Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 6, June 2021: 1747-1752

Utilizing Convoluted Neural Networks to predict COVID Infections in Individuals

Anilesh Dey[a], Swagata Dasgupta[b], Tanmoy Munshi[c], Indranil Jana[d], Sandipan Seal[e], Sangita Roy[f],Rimpi Datta[g], Surajit Bari[h], Sandhya Pattanayak[i], Kaushik Sarkar[j], Saradindu Panda[k], Pranab Hazra[l], Moupali Roy[m], Soumen Pal[n], Arpita Santra[o], Swati Barui[p], Abhijit Ghosh[q], Puspak Pain[r], Arnima Das[s].

[a*,b,c,d,e,f,g,h,I,j,k,l,m,n,o,p,q,r,s] Department of ECE, Narula Institute of Technology,Agarpara, Kolkata-700109 , West Bengal, India.

Email:[a]anileshdey@ieee.org

Abstract

The world is currently witnessing a second wave of the novel Coronavirus epidemic and countries, especially in the Indian Subcontinent have been heavily affected in its wake. With 158.4 million cases worldwide at this moment, it has not only put a tremendous pressure on the entire diagnostic process of COVID infections (using PCR, Rapid or Serology tests) in suspected individuals but has also put at risk it's availability and affordability at a community level in certain countries.Hence, suspected individuals with mild symptoms are often requested to self quarantine and not take the COVID-Diagnostic tests so that the availability of tests can be ensured for more serious patients. There is also a certain level of hesitancy in patients to go for such diagnostic tests due to factors such as overcrowding, long queues and social stigma. Since one of the major symptoms of COVID infection is coughing, a system for recognizing and diagnosing COVID positive individuals based on raw cough data would have a multitude of beneficial applications at a personal and social level. In this work, we present a system that utilizes Convolutional Neural Networks (CNNs) and audio signal spectrograms to diagnose any individual for Coronavirus infection based on their unique cough audio features. Our diagnosis model achieves an accuracy of about ___ . This result clearly shows that our single diagnosis model is capable of predicting COVID infections in suspected individuals which can serve as a pre diagnostic and preliminary assurance tool.

Keywords: Coronavirus, cough detection, machine learning, Convoluted Neural Network, spectrogram imagery, image recognition

1. Introduction

Studies being conducted in 2021[4] reveal that there are about 1.9 million new cases of coronavirus positive patients daily in India on a weekly average basis. As such there is a huge bottleneck in diagnosing and treating patients. Waiting time for diagnostic test reports are increasing whereas the number of patients in waiting are on the increase too. One of the main symptoms of COVID infection is coughing. If coughing occurs due to COVID infection it also indicates possible severities like pneumonia and respiratory tract blockage in patients. Thus it is a symptom that indicates possible infection that has spread to the lungs or respiratory tract of the patient. Thus on occurrence of such symptoms it might become too late for patients if they have to wait for diagnostic test results. Hence, an automated system that can provide these patients with a preliminary diagnosis of the possibility of COVID infections, can enable the patients to take the advice of their doctors immediately or make necessary arrangements beforehand for any possible emergencies.

Research in the field of cough event detection and diagnosis through machine learning is relatively a new domain of study and is thus not extensively common or well developed. Additionally, there are numerous approaches that can be taken to solve both the cough event detection problem as well machine learning model development. With the various network model structures, audio file sources and pre-processing strategies out there, it is difficult to identify a particular optimal strategy to approach this problem. This paper will therefore only focus on the approach of using a single Convoluted Neural Network (CNN) to identify and attempt to diagnose COVID patients through recorded cough events using image recognition technique.

Using a CNN for the purpose of audio-categorization has been attempted several times with varying success levels. The success of the CNN structure in categorizing audio into the fifty categories of the ESC-50[5] set, one of these being cough sounds, indicated that it is possible to use CNN specifically to distinguish cough sounds. However this approach has not yet been applied to detect COVID infections and there remains a huge untapped potential in this field of research. Thus our proposed solution aims to optimally identify cough events and diagnose it using a proven approach while being entirely unique in its application.

2.Methodology

2.1.. Data collection

The database used for training the CNN in cough detection is composed of various modified audio clips gathered from free online sources.The audio files of coughing are collected from two open source platforms.[1][2]

Each of these audio files originally contained at least one cough event of 2 seconds and were therefore cropped to a length of two seconds with the full cough contained in it. Audio that contained coughs of more than 2 seconds duration were separated into multiple files.The audio files were then sorted in three parts. One was used for training the CNN module, one was for testing the module and another was for checking accuracy.

The files that were used for training the module were separated in two parts- one being the coughing of COVID negative persons and another being the coughing of COVID positive persons. All audio files used had a sampling rate of 44 kHz.The final data file count amounted to 215 cough items.

