Department of Electrical and Computer Engineering
North South University

Senior Design Project

Classification of Respiratory Diseases using Machine Learning

1. Name: Md Jushraf Rahman Chowdhury **ID: 1812901042**

2. Name: Mir Sadia Afrin **ID: 1712366642**

Faculty Advisor

DR. MOHAMMAD ASHRAFUZZAMAN KHAN

Assistant Professor

Department of ECE

Fall 2021

LETTER OF TRANSMITTAL

31st December, 2021

To

Dr. Mohammad Rezaul Bari

Associate Professor and Chairman,

Department of Electrical and Computer Engineering,

North South University, Dhaka.

Subject: Submission of Final Project on "Classification of Respiratory Diseases using Convolutional Neural Network"

Dear Sir,

With due respect, we would like to submit our CSE499B Final Report on "Classification of Respiratory Diseases using Convolutional Neural Network" as a part of our BSc program. The report deals with the process of classifying respiratory diseases from audio recording of breathing sound. To perform this efficiently, we collected dataset from ICBHI 2017 database which contains 920 annotated breathing sound of patients. In this report we tried our best in explaining the entire methodology starting from data pre-processing, feature extraction, classification and results.

The final project part 'B' was very much valuable to us as it helped us to gain experience on various Neural Network Architectures like CNN, VGG16, ResNET50,etc. It was a great learning experience for us. We tried to the maximum competence to meet all the dimensions required from this report.

We will be highly obliged if you are kind enough to receive this report and provide your valuable judgment. It would be our immense pleasure if you find this report useful and informative to have an apparent perspective on the issue.

Sincerely Yours,

Md Jushraf Rahman Chowdhury

ID: 1812901042

.....

Department of ECE

North South University, Bangladesh

Mir Sadia Afrin
ID: 1712366642

.....

Department of ECE

North South University, Bangladesh

APPROVAL

The capstone project entitled "**Classification of Respiratory Diseases using Convolutional Neural Network**" by Md Jushraf Rahman Chowdhury (ID # 1812901042) and Mir Sadia Afrin (ID # 1712366642), is approved in partial fulfillment of the requirement of the Degree of Bachelor of Science in Computer Science and Engineering in December, 2021 and has been accepted as satisfactory.

Supervisor:

.....

Dr. Mohammad Ashrafuzzaman Khan

Assistant Professor

Department of Electrical and Computer Engineering

North South University

Dhaka, Bangladesh.

Department Chair:

.....

Dr. Mohammad Rezaul Bari

Associate Professor & Chairman

Department of Electrical and Computer Engineering

North South University

Dhaka, Bangladesh.

DECLARATION

This is to certify that this **"Senior Design Project"** is our original work. No part of this work has been submitted elsewhere partially or fully for the award of any other degree or diploma previously made by any other team. Any material reproduced in this project has been properly acknowledged. We also express our honest confirmation in support of the fact that the said "Senior Design Project Report" has neither been used before to fulfill any other course related purpose nor it will be submitted to any other team or authority in future.

Students' names & Signatures

Md Jushraf Rahman Chowdhury
ID: 1812901042
.....
Department of ECE
North South University, Bangladesh

Mir Sadia Afrin
ID: 1712366642
.....
Department of ECE
North South University, Bangladesh

ACKNOWLEDGEMENT

By mercy of the Almighty we have completed our Senior design project part “B” (CSE499B) entitled “**Classification of Respiratory Diseases using Convolutional Neural Network**”. Foremost, we would like to express our sincere gratitude to our advisor Dr. Mohammad Ashrafuzzaman Khan for his continuous support in our project, for his patience, motivation, enthusiasm, guidance and immense knowledge. Our sincere thanks also goes to North South University, Dhaka, Bangladesh for providing an opportunity in our curriculum which enabled us to have a professional level experience as part of our academics. We are also thankful to our team members for their help and support in this project. Every member did their best to fulfil their target.

ABSTRACT

Respiratory diseases are leading causes of death and disability in the world. One of the most important indicators of respiratory disorders and health is respiratory sound. There is a direct relation between the sound emitted by a person and the air movement, lung secretions and aerobic and anaerobic changes within the lung. In this research we have tried to automate the diagnosis of respiratory diseases from respiratory sound. We have used the ICBHI 2017 Respiratory Sound Database where there are 920 recordings acquired from 126 individual resulting in a total of 6898 respiration cycles. In this course we have successfully classified the breathing recordings and achieved an accuracy of 86.95%. This classification has the potential to detect respiratory abnormalities in early stages of respiratory diseases which can improve the outcome of decision making. It also has the potential to serve the medically underserved population by being used on areas lacking skilled medical staff.

Keywords: *Deep Learning, Convolutional Neural Network (CNN), Respiratory Disease, Respiratory Sound Dataset, VGG16, Respiratory sound classification.*

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Respiratory diseases, or lung diseases, are pathological conditions affecting the organs and tissues that make gas exchange difficult in air-breathing animals. The symptoms can be depended on the respiratory condition, such as - shortness of breath with activity, persistent cough, wet cough with mucus (COPD, cystic fibrosis, asthma), dry, "nonproductive" cough (pulmonary fibrosis), chest tightness, wheezing, rapid shallow breathing, hoarseness. It causes various diseases like lung cancer, asthma, Tuberculosis, Chronic Obstructive Pulmonary Disease (COPD), etc. Every year, a lot of people die from respiratory diseases. Still in Bangladesh, the lacking of immediate diagnosis of the diseases increasing the mortality rate.

