International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

AUTOMATIC COVID DETECTION USING COUGH SIGNAL ANALYSIS

P. SELVAVENI[1], G. RENISHA[2]

[1]Student, Dept of ECE, GCE-Tirunelveli, Tamil Nadu, India
[2]Assistant Professor, Dept of ECE, GCE-Tirunelveli, Tamil Nadu, India

---***---

Abstract - *Remote observation and measurements are valuable tools for medical applications and they are notably necessary inside the context of pandemic outbreaks similar to this COVID-19. COVID-19 subjects particularly together with symptomless, can be accurately discriminated from forced cough mobile transportable recordings. The information consists of a set of audio samples collected from Virufy/clinical data. Cough recordings are remodeled with Modified Mel frequency cepstral coefficient (MFCC) and inputted into a k-Nearest neighbor (k-NN) classifier during this technique was used to obtain models identifying cough with high performance classifiers were obtained for many of them, including COVD-19. These results are preliminary and there's potential to enhance, as there have been obtained from dataset. In this technique free, invisible, anytime, accessible by anyone, no laboratory testing is needed, no cost, instantly distributable, large scale COVID-19 symptomless screening tool to reinforce practical use cases can be daily screening of students, workers, and public as schools, jobs, and transport reopen, or pool testing to quickly alert of outbreaks in teams. COVID-19 positive samples in our dataset are haphazardly choose the quality of COVID 19 negative subjects for a balanced distribution. The topic forced- cough audios and diagnostic results were used to train and validate the COVID-19 discriminator.*

***Key Words*: Covid-19, MFCC, DFT, k-NN, MODIFIED MFCC, DWT, speech recognition,**

1. INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a highly contagious viral illness caused by severe acute respiratory syndrome SARAS-CoV-2. A Worldwide coordinated effort is required to prevent the further spread of the virus. A deadly disease defined as "occurring as over wide geographical region and affecting an exceptionally high proportion of the population. The virus has inflicted billion of lives across the world in some ways e.g.: physically, psychologically, socially. Compared to other diseases COVID-19 has had: significantly higher transmissibility; worst post recovery; frequent mutation; resulting in higher mortalities and uncontrolled virulence. The clinical manifestations of this particular virus have exhibited deleterious impacts on systems apart from the respiratory systems. In fact, across the world, outbreaks are threatening a second wave, which within the Spanish flu was far more dangerous than the primary one. These outbreaks are very hard to contain with current testing approaches are unless region wide confinement measures were sustained. This partly because limitations of current viral and serology tests and also the lack of complementary pre-screening methods to efficiently select who should be tested. They are expensive to creating the price of testing the entire country day by day impossible. COVID-19 symptoms may range from none to deadly. Severe illness is more likely in elderly patients and those with certain underlying medical conditions. COVID-19 is airborne, spread via air containment by microscope virions. Then risk of infection is highest among people in close proximity, but can occur over long distances, particularly indoors in poorly ventilated areas. transmission rarely occurs via containment surface or fluids. Infected peoples are typically contagious for 10 days, often beginning before or without symptoms.

1.1 Coronavirus

Coronavirus is a large family of viruses that cause illness like respiratory diseases or gastrointestinal diseases. Respiratory diseases can range from the respiratory disease to more severe diseases e.g.

- Middle East Respiratory Syndrome (MERS-CoV)
- Severe Acute Respiratory Syndrome (SARS-CoV)

A completely unique corona virus may be a new strain that has not been identified in humans previously. Once scientists determine exactly what coronavirus it is, the furnish it name as a within the case of COVID-19, the virus causing it is SARS-CoV-2).