2.2. 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.[3] 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 nonstationary or nonlinear signals. For this reason, the spectrogram is a very friendly tool for analyzing 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. These lines are then "laid side by side" to form a three-dimensional surface. This process, mathematically is the computation of the squared magnitude of the short-time Fourier transform (STFT) of the supplied signal, that is for a window width ω, spectrogram (t,ω) is $|STFT(t,\omega)|^2$.[6] There are several methods for converting the frequency scale to Spectrogram. Here, we convert audio frequency f into Mel-scale m as.[6]

$$m = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right)$$

The resulting images of pixel size 64x64 are then compiled to form the final database of COVID positive and COVID negative cough samples.

The features we used for the detection model training take the form of Mel-spectrograms. The

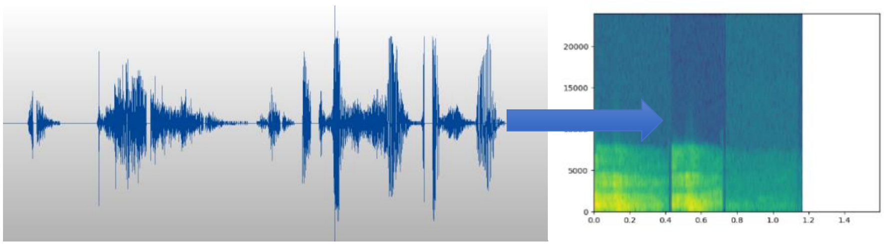

Fig II.a Coughing of COVID Negetive (Audio Wave)(Spectrum)

3. CNN structure for diagnostic model

Just as we humans can recognize objects from images using their distinguishable features, for example, if we see an animal having four legs and cone shaped face we might recognize it as a dog, in CNN, the constituent neuron network model extracts the features of the image. CNN is made up of neurons and each neuron has a learnable weight. These weights are given to each in the process of training. Our project is about diagnosing COVID Positive patients through their cough audio. It is basically a classification problem that is oddly an audio signal classification using CNN. Now, to solve this problem of classifying audio signals using a procedure that is generally used on image classification problems, we used spectrogram images of the individual audio signals.

The most important part of an image classifier is the dataset. More the data, more accurate the model. In this case we have used 215 cough samples and out of it we have 103 positive patient samples and 112 negative patient samples. Using these 215 images labeled as positive and negative 215 such spectrogram images were generated on which the CNN model was built. A CNN is formed of 2 constituent levels. The first level is for feature learning and the next level is for classification of the image based on features learnt.

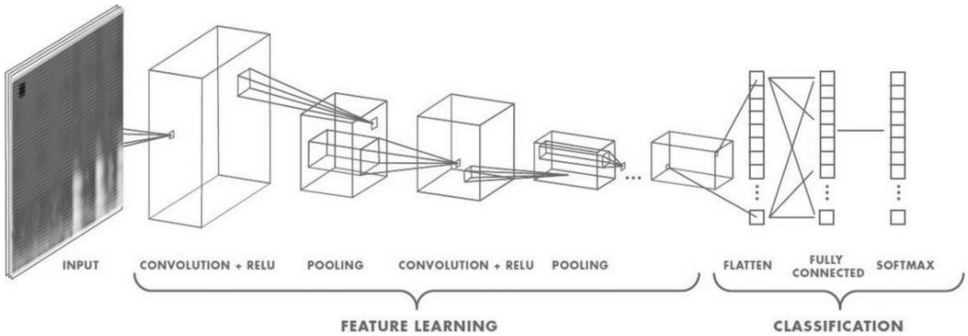

Figure B.III.1: Overall working principle of a CNN model

The first operation is convolution where the relationship between pixels of the image is preserved by using small squares of input data. Each spectrogram image of size 60x60 undergoes a convolution operation with a 3x3 matrix called the filter. Thus a convolved matrix is generated from the image which is also called a feature map. After getting the output, ReLU activation is done which omits the value which is less than 0 and stores the maximum value. Thereafter, Max-pooling is done where the size of the image is halved from 64x64 to 32x32 for easier prediction and more accurate analysis. This procedure selects the bright pixels from the 3x3 matrix. This reduces the time of estimation and processing power. These operations are shown in fig B.III.2.

