International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 5 ǁ 2022 ǁ PP. 72-80

Deep Learning for Classification of Non-Speech Human Sounds And Covid -19 Cough Detection

Mrs.I. Anee Daisy[1], Dr. R. Lavanya[2]
[1]*PG Student, Department of ECE, Sri Ramakrishna Institute of Technology, Coimbatore, India.*
[2] *Assistant Professor (Sr, Gr), Department of ECE, Sri Ramakrishna Institute of Technology, Coimbatore, India.*

ABSTRACT:
The world has been severely affected by COVID-19; an infectious disease, thus rapid non-invasive testing has become essential. Recent studies and benchmarks motivated the use of modern artificial intelligence (AI) tools that utilize audio waveform spectral features of coughing for COVID19 diagnosis. In this paper, we describe the system developed for COVID-19 coughing detection by utilizing features directly extracted from the coughing audio. The convolution neural network is widely used in the field of large-scale image recognition. This research aims to propose a deep learning-based approach that classifies COVID-19, Pneumonia, Asthma and healthy (normal cases) on open-source data available. The identification of the presence of COVID-19 is challenging as its symptoms are similar to influenza symptoms such as cough, cold, runny nose and chills. COVID-19 affects human speech sub-systems involved in respiration, phonation, and articulation. This study proposes a deep anomaly detection framework for non-speech-based detection of COVID-related anomalies in voice samples of COVID-19 affected individuals. Deep transfer learning is used to classify the data using ResNet-50 neural network architecture, thus to perform the multiple classification experiment. Experimental results show that the pre-trained model ResNet-50 achieves 0.89 for multiclass classification.

Date of Submission: 12-05-2022 Date of acceptance: 26-05-2022

I. INTRODUCTION

In the past few decades, there have been several efforts regarding the classification of the acoustic data into classes. The audio data is very informative and a rich source of extraction for the type of content involving content-based classification of the acoustic signals. Human beings use vocal tract for producing speech sounds such as talking, singing, crying, and laughing. These sounds are further classified as speech or non-speech vocalizations. Speech consists of voices that are in the form of sentences and can be understood using different Natural Language Processing (NLP) techniques. The non-speech sounds include laugh, sneeze, cough, snore, and scream. These non-speech vocalizations are sometimes segregated from speech signals to extract additional information about the context, situation, or emotional state of the Speaker. At present, deep neural network-based speech-enhancement methods have been widely adopted and have shown significant performance advantages over conventional speech enhancement techniques in complex noise environments.This Paper aims in the classification of Cough sounds which will be helpful in detecting COVID19 (**CO**rona **VI**rus Disease of 2019), caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2) virus, was declared a global pandemic on February 11, 2020 by the World Health Organization (WHO).

In the scope of this research, we however only consider acoustic sensing related to the environment. So as to recognize ambient noises during the sleep, acoustic features are first selected in both time and frequency domains including short-time energy, the loud- ness, zero crossing rate (ZCR), power spectral density (PSD), and spectral entropy. They are then used to detect acoustic events in the recordings. Finally, the detected events that depict sudden and recurrent ambient sounds are chosen by a non-linear classifier. n the scope of this research, we however only consider acoustic sensing related to the environment. So as to recognize ambient noises during the sleep, acoustic features are first selected in both time and frequency domains including short-time energy, the loud- ness, zero crossing rate (ZCR), power spectral density (PSD), and spectral entropy. They are then used to detect acoustic events in the recordings. Finally, the detected events that depict sudden and recurrent ambient sounds are chosen by a non-linear classifier. n the scope of this research, we however only consider acoustic sensing related to the environment. So as to recognize ambient noises during the sleep, acoustic features are first selected in both time and frequency domains including short-time energy, the loud-ness, zero crossing rate (ZCR), power spectral density (PSD), and spectral entropy. They are then used to detect acoustic events in the recordings. Finally, the detected events that depict sudden and recurrent ambient sounds are chosen by a non-linear classifier.

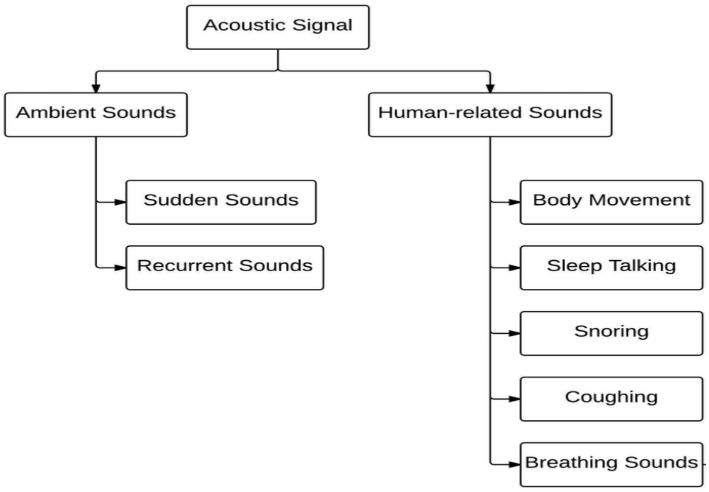

Figure 1.1: Acoustic sound Classification

Coughing is one of the predominant symptoms of COVID-19 and also a symptom of more than 100 other diseases, and its effect on the respiratory system is known to vary. Cough sounds contain underutilized pulmonary health information that can be analyzed.