Our goal is to detect respiratory anomalies in early stages which can enhance the diagnosis process and play a role in decreasing respiratory fatalities.

1.1 WHO DATA

In Bangladesh, respiratory diseases have been increasing rapidly. According to WHO, every year thousands of people die from various respiratory diseases. The latest WHO data published in 2018 -

- Lung Cancers Deaths in Bangladesh reached 13,265 or 1.71% of total deaths.
- Tuberculosis Deaths in Bangladesh reached 65,746 or 8.47% of total deaths.
- Asthma Deaths in Bangladesh reached 14,674 or 1.89% of total deaths.

1.2 LEADING CAUSES OF DEATH IN BANGLADESH

Respiratory diseases are on lead caused death in Bangladesh. And many of these happen because of misdiagnosis, late diagnosis, unserved treatment etc. Here are some lead diseases which caused deaths in Bangladesh:

- **Lung Cancer**

  Lung Cancer is one of the most common disease in Bangladesh which causes death

every year and according to newest study, the cases are rising rapidly. *According to the latest Hospital Cancer Registry Study, there is a reportedly near 200% rise in the country's lung cancer burden in just three years.*

- **Lower Respiratory Infections**

  About 7% of all death in Bangladesh are caused by lower respiratory infections and it is leading second in mortality rate in the country. 28% of children death are also caused by various respiratory infections.

- **Chronic Obstructive Pulmonary Disease**

  Chronic Obstructive Pulmonary Disease (COPD) is one of the dangerous diseases that caused death in Bangladesh very often. This is a collective term for ailments which affect the lung. These cause emphysema and bronchitis. COPD causes 7% of all deaths in this country.

- **Tuberculosis**

  According to World Health Organization, Bangladesh is the "High TB Burden Country" because of tuberculosis. 3% of all deaths in Bangladesh are caused by tuberculosis.

- **Asthma**

  Asthma is one of the most common disease that every age of people can have in Bangladesh. Even there is a huge record of children who has asthma at a very early age.

- **COVID-19**

  COVID-19 is now the most caused disease witch mainly affecting the lung and destroy the whole respiratory system and causing enormous deaths every day not just in Bangladesh, but also in the world.

1.3 ABOUT THIS PAPER

In this paper, we reviewed related literature on respiratory diseases and how to solve the issues of identifying the diseases by classifying them. We studied research papers on machine learning, convolutional neural network (CNN), etc. We chose to use the Convolutional Neural Network (CNN) to classify respiratory diseases to get better accuracy. The idea was to create a model to classify respiratory diseases from breathing sound. There was a lack of resources on this particular field of breathing sound dataset and also using ML or CNN for implementation. We collected a small audio dataset from kaggle to classify respiratory diseases which contains 920 audio recordings of breathing sound. The target was to train a cnn model for better accuracy and classification. To implement this whole project, from reviewing papers to coding, there were few tools we had to use - such as - Kaggle, Google Scholars, Medium, Youtube Sources for datasets, resources, research papers; Google Colab, Jupyter Notebook for implementation; Various machine learning libraries like Tensorflow, Keras, Pandas, Librosa, Scikit-learn, Numpy. After getting the classification we implemented different models for comparison- they are - VGG16 and ResNet50. Here are the models we used –

1.3.1 MODELS

- **Convolutional Neural Network (CNN)[10][11][12] -**

Convolutional neural network (CNN) is a deep learning neural network and it is designed to analyze visual images. It is a class of artificial neural network. Convolutional neural networks have many applications in computer vision, image and video recognition, recommender systems, image classification, image segmentation, medical image analysis, natural language processing, etc.

It has multiple layers. It processes data that has a grid-like arrangement then extracts important features. CNN doesn't need to do a lot of pre-processing on images. For image processing, CNN can filter the important characteristics and saves much time.

The convolutional neural network algorithm's main purpose is to get data into forms that are easier to process without losing the features that are important for figuring out what the data represents. This also makes them great candidates for handling huge datasets.

- **VGG16[13][14][15] -**

VGG16 model was proposed by Karen Simonyan and Andrew Zisserman at the University of Oxford (2014) in the paper "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION". VGG16 is a CNN architecture which is majorly used for ImageNet which has been identified as the best performing model dataset for this field. This model could achieve 92.7% top-5 test accuracy in ImageNet that has a dataset of over 14 million images which belonged to 1000 classes.

VGG16 has been used in many image classification techniques which is popular because of its implementation.

- **ResNet[5][16] -**

Residual Network, ResNet is a very successful deep convolutional neural network used for image classification. It is more speed/memory efficient than dense network.

We have used ResNet 50 in our implementation. This model has 48 Convolution layers, 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. It is a widely used ResNet model and we have explored ResNet50 architecture in depth. ResNet50 is a pre-trained deep learning model for image classification of CNN.

