Conversational AI: Data Science (UCS663)

Cough Audio Classification for COVID-19 Diagnosis

(Minor Project)

Submitted by:

Suvrat Arora
101903331
3CO13

Submitted to:

Dr. Sahil Sharma
Assistant Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology

May 2022

Abstract

Cough sound analysis can often prove to be valuable in the diagnosis of respiratory and lung diseases. COVID-19 (Coronavirus Disease 2019) has had devastating effects on humanity, making its early detection in patients an imperative measure for its treatment. Amongst many other symptom analyses, cough sound analysis has been revealed as a significant factor in diagnosing the virus. Applying the process of cough sound analysis for the initial assessment of COVID-19 in patients can be helpful as it is an entirely contactless way to the preliminary detection of the virus. This paper proposes a model that predicts whether a patient is COVID-19 positive or negative based on their cough sound. For this, a dataset comprising cough sound recordings of both COVID-19 positive and negative patients was taken. The audios were converted into spectrograms and certain other features of the audio were extracted. The extracted features were then used to train an ANN model that was able to perform the diagnosis of COVID-19 coughs with a test accuracy of 94.12 on the given dataset.

Keywords: *Cough Sound Classification, COVID-19 Detection*

Table of Contents

Introduction

Coronaviruses are an extensive family of viruses and a subset of Coronavirus ranging from the common cold virus to the cause of more agonizing diseases such as severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and COVID-19. The latest version of Coronavirus, SARS-CoV-2, first identified in Wuhan, China, took the shape of a pandemic by March 2020. The virus has had catastrophic impacts on the world, and medical researchers are constantly on work to counter COVID-19. It has further been established that early detection of the virus can be a crucial factor in saving the lives of patients.

An early and characteristic symptom of COVID-19 is cough - which forms the basis of sound-based study for COVID-19 detection. An amalgamation of cough samples of COVID-19 positive and COVID-19 positive can be used to train machine learning models that can classify cough audios to detect the presence of the virus. The virus spreads by means of direct contact; however, the cough analysis phase is completely contactless - making it a safety touchstone in the early assessment of COVID-19. Although cough analysis has gained great success in the respiratory disease detection sphere, the models proposed for COVID-19 detection using cough analysis are quite a few. Besides, accuracy remains a significant concern when dealing with a problem as sensitive as disease diagnosis.

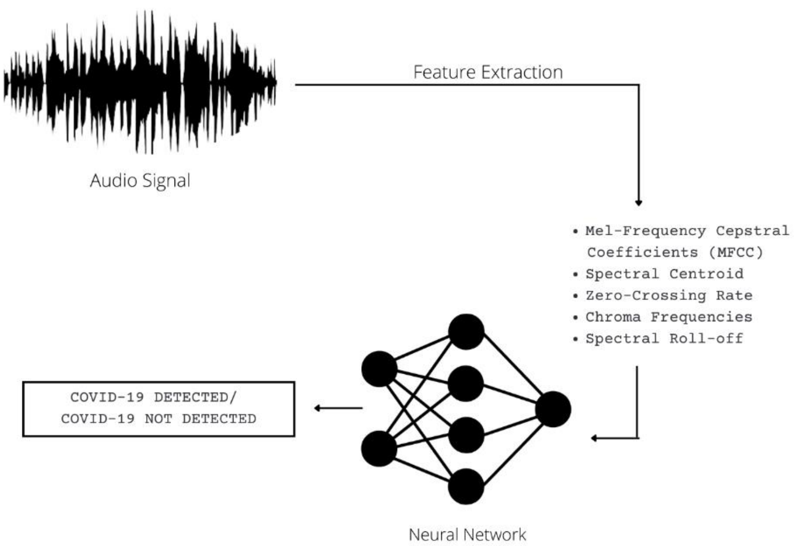

Fig. Schematic Diagram for Cough Audio Analysis for COVID-19 Detection

In this study, the classification problem of COVID-19 diagnosis was dealt with by extracting the following features from the audio signals: Mel-frequency cepstral coefficients (MFCC)(20 in number), Spectral Centroid, Zero-Crossing Rate, Chroma Frequencies, Spectral Roll-off. These extracted features were then used to train an Artificial Neural Network that has a sigmoid activation at the output layer making binary predictions.