II. LITERATURE REVIEW ON COVID 19 DETECTION:

The study [1] was sought to investigate the feasibility of using smartphone-based breathing sounds within a deep learning framework to discriminate between COVID-19, including asymptomatic, and healthy subjects. A deep learning framework was proposed herein that relies on hand-crafted features extracted from the original recordings and from the Mel-frequency cepstral coefficients (MFCC) as well as deep-activated features learned by a combination of convolutional neural network and bi-directional long short-term memory units (CNN BiLSTM) Furthermore, it detected COVID-19 subjects successfully with a maximum sensitivity of 94.21%, specificity of 94.96%, and area under the receiver operating characteristic (AUROC) curves of 0.90.

The proposed model [2] present a machine learning based COVID-19 cough classifier which can discriminate COVID-19 positive coughs from both COVID-19 negative and healthy coughs recorded on a smartphone. The publicly available Coswara dataset contains 92 COVID-19 positive and 1079 healthy subjects, while the second smaller dataset was collected mostly in South Africa and contains 18 COVID-19 positive and 26 COVID-19 negative subjects who have undergone a SARS-CoV laboratory test. Both datasets indicate that COVID-19 positive coughs are 15%-20% shorter than non-COVID coughs. Dataset skew was addressed by applying the synthetic minority oversampling technique (SMOTE). An LSTM classifier was best able to discriminate between the COVID-19 positive and COVID-19 negative coughs, with an AUC of 0.94 after selecting the best 13 features from a sequential forward selection (SFS).

In this research [3], a deep convolutional neural network (DCNN) based on a pre-trained model for the automatic detection of COVID-19 from two other classes (Viral Pneumonia and normal chest X-ray images) was proposed. For this purpose, we used a fine-tuned ResNet-50 previously trained on the ImageNet dataset in our model. For the experiment, we used chest X-ray images rather than CT scans to fine-tuned the ResNet-50 model for classification. X-rays are relatively cheaper and quicker. In the research [6], developed a computer vision solution to support diagnostic radiology in differentiating between COVID-19 pneumonia, influenza virus pneumonia, and normal biomarkers.Xu et al. [7] developed a prediction model to discriminate COVID-19 pneumonia and influenza-A viral pneumonia using deep learning techniques. The CNN model was used for prediction. The maximum accuracy obtained from prediction model was 86.7%.

Wang et al. [8] investigated the radiographic changes in CT images of infected patients.The ResNet50 pre-trained model produced accuracy of 98%, which is higher than [7,8]. Sethy et al.developed a deep learning model for detecting COVID-19 from X-ray images. They extracted deep features and transferred them to support vector machine for classification. The accuracy of 95.38% obtained from the proposed model, which is better than [7, 8].The main problem attributed to the prevalence of respiratory diseases [9]is lack of cost-effective and lab-free methods for early diagnosis. Spirometry is the standard clinical test procedure for detection of respiratory problems, but it requires repetition, and is also expensive and not available in rural areas

III. METHODOLOGY

The features of Audio, especially speech signal may be extracted using FFT (Fast Fourier Transform) and Wavelet to detect the frequency information of the signal. But it is difficult to extract the changes of small variation of speech signal with time-varying morphological characteristics. So, it is needed to be extracted by signal processing method for the graphical representation of the signal.In the first stage, the pre-processed audio signal is given to noise reduction and the feature is extracted. According to the features, the audio signal is classified as scream, shout, conversation and noises.The proposed wavelet method found to be more summarized, followed by pre-processing, creating, and training a deep learning model using Convolutional Neural Network (CNN) to perform classification.

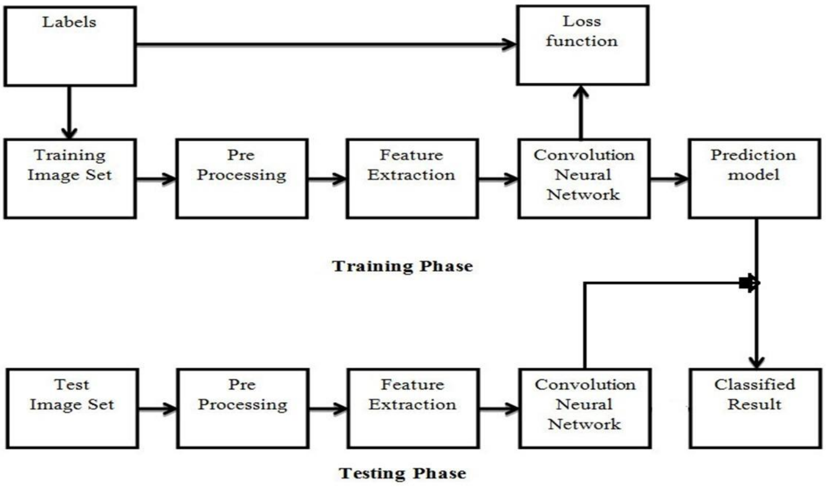

Figure 3.1 Block diagram of proposed system

(A) COUGH SEGMENTATION AND DETECTION:

One of the first tasks of cough audio analysis is to be able to detect and identify a cough signal then classify it using some biomarkers. Several research studies have addressed cough detection using different methods.For segmentation phase, predefined thresholds can be chosen based on the physiology of cough sounds. The raw recorded sounds contain a lot of silent fragments (with low intensity) and background noise. Therefore, silence removal phase is required for saving storage space. For signal quality check, an estimate of the signal-to-noise ratio (SNR) can be computed by taking the ratio of the power of the cough part of the signal to the ratio of the rest of the signal. This process consists of cleaning the cough sounds dataset by filtering out all the interferences and the environmental noise from the audio frames and keeping only the relevant frames to the cough of audio separation consists of extracting the cough sounds.

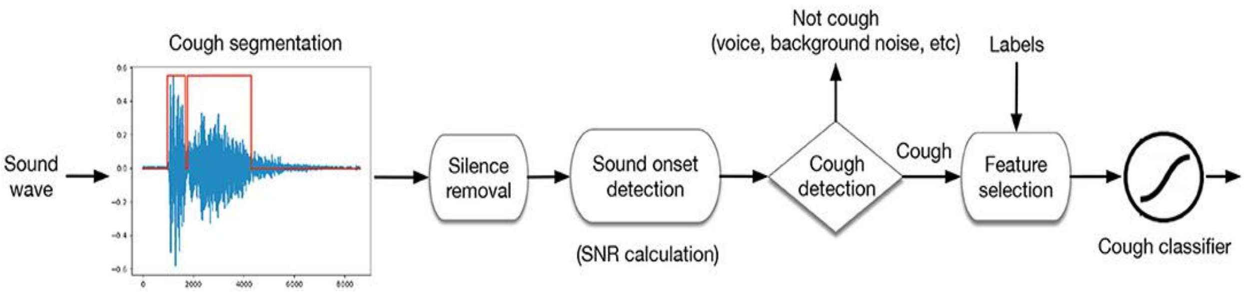

FIGURE 3.2 Cough segmentation

(B) SPECTRUM ANALYSIS:

A spectrogram is a visual representation of the spectrum of frequencies of a signal as it varies with time which is usually depicted as a heat map. Spectrograms can be used to locate strong signals and determine how frequencies change over time. As a collection of time-frequency analysis, the spectrogram is used to identify the property of nonlinear signals. For this reason, the spectrogram is a very friendly tool for analysing real-world data where there are different kinds of frequency components and mechanical and electrical noise. The collected audio recordings, in the time domain, are broken up into chunks and Fourier transformed to calculate the magnitude of the frequency spectrum for each chunk. Each chunk can therefore be considered as a measurement of magnitude versus frequency for a specific moment in time.

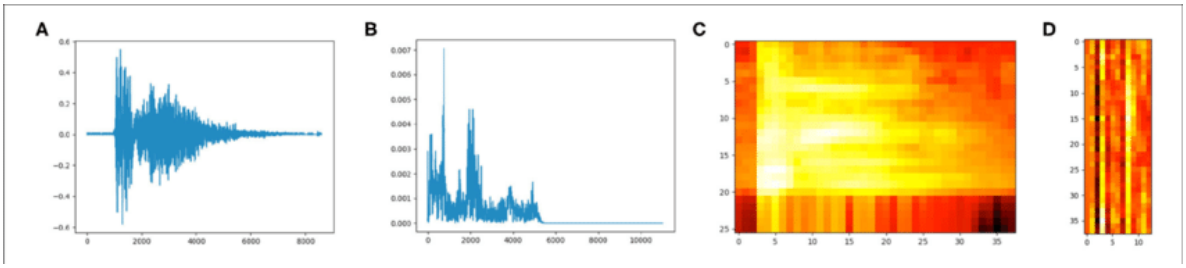

Figure 3.3Cough time-frequency representation analyses.

(C)CONVOLUTIONAL NEURAL NETWORK:
Convolutional Neural Network (CNN) is a class of Deep Learning, mainly use for Computer Vision. It is similar to artificial neural network, only difference is it uses convolutional mathematical linear operation instead of simple matrix multiplication in at-least one of its layers. Building a Convolutional Neural Network is nothing but building a human eye, how human see the world and recognize the pattern.