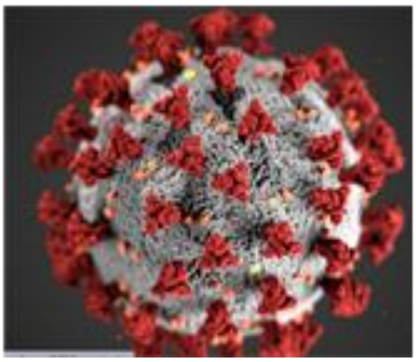

Fig -1: Microscopic picture of coronavirus

Coronaviruses got their name from the way that they appear under a microscope. The virus consists of surrounded by an envelope with protein spikes. The gives it the look of a crown. The word corona means "crown'" in Latin. Coronavirus are transmitted between animals and humans. It's been determined that MERS-COV was transmitted from dromedary camels to humans and SARS from civet cat to humans. The source of SARS-COV2 is yet to be determined, but investigations are ongoing to spot the zoonotic source to the outbreak.

2. DATABASE USED

Recording was available in various browser and devices (virufy/clinical data). data was anonymized before being collected on our secure samples were saved without compression in WAV format. Samples that had no audio content were removed. Segmentation was performed on the cough recordings used to train and test. COVID- 19 positive samples in our dataset are randomly selected the identical range of COVID 19 negative subjects for a balanced distribution. The topic forced- cough audios and diagnostic results were used to train and validate the COVID-19 discriminator.

3.FEATURE EXTRACTION

Feature extraction is an important and basic step of speech recognition. It is associate in nursing exceptional kind of property decreases technique that is used to reduce the data which is intensive to be ready by calculation. In speech recognition, Feature extraction is that the way towards holding valuable data of the signal whereas eliminating of repetitive and undesirable data. It's the parameterization of the speech signal. It is the procedure of fixing the signal to digital type, measuring some imperative character of the signal, for example, energy and frequency responses. Speech recognition could be a supervised learning task. In speech recognition problem input will be the audio signal and we got to predict the text from the audio signals. take the raw audio signal as input to our model be as a result of there will be a lot of noise within the audio signal. It is determined that extracting options from the audio signal and using it as input to the bottom model can manufacture much better performance than directly considering raw audio signal as input

MFCC is that the wide used technique for extracting the options from the audio signal. Instead of Mel frequency cepstral coefficient using (MFCC) using Discrete Fourier Transform (DFT) here modified MFCC with Discrete wavelet transform (DWT) is used to get better accuracy then MFCC using DFT.

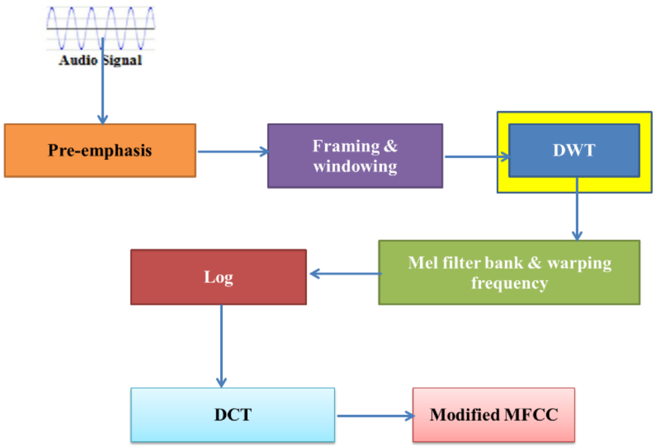

Fig -2: Block diagram of modified MFCC

4.RESULTS

Recording was out there in numerous browser and devices (virufy/clinical data). data was anonymized before being collected on our secure samples were saved without compression in WAV format. Samples that had no audio content were removed. Segmentation was performed on the cough recordings used to train and check.

COVID-19 positive samples in our dataset are randomly selected the same number of COVID 19 negative subjects for a balanced distribution. The topic forced- cough audios and diagnostic results were used to train and validate the COVID-19.

4.1 Input signal

Fig -3 is the plot of the cough input of a 53 old male, who is diagnosis in covid-19.