The formation of work is as follows: Section 2 explores the previous studies in cough analysis for disease detection. Section 3 discusses the methodology employed to counter the problem statement. Section 4 presents the proposed solution architecture. Section 5 represents an experimental result. In Section 6, the relevant links to the model are presented. Finally, Section 7 concludes the paper.

Related Work

Literature [1] also cough classification for Tuberculosis detection. They performed the study by taking various models into account, namely, K-Nearest Neighbour, Support Vector Machine, MLP, and Convolutional Neural Network. The models were evaluated for different hyperparameters on various evaluation metrics, and a comparative study was carried out.

Literature [2] puts forward a Convolutional Neural Network model for cough classification for COVID-19. The model comprises three convolutional layers with dropouts followed by three fully connected layers with 128, 64 and 32 neurons, respectively. They used the sigmoid activation function at the output layer and achieved a decent accuracy score.

Literature [3] proposed a model for cough sounds classification in noisy environments wherein they extracted several features from the audio signals and then used them for training a Support Vector Machine (SVM) classifier. They were able to achieve a high level of accuracy; however, the model was not specific to the diagnosis of COVID-19.

Literature [4] presented a model for Croup diagnosis (a respiratory disease usually found in children). The proposed to use mathematical features inspired by the human auditory system i.e cochleagram for feature extraction; further Logistic Regression and KNN were applied for classification and a decent accuracy was achieved.

Literature [5] presented a model for Tuberculosis cough sound classification and employed Logistic Regression as the technique for model training. As a part of the pre-processing data stage, they extracted several numerical features from the audio signals and fed them to the Logistic Regression classifier and obtained a moderate accuracy.

Literature [6] presented a study for the detection of Chronic Obstructive Pulmonary Disease (CPOD) using audios recoded using a smartphone's microphone. After having performed the mathematical feature extraction, a comparative study amongst various Machine Learning models was carried out. The best accuracy was obtained by using a Random Forest Classifier.

Literature [7] proposes a cough detection system that constitutes its feature extraction upon STFT. They put forward a Convolutional Neural Network Model that 2 Convolutional Layers, 1 Max Pooling Layer and 2 fully connected layer with which the classification was performed.

Year	Citation	Pre-Processing	Model Basis	Evaluation Metrics	Remarks
2021	[1]	MFCC Feature Extraction	KNN, SVM, MLP, CNN	ROC Curve based metrics	The model focused on Tuberculosis diagnosis.
2020	[2]	Mel Spectrogram Generation	CNN	Accuracy	A decent accuracy score was achieved for COVID-19 diagnosis.
2019	[3]	MFCC Feature Extraction	SVM	Accuracy	The study was not specific to COVID-19 detection.
2019	[4]	MFCC Feature Extraction	Logistic Regression, SVM	Sensitivity, Specificity	The model focused on Croup diagnosis in children.
2018	[5]	MFCC Feature Extraction	Logistic Regression	Sensitivity, Specificity and Accuracy	The study too focused on Tuberculosis diagnosis.
2018	[6]	MFCC Feature Extraction	Random Forest	Accuracy	The study focused on of Chronic Obstructive Pulmonary Disease (CPOD) diagnosis.
2015	[7]	STFT Generation	CNN	Accuracy	The study was confined to cough detection.

Table 1. Previous Studies and Remarks

Methodology

Audio Signals that were present in the dataset in the form of .wav files. These audio files could not be directly fed to a model. This left us with two major approaches – plotting the audio signal or extracting certain characteristic features from the signals. For the purpose of this study, the latter approach was employed.

