Deep-learning Based Approach to Identify Covid-19

Ke Feng
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, FL, USA
fengk2@my.erau.edu

Fengyu He
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, FL, USA
hef1@my.erau.edu

Jessica Steinmann
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, FL, USA
steinmaj@my.erau.edu

Ilteris Demirkiran
Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, FL, USA
demir4a4@erau.edu

***Abstract*— COVID-19 is a large-scale contagious respiratory disease that has spread across the world in 2020. Therefore, a low-cost, fast, and easily available solution is needed to provide a COVID-19 diagnosis to curb the outbreak. According to recent studies, one of the main symptoms of COVID-19 is coughing. The goal of this research effort is to develop a method for the automatic diagnosis of COVID-19 by detecting cough during recorded conversations. The method is composed of five main modules: sound extraction, sound feature extraction, cough detection, cough classification, and COVID-19 diagnosis. The method extracts relevant features from the audio signal and then uses machine learning and deep learning models, like SVM, KNN, and RNN, to classify them, which can successfully diagnose COVID-19 from audio recordings. Our method has relatively high accuracy when dealing with completely unfamiliar cough samples. When the training set and the test set are from two different databases, it still achieves an accuracy of 81.25% (AUC of 0.79). As more data sets are collected, the model can be further developed and improved to create a machine learning solution based on cough analysis for COVID-19 detection, which may be promoted as a non-clinical self-inspection solution.**

Keywords*— COVID-19*, Recurrent neural network, Deep Learning

I. Introduction & background

Since the outbreak of the new crown pneumonia in Wuhan, China in December 2019, the epidemic has appeared and spread rapidly in many countries around the world. As of mid-February 2021, the new crown virus has infected 111M confirmed patients worldwide (2.47M deaths were reported and 62.9M recovered), and it has become a major global health problem [1]. With the deepening of research around the world, more and more studies have shown that COVID-19 spreads from person to person through droplets or direct contact [2]. The spread of the virus is accelerated through human-to-human transmission [3]. According to recent research, the load of this virus is at its peak in the early stages of the illness, making COVID-19 infectious even before the symptom onset[3][4][5], so it is crucial to limit transmission by different measures, including screening individuals with or without the symptoms [6]. Due to the sudden emergence of the virus, companies around the world have developed different vaccines, but stock is limited. Therefore, the world has not yet reached the number of people vaccinated needed in order to slow down or stop the spread of the virus [7]. Currently, preventive measures are still an important procedure to limit the spread of COVID-19. There is an urgent need for a method that can quickly determine whether people are infected with COVID-19. At present, the detection of COVID-19 is mainly conducted by medical staff and patients through face-to-face collection of biological samples. This method has many disadvantages, such as spread risk, high detection cost, and long detection cycle. The cost of testing is between \$23 to \$2315. Even with the lowest cost testing method, it will still cost over \$8 billion [8]. In some underdeveloped areas, it is difficult to diagnose the virus due to lack of medical facilities and medical professionals. Therefore, a low-cost, fast, and easily available solution is needed.

It is confirmed that fever and cough are the most common symptoms of a COVID-19 infected person [9] and in most cases, dry cough is a distinctive feature, usually coupled with other severe symptoms including ground-glass opacities and pneumonia[10] Recently, many different groups have proposed AI approaches to support pandemic management [11][12][13]. Renard et al. proposed an algorithm for automatically diagnosing whooping cough by analyzing cough sounds through machine learning, indicating that machine learning has the prospect of diagnosing diseases through cough sounds [14]. Research from the Massachusetts Institute of Technology [15] also shows that machine learning has the potential to detect COVID-19. Cambridge uses Mel Frequency Cepstral Coefficient (MFCC) and other audio statistics [16]. Then these features are further processed by principal component analysis and used as input to a simple binary classifier to achieve an AUC of 0.82 [16]. But the entire data set only has 86 cough samples. The Massachusetts Institute of Technology[16] collects cough audio samples from smartphones around the world, but the database is not open source, and depends on volunteers' self-reporting of their COVID-19 results as a labeling standard, the reliability of the data is difficult to determine. Virufy AI Research Group issued an open clinical dataset in order to provide a dataset with high reliability, but the database samples are currently limited to only 16 samples [17].