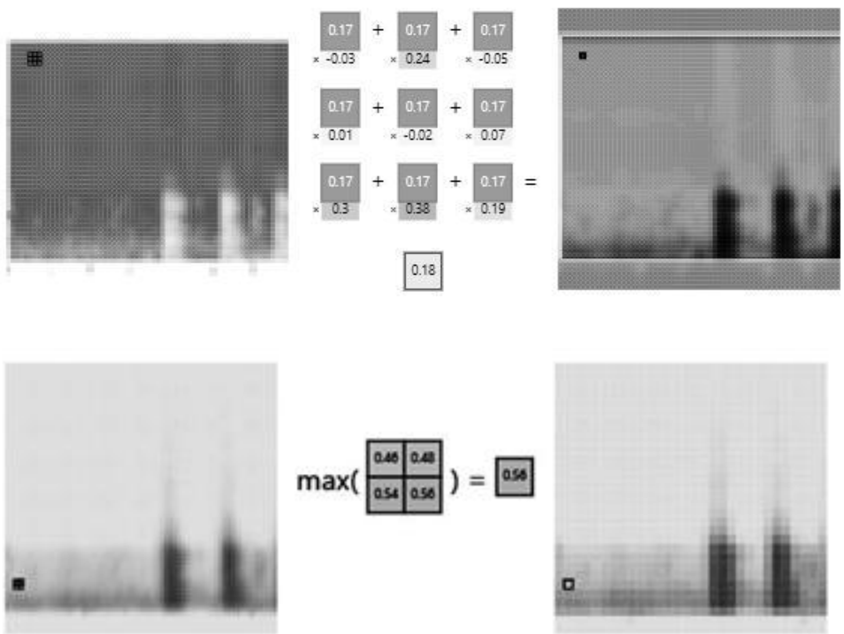

Figure B.III.2 : Max Pooling operation

This explains the working principle of one layer. Several such layers form a single CNN. In our application of the CNN we have used 3 hidden layers. In the last layer, the image size came down to 13x13, after which it was flattened.For prediction, dense operation was done on 2 parts- Positive and Negative, for easier prediction.

4. Diagnostic model training

The CNN was constructed using Keras machine learning library. For the 215 cough samples, 112 of which were COVID negative and 103 of which were COVID positive- a validation split of 80% training and 20% test data was done. Thus 43 samples were reserved for testing out of which 22 were COVID negative and 20 were COVID positive. After flattening the layers, the data is passed on to the dense layer where the final prediction is done. The model was trained for 25 epochs, run several times to assist in tuning the hyperparameters for the detection task. It was found that around 22 epochs were more than enough for training purposes based on the graphs shown in figure B.IV.1 and B.IV.2. The accuracy and loss metrics based on 20 epochs of training thus form the basis of the success of our model.