Dataset:

The dataset used for the model was obtained from Kaggle. It comprised of cough audio recordings of a pool of 170 COVID-19 positive and negative patients. The audio file paths were stored in a .csv file and were tagged as 'covid' and not_covid'.

Data Pre-Processing:

The following features were extracted from all the audio signals and stored into a new csv file with the output labels retained.

- **Mel-frequency cepstral coefficients (MFCC) (20 in number):** Mel-frequency cepstral coefficients (MFCCs) are coefficients that collectively make up an MFC.[1] They are derived from a type of cepstral representation of the audio clip (a nonlinear "spectrum-of-a-spectrum"). The difference between the cepstrum and the mel-frequency cepstrum is that in the MFC, the frequency bands are equally spaced on the mel scale, which approximates the human auditory system's response more closely than the linearly-spaced frequency bands used in the normal spectrum. This frequency warping can allow for better representation of sound, for example, in audio compression.
- **Spectral Centroid**: It is a measure of the amplitude at the centre of the spectrum of the signal distribution over a window calculated from the Fourier transform frequency and amplitude information.
- **Zero Crossing Rate:** The zero-crossing rate (ZCR) is the rate at which a signal transitions from positive to zero to negative or negative to zero to positive. Its value has been extensively used in both speech recognition and music information retrieval for classifying percussive sounds.
- **Chroma Frequencies:** Chroma features are an interesting and powerful representation for audio in which the entire spectrum is projected onto 12 bins representing the 12 distinct semitones (or chroma).
- **Spectral Roll-off:** Spectral roll off is the frequency below which a specified percentage of the total spectral energy lies.

Further as a part of the pre-processing phase, the output labels 'covid' and not_covid' were label encoded to 1s and 0s respectively – making the data fit for feeding to any classification model.

The newly formed dataset comprising extracted features was then used to train an Artificial Neural Network model for classification purpose. Nearly 80% of the data was used for training the neural network and remaining 20% was used as test data. The following diagrammatic representation summaries the methodology employed for the cough audio classification:

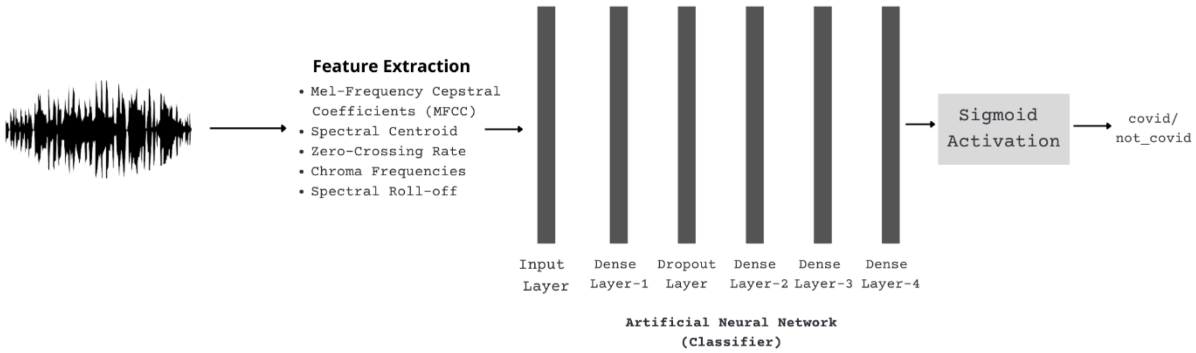

Fig 2. Schematic flowchart for COVID-19 Cough audio classification

The so proposed methodology was put forth after having gauged the various state-of-art models. The MFCC feature extraction has proved to be quite impressive a technique – though usually it is coupled with standard Machine Learning models. Besides, this technique, as such, has hardly ever been employed in the building of a COVID-19 cough classification model. Our study taps all these potential points and attempts to present a model that applies the MFCC extraction technique followed by a Neural Network classifier that gives a more promising expression of higher accuracy.