II. Method

A. Dataset Description

Data from two open-source databases were used for feature extraction and model training. The first database is Coswara[18], which currently has open source data from a total of 1433 participants. Coswara collected multiple voice samples for each participant, but for this study only shallow cough recordings were used. To make the training data set more balanced, thus improving efficiency and accuracy when training the model, only 200 samples were randomly selected from 1000 healthy individuals, and all of the data from patients with a confirmed COVID-19 diagnosis.

The second dataset is Virufy, which is very accurate because it was collected at a hospital under supervision by physicians following Standard Operating Procedures (SOP) and informed patient consent [17]. This dataset collected a

SoutheastCon 2021 | 978-1-6654-0379-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/SOUTHEASTCON45413.2021.9401826

total of 16 recordings of cough sounds from 7 patients with confirmed COVID-19 and 9 from healthy individuals.

Given the high number of recordings in the Coswara database, it was used as the training dataset. In general, deep learning is prone to performs well with training datasets, but often sees a significant drop in accuracy when used with other datasets which is a problem that was considered. To prove the generalizability of our proposed algorithm, only the Coswara database was used for training and the full Virufy dataset for testing.

B. Algorithm Composition

The algorithm is composed of the five main modules: sound extraction, sound feature extraction, cough detection, cough classification, and COVID-19 diagnosis. The first step is to extract possible cough sounds from recorded audio and remove the silent part. After removing the silent part of the audio signal, in the extracted signal, coughs are mixed in with people's conversations. In such a situation, high accuracy machine learning models were utilized to distinguish between normal conversation sounds and coughing. In order to do that, some effective features were extracted to train the model, such as Mel-Frequency Cepstral Coefficient, Entropy of Energy and Zero Crossing Rate. The third step is to train the unsupervised learning models to identify the real cough and delete the non-cough sounds. The fourth step is to build a deep learning model with more input features to focus on identifying whether the cough is caused by COVID-19 or not. Then finally provide the diagnoses. All programs are based on the python libraries： pyaudio, pyAudioAnalysis, sklearn, NumPy.

a) Sound Extraction:

Before starting speech recognition, the silent part of the recorded audio needs to be removed in order to reduce the interference caused to the subsequent steps and to improve the computation speed.

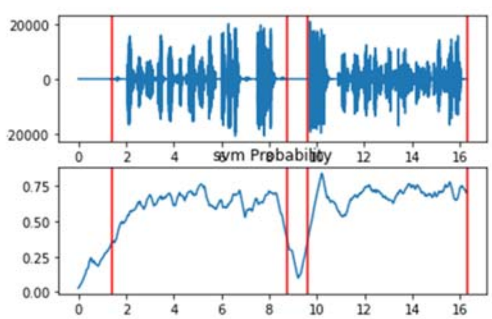

Figure 1: The top half of Figure 1 is the original signal of an audio recording, and the bottom half is the possibility that the SVM model returns the current frame as a human conversation.

In order to eliminate the silent part, the Energy feature of each audio recording was first extracted. The energy is calculated by summing the squares of the signal and then normalizing it by the frame length. Next, an SVM model was trained with the audio recordings whose Energy features were within the top 10% and bottom 10%. This SVM model is a binary classifier which can distinguish between Silent recordings and noise/speaking recordings. Finally, we remove the quiet parts of the recordings and keep only the parts with conversation activity.

b) Sound Feature Extraction:

We extracted a total of three types of features: short-term, medium-term and long term-features. According to this paper about pyAudioAnalysis[19], the input audio recordings are first divided into segments of fixed length of 1 second, and then each segment is divided into frames of 50 ms. The features extracted for each frame are called short-term features, and the mean and standard deviation of short-term features extracted for all frames of each segment are the medium-term features. Lastly, features of each segment are averaged to obtain the long-time features of each audio recording.

