

COVID-19 Diagnosis from Crowdsourced Cough Sound Data

Myoung-Jin Son ¹  and Seok-Pil Lee ^{2,*} 

¹ Department of Computer Science, Graduate School, SangMyung University, Seoul 03016, Korea; tbvltm0306@naver.com

² Department of Electronic Engineering, SangMyung University, Seoul 03016, Korea

* Correspondence: esprit@smu.ac.kr

Abstract: The highly contagious and rapidly mutating COVID-19 virus is affecting individuals worldwide. A rapid and large-scale method for COVID-19 testing is needed to prevent infection. Cough testing using AI has been shown to be potentially valuable. In this paper, we propose a COVID-19 diagnostic method based on an AI cough test. We used only crowdsourced cough sound data to distinguish between the cough sound of COVID-19-positive people and that of healthy people. First, we used the COUGHVID cough database to segment only the cough sound from the original cough data. An effective audio feature set was then extracted from the segmented cough sounds. A deep learning model was trained on the extracted feature set. The COVID-19 diagnostic system constructed using this method had a sensitivity of 93% and a specificity of 94%, and achieved better results than models trained by other existing methods.

Keywords: AI diagnostics; COVID-19 screening; deep learning; speech recognition



Citation: Son, M.-J.; Lee, S.-P.

COVID-19 Diagnosis from
Crowdsourced Cough Sound Data.
Appl. Sci. **2022**, *12*, 1795. <https://doi.org/10.3390/app12041795>

Academic Editor: Namkug Kim

Received: 8 January 2022

Accepted: 7 February 2022

Published: 9 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

COVID-19, a disease caused by the coronavirus SARS-CoV-2, was declared a pandemic by the World Health Organization (WHO, Geneva, Switzerland) on 11 March 2020. The virus is spread in the form of small particles expelled from the mouth or nose when an infected person coughs, sneezes, speaks, sings, or breathes. These particles take various forms, from respiratory droplets to aerosols. Therefore, individuals adjacent to a person infected with the COVID-19 virus, or who touch the eyes, nose, or mouth with hands that have come into contact with contaminated surfaces, may get infected. The infection therefore spreads more easily indoors or in crowded places.

The symptoms of COVID-19 range from none to mild, moderate, or severe, with major symptoms including fever of 37.5 °C or higher, fatigue and cough, and dyspnea (shortness of breath), while other symptoms include chills, muscle pain, headache, and olfactory, or taste loss. Other symptoms, such as loss of appetite, production of phlegm, digestive symptoms, confusion, dizziness, runny nose, or stuffy nose may occur. For most people, symptoms are mild. However, in some people, the virus affects the respiratory system and changes the quality of voice, coughing sound, and breathing sound. The trend of COVID-19 confirmed patients around the world is shown in Figure 1.

As shown in Figure 1, the total number of confirmed infections worldwide is about 250 million, and the death toll is about 5.1 million. As the number of confirmed cases increases, the number of patients who have recovered is increasing, but there have been many deaths.

Many epidemiological experts argue that a large-scale coronavirus testing is essential. Reverse transcription–polymerase chain reaction (RT-PCR), which is currently used all over the world, is currently the standard approach to diagnosing COVID-19 with high accuracy [2]. The RT-PCR test is reliable, but the process of collecting a sample using a long swab in the nose can be painful, and is costly to the individual in some countries [3]. Rapid diagnosis is also difficult, because the test results can be received within a few hours to

a few days after the test. With the advent of the new virus, many experts are uncertain as to when herd immunity will be reached. There are four recently developed vaccines: AstraZeneca, Janssen, Pfizer, and Moderna. Vaccination helps protect individuals from COVID-19. Vaccination is recommended at the national level, and many people in South Korea have been vaccinated (Figure 2).