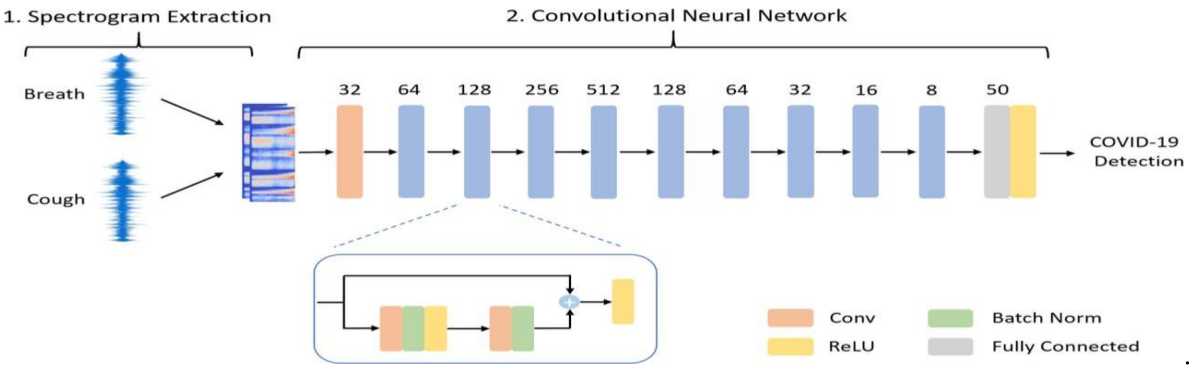

Figure 3.4 CNN Architecture

COVID-19 Identificationis based on ResNet 50a variant of the CNN architecture, which uses *residual* blocks. As shown in figure 1a residual block consists of two convolutions, batch normalisation[15] and a rectified linear unit non-linearity. These blocks use 'skip' connections which add the output from these operations to the input activations for this layer. This alleviates the vanishing gradient problem, facilitating deeper architectures with more layers, thereby permitting richer hierarchical learnt representations. The CNN outputs a single logit which is then passed through a softmax layer to obtain a (0,1) score, representing the probability of a COVID-positive sample.

(D) TRAINING THE NETWORK:
Initially network is fine-tuned by resizing the images. Furthermore, ImageNet dataset constantly increases training set. Instead of manually separating learning rate of images technique called Cyclical Learning Rate is used. For maximum optimization of learning rate this technique is used. The images taken from input dataset are resized to 50x50x3. The entire network is fine-tuned by applying discriminative learning rate for 20 epochs. Progressive resizing technique is most beneficial to train the model iteratively.

(E) Overall Performance Analysis
The most popular performance measures such as accuracy, precision, sensitivity, F1 Score are used to evaluate the performance of proposed model. By assessing the performance of our deep transfer learning model for validation and testing dataset considering the following evaluation metrics: accuracy (ACC), precision (PPR), sensitivity or recall (SN), specificity (SP), and F1- score. The following equation measures the performance metrics

$$\text{Accuracy (ACC\%)} = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

$$\text{Precision (PPR)} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Sensitivity (SN)} = \frac{TP}{TP + FN} \quad (4)$$

(F) CONFUSION MATRIX:

A Confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the number of target classes. The matrix compares the actual target values with those predicted by the machine learning model.

For a binary classification problem, we would have a 2 x 2 matrix as shown below with 4 values:The model's classification performance can be observed by the confusion matrices provided in below Figure 3.16.