Next multiple short-term features of the audio signal in the time domain and frequency domain were extracted. Time domain features are extracted directly from the original signal. Most of the frequency domain features used are based on the Discrete Fourier Transform (DFT).

The frequency domain feature used is Mel-Frequency Cepstral Coefficients (MFCCs), which is the cepstral domain feature generated by applying Inverse DFT on the spectrum. A commonly used audio feature derived from the short-term power spectrum [20].

The complete list of features extracted is presented in Table 1 below.

Index	Feature	Description
1	Zero Crossing Rate	The sign change rate of the signal within a particular frame duration
2	Energy	The sum of the signal values squares
3	Entropy of Energy	The sub-frames' normalized energies entropy, which can be interpreted as a measure of abrupt changes of energies.
4	Spectral Centroid	The gravity center of the frequency spectrum.
5	Spectral Spread	The second gravity center of the frequency spectrum.
6	Spectral Entropy	the normalized spectral energies entropy for a set of sub-frames.
7	Spectral Flux	The squared difference between the frequency spectra of the two successive frames.
8	MFCCs	a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency.

Table 1: List of features extracted from the audio signal, index 1-3 are the time-domain features, index 4-8 are the frequency-domain features.

c) Cough Detection:

The purpose of cough detection is to distinguish between conversational activity and coughing. In the previous step, short time features and mid-term features for all recordings were extracted. Since the mid-term features are the features of each segment, which are calculated as the average and standard deviation of the short-time features extracted from each frame, in order to reduce the computational effort and speed up the computation, the analysis was done at the segment level when performing the clustering analysis of cough detection. Now the extracted mid-term features are passed together with the recordings to a simple KNN model that automatically clusters the human conversation and cough in each recording, and then connects the clustered recordings to generate a new recording. An example is shown in Figure 2 below.

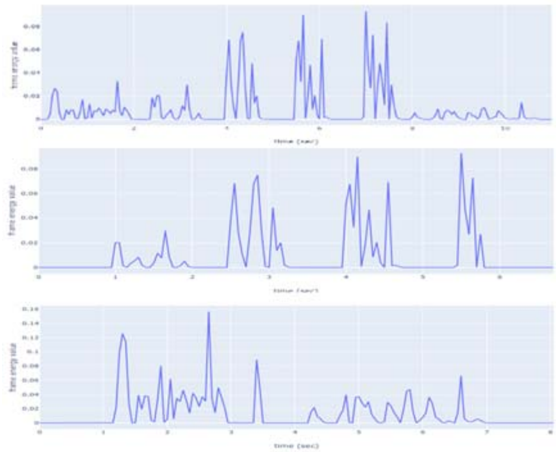

Figure 2: (a) Original audio signal which contains both cough and conversation; (b) Audio signal after clustering with only the cough; (c) Audio signal after clustering with only the conversation.

d) Cough Classification:

Now that all coughs have been extracted from the recordings, the purpose of the cough classification is to distinguish whether a participant's cough is of COVID-19 or a health individual.

Four models were trained, KNN, SVM, Random Forest, and RNN, but firstly, the features in section b were re-extracted from the training data since the conversation part of the recordings were removed and only used the long-term features of each recording. The optimal parameters for the KNN model are the number of neighbors, the optimal parameters for the SVM are the cost of constraints violation, and the optimal parameters for the Random Forest is the number of trees in the forest. Table 2, below shows the best parameter selection and accuracy of the four models on the training set.

Model	Parameter	Accuracy
KNN	13	71.4
SVM	0.01	73.2
Random Forest	100	72.5
RNN		0.9956

Table 2: Best training results for various models.