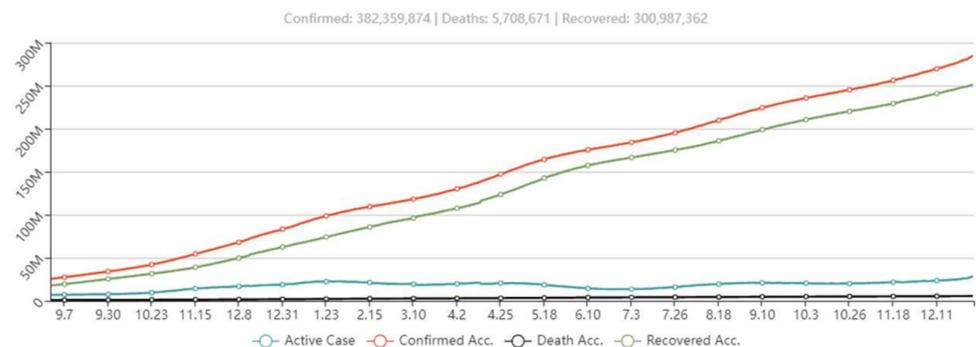


Figure 1. The trend of COVID-19 confirmed patients around the world [1].

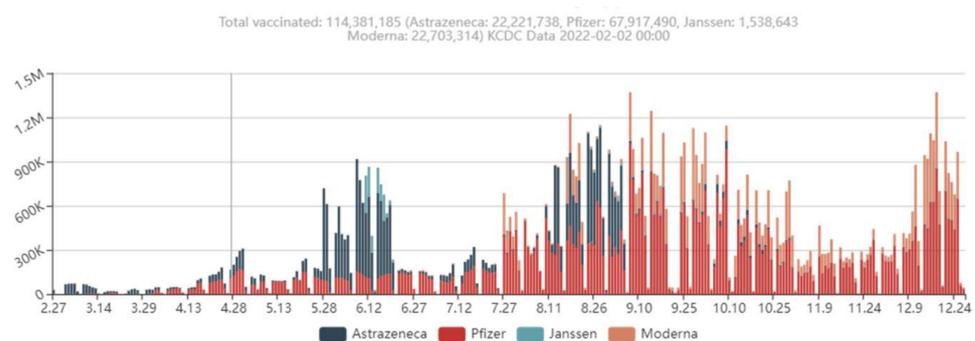


Figure 2. Extent of vaccination in South Korea by type of vaccine [1].

As shown in Figure 2, more and more people have been vaccinated over time, and some experience side effects from the vaccines. Side effects may affect the ability to perform daily activities. They may disappear within a few days, or cause an allergic reaction. Notably, vaccine-associated adverse reactions were more commonly reported in the AstraZeneca vaccination group than in the Pfizer vaccination group [4]. Above all, with the advent of new strains and breakthrough infections, many experts are uncertain as to when herd immunity can be reached. Rapid and scalable diagnostic testing technologies are required to solve this problem.

Research into diagnostic test technology has been actively conducted in many fields. The fact that the COVID-19 virus affects the human lungs has made it possible to study diagnostic testing techniques in a variety of ways. Studies have been conducted to distinguish between the lungs of COVID-19 positive people and the lungs of healthy people using lung CT scans and X-rays. These studies have more than 90% accuracy, and the results are constantly improving [5,6]. Diagnostic test studies using artificial intelligence (AI) to interpret audio signals generated from the human body, such as cough sounds, are ongoing [7–15]. The flowchart of the basic diagnostic process is shown in Figure 3.

As shown in Figure 3, cough sound data is generally collected using crowdsourcing. An AI model is then created, which analyzes and classifies the collected data and uses the results for diagnosis. Users of the application can record cough sounds, and be informed of the probability of being positive for COVID-19. Acoustic features are extracted from the cough sound data, and the model is trained based on these features. Diagnostic studies into the diagnosis of COVID-19 based on cough sounds have attempted to classify the COVID-19 cough sounds based on preliminary studies involving diagnosing asthma [12,16],

pneumonia [17], or Alzheimer’s disease [18] using cough sounds. The success of this work indicates that there are specific characteristics of cough produced by the way COVID-19 affects the respiratory system.

