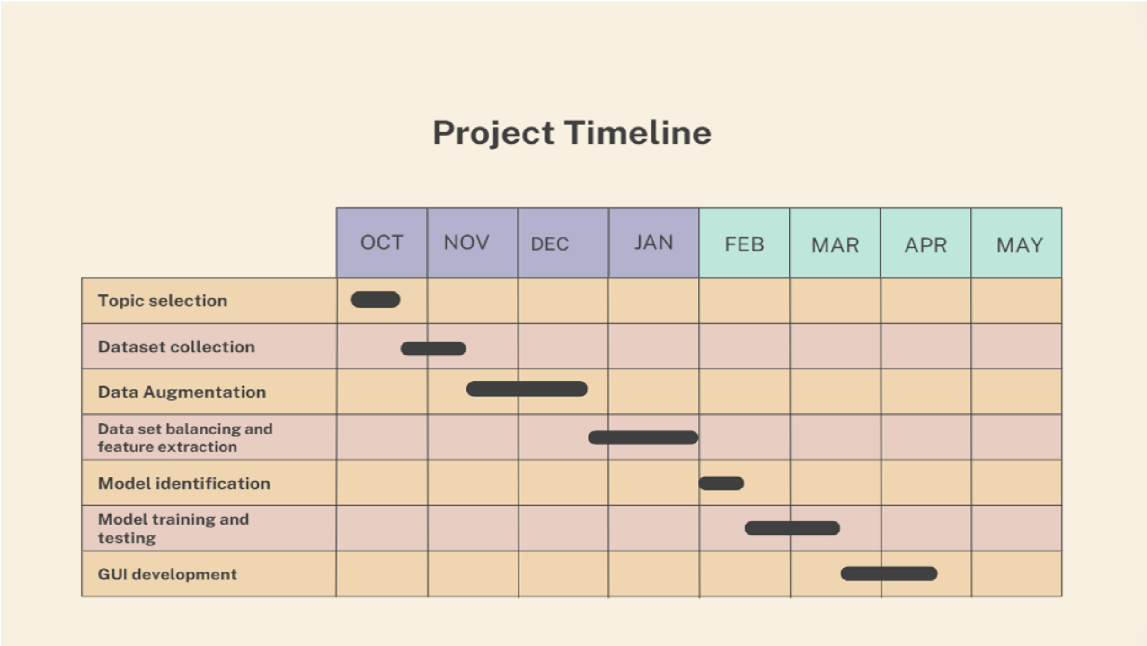
Training and Classification:

- Train the CNN-GRU model based on the pre processed audio data.
- Using Adam optimizer with a learning rate of 0.0001
- Using categorical cross-entropy loss function and accuracy with two callback functions; early stopping and model checkpointing

Inference:

- Once the model is trained, it can be used to infer the pulmonary disease of a patient from their lung sound.
- Load the patient's lung sound audio file
- Extract the Mel-frequency cepstral coefficients from the audio file
- Pass the extracted features to the trained model
- The model will predict the pulmonary disease of the patient.

PROJECT GANTT CHART



HARDWARE AND SOFTWARE REQUIREMENTS

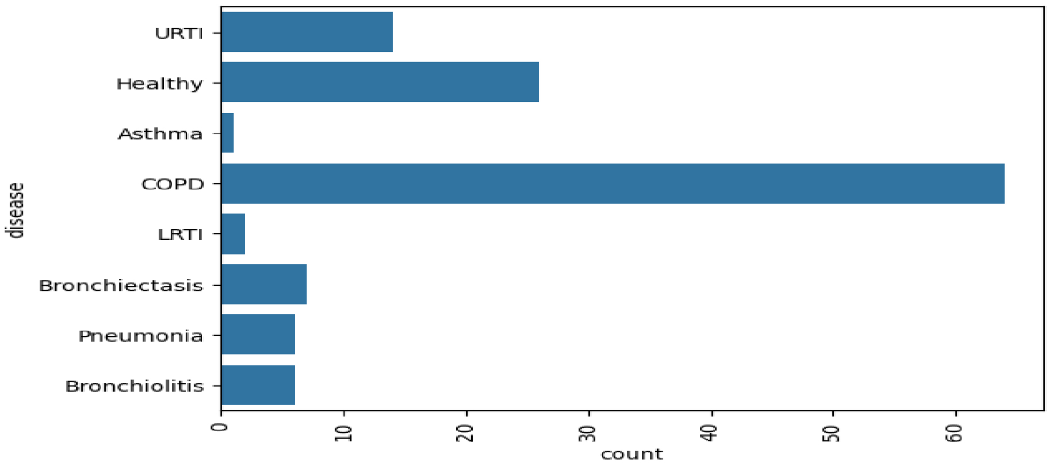
- CPUs and GPUs required to train the deep learning model
- Operating System - software should be compatible with targeting operating system such as Windows or MacOS
- Programming Language - Python
- Development Environment - IDEs should be chosen such as Jupyter Notebook or Anaconda and VS Code
- Signal Processing Libraries - Librosa, Matplotlib, Keras and TensorFlow to perform preprocessing of lung sound data

WORK DONE (30% EVALUATION)

