2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

COVID-19 Respiratory Symptoms Detection using Speech Processing Methods

I. Kassis and Y. Zigel, *Senior Member, IEEE*

***Abstract*— This work presents an SVM-CNN-based system for detecting COVID-19 using speech signals in a non-invasive and remote way. Speech recordings of 42 diagnosed and control subjects from the Voca.ai database were analyzed. The system achieved an average accuracy of 74.7%.**

***Clinical Relevance*— This work proposes a system that classifies the presence of respiratory symptoms of COVID-19 from speech recordings remotely made on simple mobile devices. This approach has the potential of screening for a large population and reduce the spread of the pandemic.**

I. INTRODUCTION

SARS-CoV-2, a newly emergent coronavirus, caused coronavirus disease 2019 (COVID-19), was first recognized in Wuhan, China, in December 2019 [1]. The world population is suffering from more than 4 million registered deaths induced by the virus since the outbreak of the epidemic. COVID-19 is a respiratory disease, that reaches specifically to the airways, with a wide range of symptoms such as dry cough, shortness of breath, and mucus production [2]. In speech production, during the exhalation stage, air from the lungs moves through the trachea and larynx and into the vocal tract pharyngeal, oral and nasal cavities. How we breathe during speech has much influence on our voice; therefore, a condition involving inflammation can affect vocal quality [3].

II. METHODS

In this work, we processed speech recordings of 42 subjects in the 17-58 age range from the Voca.ai database (21 were diagnosed as positive for COVID-19 with respiratory symptoms, and 21 were diagnosed as negative with no symptoms). Each subject recorded several audio files such as vowels for the longest duration possible without breathing, pronouncing the Alphabet, and counting in English. The proposed system, shown in Fig. 1, includes a pre-processing step (down-sampling and framing). Next, speech event detection was performed using an energy threshold and voicing level of each frame. The first part of the system uses a Support Vector Machine (SVM) classifier to perform the classification of elongated vowels recordings. We extracted a set of 9 features from the vowels using custom-written code such as maximum phonation time, energy ratio and standard deviation of the pitch. We used an SVM classifier with Radial Basis Function kernel and performed the classification 15 times with different data divisions. The second part is a Convolutional Neural Network (CNN) and SVM combined model to classify the continuous speech recordings. We used the CNN to extract 5 deep-spectrum features from the mel-scaled spectrograms and extracted 5 short-time features from the raw speech signal. An SVM classifier was used to classify the segments into two classes. At last, the two sub-systems were fused, and a binary decision was made for each subject based on 5 speech recordings.

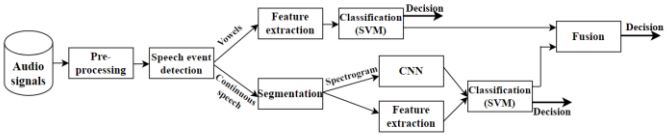

Figure 1. A block diagram of the proposed analysis system

III. RESULTS

To test the performance of the proposed system, we divided our dataset into 15 different train-test groups (75% and 25%, respectively) while balancing the number of positive and negative in each group. The mean and standard deviation of each performance measure on the different data divisions are presented in Table 1.

TABLE I. CLASSIFICATION RESULTS FOR 15 DIFFERENT TRAIN-TEST DATASETS: *MEAN ± STD*

Measure [%]	Elongated vowels	Continuous speech	Fusion
Accuracy	74.7±12.5	51.3±14.6	68.0±11.4
Sensitivity	76.0±21.6	52.0±24.8	69.3±22.5
Specificity	73.3±30.8	50.7±24.9	66.7±28.9
PPV	80.3±18.4	54.7±23.3	72.8±19.3

IV. DISCUSSION & CONCLUSION

The results show that system performances are higher in vowels than in continuous speech or fusion. This study shows high potential for a speech signal processing-based system for remote detection of COVID-19 symptoms.

REFERENCES

[1] World Health Organization, 'Clinical management of COVID-19'. https://www.who.int/publications-detail-redirect/clinical-management-of-covid-19 (accessed Oct. 15, 2020).

[2] A. Bernheim *et al.*, 'Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection', *Radiology*, vol. 295, no. 3, p. 200463, Feb. 2020, doi: 10.1148/radiol.2020200463.

[3] T. F. Quatieri, T. Talkar, and J. S. Palmer, 'A Framework for Biomarkers of COVID-19 Based on Coordination of Speech-Production Subsystems', *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 1, pp. 203–206, 2020, doi: 10.1109/OJEMB.2020.2998051.

*Research supported by the Data Science Research Center, BGU.

I. Kassis is with the Ben-Gurion University of the Negev, Beer-Sheva, Israel (e-mail: idankas@post.bgu.ac.il).