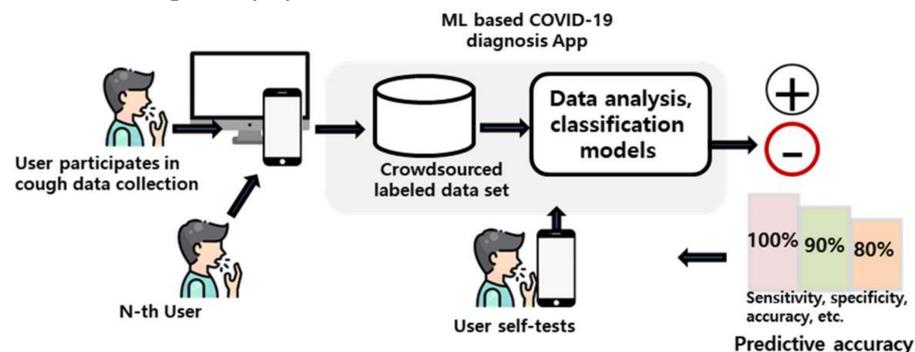


Figure 3. COVID-19 basic diagnostic flow using the app.

In this paper, we propose a method of diagnosing COVID-19-infected people using cough sound data. To extract more information from the cough sound data, we created a feature set by extracting 36 audio features. As well as mel-frequency cepstral coefficients (MFCC) [19] and spectrograms, which have been widely used in the past, spectral centroids, spectral bandwidth, seven spectral contrast features, spectral flatness, spectral roll-off, and 12 chroma features were extracted from cough sound data. Then, using a model based on [13], the spectrogram image was presented as an input to ResNet-50, the feature set was presented as an input to DNN, and then the output values of each model were connected to derive a result. In comparison with other methods, the method proposed in this paper showed the highest performance, with a sensitivity of 93% and a specificity of 94%.

The paper is structured as follows. Section 2 introduces related work and Section 3 describes the database and the proposed method. Section 4 describes the experiment method and compares the performance with existing methods to evaluate the proposed method. Section 5 ends with conclusions.

2. Related work

2.1. Cough Data

There have been several studies which have collected cough data from COVID-19-positive patients. The Cambridge Dataset [12], Coswara [20] and COUGHVID [21] are accessible means of data. For the diagnosis of COVID-19 using cough sound, it is important to collect a large amount of data to increase the diagnostic accuracy. However, data have mainly been collected through crowdsourcing, and through applications, as contacting COVID-19-positive patients is difficult. The cough sound data were recorded using built-in microphones of phones or computers. Precautions were also required with respect to security, since the collected data included personal data such as the region, gender, and respiratory disease status of the people who provided the data, as well as the cough sound.

The Cambridge Dataset [12] was created using web applications and Android apps to collect data using crowdsourcing. Users enter their symptoms, record a cough three times, record a breathing sound three times, and enter their COVID-19 test results. Because the data contains user information, a unique ID is created and stored, without collecting personal identifiers or users’ e-mail addresses. Healthy data were defined as the data of users who did not have a history of smoking, had not tested positive for COVID-19, and did not have any symptoms. Among the first data shared by the Cambridge researchers, 93 people were healthy and 46 people were COVID-19 positive. The second dataset shared by the Cambridge researchers was divided into train, dev, and test data. The dataset contained data from 725 individuals, 567 of whom were healthy, and 158 who were COVID-19 positive. The length of the recording was at least 2 to 5 s, and the sampling frequency was 16 kHz.

Coswara [20] was created in India, and collected breathing sounds, cough sounds, and the voices of healthy and unhealthy people using a website application to diagnose COVID-19. For the breathing and cough sounds, shallow sounds and deep sounds were collected, respectively. The pronunciation sounds of three vowels ('eu', 'i', 'u'), and the voices used when counting numbers from 1 to 20 were collected, and the sounds pronounced at normal speed and at high speed were collected. Nine audio files were recorded per participant, and the number of participants in the currently published dataset was 2131 as of 14 September 2021. Among them, 1372 people were healthy, 314 people had respiratory disease, and 99 people were completely recovered. A total of 346 people tested positive for COVID-19, of which 72 people had severe symptoms and 231 people had mild symptoms. In addition, there were 42 asymptomatic patients. The length of the recording was at least 2 to 5 s, and the sampling frequency was 44.1 kHz.

