

Detection of Covid-19 using Cough Sound

Dristanta Das , Dr. Cheruvu Siva Kumar

Ramakrishna Mission Vivekananda Educational and Research Institute
Big Data Analytics , 2nd Year
IIT Kharagpur
Dept. of Mechanical Engineering

Abstract

The goal of this study was to see if it was possible to use breathing sounds inside a deep learning framework to distinguish between COVID-19, including asymptomatic, and healthy patients. A total of 1611 breathing sounds were gathered from the Coswara database, which is freely accessible to the public. These noises were captured using a smartphone microphone and an internet application by COVID-19 and healthy participants. A deep learning system based on hand-crafted features taken from original recordings and mel-frequency cepstral coefficients is proposed here (MFCC).

Introduction

Corona virus 2019 (COVID-19), a new pathogen of the severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) that first surfaced in late November 2019, has triggered a worldwide pandemic by spreading all over the world. According to the World Health Organization's (WHO) April 2021 report almost 150 million confirmed cases and over 3 million fatalities have occurred since the epidemic began in 2019. Furthermore, with over 32.5 million cases and 500,000 fatalities, the United States (US) has recorded the largest number of cumulative cases and deaths. These massive numbers have put a significant strain on many health-care facilities, especially given the virus's potential to create additional genetic variations and transmit more easily among individuals.

COVID-19 individuals often exhibit a variety of symptoms, from asymptomatic to developing pneumonia and, in severe circumstances, death. In the majority of reported instances, the virus lies dormant for 1 to 14 days before signs of infection appear. Cough, shortness of breath, fever, tiredness, and other acute respiratory distress syndromes have been observed in patients with COVID-19 (ARDS). Most infected patients experience mild to severe viral symptoms, but they eventually recover. Patients who acquire severe symptoms such as severe pneumonia, on the other hand, are

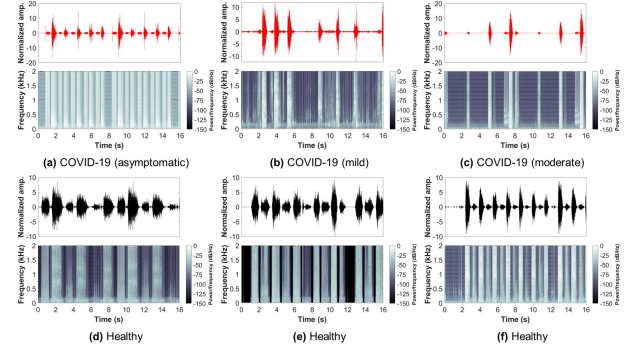


Figure 1: Examples from the shallow breathing sounds along with their corresponding spectrograms

generally over 60 years old and have disorders such as diabetes, cardiovascular disease (CVD), hypertension, and cancer.

Materials

Dataset collection and subjects information

The dataset utilised in this investigation was collected from Coswara, a project that aims to provide an open-access database for respiratory sounds of healthy and ill persons, including those with COVID-19. The project is a global respiratory data gathering endeavour that began on August 7th, 2020. Since then, it has collected data from over 1,611 people (1185 men and 426 women) from all across the world (mostly Indian population). The database was accepted by the Indian Institute of Science's (IISc) human ethics council in Bangalore, India, and adheres to the Helsinki Declaration's ethical guidelines. Participants' personally identifying information was not gathered, and their data was thoroughly anonymised before being stored in the database.

The database contains breath, cough, and speech sounds gathered by crowdsourcing via an interactive online application designed for smartphone devices. All noises were recorded with a smartphone microphone and sampled at a sampling frequency of 48 kHz. The participants were able to use any equipment to record their breathing sounds, which reduced device bias in the data. All participants' audio sam-

ples were manually selected using a web interface that allows numerous annotators to look over each audio file and evaluate the quality as well as the validity of labelling. Before beginning the recording, all participants were asked to preserve a 10 cm gap between their faces and the instrument.

The database currently has 250 COVID-19 patients, which is nearly a one-to-four ratio when compared to healthy (control) people. Furthermore, just two types of breathing sounds, shallow and deep, were recorded and analysed from each patient. Figures 1 exhibit spectrogram representations for samples from the shallow and deep breathing datasets, respectively. All recordings less than 16 seconds were padded with zeros. Furthermore, the final signals were resampled with a sampling frequency of 4 kHz.