We converted audio files to image data (MFCC) and used them as input to our CNN and VGG16 models. During pre-processing, audio files Features have been extracted and converted into spectrogram and MFCC -

- **MFCC[17] -**

Mel Frequency Cepstral Co-efficients (MFCC) is an internal audio representation format which is easy to work on. It is a mathematical coefficient for sound modeling. It is similar to the JPG format for images.

- **Audio Spectrogram[18] -**

An audio spectrogram is basically a visual representation of the spectrum of frequencies of an audio signal as it varies with time. It basically represents time, frequency, and amplitude all on one graph.

1.4 AIM

We aim to automate the diagnosis of respiratory diseases in early stages by classifying them from breathing recordings by the help of deep learning. We will use sound data sample and extract the key features and pass it to our deep learning model which will then classify the sound into a class of disease or healthy. We plan on improving the accuracy and efficiency of the existing architecture.

CHAPTER 2: LITERATURE REVIEW

While working on this research project, we studied various research papers related to our project topic and picked a few papers from there which were based on predicting Respiratory Diseases with better accuracy. We chose these particular papers because their working approach is closely related to our research work. We also deduced some ideas from there for doing our work.

2.1 LITERATURE EXPLANATION

These following research works are based on different types of respiratory diseases to predict them in various ways. Scientists have been studying these for years to improve the prediction accuracy.

2.1.1. Lung Disease Classification using Deep Convolutional Neural Network

This is a study by "Zeenat Tariq, Sayed Khushal Shah, Yugyung Lee" from *"School of Computing and Engineering, University of Missouri-Kansas City, USA"* in November 2019. While modern technologies are improving on medicine, researchers are working on more precise and customized treatment strategies for various diseases. This paper aimed for making a better accuracy on the medical field by using deep learning with available datasets. By extracting spectrogram features and labels of the lung sounds sample data to 2D Convolutional Neural Network (CNN) model. Then, normalizing the sound data by removing the peak values and noise, the classifications are made using the deep learning process.

The final accuracy has been gained approximately 97%. This model helps to get an accuracy which will ben accepted in the medical field for better performance.

2.1.2. Asthma and Chronic Obstructive Pulmonary Disease (COPD) – Differences and Similarities

This is a study by "Vesna Cukic, Vladimir Lovre, Dejan Dragisic, Aida Ustamujic" from *"Clinic for Pulmonary Diseases and TB "Podhrastovi", Clinical center of Sarajevo University, Bosnia and Herzegovina"* in 2012.

This paper is focused on Asthma and Chronic Obstructive Pulmonary Disease (COPD). These diseases are highly common in todays world, affecting millions of people. Where Asthma is very common and serious problem, COPD has been causing serious damages like chronic morbidity and mortality. As there are many similarities between Asthma and COPD, there are many differences too such as, differences in etiology, symptoms, type of airway inflammation, inflammatory cells, mediators, consequences of inflammation, response to therapy, course.

So, the aim is to discuss the similarities and differences of these diseases to better understanding pf pathology to do the correct diagnosis for better treatment to reduce the morbidity and mortality.

2.1.3. LungBRN: A Smart Digital Stethoscope for Detecting Respiratory Disease Using bi-ResNet Deep Learning Algorithm

This is a conference paper by "Yi Ma, Xinzi Xu, Qing Yu, Yuhang Zhang, Yongfu Li*, Jian Zhao and Guoxing Wang" from *"Department of Micro-Nano Electronics and MoE Key Lab of Artificial Intelligence Shanghai Jiao Tong University, Shanghai, China"* in Octiober 2019.

To improve health care services for immediate diagnosis, this study is to help the medical staff. The idea is of a digital stethoscope to predict the disease with better accuracy from the basic lung sound to have the immediate diagnosis for respiratory related problems.

The idea is to develop an improved bi-ResNet deep learning architecture, LungBRN, which uses STFT and wavelet feature extraction techniques to improve the accuracy.

This technique has achieved a performance of 50.16% accuracy. Compared to all participating teams from ICBHI 2019.

2.1.4. A Respiratory Sound Database for the Development of Automated Classification

This is a conference paper by "B. M. Rocha[1], D. Filos[2], L. Mendes[1], I. Vogiatzis[2], E. Perantoni[2], E. Kaimakamis[2], P. Natsiavas[2], A. Oliveira[3,4] , C. Jácome[3], A. Marques[3,4], R. P. Paiva[1], I. Chouvarda[2], P. Carvalho[1], N. Maglaveras[2]" from *" [1] Centre for Informatics and Systems, Department of Informatics Engineering, Faculty of Sciences and Technology, University of Coimbra, Coimbra, Portugal; [2] Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece; [3] Lab3R - Respiratory Research and Rehabilitation Laboratory, School of Health Sciences, University of Aveiro, Aveiro, Portugal; [4] Institute for Biomedicine (iBiMED), University de Aveiro, Portugal "* in November 2017.

This paper has focused on the development of the automatic analysis of respiratory diseases by respiratory sound database. As the publicly available large datasets have been used in various new algorithms to evaluate, still there is lacking of better result.

So, in this paper, the idea is to collect respiratory sound data and use it to develop an automated classification of the diseases.