COUGHVID [21] collected cough sounds from 1 April 2020 to 10 September 2020 using a web application. When the user finished recording the cough sound, the user's age, gender, and current condition were entered. The status of the cough sound data was recorded as COVID-19, symptomatic, or healthy. All data in COUGHVID are in .webm or .ogg format, with a sampling frequency of 48 kHz. The length of the recording was at least 2 to 9 s. There are more than 20,000 pieces of data, and the datasets were filtered using a cough detection algorithm. As there is a risk of collecting the wrong sample using crowdsourced data, the COUGHVID team developed a classifier which analyzes the score of cough sound detection in the data, so that non-cough data can be excluded. The cough sound score is included in the metadata as an item called cough_detected.

2.2. Cough Testing

Brown et al. [12] collected cough and breath sound data, as described in Section 2.1, to distinguish between the cough sounds of COVID-19-positive individuals and those of healthy people. Brown et al. trained a classification model with 477 handcrafted features and features extracted using VGGish [22]. The handcrafted features consisted of features extracted at the segment level and features extracted at the frame level. Duration, onset, tempo, and period were extracted at the segment level, and RMS energy, spectral centroid, roll-off frequency, zero-crossing, MFCC, Δ -MFCC, and Δ^2 -MFCC were extracted at frame level. Therefore, a total of 477 handcrafted features were extracted to train the classification model. The researchers also used VGGish to extract features. With a sampling frequency of 16 kHz, a pre-trained VGGish model returns 128-dimensional features every 0.96 s. The researchers trained a classification model by extracting 256 features using VGGish. The classification model used a support vector machine. In an experiment by Brown et al. extracting and training the VGGish feature and segment level features using only cough sound data produced an area under the receiver operating curve (AUC) of 0.82 and a sensitivity of 0.72. These were the best results of their experiments.

Ahmed et al. [13] used the COUGHVID dataset to identify the cough sound of a person who was COVID-19 positive. They extracted MFCC and spectrogram images from the data to train a multi-branch network. The spectrogram images were input into a ResNet50 model, and the 13 MFCCs extracted were input to the fully connected layer. Clinical features such as fever symptoms and respiratory diseases were input to the fully connected layer, and combined with the MFCC model. Then, the model was combined with the spectrogram model and trained. The multi-branch network had a sensitivity of 85% and a specificity of 99.2%. The model excluding clinical features had a sensitivity of 93% and a specificity of 86%.

3. Proposed Method

3.1. Data

The experimental data used in this paper came from COUGHVID [21]. Crowdsourced audio data usually contains unnecessary content. Therefore, Lara et al. [21] developed a classifier that analyzes the score at which cough sounds are detected. The cough score for

each .wav file is included as metadata. Of the cough sound data, the total number of data with a cough score of 0.9 or higher was 6092. Table 1 shows the number of data with a cough sound detection score of 0.8 or more, and the number of data with a cough sound detection score of 0.9 or more in COUGHVID.

Table 1. Number of data having a coughing sound detection score of 0.8 or higher, and 0.9 or higher in COUGHVID.

	Cough_Detection_Score \geq 0.8	Cough_Detection_Score \geq 0.9
Healthy	5608	4702
Symptomatic	1135	949
COVID-19	547	441
Total	8295	6092

There were 5608 cough sound data from healthy people with a cough detection score of 0.8 or higher, 1135 cough sound data from symptomatic people, and 547 cough sound data from COVID-19-positive people. In this study, we used only cough data with a cough detection score of 0.9 or higher, to minimize noise and train the model with more accurate cough sounds. Therefore, we used 4702 cough sound data for healthy people, 949 cough sound data for symptomatic people, and 441 cough sound data for COVID-19 positive people.

3.2. Preprocessing

Cough sound data included unnecessary sounds between coughs, and the number of coughs varied between recordings. These inconsistencies could reduce the performance of the model, so the process of segmenting only the cough was essential. In this study, the cough sound was segmented from the cough sound data using a method published by Lara et al. [21].