Acoustic Features from Dataset

Displaying some acoustic characteristics of sound samples from the COSWARA database. We shall visualise several characteristics by grouping sound identification and gender. The sustained phonation (vowel) sound samples will be the focus here. The current vowel phonetics literature will assist us in interpreting the estimated feature values. The various acoustic characteristics are as follows:

- Intensity
- Pitch or fundamental frequency
- Formant frequencies
- Harmonic-to-noise ratio (HNR)

File duration before and after VAD

The figure below depicts the file time of sound samples from around 650 people. We only evaluated clean audio files in this case. The shortened file duration after VAD is also visible in the scatter plot. The thresholds for intensity, pitch, and HNR are shown in the figure and were determined experimentally based on observations on a set of files and phonetic expertise.

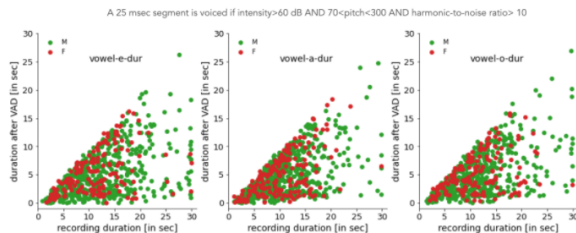


Figure 2: File duration before and after VAD

Because every sound file has some voiceless activity, the data points are located below the $y=x$ line. We can notice that some sound files are longer than 25 seconds!

Distribution of Pitch values

The fundamental frequency of the supposed harmonic complex in the signal is computed to provide the signal processing quantification of perceived pitch of a voice signal.

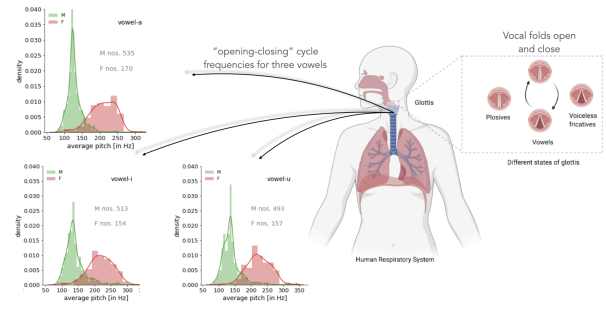


Figure 3: Pitch Distribution

The vibrations (opening and closing) of the vocal folds in the glottis are responsible for the creation of this harmonic complex in the speech signal (shown in the figure below). The average pitch distribution determined by pooling all of the sound sample files is shown below (each file belongs to a different individual). The peaks of the distribution for the two genders (male/female) are clearly separated. Males have a lower pitch than females.

Related Works

There were various approaches previously in the past years to build a classifier that can give an indication towards Covid status from cough sound. A team in Stanford University tried various approaches but those were not significant to be called ground breaking. Amongst those approaches MIT got a breakthrough in this work. The MIT Open Voice Medicine architecture employs the same four biomarkers that were previously studied for Alzheimer's detection and obtained above-state-of-the-art accuracy. These four biomarkers, which were chosen by the medical community, are: muscular degradation, changes in vocal cords, changes in sentiment/mood, and alterations in the lungs and respiratory tract.

Method and Workings

Model Description

For the working purpose I chose a pre-trained ResNet-34 model and made modification to this model for my requirement.

Conv Layer Name	Output Size	Layers
Conv1	112 x 112	[7 x 7, 64]
Conv2_x	56 x 56	[[[3 x 3, 64] [3 x 3, 64]]x6]
Conv3_x	28 x 28	[[[3 x 3, 128] [3 x 3, 128]]x7]
Conv4_x	14 x 14	[[[3 x 3, 256] [3 x 3, 256]]x11]
Conv5_x	7 x 7	[[[3 x 3, 512] [3 x 3, 512]]x5]

Working and Experimental Setup

The deep learning architecture described in this study combines hand-crafted features with deep-activated fea-