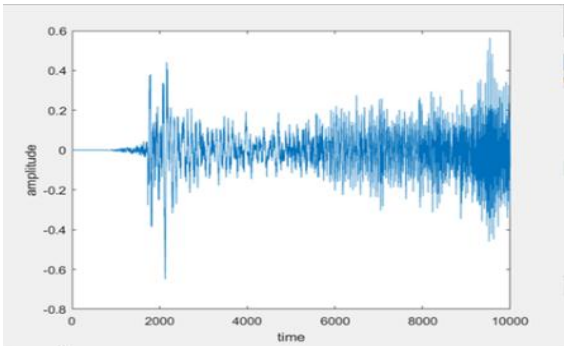

Fig -3: Input signal for covid 19 positive

Fig -3 is the input of a 43 years old normal man.

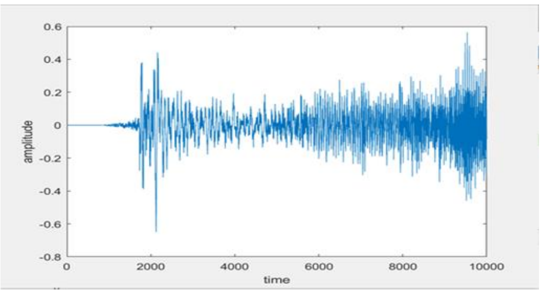

Fig -4: Input signal for covid 19 negative

4.2 Output for Mel filter bank

The final step to computing filter banks is applying triangular filters, generally 40 filters, nfilter=40 on a Mel-scale to the facility spectrum to extract frequency bands. Mel scale accustomed convert between Hertz (f) and Mel (m) using the subsequent equations:

$$f= 700(^{10m/2595}-1)$$

Each filter within the filter bank is triangular having a response of 1 at the center frequency and reduce linearly towards zero until it reaches the middle frequencies of the two adjacent filters wherever the response is 0, as shown during this **Fig -5**

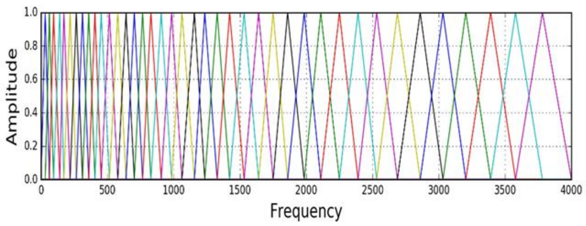

Fig -5: Output for Mel filter bank

4.3 PERFORMANCE METRICS:

Performance metrices are outlined as figures and information representative of an organization's actions, abilities, and overall quality. Performance metrics will vary significantly once when viewed through completely different industries. Performance metrics are integral to an organization's success.

4.3.1 Confusion Matrix

The confusion matrix is an m x m matrix. m is that the no of variables concerned. Diagonal parts represent the true values and false values are expressed by non- diagonal parts. There are four parameters True positive, True negative, false positive, false negative.

4.3.2 Accuracy

Accuracy is decided as a magnitude relation of the events properly classified to the whole events.

Accuracy= TP+TN/TP+TN+FP+FN

4.3.3 Sensitivity

Sensitivity measures the degree of completely classified positive event.

Sensitivity= TP/TP+FN

4.3.4 Precision

Precision is decided because the magnitude relation of range the amount of quality of positive samples properly known to the whole number of positive samples either properly or incorrectly.

Precision=TP/TP+FP.

4.3.5 F-score

It is the common of all recall (sensitivity) and accuracy.

F-score=2*((precision*Recall)/ (precision + Recall)).

Table -1: Performance evaluation report

Performance Metrics	MFCC	Modified MFCC
Accuracy	0.6429	0.8846
Sensitivity	0.2857	0.8846
Specificity	0.64285	0.8846
Precision	0.79165	0.90625
F-score	0.5906	0.8696

From **Table -1** it is inferred that the performance measures using Modified MFCC are better than Mel frequency cepstral coefficient.