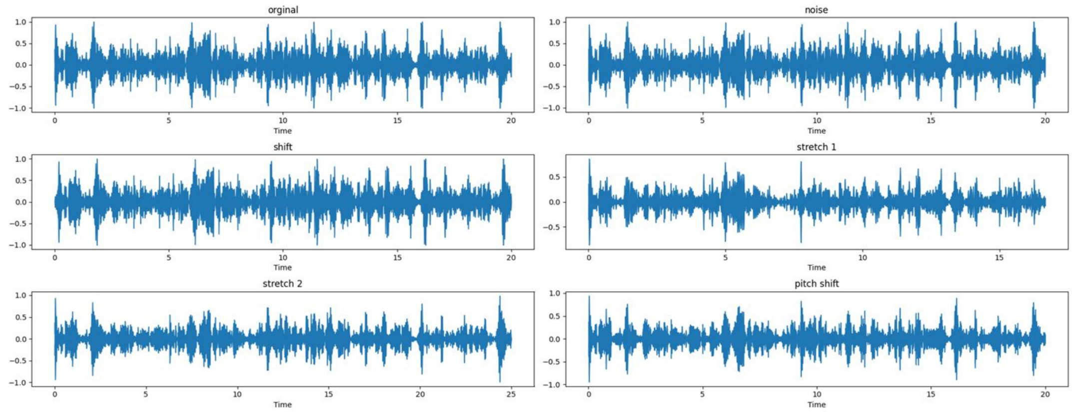
During the 30% completion of this project, the dataset was found and data augmentation was performed. The different data augmentation techniques used were time stretching, shifting, noise addition and pitch shifting.

We also completed the data preprocessing, feature extraction was performed on the data. The features extracted were 52 features derived from the MFCCs from each audio after augmentation along with additional features.

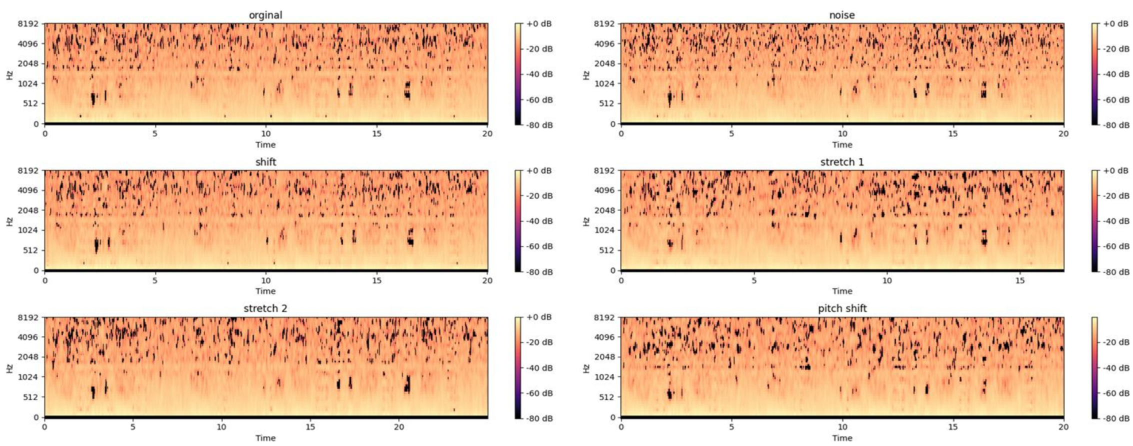
DATA DISTRIBUTION



AUGMENTED AUDIO FORMS



SPECTROGRAM REPRESENTATION OF MFCC FEATURES



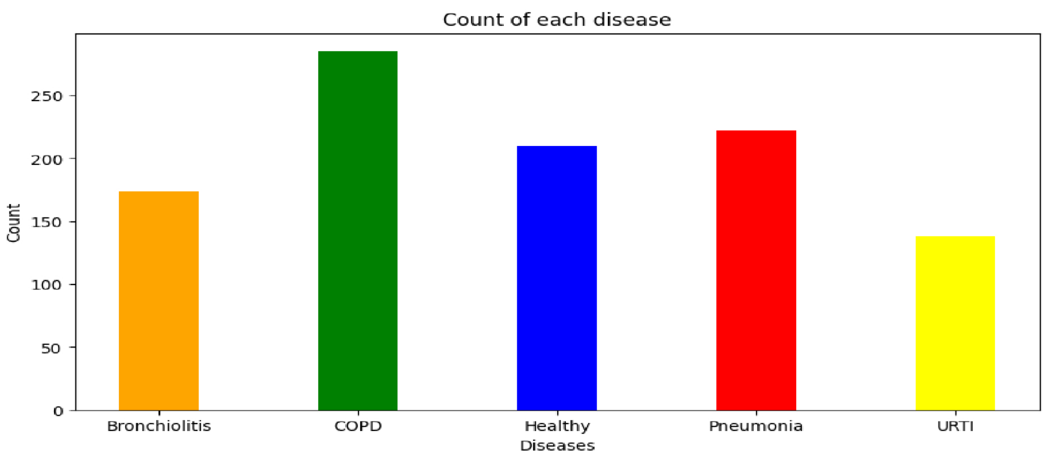
WORK DONE (60% EVALUATION)

After 30% evaluation ,we had mitigated the possibility of overfitting through data augmentation and extracted temporal features from audio using MFCC feature extraction procedure.