Figure 4 is an example of the original cough sound data. When we actually listened to the original data, the first part of the data was a coughing sound, and the last part of the data was a small cough sound mixed with noise. Therefore, only the first part of this data should be segmented and used. In this case, the sampling frequency was set to 24,000 Hz, the minimum length of the cough sound was set to 200 ms, and a sample signal of 200 ms length was added before and after the cough was detected. Here, the sample signal was the number of seconds added to the beginning and end of each detected cough to make sure the coughs were not cut short.

In the part recognized as a cough sound, there were cases where the coughing sound was recorded once as “Cough!”, or twice as “Cough! Cough!”, and so was segmented based on the coughing sound, and used as shown in Figure 5.

3.3. Audio Features

The feature set was created by extracting audio features from an audio chunk file by cutting only the cough from the original data. MFCC is one of the most commonly used features in the field of speech recognition. It has been widely used in studies into the diagnosis of COVID-19 from cough sounds [11,13–15,23]. Spectrograms were also frequently used in this study, and were considered necessary for high accuracy. In this study, in addition to the MFCC and spectrograms used primarily in existing papers, we added the following audio features:

- 13 MFCCs;
- 5 spectral features: spectral centroid, spectral bandwidth, 7 spectral contrast features, spectral flatness, and spectral roll-off;
- 12 chroma features: 12-dimensional chroma vector.

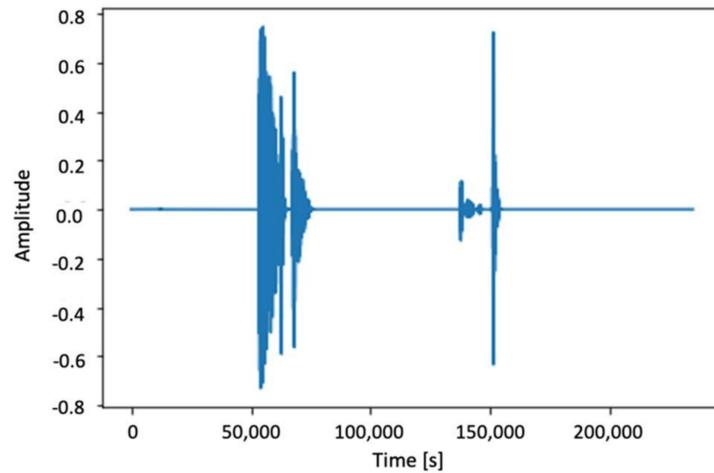


Figure 4. An example of the original data.