The detailed architecture and description of the model will be discussed in the subsequent section of the paper.

Proposed Solution Architecture

In order to classify the pre-processed audio data, one has a multitude of models and options available; however, determining the best model is the key. Since we had numerical data at hand (that was obtained after having pre-processed the audio signals) – in this study, we made use of an ANN (Artificial Neural Network) that was trained on the 30% of the data – the remaining 30% was used for testing.

The following diagram depicts in great detail, the architecture of the model that was employed to accomplish the classification:

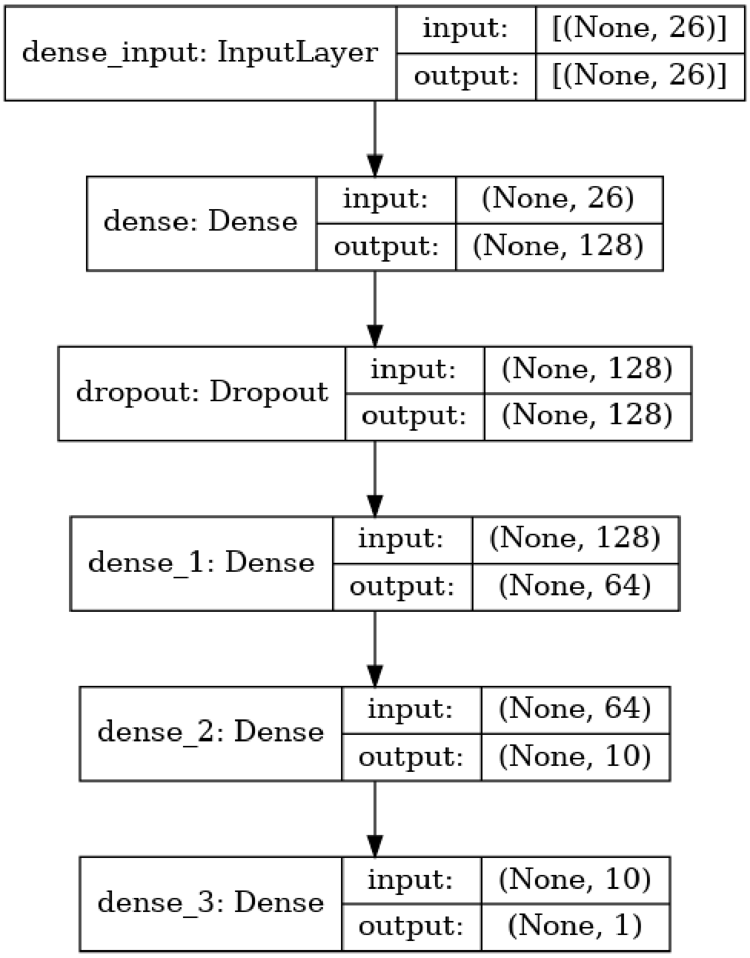

Fig 3. Proposed Model Architecture

All the layers operate in a single dimension since we are dealing with simplified numeric data. The first layer is the input layer. Following which is the first dense layer with 128 neurons. In order to counter overfitting on training data, the mode makes use of dropout regularization. For this the third layer is a dropout layer. Following this, in order to ensure that the data is deciphered better, there are three more dense layers comprising 128, 64 and 10 neurons each.

Activation Function: At the output layer the binary classification task is accomplished by the use of a sigmoid activation function defined as follows.

$$S(x) = \frac{1}{1 + e^{-x}}$$

For the sigmoid function, that results in a value between 0 and 1, **0.5** was taken as the threshold value. For all dense and dropout layers – ReLU Activation function was used.

The overall model can be summarised as the following tabular representation:

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 128)	3456
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8256
dense_2 (Dense)	(None, 10)	650
dense_3 (Dense)	(None, 1)	11

Total params: 12,373
Trainable params: 12,373
Non-trainable params: 0

In total, the model comprises 12, 373 trainable parameters.