The best performance comes from recurrent neural network (RNN) model. The structure of this model is as follows, the input layer is a SimpleRNN with 20 units and the input shape is (383, 1, 136). The first hidden layer is a dense layer with 512 units, the second hidden layer is a SimpleRNN with 10 units, the third hidden layer is also a dense layer with 256 units. All the hidden layers use the RELU activation function. The output layer is a SimpleRNN with only 1 unit, and since the objective is to determine whether the cough is COVID-19 or not, which is a binary problem, the sigmoid activation function was used.

e) COVID-19 Diagnosis

The last step is to give a possible diagnosis based on the predictions given by the model. However, the diagnosis given is only based on the patient's cough sound. According to previous studies, COVID-19 will also have other symptoms besides cough, the most common of which are fever and myalgia (muscle pain) [21]. Therefore, in order to give a more accurate diagnosis, test subjects should be remined to pay attention to their symptoms, get a biologically tested, and seek medical attention when they occur.

III. RESULT

To evaluate the predictive power of the model, the ROC curve, area under the ROC Curve (AUC), and accuracy as evaluation metrics were used. First, recordings were randomly divided from the pre-processed Coswara database into a training set and a test set, with a ratio of 80% and 20%. The accuracy of the RNN model on the test set is about 0.9, and the ROC curve is shown in the Figure 3 below, with an AUC of 0.9282. In order to test the generalizability of the model, the second test set used data that the model had never seen before. Since the Virufy dataset has only 16 data items in total, all samples were used as a test set to evaluate the performance of the RNN model. The accuracy of the model on this dataset is about 0.8125, and the ROC curve is shown below in Figure 4 below, with an AUC of 0.79.

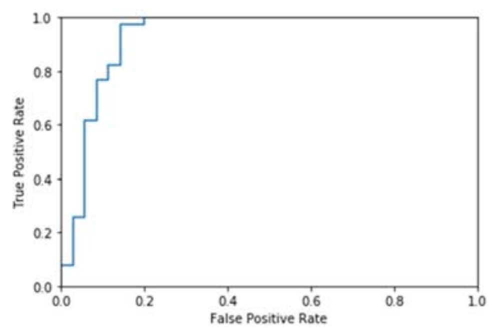

Figure 3: ROC curve when both train set and test set were from the Coswara dataset.

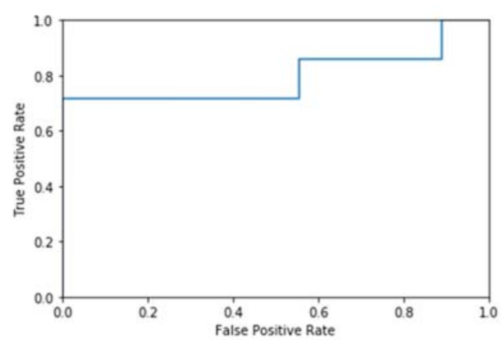

Figure 4: ROC curve with train set from Coswara dataset and test set from Virufy dataset.

Comparing the performance of RNN on the Coswara and Virufy test sets, this RNN model has some generalizability to different datasets, and the AUC only decreases by about 10% on the new dataset, which is about 0.8

IV. Conclusion & Discussion

In this paper, a method was proposed that detects whether or not a cough is due to COVID-19 from recordings of conversations. This is done by first removing the silent part of conversational recordings in order to extract the sound activity, then multiple signal features were extracted, clustered using a KNN model to extract coughs from the conversation, then a RNN model was trained to classify whether the cough is due to COVID-19 or not, this is a quick way to prescreen for COVID-19.

The model mentioned in the paper achieves an AUC of 0.9282 when classifying data from the same sources as those used in training, and the model also has good generalizability, with an AUC of about 0.8 for databases from different sources. Thus, it can increase the confidence of this hypothesis that COVID-19 can be detected through people's coughing sounds in conversations and provides a cheap and fast detection method to curb the development of this infectious disease.