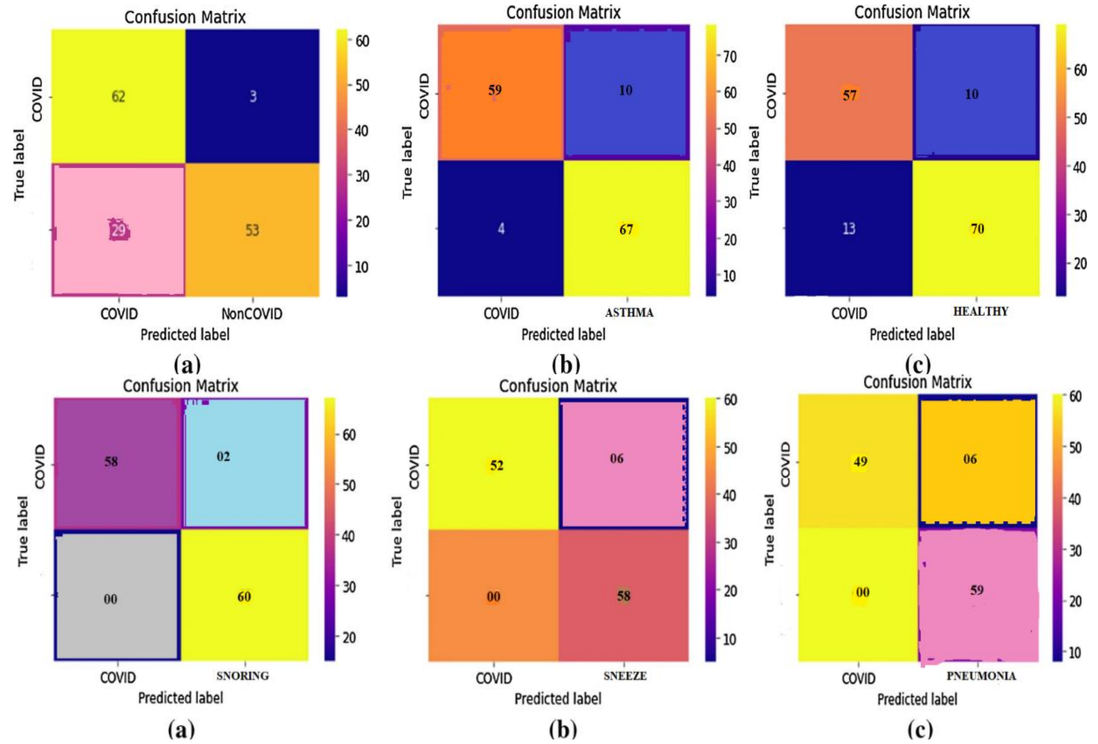

Figure 3.5 Confusion matrix of the proposed system

IV. SOFTWARE DESCRIPTION

MATLAB, software is a technical matrix manipulating based computation software manipulating matrices leads to big data analysis. The things to learn in MATLAB are entering matrices, usage of the: (colon) operator, invoking functions. At the heart of MATLAB is a new high-level language due to its multilanguage inheritance fully exploits its power. Matrix manipulation and function working will be the basics of MATLAB and. Users will be rewarded with high productivity, high- creativity, and strong computing power that will change the way us work.

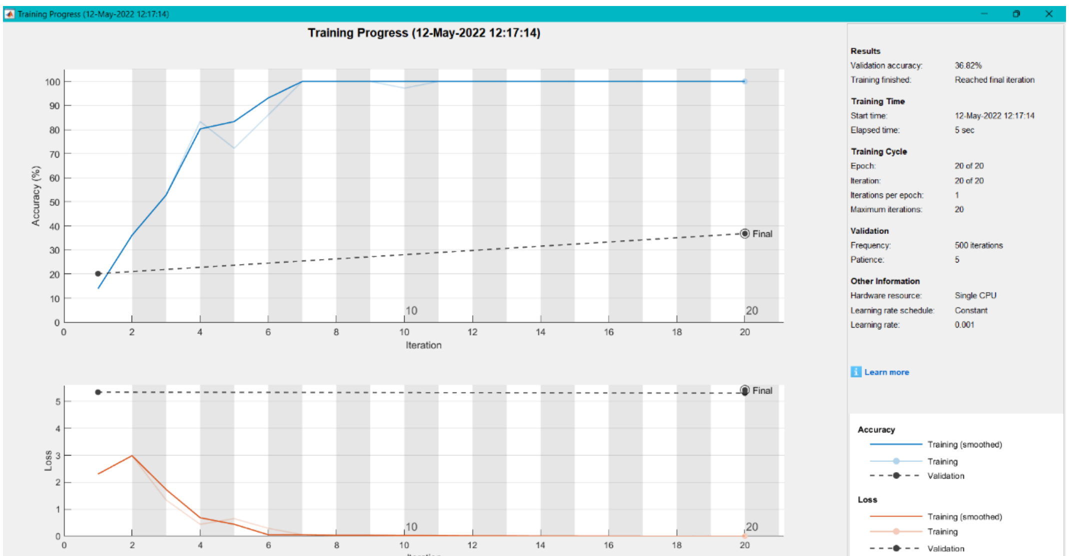

Figure 4.1: **Training Progress [Accuracy / Loss]**

V. RESULTS AND DECLARATION:

The typical cough sound is usually divided into three phases (in Figure 2) [3]: (1) an explosive expiration due to the glottis suddenly opening, (2) the intermediate phase with the attenuation of cough sounds, and (3) the voiced phase due to the closing of the vocal cord. In fact, there are a variety of patterns of cough that occur; for example, some cough sounds only have two phases (the intermediate phase and the voiced phase) and the explosive phase usually prolonged because of some diseases. However, the cough sound sample of the COVID-19 patient varies significantly from the typical human cough sound sample. For example, both the intermediate and voiced phases are longer for the COVID-positive patient than for the healthy subject

Thus, using the hyperparameter tuning we achieved the output waveform for the proposed system as shown below in figure