Experimentation and Results

Model Training and Hyperparameter Tuning:

For training the model it was highly imperative for us to determine the optimal values of the parameters. The same was accomplished by using the trail and error method. The issues of overfitting, and inadequate accuracy kept on arising. However, the best and the most optimal results were obtained using the following hyperparameters.:

Parameter	Value
Dropout Rate	0.3
Activations in Hidden Layers	ReLU
Activation in Output Layer	Sigmoid
Number of Epochs	15
Optimizer	Adam
Loss	'binary_crossentropy'
Metrics	'accuracy'

Table 2. Hyper Parameters for the Proposed Model

Model Evaluation:

As discussed earlier, accuracy was used as an evaluation metric. The following were the training and testing accuracies obtained:

Training Accuracy: 94.5%

Testing Accuracy: 97.35%

Link to Kaggle Notebook: The Kaggle notebook wherein the project has been implemented can be found here: https://www.kaggle.com/code/suvratarora/cough-classification-for-covid-19-final

Conclusion and Future Scope:

Based on the performance of our technology on both the training/validation and prospective data sets, we conclude that it is indeed possible to accurately and objectively diagnose COVID-19 cough sounds alone. It is possible to augment cough-based features with simple symptoms observable by parents targeting further improvement in performance. Though the accuracy level is quite high and acceptable – it must be noted that the study proposes the model only for preliminary diagnosis of COVID-19 and that the results must be taken as a mere indication. Though the results might vary on datasets, the model lays a groundwork for the imperative task of COVID-19 cough classification using MFCC feature extraction and ANN modelling. This has immense future scope as it can form the basis for classification/diagnosis for any respiratory ailment with characteristics cough audio which is usually the case.

References:

[1] A. Windmon, M. Minakshi, S. Chellappan, P. R. Athilingam, M. Johansson, and B. A. Jenkins, "On detecting chronic obstructive pulmonary disease (COPD) cough using audio signals recorded from smart-phones," in *HEALTHINF 2018 - 11th International Conference on Health Informatics, Proceedings; Part of 11th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2018*, 2018, vol. 5, pp. 329–338. doi: 10.5220/0006549603290338.

[2] J. Amoh and K. Odame, "DeepCough: A deep convolutional neural network in a wearable cough detection system," in *IEEE Biomedical Circuits and Systems Conference: Engineering for Healthy Minds and Able Bodies, BioCAS 2015 - Proceedings*, Dec. 2015, vol. 2015-January. doi: 10.1109/BioCAS.2015.7348395.

[3] R. v. Sharan, U. R. Abeyratne, V. R. Swarnkar, and P. Porter, "Automatic croup diagnosis using cough sound recognition," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 2, pp. 485–495, Feb. 2019, doi: 10.1109/TBME.2018.2849502.

[4] G. H. R. Botha *et al.*, "Detection of tuberculosis by automatic cough sound analysis," *Physiological Measurement*, vol. 39, no. 4, Apr. 2018, doi: 10.1088/1361-6579/aab6d0.

[5] M. Pahar, M. Klopper, B. Reeve, R. Warren, G. Theron, and T. Niesler, “Automatic cough classification for tuberculosis screening in a real-world environment,” *Physiological Measurement,* vol. 42, no. 10, Oct. 2021, doi: 10.1088/1361-6579/ac2fb8.

[6] Bhateja V, Ahmad T, and Sharma DK, *2019 4th International Conference on Information Systems and Computer Networks (ISCON)*. IEEE, 2019.

[7] V. Bansal, G. Pahwa, and N. Kannan, “Cough classification for COVID-19 based on audio mfcc features using convolutional neural networks,” in *2020 IEEE International Conference on Computing, Power and Communication Technologies, GUCON 2020*, Oct. 2020, pp. 604–608. doi: 10.1109/GUCON48875.2020.9231094.