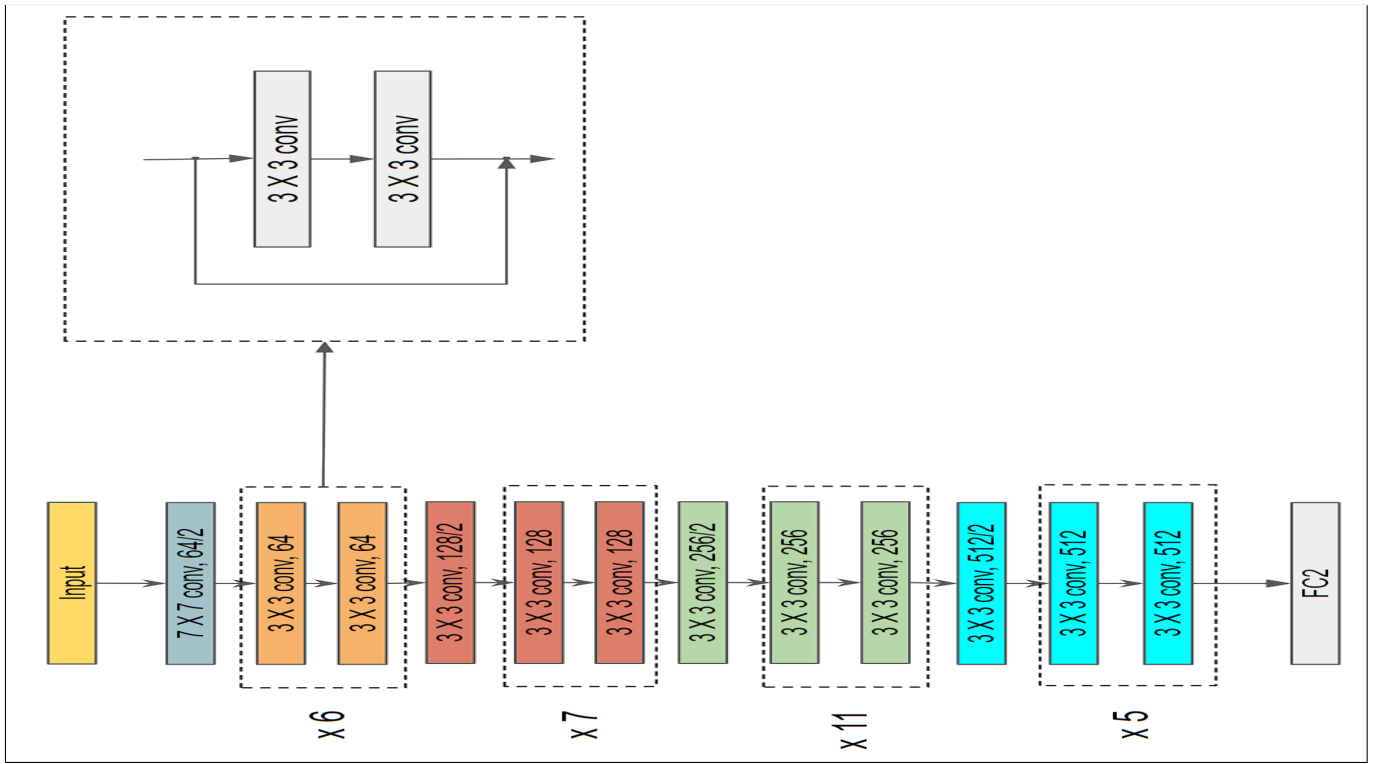


Figure 2: Model Diagram

tures gained through model training and reflected as time-activations of input. Various algorithms and functions were utilised to collect signal properties from the original breathing recording and its accompanying mel-frequency cepstral coefficients (MFCC). Primarily from all the audio samples 13-MFCC features were extracted. And then the MFCCs are passed onto the model to get at the very end the probabilities of that sample being a Covid positive or negative.

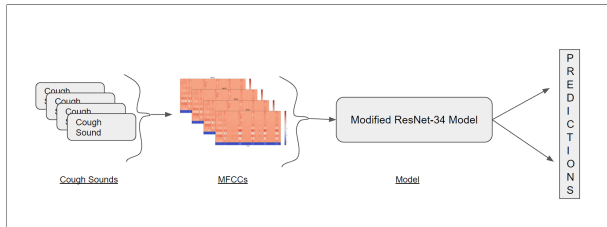


Figure 5: Working of the process

Figure below show examples of the 13 MFCC derived from the original shallow and deep breathing signals for COVID-19 and healthy patients, respectively.