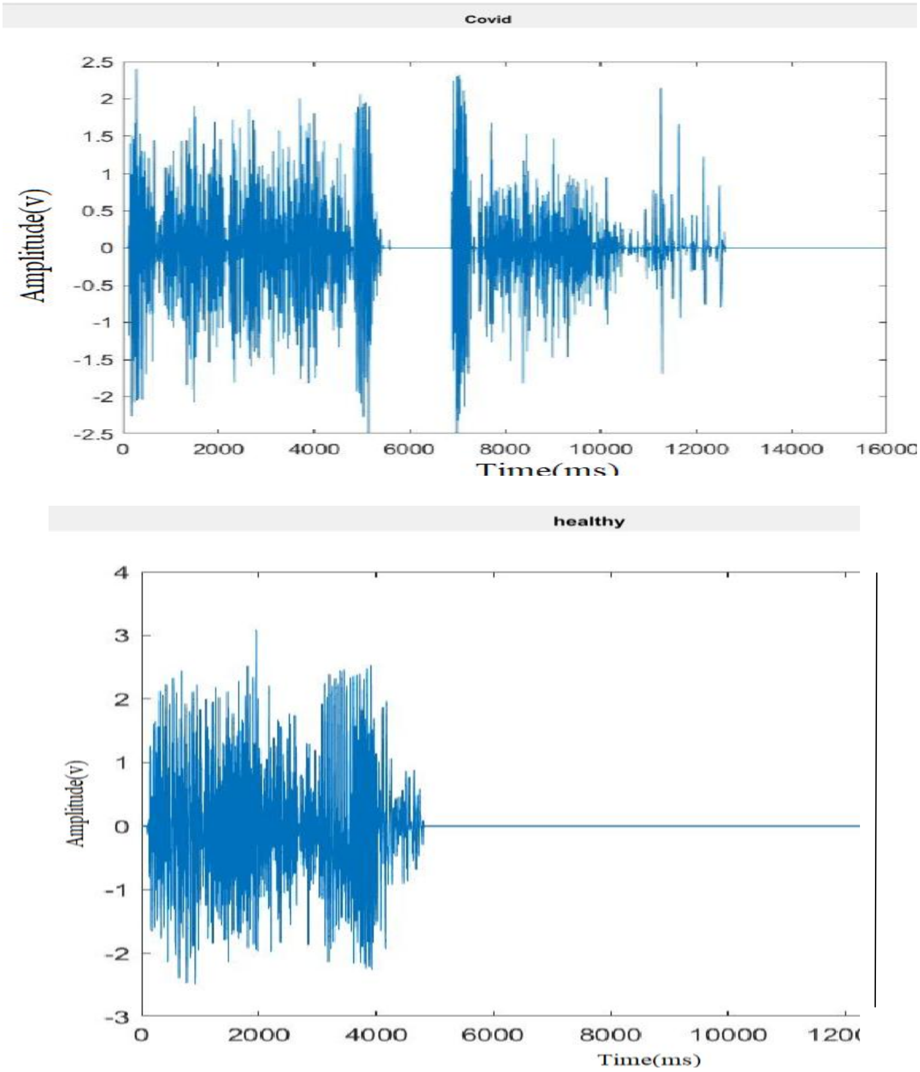

Figure 5.1: Covid 19 Cough versus Healthy Cough

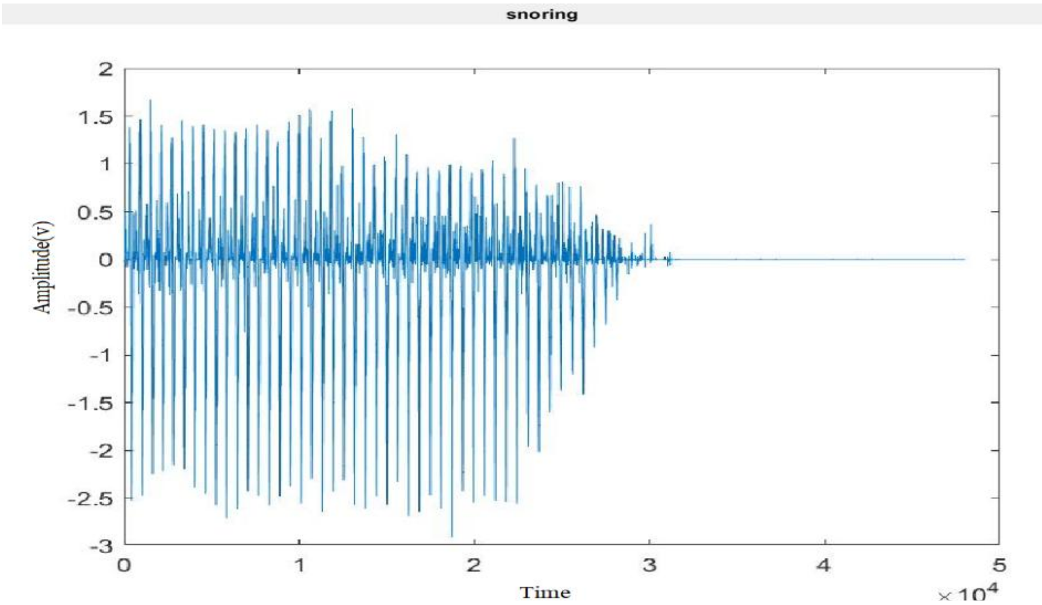

Figure 5.2: Amplitude variations in snoring

A lower respiratory tract infection (RTI) occurs when there is an infection of the lungs, specifically in the lower airways. This infection is usually caused by a virus, but it can also be caused by bacteria or other less common organisms. Common lower RTIs includes:

Pneumonia. Pneumonia is an infection that causes inflammation of the air sacs in one or both of the lungs. Its symptoms can range from mild to severe enough to require hospitalization.