The next step was to identify a model suitable for classification, finally we chose GRU model. GRU (Gated Recurrent Unit) is a type of RNN which helps in solving the vanishing gradient problem.

Then we trained the given model along with the extracted features and got an accuracy of 83%.

AUGMENTED DATA DISTRIBUTION



MODEL TRAINING RESULTS

```
[54] Model_Results = gru_model.evaluate(x_train_gru, y_train_gru)
     print("LOSS:  " + "%.4f" % Model_Results[0])
     print("ACCURACY:  " + "%.4f" % Model_Results[1])
```

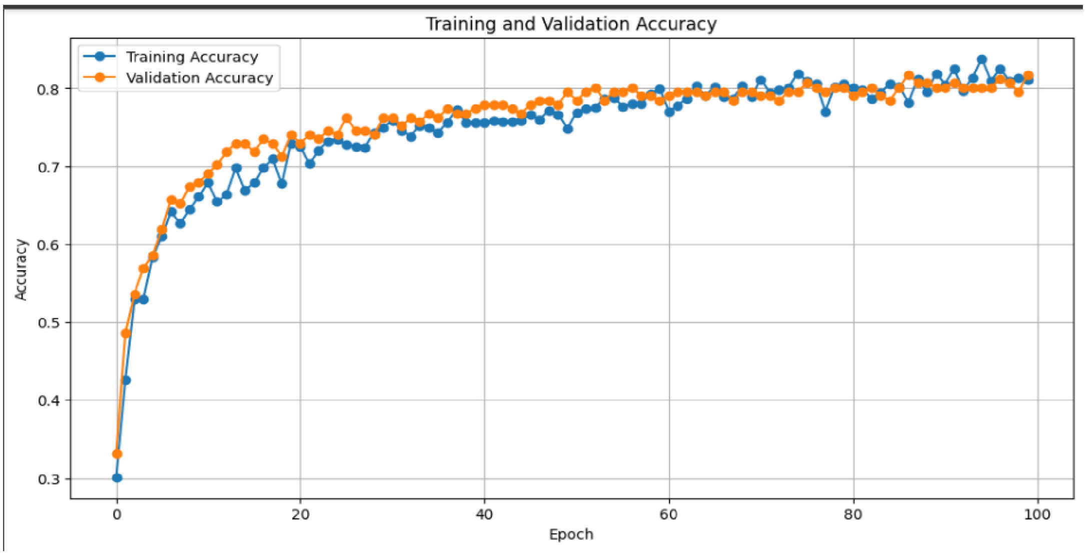
```
25/25 ━━━━━━━━━━━━━━━━━━━━ 0s 5ms/step - accuracy: 0.8888 - loss: 0.2500
LOSS:  0.2816
ACCURACY:  0.8776
```

MODEL PREDICTION SCORES

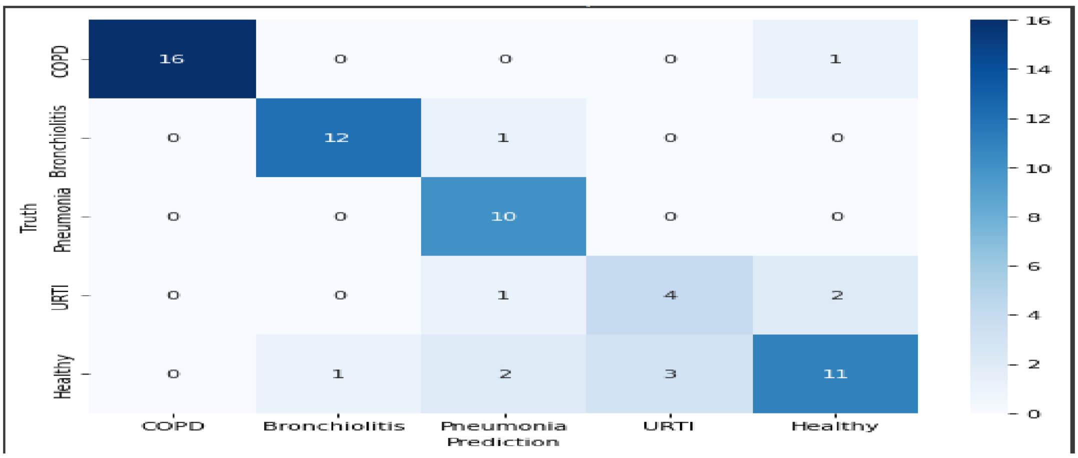
```
print(classification_report(y_testclass, classpreds, target_names=classes))
```

```
              precision    recall  f1-score   support

        COPD       1.00      0.94      0.97        17
Bronchiolitis      0.92      0.92      0.92        13
   Pneumonia       0.71      1.00      0.83        10
        URTI       0.57      0.57      0.57         7
     Healthy       0.79      0.65      0.71        17

    accuracy                           0.83        64
   macro avg       0.80      0.82      0.80        64
weighted avg       0.84      0.83      0.83        64
```

TRAINING VALIDATION GRAPH:



CONFUSION MATRIX



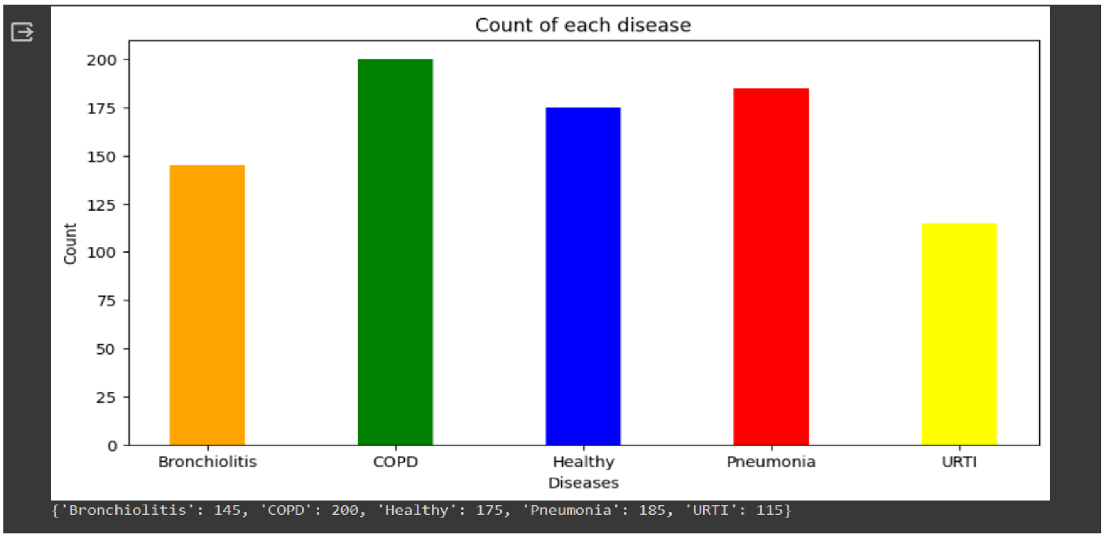
WORK DONE (100% EVALUATION)

Following the work completed during 60% evaluation in order to increase the accuracy of training and validation, the augmentation process was again modified by randomizing the features selected from the class COPD for improving the balance of the data set.

Then the model was modified by adding convolutional layers with the existing GRU model for better accuracy.

A function for diagnosis prediction was developed which seeks audio path as input and predicts the disease corresponding to the audio along with its prediction accuracy.

NEW AUGMENTED DATA DISTRIBUTION



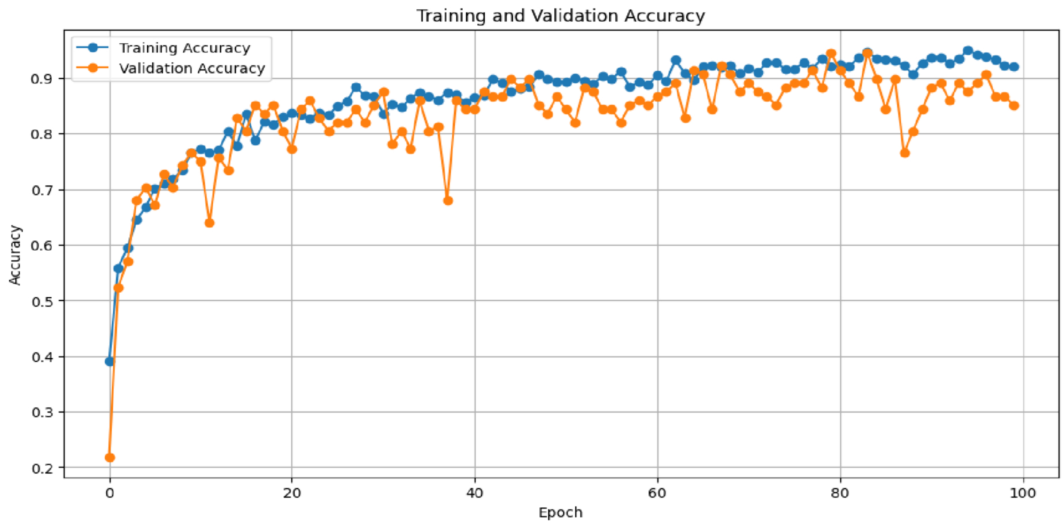
NEW MODEL ACCURACY ON TRAINING:

```
[41] Model_Results = gru_model.evaluate(x_train_gru, y_train_gru)
     print("LOSS:  " + "%.4f" % Model_Results[0])
     print("ACCURACY:  " + "%.4f" % Model_Results[1])

21/21 [==============================] - 4s 10ms/step - loss: 0.1257 - accuracy: 0.9632
LOSS:  0.1257
ACCURACY:  0.9632
```