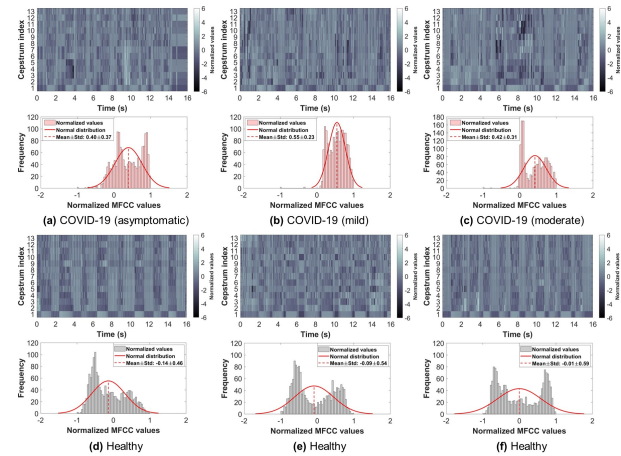


Figure 6: Examples of the mel-frequency cepstral coefficients (MFCC) extracted from the deep breathing dataset and illustrated as a normal distribution of summed coefficients

Furthermore, the pictures depict MFCC values distributed as a normal distribution (after summing all coefficients). According to the figure, the normal distribution of COVID-19 participants was somewhat skewed to the right side of the mean, whereas the normal distribution of healthy subjects was more towards the zero mean, indicating that it represents a normal distribution better. It is worth noting that shallow breathing had lower MFCC values than deep breathing in both COVID-19 and healthy participants.

Results

I have initially used the MIT's pre-trained model as baseline and the results on this dataset was not good enough. Then I trained my model and the diagram below shows the training loss and validation loss, when it is trained for 60 epochs.

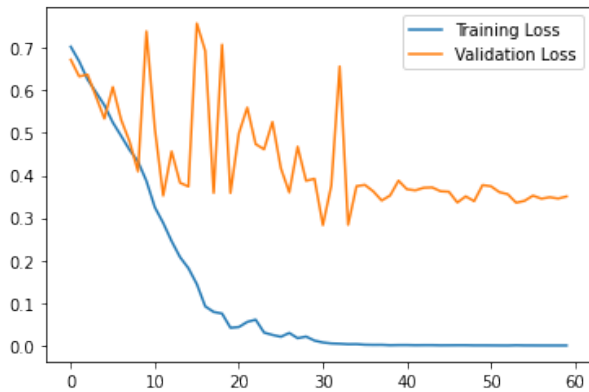


Figure 7: Visualisation of Training and Validation Loss

This is the confusion metric on the testing data.

	Pred Positive	Pred Negative
Positive	320	14
Negative	5	68

Figure 8: Confusion Metric

From the confusion metrics we can get some of the other informations too.

	Values
Precision	0.958
Recall	0.985
F1	0.971
Specificity	0.83
Accuracy	95%

This shows my model has given significantly good result. Although since this whole thing is data dependent so with different data this results may differ.

Room Of Improvement

Although I tried this model. With large amount of data which can be gathered in future one can try to use Recurrent Neural Network approach to get results differently. To deploy the whole thing inference time matters so RNN models might help in that.

ACKNOWLEDGMENT

I want to convey my gratitude to the authors of the reference papers. I would like to express my sincere thanks to my university (Ramakrishna Mission Vivekananda Educational and Research Institute), the Computer Science Department,

HOD Swami Dhyaganamyananda Maharaj, and my project guide Dr. Cheruvu Siva Kumar for their assistance and for providing a suitable habitat and Research Laboratory.

References

- COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings*
By Jordi Laguarda , Ferran Hueto, and Brian Subirana
- Covid-19 Detection System Using Recurrent Neural Networks*
By Abdelfatah Hassan, Ismail Shahin, Mohamed Bader Alsabek
- Early Detection of COVID-19 from Cough Sounds, Symptoms, and Context*
Bernard Lange, Derrick Li, Effie Nehoran, Elena Tuzhilina, Mandy Lu
- AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app*
By Ali Imran , Iryna Posokhova , Haneya N. Qureshi , Usama Masood Muhammad Sajid Riaz , Kamran Ali , Charles N. John, MD Iftikhar Hussain, Muhammad Nabeel
- Coswara - A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis*
By Neeraj Sharma, Prashant Krishnan, Rohit Kumar, Shreyas Ramoji, Srikanth Raj Chetupalli, Nirmala R., Prasanta Kumar Ghosh, and Sriram Ganapathy
- COVID-19 Cough Classification using Machine Learning and Global Smartphone Recordings*
By Madhurananda Pahar, Marisa Kloppe, Robin Warren, and Thomas Niesler
- Robust Detection of COVID-19 in Cough Sounds Using Recurrence Dynamics and Variable Markov Model*
By Pauline Mouawad , Tammuz Dubnov, Shlomo Dubnov