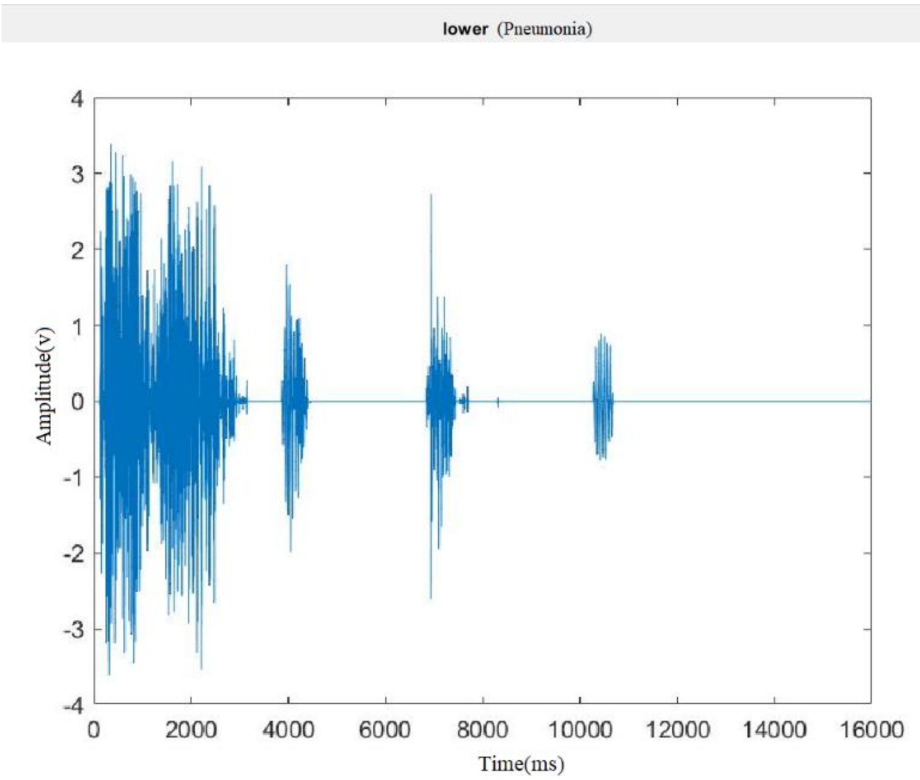

Figure 5.3: Pneumonia affected Cough Analysis

Upper respiratory infections (URIs) are one of the most common reasons for doctor visits. Symptoms of upper respiratory infection include cough, sneezing. nasal discharge, nasal congestion, runny nose,fever, sore throat, nasal breathing

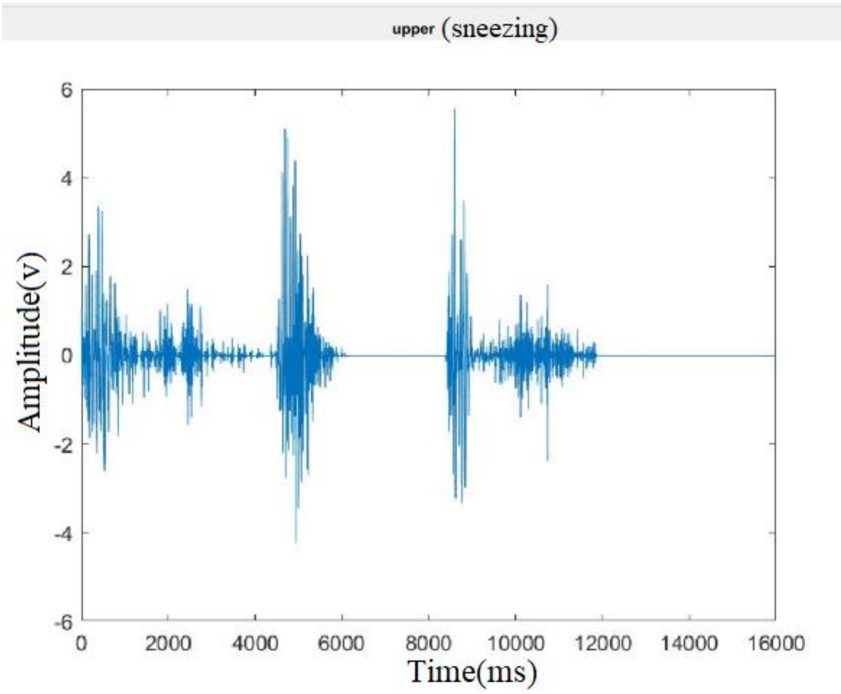

Figure 5.4 Amplitude analysis in Sneezing

Chronic obstructive pulmonary disease (COPD) is the name for a group of lung conditions that cause breathing difficulties. The main symptoms of COPD are:
persistent wheezing/Asthma, increasing breathlessness, particularly when you're activea persistent chesty cough with phlegm