In the future, there are plans to do more work to overcome the limitation of the current method. First, more data to train the model needs to be collected. It is not only important to increase the number of collected recordings, but there is also a need to collect samples from different countries, regions, ages, genders, and ethnicities to improve the generalizability of the model, and to further improve the accuracy by training the model with more data. Second, a total of eight signal features were used to train the model, but the PCA analysis can be used to select the highest variance features in order to improve the training efficiency and avoid overtraining. Finally, when selecting features, in addition to the time and frequency domain features of the signal, some COVID-19 related symptoms could also be recorded while collecting participants' coughs. This can be used to train the model and improve the performance of the model if the model was improved to look for keywords in conversations that describe COVID-19 symptoms.

References

[1] News.google.com, 2021. [Online]. Available: https://news.google.com/covid19/map.

[2] R. Wölfel et al., "Author Correction: Virological assessment of hospitalized patients with COVID-2019", Nature, vol. 588, no. 7839, pp. E35-E35, 2020.

[3] "Coronavirus Disease 2019 (COVID-19)", Centers for Disease Control and Prevention, 2021. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/index.html.

[4] Woelfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Mueller MA, et al. Clinical presentation and virological assessment of hospitalized cases of coronavirus disease 2019 in a travel-associated transmission cluster.

[5] J. Yang et al., "Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis", International Journal of Infectious Diseases, vol. 94, pp. 91-95, 2020.

[6] K. Gostic, A. Gomez, R. Mummah, A. Kucharski and J. Lloyd-Smith, "Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19", eLife, vol. 9, 2020.

[7] "COVID-19 Vaccine Key to Reaching 'Herd Immunity'", Muhealth.org, 2021. [Online]. Available: https://www.muhealth.org/our-stories/covid-19-vaccine-key-reaching-herd-immunity.

[8] "covid-test-cost", Advisory.com, 2021. [Online]. Available: https://www.advisory.com/daily-briefing/2020/06/17/covid-test-cost.

[9] M. Grant et al., "The Prevalence of Symptoms in 24,410 Adults Infected by the Novel Coronavirus (SARS-CoV-2; COVID-19): A Systematic Review and Meta-Analysis of 148 Studies from 9 Countries", SSRN Electronic Journal, 2020.

[10] A. Mahashur, "Chronic dry cough: Diagnostic and management approaches", Lung India, vol. 32, no. 1, p. 44, 2015.

[11] L. Li et al., "Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy", Radiology, vol. 296, no. 2, pp. E65-E71, 2020.

[12] X. Mei et al., "Artificial intelligence–enabled rapid diagnosis of patients with COVID-19", Nature Medicine, vol. 26, no. 8, pp. 1224-1228, 2020.

[13] A. Imran et al., "AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app", Informatics in Medicine Unlocked, vol. 20, p. 100378, 2020.

[14] A. P. Renard Xaviero, A. I. Syed and E. Rodriguez-Villegas, "A Cough-Based Algorithm for Automatic Diagnosis of Pertussis," PLoS One, vol. 11, (9), 2016.

[15] C. Brown et al., "Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data," in Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, New York, NY, USA, Aug. 2020, pp. 3474–3484.

[16] Laguarta, J., Hueto, F., & Subirana, B. (2020). COVID-19 Artificial Intelligence Diagnosis using only Cough Recordings. IEEE Open Journal of Engineering in Medicine and Biology., 1–1.

[17] Chaudhari, G., "Virufy: Global Applicability of Crowdsourced and Clinical Datasets for AI Detection of COVID-19 from Cough", 2020.

[18] N. Sharma et al., "Coswara -- A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis," ArXiv200510548 Cs Eess, Aug. 2020.

[19] Giannakopoulos T (2015) pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis.

[20] S. W. Davis and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Se," 1980.

[21] S. A. Gómez-Ochoa et al., "COVID-19 in Health-Care Workers: A Living Systematic Review and Meta-Analysis of Prevalence, Risk Factors, Clinical Characteristics, and Outcomes," Am. J. Epidemiol., no. kwaa191.