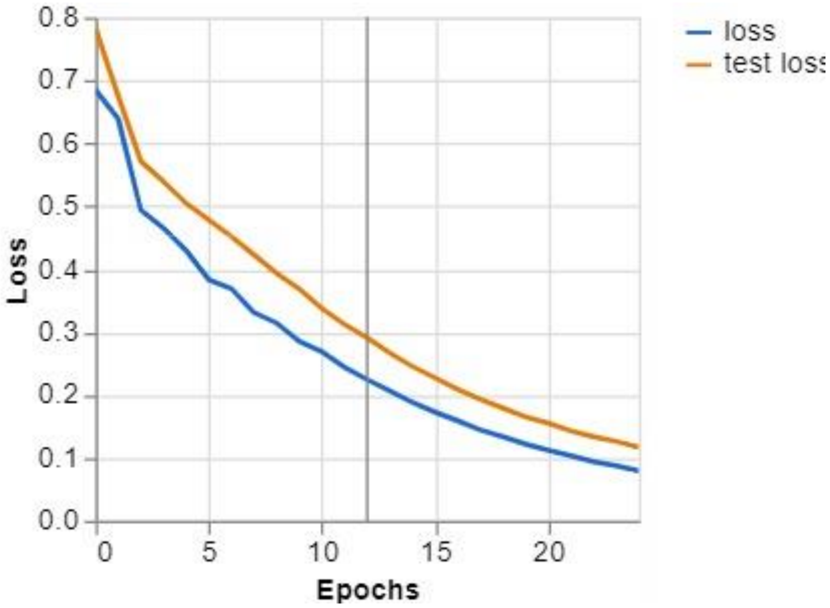

Figure B.IV.1 : Loss vs No. of Epochs graph

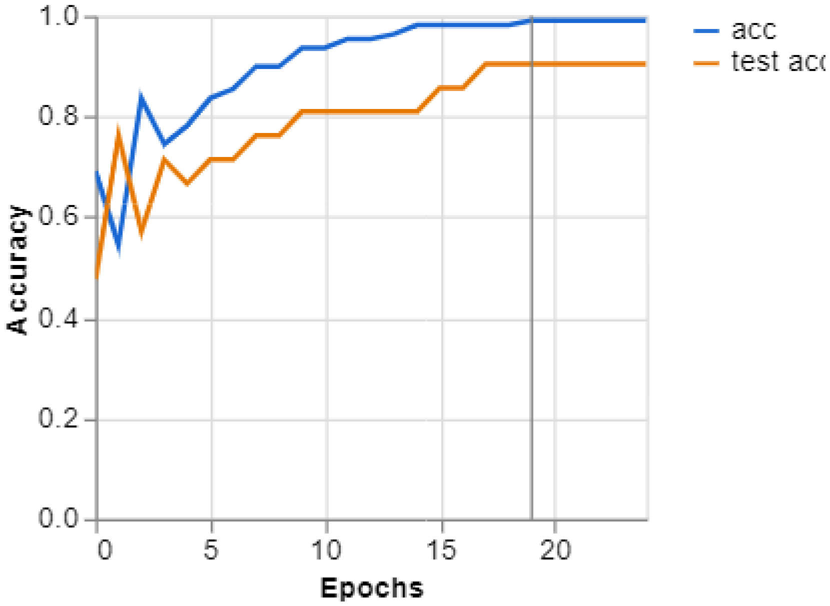

Figure B.IV.2 : Accuracy of model vs No. of Epochs graph

4.1.Observations and Results

The model after training is tested for accuracy using various metrics. We can observe an accuracy of 91% while detecting negative samples.

Accuracy per class

Class	Accuracy	#Samples
Positive	0.80	20
negative	0.91	22

Table C.1 : Accuracy obtained on test set per class

The class vs prediction graph as shown in figure C.1 shows the number of actual positive and negative test samples and the corresponding prediction made against them. Here we observe that out of 22 negative test samples, 20 were predicted to be negative and only 2 were predicted to be positive. Hence this model generated 20 true negatives and 2 false negatives. Also while predicting positive test samples it labelled 4 as negative and 16 as positive. Hence this model generated 4 false negatives and 16 true positives. From this it can be observed that our model has an accuracy of 80% while predicting positive cases and 91% accuracy while predicting samples to be negative cases.

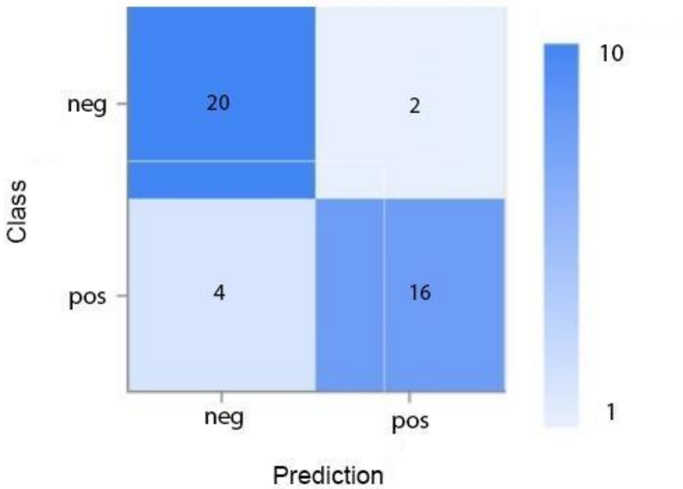

Figure C.1 : Class vs Prediction graph.

5.Conclusion

Our work uses image recognition using CNN to approach the goal of COVID 19 diagnosis by using cough audio data of suspect individuals with cough symptoms.

From such results that we have obtained, it becomes clear that this method of using a Convoluted Neural Network on cough samples of individuals can be a valid method of building a preliminary diagnostic tool for COVID 19 diagnosis. Using limited samples of cough data that we could source, we can still get an accuracy of 80% and 90% while diagnosing positive and negative cases respectively. This model of diagnosing, though based on a limited database, still performs well enough with moderately high accuracy. Given a much larger database for model creation, accuracy can only increase. Thus this research can prove to be not only very relevant at this time but also very much applicable to the field of operation. In future work, we would like to focus on implementing the detection model on mobile devices for a more consistent data collection and extensive database formation for further improvement of accuracy. The datasets acquired would be a community resource for further research and progress.

References

[1] Coughing Dataset of COVID patient by Cosware, IISc
[2] Corona Hack Respiratory Sound dataset, Kaggle
[3] S. Braun, "Ultra-fast real-time spectrum analysis for investigation of communication signals," 2013 European Microwave Conference, 2013, pp. 1119-1122, doi: 10.23919/EuMC.2013.6686858.
[4] Saleem S M, Detecting COVID-19 infection hotspots using Arogya Setu mobile App. Indian J Forensic Community Med 2020;7(4):158-159
[5] Wang, Hui-Hui & Liu, Jia-Ming & You, Mingyu & Li, Guo-Zheng. (2015). Audio signals encoding for cough classification using convolutional neural networks: A comparative study. 442-445. 10.1109/BIBM.2015.7359724.
[6] Bales, Charles & Nabeel, Muhammad & John, Charles & Masood, Usama & Qureshi, Haneya & Farooq, Hasan & Posokhova, Iryna & Imran, Ali. (2020). Can Machine Learning Be Used to Recognize and Diagnose Coughs?.