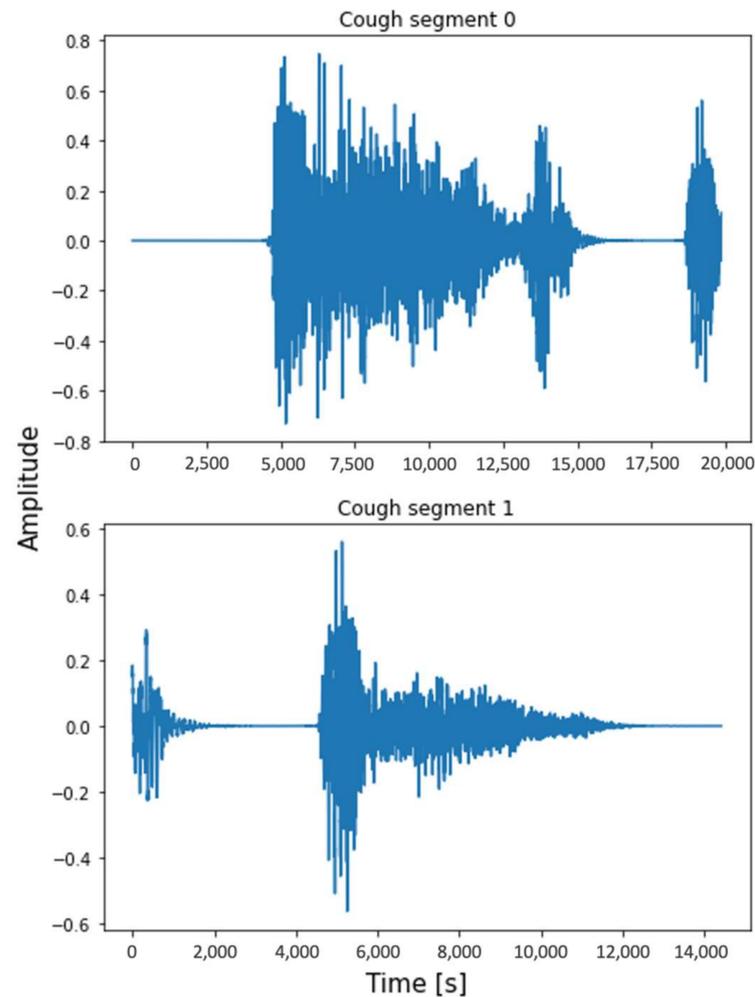


Figure 5. Segmented cough sounds.

Both the spectral features and the chroma features have been widely used as effective features in studies related to speech. Therefore, the feature sets were combined by selecting features based on preliminary research. All features were extracted using the librosa [24] package with a sampling frequency of 24 kHz. In addition, 36-dimensional feature sets were constructed using the average value of each extracted feature. Figure 6 shows spectrogram

images extracted from the segmented cough sound data of male and female participants who were positive and negative for COVID-19.

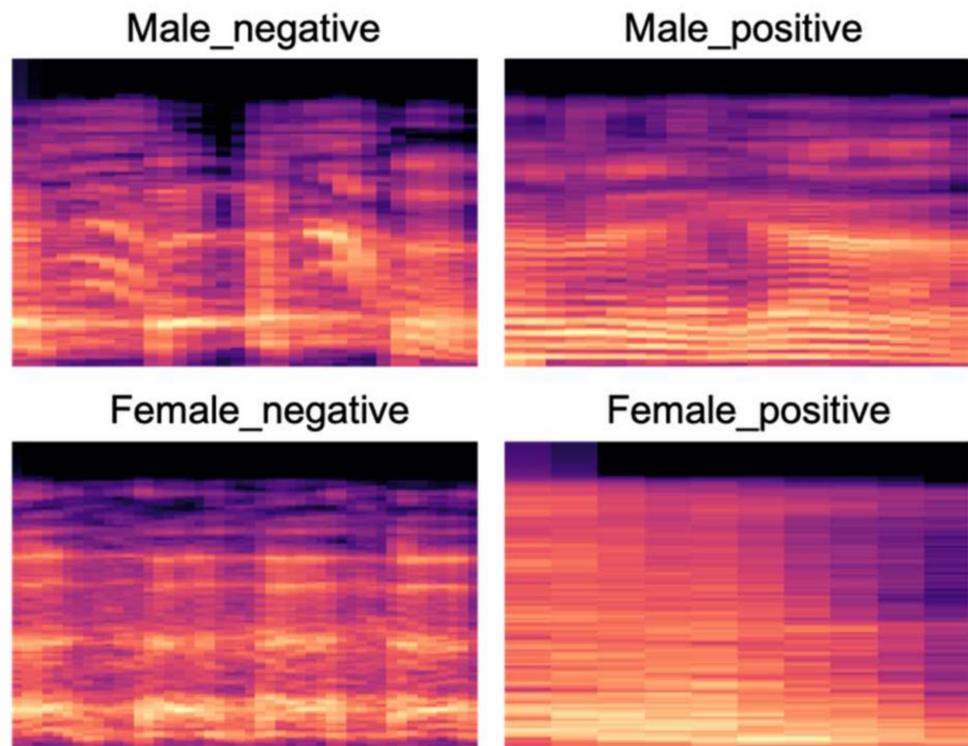


Figure 6. Spectrogram images extracted from the segmented cough sound data of male and female participants who were positive and negative for COVID-19.

3.4. Model

The model in this study used a combination of ResNet-50 [25] and a deep neural network (DNN) to distinguish between the cough sounds of COVID-19 positive people and the cough sounds of healthy people. This model was constructed based on the model proposed in [13]. The ResNet-50 was trained with spectrogram images (224, 224, 3) extracted from the audio chunk file. The ResNet-50 model was divided into a Global Average Pooling layer and a Global Max Pooling layer, and was reconnected after performing batch normalization and dropout. The DNN was trained with the 36-dimensional feature set configured in this study as an input. It was divided into two layers of 256 node layers and two layers of 64 node layers, respectively, and was connected after dropout was performed on each layer. GlorotUniform was used for the kernel initializer of each layer and Relu was used for the active function. In all models, the dropout was 0.5. The output values from ResNet-50 and DNN were connected to each other (Figure 7).