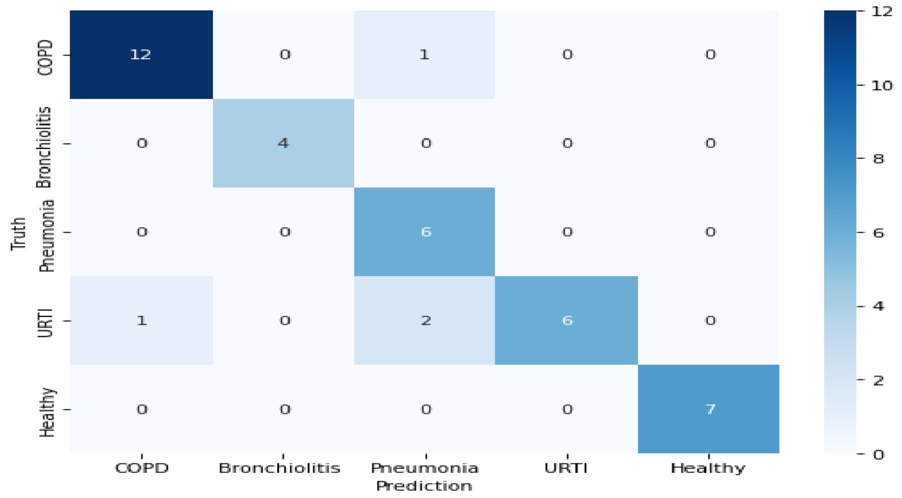
MODEL ACCURACY DATA

```
print(classification_report(y_testclass, classpreds, target_names=classes))

               precision    recall  f1-score   support

         COPD       1.00      0.95      0.97        20
Bronchiolitis       1.00      1.00      1.00        12
    Pneumonia       1.00      0.88      0.93         8
         URTI       1.00      1.00      1.00         3
      Healthy       0.80      1.00      0.89         8

     accuracy                           0.96        51
    macro avg       0.96      0.97      0.96        51
 weighted avg       0.97      0.96      0.96        51
```

NEW TRAINING VALIDATION GRAPH



CONFUSION MATRIX



PREDICTION RESULTS

```
file1 = '/content/drive/MyDrive/Respiratory_Sound_Database/audio_and_txt_files/143_1b1_Al_sc_Meditron.wav'
gru_diagnosis_prediction(file1)
file2 = '/content/drive/MyDrive/Respiratory_Sound_Database/audio_and_txt_files/168_1b1_Al_sc_Meditron.wav'
gru_diagnosis_prediction(file2)
file3 = '/content/drive/MyDrive/Respiratory_Sound_Database/audio_and_txt_files/163_2b2_Ar_mc_AKGC417L.wav'
gru_diagnosis_prediction(file3)
file4 = '/content/drive/MyDrive/Respiratory_Sound_Database/audio_and_txt_files/135_2b3_Pr_mc_LittC2SE.wav'
gru_diagnosis_prediction(file4)
```

```
1/1 [==============================] - 0s 53ms/step
Healthy 0.5463287
1/1 [==============================] - 0s 43ms/step
Bronchiolitis  0.69719744
1/1 [==============================] - 0s 42ms/step
COPD 0.99986315
1/1 [==============================] - 0s 35ms/step
Pneumonia 0.53143257
```

FUTURE SCOPE

- **Telemedicine and Remote Monitoring:** Remote monitoring of lung sounds can facilitate early detection and intervention for pulmonary diseases, particularly in remote or underserved areas where access to healthcare is limited.
- **Integration with Imaging Technologies:** Combining lung sound analysis with imaging technologies such as chest X-rays or CT scans can provide a more comprehensive assessment of pulmonary health. This multimodal approach can improve diagnostic accuracy and aid in treatment planning.
- **Personalized Medicine:** Tailoring pulmonary disease detection and management strategies to individual patients based on their unique lung sound patterns and medical history can optimize treatment outcomes.
- **Population Health Monitoring:** Aggregating and analyzing data from large populations can help identify trends and risk factors associated with pulmonary diseases, leading to better public health interventions and preventive measures.

TASK DISTRIBUTION

Abu Jose	• **Data Augmentation** • **Data Processing**
Azmina Iqbal	• **Data Processing** • **Feature Extraction**
Cathrin Raju	• **Model Designing** • **Model Training**
Dona Francis	• **Model Training** • **Epoch monitoring and reviewing performance analysis (Accuracy)**

CONCLUSION

- Detection of lung diseases using respiratory sounds represents a promising and innovative approach in the field of healthcare by highlighting the significance of technology to analyze respiratory sounds for early diagnosis and detection of lung diseases.
- It offers several advantages, including non-invasiveness, cost-effectiveness, and accessibility, making it a valuable tool for both healthcare professionals and patients.
- Lung disease detection using respiratory sounds represents a promising frontier in healthcare, offering the potential to improve early diagnosis, treatment, and overall patient outcomes.
- For 30% project evaluation, we completed feature extraction as well as conversion to MFCC spectrogram. For 60% project evaluation, we have created the ML model and performed training on the model.
- For completion of the project, testing of the model is being completed. The epoch accuracy is being monitored and performance analysis is being calculated.
- The work of the article for paper publication has been commenced.

REFERENCES

- S. Jayalakshmy and G. F. Sudha, “GTCC-based BiLSTM deep-learning framework for respiratory sound classification using empirical mode decomposition,” Neural Computing and Applications, vol. 33, no. 24, pp. 17029–17040, Jul. 2021, doi: https://doi.org/10.1007/s00521-021-06295-x.