2.1.5. XRAY AI: Lung Disease Prediction Using Machine Learning

This is an International Journal of Information Systems and Computer Sciences by "Justin Monsi[1], Justine Saji[2], Keerthy Vinod[3], Liya Joy[4], Jis Joe Mathew[5]" from *" 1 Amal Jyothi*

College of Engineering, Kottayam, India; 2 Amal Jyothi College of Engineering, Kottayam, India; 3 Amal Jyothi College of Engineering, Kottayam, India; 4Amal Jyothi College of Engineering, Kottayam, India; 5 Amal Jyothi College of Engineering, Kottayam, India " in March-April 2019.

In this paper, the NIH chest X-ray dataset comprised of 112,120X-ray images with disease labels from 30,000 unique patients has been presented. The accuracy is expected to achieve more than 90%. The deep convolutional neural network has been used to diagnosis the disease patterns and locations. Also, the automated detection system is being used.

The goal is to create a system to get better accuracy and do the correct diagnosis without getting a second opinion by reducing misdiagnosis and fatal consequences.

2.1.6. Improving Disease Prediction by Machine Learning

This is an International Research Journal of Engineering and Technology (IRJET) by "Smriti Mukesh Singh[1], Dr. Dinesh B. Hanchate[2]" from *"1 Student of Computer Engineering, Pune University, VPKBIET, Baramati, India; 2 Professor of Computer Engineering, Pune University, VPKBIET, Baramati, India"* in June 2018.

As now biomedical information has a huge data collection. From that big data, medical services need to investigate medicinal information, diagnosis different types of diseases and also maintain the administration. All of these become very difficult with the big amount of data with a proper model, especially disease prediction. As the importance of diagnosis is very high, this paper discussed how to improve the disease prediction using machine learning. The goal of this paper is to improve the disease prediction with better accuracy and correct diagnosis using machine learning model.

2.1.7 Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease [8]

This is a study by "Arpan Srivastava[1] , Sonakshi Jain[1] , Ryan Miranda[1] , Shruti Patil[2] , Sharnil Pandya[2] and Ketan Kotecha[2]" from *"1 CS&IT Dept, Symbiosis Insitute of Technology, Symbiosis International (Deemed University), Pune, Maharastra, India;* published in 11 February, 2021.

As the recent technologies such as machine learning and deep learning are assisting by solving issues of many medical challenges, this paper provides a study to apply CNN based on deep learning methods by giving a detailed and proper analysis from the audio data of COPD. They have used Librosa machine learning library features such as MFCC, Mel-Spectrogram, Chroma, Chroma (Constant-Q) and Chroma CENS. The method of classification accuracy has been enhanced to an ICBHI score of 93%.

2.1.8 Convolutional neural networks based efficient approach for classification of lung diseases [9]

This is a study by "Fatih Demir[1], Abdulkadir Sengur[1] and Varun Bajaj[2]" from *"1 Electrical and Electronics Engineering Dept., Technology Faculty, Firat University, Elazig, Turkey; 2 Discipline of ECE, IIITDM, Jabalpur, India"*, published in 23 December, 2019.

Lung sound analysis using artificial intelligence has been the best treatment to help medical experts as lung diseases are one of the most common causes of death. This paper has used different types of lung sounds, frequencies, noise and background sounds for the classification of the diseases. The short time Fourier transform (STFT) method was considered as time–frequency transformation and also two deep learning based approaches for lung sound classification. First, they pre-trained CNN model for feature extraction, an SVM classifier for classification of lung sounds. Then, they pre-trained CNN model for fine-tuning with spectrogram images and was tested by using ten-fold cross validation. The accuracies for the first and second proposed methods were 65.5% and 63.09%, respectively.

CHAPTER 3: METHODOLOGY

This chapter provides a chronological outline of the different sections of the work. We will talk about the theories, techniques and step by step workflow of our work.

WORKFLOW:

A rudimentary diagram of our work method is shown below:

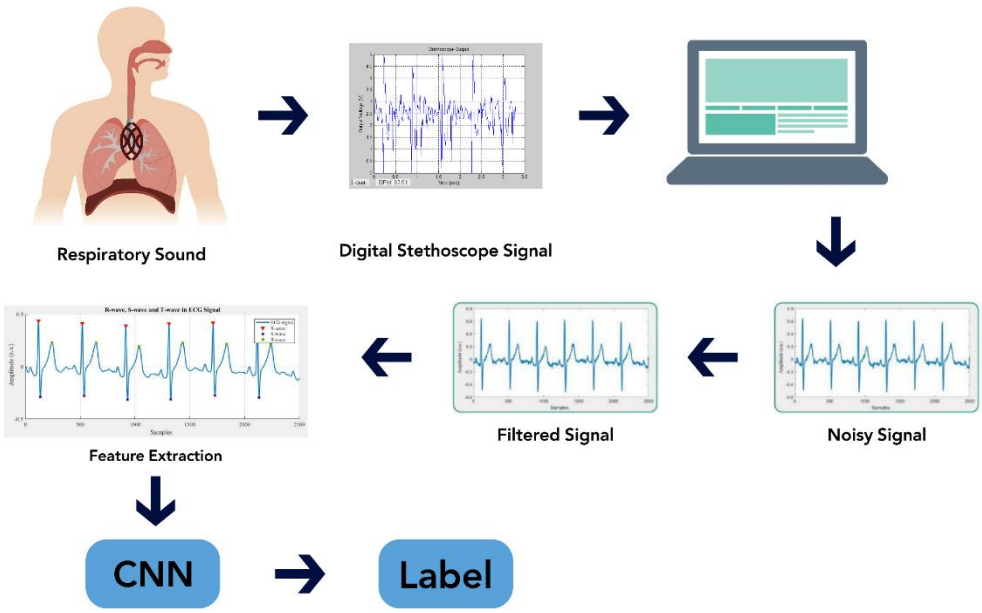


Figure: Fundamental diagram of our work

The above diagram portrays the overall steps of our work for a better understanding. We will record respiratory or breathing sounds of individuals through a digital stethoscope or the listed recording equipment. From there the audio files (.wav format) will stored in a computer and the work of our model starts from there. The noise from the audio files will be filtered out and then will be converted to digital spectrogram and MFCC (mel-frequency cepstrum coefficient) after which it will be passed on to our CNN model. Our model will finally put a label as to which category of the respiratory diseases a certain audio file belonged to. In this way we will classify audio files into different respiratory diseases.

DATASET:

We will use the ICBHI 2017 Respiratory Sound Database. This was created by two research teams in Portugal and Greece. It includes 920 annotated recordings of varying length - 10s to 90s. These recordings were taken from 126 patients. There are a total of 5.5 hours of recordings containing 6898 respiratory cycles - 1864 contain crackles, 886 contain wheezes and 506 contain both crackles and wheezes. The data includes both clean respiratory sounds as well as noisy recordings that simulate real life conditions. The patients span all age groups - children, adults and the elderly.

This dataset includes:

- 920 .wav sound files
- 920 annotation .txt files
- A text file listing the diagnosis for each patient
- A text file explaining the file naming format
- A text file listing 91 names (filename_differences.txt)
- A text file containing demographic information for each patient

PREPROCESSING:

- Since our dataset is of audio type, it is a complicated process for the computer to understand it. So all the audio files has been converted to a visual files which are images. We were successful in converting the audio files to both digital spectrogram and MFCC (mel-frequency cepstrum coefficient) which will then work as input for our CNN models.
- We removed the very rare diseases. In our dataset there were only one patient each for the ‘asthma’ and ‘LRTI’ disease class out of 920 patients. Since they were outliers we removed these classes in order to have a better distribution of our data.

FEATURE EXTRACTION:

So far the features that we extracted from the audio files through our digital spectrogram and MFCC are:

Feature Name	Description
Frequency	From the digital spectrogram we were able to see the intensity of sound at different frequencies
Pitch	From the MFCC we were able to figure out the nature of pitch of the audio file. Higher than origin meant high pitch sound and lower meant low pitch.

DEEP LEARNING MODELS:

We will be using the following models:

CNN: A convolutional neural network (CNN) is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data. CNNs are powerful image processing, artificial intelligence (AI) that use deep learning to perform both generative and descriptive tasks, often using machine vison that includes image and video recognition, along with recommender systems and natural language processing. There are two main parts to a CNN architecture:

- A convolution tool that separates and identifies the various features of the image for analysis in a process called as Feature Extraction
- A fully connected layer that utilizes the output from the convolution process and predicts the class of the image based on the features extracted in previous stages.

VGG16

VGGNet-16 consists of 16 convolutional layers and is very appealing because of its very uniform Architecture. Similar to AlexNet, it has only 3x3 convolutions, but lots of filters. It can be trained on 4 GPUs for 2–3 weeks. It is currently the most preferred choice in the community for extracting features from images. The weight configuration of the VGGNet is publicly available and has been used in many other applications and challenges as a baseline feature extractor. However, VGGNet consists of 138 million parameters, which can be a bit challenging to handle. VGG can be achieved through transfer Learning

ResNet50

Residual Network, ResNet is a very successful deep convolutional neural network used for image classification. It is more speed/memory efficient than dense network. This model has 48 Convolution layers, 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. It is a widely used ResNet model and we have explored ResNet50 architecture in depth. ResNet50 is a pre-trained deep learning model for image classification of CNN.

CHAPTER 4: RESULTS

Initially we converted the audio files to visual images. We converted some of the audio files to digital spectrogram and MFCC where we found the following results:

Digital Spectrogram:

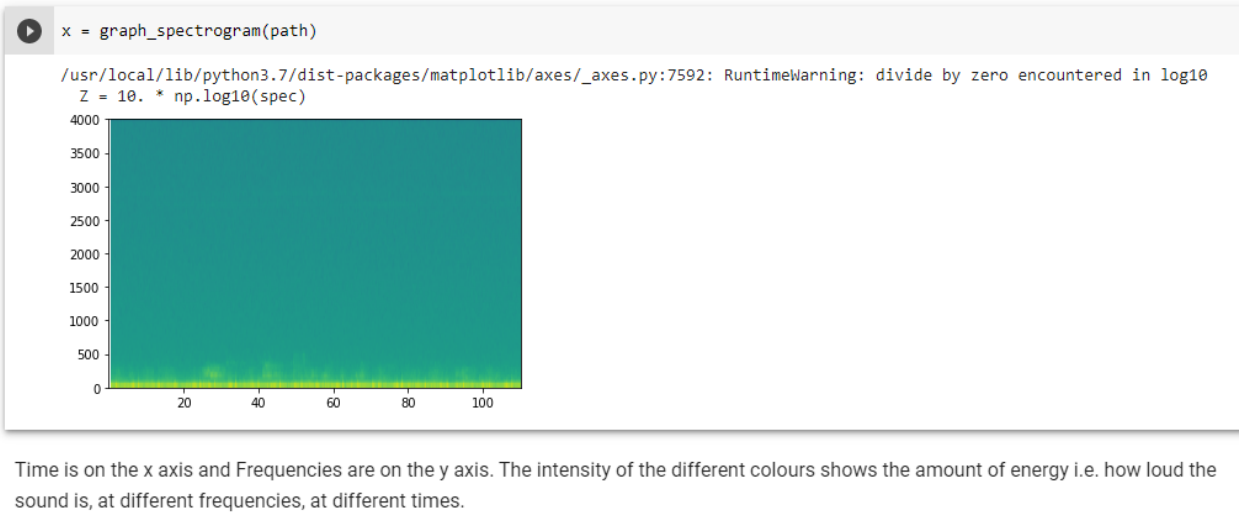


Figure: Conversion of sound to digital spectrogram

As we can see, the loudness of the sound can be identified through this image. At x-axis there is time and on y-axis there is frequency. At different time we can see the frequency of the sound and can judge its intensity or loudness.

MFCC:

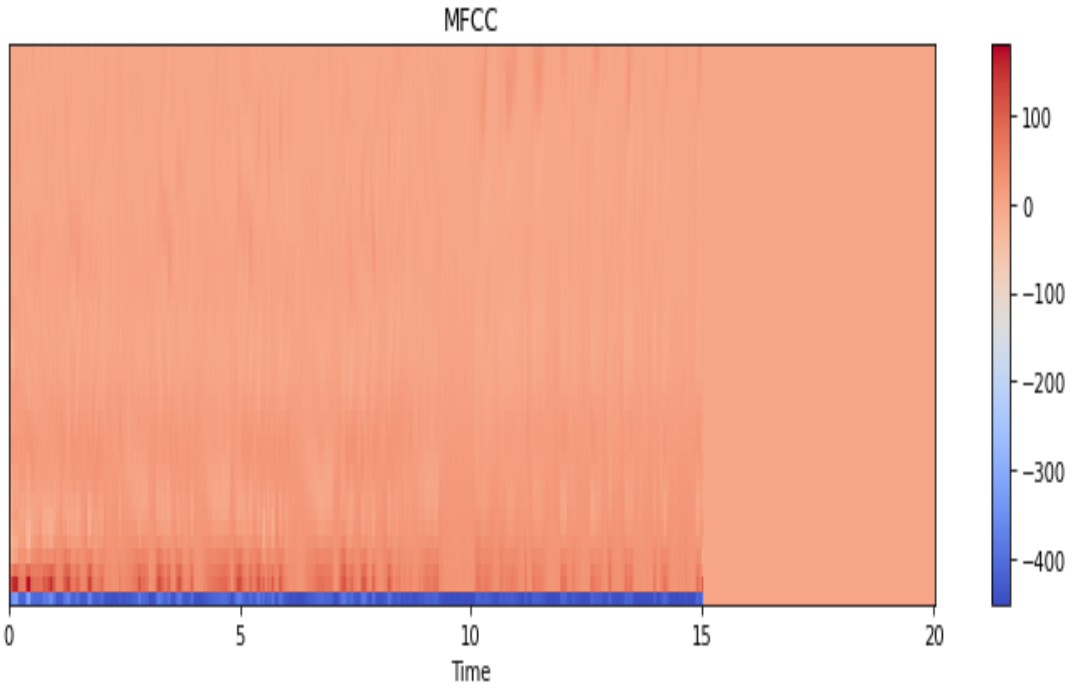
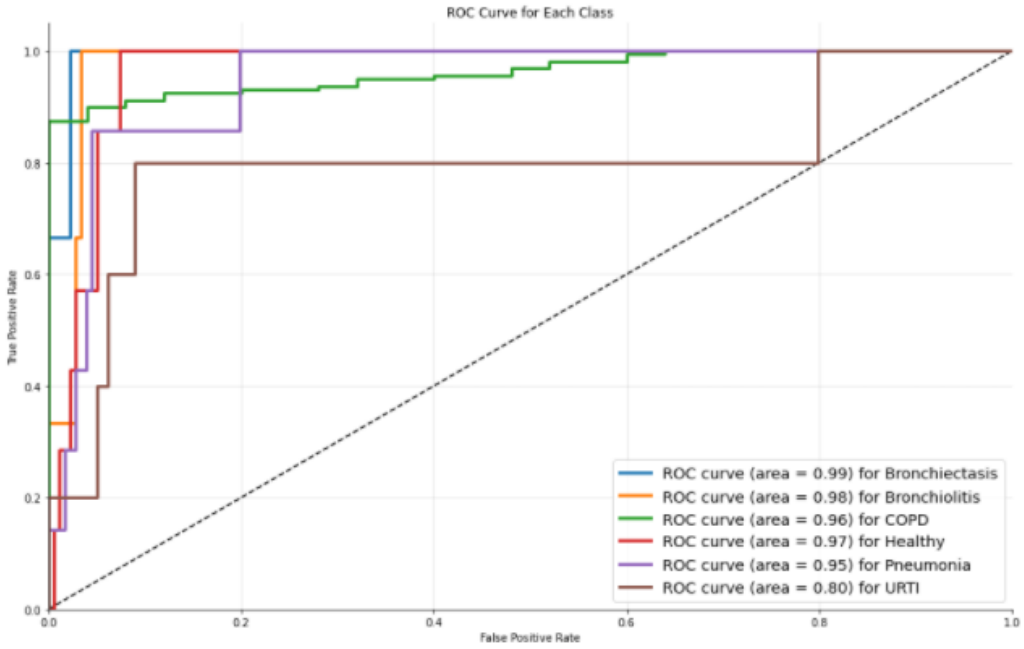