This is graphical representation of performance metrics comparison of MFCC and Modified MFCC

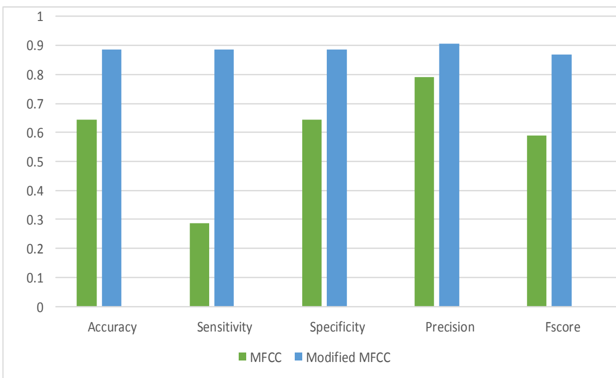

Chart -1: Performance metrics comparison of MFCC and Modified MFCC

Instead of Mel frequency cepstral coefficient using (MFCC) using Discrete Fourier Transform (DFT) here modified MFCC with Discrete wavelet transform (DWT) is used to get 88. % Accuracy. This accuracy is 20% better than MFCC using DFT.

5.CONCLUSIONS

Sound recordings of cough events contain prognosticative information with relevancy the identification of COVID-19 positive cases from a forced- cough recordings, together with 100% of symptomless, at primarily no cost. KNN classifier methods revealed a complex structure for the dynamic time deformation distance between the audio signal clusters. However, there are distribution patterns within the signal space with enough structure on change KNN classifier at high performance to spot COVID-19 positive cases.

The result conferred here should be thought about baseline since they were obtained directly from the raw audio signals of dataset, for classifying a replacement sample signal, a feature vector is built by computing the dissimilarities between the new signal and training set paradigm.

Then, the classifier is applied to the feature vector to get the expected class. There is potential to boost the characterization and the classification result obtained, together with an additional comprehensive analysis of the ways. These should be aspects to be thought about in future studies.

REFERENCES

[1] L. Li et al., "Using artificial intelligence to detect COVID-19 and community acquired pneumonia based pulmonary CT: Evaluation of the diagnostic accuracy", Radiology, Vol.296, no.2, pp. E65-E71, Aug. 2020.

[2] X. Mei et al., "Artificial intelligence-enabled rapid diagnosis of patients with COVID-19," Nature Med., vol 26, pp.1224-1228,2020

[3] A. Imran et al., "AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app,"2020, ArXiv :2004.01275.

[4] G. Botha *et al.*, "Detection of tuberculosis by automatic cough sound analysis," *Physiological Meas.*, vol. 39, no. 4, 2018, Art. no. 045005.

[5] A. Windmon *et al.*, "Tussis watch: A smartphone system to identify cough episodes as early symptoms of chronic obstructive pulmonary disease and congestive heart failure," *IEEE J. Biomed. Health In format.*, vol. 23, no. 4, pp. 1566–1573, Jul. 2019.

[6] C. Pham, "Mob cough: Real-time cough detection and monitoring using low-cost mobile devices," in *Proc. Asian Conf. Intel. Inf. Database Syst.*, 2016, pp. 300–309.

[7] U.R. Abeyratne, V. Swanker, A. Satyarthi, and R. Triasih, "Cough sound analysis can rapidly diagnose childhood pneumonia," *Ann. Biomed. Eng.*, vol. 41, no. 11, pp. 2448–2462, 2013.

[8] R.X.A. Pramono, S. A. Imtiaz, and E. Rodriguez-Villegas, "A Cough based algorithm for automatic diagnosis of pertussis," vol. 11, no. 9, 2016.

[9] S.T. Moein, S.M. Hashemian, B. Mansour Afshar, A. Khorram Tousi, P. Tamaris, and R.L. Doty, "Smell dysfunction: AbiomarkerforCOVID-19,"in *International Forum of Allergy and Rhinology*. Hoboken, NJ, USA: Wiley Online Library, 2020

[10] T.T.T.F. Quartier and J. S. Palmer, "A framework for biomarkers of COVID-19 based on coordination of speech-production subsystems," *IEEE Open J. Eng. Med. Biol.*, vol. 1, pp. 203–206, 2020.

[11] T. Higginbottom and J. Payne, "Glottis narrowing in lung disease," *Amer. Rev. Respiratory Disease*, vol. 125, no. 6, pp. 746–750, 1982