- “Characteristic of breath sounds and adventitious respiratory sounds – ScienceOpen,” www.scienceopen.com. https://www.scienceopen.com/document?vid=98eda31a-ab42-487b-98a3-ab0515dd6cdb (accessed Oct. 10, 2023)

- A. Kandaswamy, C. Sathish. Kumar, Rm. Pl. Ramanathan, S. Jayaraman, and N. Malmurugan, “Neural classification of lung sounds using wavelet coefficients,” Computers in Biology and Medicine, vol. 34, no. 6, pp. 523–537, Sep. 2004, doi: https://doi.org/10.1016/s0010-4825(03)00092-1.

- M. Aykanat, Ö. Kılıç, B. Kurt, and S. Saryal, “Classification of lung sounds using convolutional neural networks,” EURASIP Journal on Image and Video Processing, vol. 2017, no. 1, Sep. 2017, doi: https://doi.org/10.1186/s13640-017-0213-2.

- A. Çınar and S. A. Tuncer, “Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks,” Computer Methods in Biomechanics and Biomedical Engineering, pp. 1–12, Sep. 2020, doi: https://doi.org/10.1080/10255842.2020.1821192.

- A. M. Alqudah, S. Qazan, and Y. M. Obeidat, “Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds,” Soft Computing, Sep. 2022, doi: https://doi.org/10.1007/s00500-022-07499-6.

- A. M. Alqudah, S. Qazan, and I. S. Masad, “Artificial Intelligence Framework for Efficient Detection and Classification of Pneumonia Using Chest Radiography Images,” Journal of Medical and Biological Engineering, Jun. 2021, doi: https://doi.org/10.1007/s40846-021-00631-1.

- M. T. García-Ordás, J. A. Benítez-Andrades, I. García-Rodríguez, C. Benavides, and H. Alaiz-Moretón, “Detecting Respiratory Pathologies Using Convolutional Neural Networks and Variational Autoencoders for Unbalancing Data,” Sensors, vol. 20, no. 4, p. 1214, Feb. 2020, doi: https://doi.org/10.3390/s20041214

- A. Çınar and S. A. Tuncer, “Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks,” Computer Methods in Biomechanics and Biomedical Engineering, pp. 1–12, Sep. 2020, doi: https://doi.org/10.1080/10255842.2020.1821192.

- [1]A. M. Alqudah, S. Qazan, L. Al-Ebbini, H. Alquran, and I. A. Qasmieh, “ECG heartbeat arrhythmias classification: a comparison study between different types of spectrum representation and convolutional neural networks architectures,” Journal of Ambient Intelligence and Humanized Computing, Apr. 2021, doi: https://doi.org/10.1007/s12652-021-03247-0.

- N. Baghel, V. Nangia, and M. K. Dutta, “ALSD-Net: Automatic lung sounds diagnosis network from pulmonary signals,” Neural Computing and Applications, vol. 33, no. 24, pp. 17103–17118, Jul. 2021, doi: https://doi.org/10.1007/s00521-021-06302-1.

- E. Messner et al., “Multi-channel lung sound classification with convolutional recurrent neural networks,” Computers in Biology and Medicine, vol. 122, p. 103831, Jul. 2020, doi: https://doi.org/10.1016/j.compbiomed.2020.103831.

- A. Kandaswamy, C. Sathish. Kumar, Rm. Pl. Ramanathan, S. Jayaraman, and N. Malmurugan, “Neural classification of lung sounds using wavelet coefficients,” Computers in Biology and Medicine, vol. 34, no. 6, pp. 523–537, Sep. 2004, doi: https://doi.org/10.1016/s0010-4825(03)00092-1.

- M. Bahoura, “Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes,” Computers in Biology and Medicine, vol. 39, no. 9, pp. 824–843, Sep. 2009, doi: https://doi.org/10.1016/j.compbiomed.2009.06.011.

- “diffGrad: An Optimization Method for Convolutional Neural Networks,” ieeexplore.ieee.org. https://ieeexplore.ieee.org/abstract/document/8939562?casa_token=GmugfU4ROtQAAAAA:GHzVujV_jQlQXDcAdcEFdTIZZv4hbFMLOyFGzR8KRVXq7gslvQphmQ-ZNmnBHDBT0d6s7XpAUWRTxA (accessed Oct. 10, 2023).

- J. Sharma, O.-C. Granmo, and M. Goodwin, “Environment Sound Classification using Multiple Feature Channels and Attention based Deep Convolutional Neural Network,” 1186-1190, 2020, Accessed: Oct. 10, 2023. [Online]. Available: https://uia.brage.unit.no/uia-xmlui/handle/11250/3056508

- M. Slaney, “An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank Perception Group-Advanced Technology Group.” Accessed: Oct. 10, 2023. [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=07eab74512fa7d1b9554274a61a0d077d4a6ecc

- F. Li et al., “Feature extraction and classification of heart sound using 1D convolutional neural networks,” EURASIP Journal on Advances in Signal Processing, vol. 2019, no. 1, Dec. 2019, doi: https://doi.org/10.1186/s13634-019-0651-3.