Figure: Conversion of sound to MFCC

From the MFCC we can judge the quality of pitch. The higher the y-coordinates are from origin the higher is the pitch and the lower it is from the origin the lower is the pitch.

CNN: After full completion of 250 epochs our CNN model achieved a training accuracy of 92.9% and testing accuracy of 86.95%. The ROC curve has been given below:



FigL ROC Curve

The complete classification report for the CNN model is given below:

	precision	recall	f1-score	support
Bronchiectasis	1.00	0.67	0.80	3
Bronchiolitis	0.25	0.67	0.36	3
COPD	0.94	0.95	0.95	159
Healthy	0.50	0.14	0.22	7
Pneumonia	0.33	0.43	0.38	7
URTI	0.33	0.20	0.25	5
accuracy			0.87	184
macro avg	0.56	0.51	0.49	184
weighted avg	0.88	0.87	0.87	184

Fig: Classification Report

VGG16: After completion of 2/100 epochs our VGG16 model achieved a training accuracy of 83.71%. We could not run all the epochs due to lack of computational power as each epoch was taking approximately 4 hours of time.

Ref. No.	*Classifier*	*Accuracy*
19	CNN	93%
Ours	**CNN**	**86.95%**
20	Decision Tree	85%
21	CNN	83.7%
22	CNN	65.5%
3	bi-ResNet	60.16%

Table: Comparison of our performance with other related work using the same dataset

CHAPTER 5: CONCLUSION

In this chapter, we are going to discuss how our work can serve as a benchmark in the medical industry and pave path for future significant works.

5.1 DISCUSSION

In this course first our aim was to convert the audio file to a visual image. We achieved that by converting audio files (.wav format) to a digital spectrogram and MFCC (mel-frequency cepstrum coefficient). Another step of pre-processing was to remove the very rare diseases like asthma and LRTI which occurred only among 2 patients only. We then fed the MFCC images as input to our deep learning models. We successfully ran the CNN model and achieved a testing accuracy of 86.95%. Unfortunately we could not completely run our VGG16 and ResNet50 models due to lack of computational power. We used google colab for the project and it took approximately 4 hours to run each epoch of our VGG16 model. After running 2 epochs it achieved a training accuracy of 83.71%. We also plan on equipping us with more computational resources to improve and take our accuracy above 90% and us also intent on broadening the dataset to dive deep into the causes and hope for better findings.

5.2 SUMMARY

The climate is deteriorating faster than ever. The air quality that we have today has degraded many times compared to the air quality we had 10 years earlier. As a result the number of deaths and sickness due to respiratory disease have increased manifold. But unfortunately this is an area very under looked upon as there are very few papers and works related to serve this sector. Our work can play a significant role in this sector. There are a lot of areas in the world where there is huge lacking of skilled medical staff. Highly experienced doctors are very expensive and not everyone can afford them. Also many doctors hesitate to serve in remote areas. As a result the people living in those areas are medically underserved. In such circumstances just a recording though a digital stethoscope and run-through via our model can provide immediate diagnostic reports. It can tell what type of respiratory disease that person has.

It does not require any skilled medical staff to record audio and to pass it into our model.

5.3 FUTURE WORK

In future we would like to gain more computational resources and run more advanced and latest deep learning models like CoAtNet-7, ViT-G/14, etc. and we are confident that we can take our accuracy above 90%. We would also like to add more data to test our algorithm to make it more credible for real-world use.

REFERENCES

[1] Zeenat Tariq, Sayed Khushal Shah, Yugyung Lee, "Lung Disease Classification using Deep Convolutional Neural Network" *IEEE International Conference on Bioinformatics and Bio-medicine 2019, At: San Diego, California*, Nov. 2019. [online]. Available at: (PDF) Lung Disease Classification using Deep Convolutional Neural Network (researchgate.net)

[2] Vesna Cukic, Vladimir Lovre, Dejan Dragisic, and Aida Ustamujic,"Asthma and Chronic Obstructive Pulmonary Disease (COPD) – Differences and Similarities", v.24(2); 2012. [online]. Available at: Asthma and Chronic Obstructive Pulmonary Disease (COPD) – Differences and Similarities (nih.gov)

[3] Yi Ma, Xinzi Xu, Qing Yu, Yuhang Zhang, Yongfu Li*, Jian Zhao and Guoxing Wang, "LungBRN: A Smart Digital Stethoscope for Detecting Respiratory Disease Using bi-ResNet Deep Learning Algorithm", *BioCAS*, Oct. 2019. [online]. Available at: (PDF) LungBRN: A Smart Digital Stethoscope for Detecting Respiratory Disease Using bi-ResNet Deep Learning Algorithm (researchgate.net)

[4] B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho, N. Maglaveras, "A Respiratory Sound Database for the Development of Automated Classification" *ICBHI 2017*, Nov. 2017. [online]. Available at: (PDF) A Respiratory Sound Database for the Development of Automated Classification (researchgate.net) [Accessed Nov- 2017].