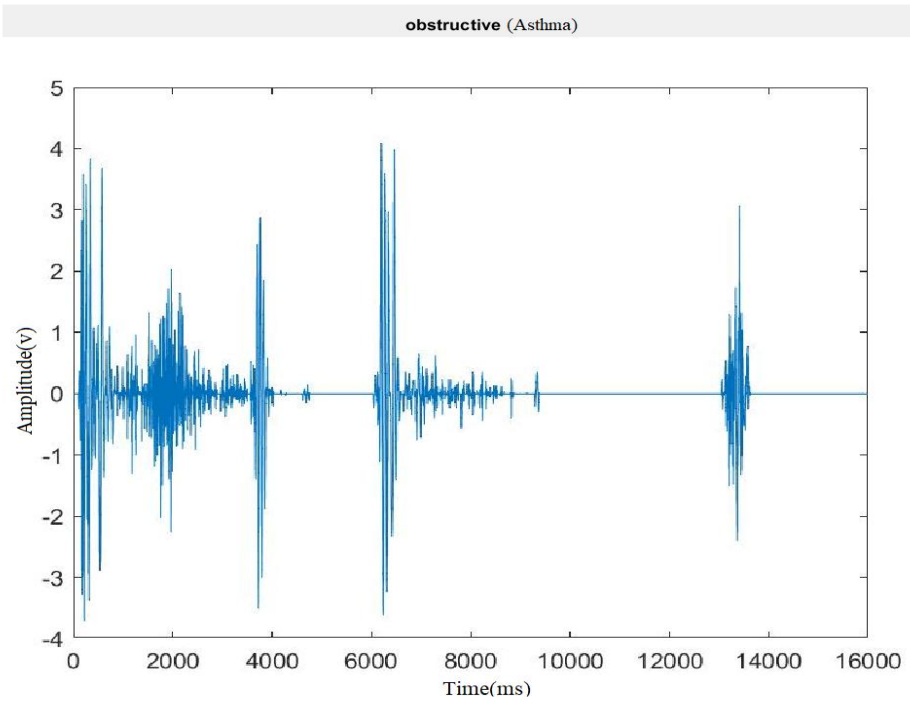

Figure 5.5 Asthma affected Cough Analysis

VI. CONCLUSION

To diagnosis the COVID-19 disease, it is essential to have a cost-effective, fast, easy, and accurate method, considering the high cost of clinical tests, long turnaround time, and lack of equal access around the world. Therefore, it is quite interesting and essential to distinguish COVID-19 patients from non-COVID-19 ones by evaluating their cough.The proposed work showcases the possibility of using Convolutional Neural

Networks as a testing method for Covid-19. An encouraging result was obtained in the paper by taking cough, snoring, sneezing and breathing sounds as inputs. MFCCs and Mel spectrogram images were obtained to extract features from the audio samples. Moreover, in this study, the used data set includes the various types of audio sounds such as cough sounds, snoring, sneezing and COPD. Thus, for future studies, it is planned to test the system on a larger number of samples and an online platform by adding serious pre-processing steps such as systems to distinguish cough sound from other sounds. These limitations can be minimized in future work, as the application is being proposed so that the network can be trained with a more diverse dataset to improve itself.This will improve the robustness of the proposed system.

REFERENCES

[1]. Alkhodari M, Khandoker AH Detection of COVID-19 in smartphone-based breathing recordings: A pre-screening deep learning tool. PLoS ONE 17(1): e0262448. https:// doi.org/10.1371/journal.pone.0262448, Jan 2021

[2]. MadhuranandaPahar, Marisa Klopper, Robin Warren, and Thomas Niesler COVID-19 Cough Classification using Machine Learning and Global Smartphone Recordings, https://arxiv:2012.01926v2 [cs.SD] Jun 2021

[3]. Shadman Sakib1 , Md. Abu Bakr Siddique2 , Mohammad Mahmudur Rahman Khan3 , Nowrin Yasmin4 , Anas Aziz5 , Madiha Chowdhury6 , Ihtyaz Kader Tasawar7 t Detection of COVID-19 Disease from Chest X-Ray Images: A Deep Transfer Learning Framework https://doi.org/10.1101/2020.11.08.20227819; this version posted November 12, 2020

[4]. Jordi Laguarta, FerranHueto, Brian Subirana. COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings. 29 September 2020.

[5]. Abdelkader Nasreddine Belkacem1, ∗ Sofia Ouhbi2 ,Abderrahmane Lakas1 , Elhadj Benkhelifa3 , and Chao Chen4 End-to-End AI-Based Point-of-Care Diagnosis System for Classifying Respiratory Illnesses and Early Detection of COVID-19 . June 30, 2020

[6]. Julianna Antonchuk1, Benjamin Prescott1, Philip Melanchthon1, Robin Singh1: COVID-19 Pneumonia and Influenza Pneumonia Detection Using Convolutional Neural Networks https://arxiv.org/pdf/2112.07102 dec 2021

[7]. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Chen Y, Su J, Lang G, Li Y, Zhao H, Xu K, Ruan L, Wu W (2020) Deep learning system to screen coronavirus disease 2019 pneumonia. arXiv preprintarXiv:2002.09334, 1–29

[8]. Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J,Cai M, Yang J, Li Y, Meng X, Xu B (2020) A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). medRxiv preprint. https://doi.org/10.1101/2020.02.14.20023028,

[9]. G. Rudraraju, S. Palreddy, B. Mamidgi, N. R. Sripada, Y. P. Sai, N. K. Vodnala, and S. P. Haranath, “Cough sound analysis and objective correlation with spirometry and clinical diagnosis,” Informatics in Medicine Unlocked, p. 100319, 2020.

[10]. G. Deshpande and B. Schuller, “An Overview on Audio, Signal, Speech, & Language Processing for 19,” arXiv preprint arXiv:2005.08579, 2020.