- J. Sharma, O.-C. Granmo, and M. Goodwin, “Environment Sound Classification using Multiple Feature Channels and Attention based Deep Convolutional Neural Network,” 1186-1190, 2020, Available: https://uia.brage.unit.no/uia-xmlui/handle/11250/3056508

- K. K. Lella and A. Pja, “Automatic diagnosis of COVID-19 disease using deep convolutional neural network with multi-feature channel from respiratory sound data: Cough, voice, and breath,” Alexandria Engineering Journal, Jun. 2021, doi: https://doi.org/10.1016/j.aej.2021.06.024

PAPER PUBLICATION

The article of this project titled ‘Pulmonary Disease Detection from Lung Sounds using CNN-GRU’ has been communicated to three conferences; International Conference on Advanced Research in Computer Science and Information Technology (ICARCSIT) to be held on June 2nd 2024 and June 14th 2024 in Kochi, India, and International Conference on Artificial Intelligence, Machine Learning and Big Data Engineering (ICAIMLBDE) to be held on June 7th 2024 in Kochi, India.

Thank You

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes

Vision, Mission, Programme Outcomes and Course Outcomes

Institute Vision

To evolve into a premier technological institution, moulding eminent professionals with creative minds, innovative ideas and sound practical skill, and to shape a future where technology works for the enrichment of mankind.

Institute Mission

To impart state-of-the-art knowledge to individuals in various technological disciplines and to inculcate in them a high degree of social consciousness and human values, thereby enabling them to face the challenges of life with courage and conviction.

Department Vision

To become a centre of excellence in Computer Science and Engineering, moulding professionals catering to the research and professional needs of national and international organizations.

Department Mission

To inspire and nurture students, with up-to-date knowledge in Computer Science and Engineering, ethics, team spirit, leadership abilities, innovation and creativity to come out with solutions meeting societal needs.

Programme Outcomes (PO)

Engineering Graduates will be able to:

1. Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. Conduct investigations of complex problems: Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. Individual and Team work: Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.

10. Communication: Communicate effectively with the engineering community and with society at large. Be able to comprehend and write effective reports documentation. Make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team. Manage projects in multidisciplinary environments.

12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change.

Programme Specific Outcomes (PSO)

A graduate of the Computer Science and Engineering Program will demonstrate:

PSO1: Computer Science Specific Skills
The ability to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas by understanding the core principles and concepts of computer science and thereby engage in national grand challenges.

PSO2: Programming and Software Development Skills
The ability to acquire programming efficiency by designing algorithms and applying standard practices in software project development to deliver quality software products meeting the demands of the industry.

PSO3: Professional Skills
The ability to apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs thereby evolving as an eminent researcher and entrepreneur.

Course Outcomes (CO)

Course Outcome 1: Model and solve real world problems by applying knowledge across domains (Cognitive knowledge level: Apply).

Course Outcome 2: Develop products, processes or technologies for sustainable and socially relevant applications (Cognitive knowledge level: Apply).

Course Outcome 3: Function effectively as an individual and as a leader in diverse teams and to comprehend and execute designated tasks (Cognitive knowledge level: Apply).

Course Outcome 4: Plan and execute tasks utilizing available resources within timelines, following ethical and professional norms (Cognitive knowledge level: Apply).

Course Outcome 5: Identify technology/research gaps and propose innovative/creative solutions (Cognitive knowledge level: Analyze).

Course Outcome 6: Organize and communicate technical and scientific findings effectively in written and oral forms (Cognitive knowledge level: Apply).

Appendix C: CO-PO-PSO Mapping

CO-PO AND CO-PSO MAPPING

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO 3
CO 1	2	2	2	1	2	2	2	1	1	1	1	2	3		
CO 2	2	2	2		1	3	3	1	1		1	1		2	
CO 3									3	2	2	1			3
CO 4					2			3	2	2	3	2			3
CO 5	2	3	3	1	2							1	3		
CO 6					2			2	2	3	1	1			3

3/2/1: high/medium/low

JUSTIFICATIONS FOR CO-PO MAPPING

MAPPING	LOW/MEDIUM/ HIGH	JUSTIFICATION
100003/ CS722U.1- PO1	M	Knowledge in the area of technology for project development using various tools results in better modeling.
100003/ CS722U.1- PO2	M	Knowledge acquired in the selected area of project development can be used to identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions.

100003/ CS722U.1-PO3	M	Can use the acquired knowledge in designing solutions to complex problems.
100003/ CS722U.1-PO4	M	Can use the acquired knowledge in designing solutions to complex problems.
100003/ CS722U.1-PO5	H	Students are able to interpret, improve and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
100003/ CS722U.1-PO6	M	Students are able to interpret, improve and redefine technical aspects by applying contextual knowledge to assess societal, health and consequential responsibilities relevant to professional engineering practices.
100003/ CS722U.1-PO7	M	Project development based on societal and environmental context solution identification is the need for sustainable development.
100003/ CS722U.1-PO8	L	Project development should be based on professional ethics and responsibilities.
100003/ CS722U.1-PO9	L	Project development using a systematic approach based on well defined principles will result in teamwork.
100003/ CS722U.1-PO10	M	Project brings technological changes in society.