[5] Justin Monsi, Justine Saji, Keerthy Vinod, Liya Joy, Jis Joe Mathew, "XRAY AI: Lung Disease Prediction Using Machine Learning", Volume 8, No.2, March - April 2019. [online]. Available at: http://warse.org/IJISCS/static/pdf/file/ijiscs12822019.pdf

[6] Smriti Mukesh Singh, Dr. Dinesh B. Hanchate, "Improving Disease Prediction by Machine Learning", Volume: 05 Issue: 06 | June-2018. [online]. Available at: (PDF) Improving Disease Prediction by Machine Learning (researchgate.net)

[7] Rohan Tan Bhowmik, "A Multi-Modal Respiratory Disease Exacerbation Prediction Technique Based on a Novel Spatio-Temporal Machine Learning Architecture and Local Environmental Sensor Networks" *et al. · The Harker School*, Apr. 3, 2021. [online]. Available at: 2103.03086.pdf (arxiv.org)

[8] Srivastava A, Jain S, Miranda R, Patil S, Pandya S, Kotecha K. 2021. "Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease". *PeerJ Computer Science* 7:e369 https://doi.org/10.7717/peerj-cs.369

[9] Demir, F., Sengur, A. & Bajaj, V. "Convolutional neural networks based efficient approach for classification of lung diseases". *Health Inf Sci Syst* 8, 4 (2020). https://doi.org/10.1007/s13755-019-0091-3

[10] Milecia McGregor, "What Is a Convolutional Neural Network? A Beginner's Tutorial for Machine Learning and Deep Learning", freeCodeCamp.org, February 4, 2021. Accessed on: Dec. 29, 2021. [Online]. Available at: https://www.freecodecamp.org/news/convolutional-neural-network-tutorial-for-beginners/

[11] "Convolutional neural network". Accessed on: Dec. 29, 2021. Available at: https://en.wikipedia.org/wiki/Convolutional_neural_network

[12] Thomas Wood, "Convolutional Neural Network", Deep AI. Accessed on: Dec. 29, 2021. Available at: https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network

[13] Tinsy John Perumanoor, "What is VGG16? — Introduction to VGG16", Medium, Sep 23. Accessed on: Dec. 29, 2021. Available at: https://medium.com/@mygreatlearning/what-is-vgg16-introduction-to-vgg16-f2d63849f615

[14] Pawan Saxena, "VGG-16 | CNN model", GeeksforGeeks, Feb. 27, 2020. Accessed on: Dec. 29, 2021. Available at: https://www.geeksforgeeks.org/vgg-16-cnn-model/

[15] Avinash Thite, "Introduction to VGG16 | What is VGG16?", GreatLearning Blog, Oct. 1, 2021. Accessed on: Dec. 29, 2021. Available at: https://www.mygreatlearning.com/blog/introduction-to-vgg16/

[16] Introduction to Resnet or Residual Network, GreatLearning Blog, Sep. 28, 2020. Accessed on: Dec. 29, 2021. Available at: https://www.mygreatlearning.com/blog/resnet/#sh1

[17] "MFCC (Mel Frequency Cepstral Coefficients) for Audio format", OpenGenus IQ, 2021. Accessed on: Dec. 29, 2021. Available at: https://iq.opengenus.org/mfcc-audio/

[18] "Understanding Spectrograms", iZotope, April 3, 2020. Accessed on: Dec. 29, 2021. Available at: https://www.izotope.com/en/learn/understanding-spectrograms.html

[19] Arpan Srivastava1, Sonakshi Jain1, Ryan Miranda1, Shruti Patil2, Sharnil Pandya2, Ketan Kotecha2 "Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease". Accessed on: Dec. 29, 2021. Available at: https://peerj.com/articles/cs-369/

[20] Gaëtan Chambres; Pierre Hanna; Myriam Desainte-Catherine, "Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis". Accessed on: Dec. 29, 2021. Available at: https://ieeexplore.ieee.org/document/8516489

[21] Truc Nguyen; Franz Pernkopf, "Lung Sound Classification Using Snapshot Ensemble of Convolutional Neural Networks" Accessed on: Dec. 29, 2021. Available at: https://ieeexplore.ieee.org/abstract/document/9176076

[22] Fatih Demir, Abdulkadir Sengur & Varun Bajaj "Convolutional neural networks based efficient approach for classification of lung diseases" Sep. 28, 2020. Accessed on: Dec. 29, 2021. Available at: https://link.springer.com/article/10.1007/s13755-019-0091-3