100003/ CS722U.1-PO11	H	Acquiring knowledge for project development gathers skills in design, analysis, development and implementation of algorithms.
100003/ CS722U.1-PO12	H	Knowledge for project development contributes engineering skills in computing & information gatherings.
100003/ CS722U.2-PO1	H	Knowledge acquired for project development will also include systematic planning, developing, testing and implementation in computer science solutions in various domains.
100003/ CS722U.2-PO2	H	Project design and development using a systematic approach brings knowledge in mathematics and engineering fundamentals.
100003/ CS722U.2-PO3	H	Identifying, formulating and analyzing the project results in a systematic approach.
100003/ CS722U.2-PO5	H	Systematic approach is the tip for solving complex problems in various domains.
100003/ CS722U.2-PO6	H	Systematic approach in the technical and design aspects provide valid conclusions.
100003/ CS722U.2-PO7	H	Systematic approach in the technical and design aspects demonstrate the knowledge of sustainable development.

100003/ CS722U.2-PO8	M	Identification and justification of technical aspects of project development demonstrates the need for sustainable development.
100003/ CS722U.2-PO9	H	Apply professional ethics and responsibilities in engineering practice of development.
100003/ CS722U.2-PO11	H	Systematic approach also includes effective reporting and documentation which gives clear instructions.
100003/ CS722U.2-PO12	M	Project development using a systematic approach based on well defined principles will result in better teamwork.
100003/ CS722U.3-PO9	H	Project development as a team brings the ability to engage in independent and lifelong learning.
100003/ CS722U.3-PO10	H	Identification, formulation and justification in technical aspects will be based on acquiring skills in design and development of algorithms.
100003/ CS722U.3-PO11	H	Identification, formulation and justification in technical aspects provides the betterment of life in various domains.
100003/ CS722U.3-PO12	H	Students are able to interpret, improve and redefine technical aspects with mathematics, science and engineering fundamentals for the solutions of complex problems.

100003/ CS722U.4-PO5	H	Students are able to interpret, improve and redefine technical aspects with identification formulation and analysis of complex problems.
100003/ CS722U.4-PO8	H	Students are able to interpret, improve and redefine technical aspects to meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
100003/ CS722U.4-PO9	H	Students are able to interpret, improve and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
100003/ CS722U.4-PO10	H	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools for better products.
100003/ CS722U.4-PO11	M	Students are able to interpret, improve and redefine technical aspects by applying contextual knowledge to assess societal, health and consequential responsibilities relevant to professional engineering practices.
100003/ CS722U.4-PO12	H	Students are able to interpret, improve and redefine technical aspects for demonstrating the knowledge of, and need for sustainable development.
100003/ CS722U.5-PO1	H	Students are able to interpret, improve and redefine technical aspects, apply ethical principles and commit to

		professional ethics and responsibilities and norms of the engineering practice.
100003/ CS722U.5-PO2	M	Students are able to interpret, improve and redefine technical aspects, communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
100003/ CS722U.5-PO3	H	Students are able to interpret, improve and redefine technical aspects to demonstrate knowledge and understanding of the engineering and management principle in multidisciplinary environments.
100003/ CS722U.5-PO4	H	Students are able to interpret, improve and redefine technical aspects, recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.
100003/ CS722U.5-PO5	M	Students are able to interpret, improve and redefine technical aspects in acquiring skills to design, analyze and develop algorithms and implement those using high-level programming languages.
100003/ CS722U.5-PO12	M	Students are able to interpret, improve and redefine technical aspects and contribute their engineering skills in computing and information engineering domains like network design and administration, database design and

		knowledge engineering.
100003/ CS722U.6-PO5	M	Students are able to interpret, improve and redefine technical aspects and develop strong skills in systematic planning, developing, testing, implementing and providing IT solutions for different domains which helps in the betterment of life.
100003/ CS722U.6-PO8	H	Students will be able to associate with a team as an effective team player for the development of technical projects by applying the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
100003/ CS722U.6-PO9	H	Students will be able to associate with a team as an effective team player to Identify, formulate, review research literature, and analyze complex engineering problems
100003/ CS722U.6-PO10	M	Students will be able to associate with a team as an effective team player for designing solutions to complex engineering problems and design system components.
100003/ CS722U.6-PO11	M	Students will be able to associate with a team as an effective team player, use research-based knowledge and research methods including design of experiments, analysis and interpretation of data.
100003/ CS722U.6-PO12	H	Students will be able to associate with a team as an effective team player, applying ethical principles and

		commit to professional ethics and responsibilities and norms of the engineering practice.
100003/ CS722U.1-PSO1	H	Students are able to develop Computer Science Specific Skills by modeling and solving problems.
100003/ CS722U.2-PSO2	M	Developing products, processes or technologies for sustainable and socially relevant applications can promote Programming and Software Development Skills.
100003/ CS722U.3-PSO3	H	Working in a team can result in the effective development of Professional Skills.
100003/ CS722U.4-PSO3	H	Planning and scheduling can result in the effective development of Professional Skills.
100003/ CS722U.5-PSO1	H	Students are able to develop Computer Science Specific Skills by creating innovative solutions to problems.
100003/ CS722U.6-PSO3	H	Organizing and communicating technical and scientific findings can help in the effective development of Professional Skills.