The values output by the ResNet-50 and DNN were connected. Then, an output value was calculated using the sigmoid function after passing through the dense layer, batch normalization layer, and dropout layer. Using this value, the cough sound of COVID-19-positive people was distinguished from the cough sound of healthy people.

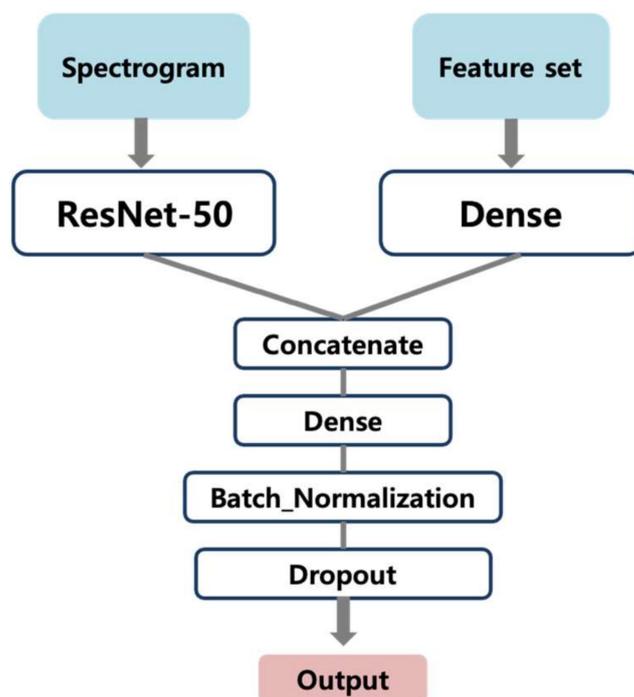


Figure 7. Structure of the model combining a ResNet-50 and a DNN.

4. Experiments

4.1. Experimental Design

In this study, we used only cough sound data to distinguish between the cough sound of COVID-19 positive people and the cough sound of healthy people. To achieve this aim, a more precise preprocessing method was added, and a new feature set was constructed. Table 2 shows the databases used in the experiment, the number of cough sound data points of COVID-19 positive people and healthy people in each database, and the number of segmented cough sounds.

Table 2. Counts of original data in each database and the number of segmented cough data.

Database	Negative	Positive	Negative Segmented File	Positive Segmented File
COUGHVID [21]	5651	441	11,981	1140
Cambridge [12]	660	204	1634	586
Coswara [20]	1372	303	2353	448

The data had an imbalance between negative and positive data, as shown in Table 2. So, only 1200, 1000, and 1000 of the negative segmented data in each database in Table 2 were used, respectively. The experiments were designed to compare the accuracy, sensitivity, and specificity of each case. There were cases in which only MFCC was extracted from each of the three databases and analyzed with the long short-term memory (LSTM) [26] model. The second case was when only the spectrogram was extracted from COUGHVID, and analyzed only with the ResNet-50 model. The third case was where both MFCC and spectrogram data were extracted from COUGHVID and analyzed with the ResNet-50+DNN model. Then, in the fourth case, the ResNet-50+DNN model was trained by extracting a new 36-dimensional feature set and spectrogram from COUGHVID: the proposed method. For validation, the data were randomly grouped into training, validation, and testing sets in 70–15–15 splits. So, the number of train–validation–test datasets of COUGHVID were 9185, 1968, and 1968, respectively. The experiment was conducted in the same environment for each model, and the software and hardware specifications used are shown in Table 3.

Table 3. Specifications of software and hardware used in the experiment.

	Specifications
Operation system	Ubuntu 18.04 LTS
Tensorflow	2.4.1
Cuda	10.1
CPU	intel Core i7-4770
GPU	GeForce GTX 1080Ti × 1
RAM	16 GB

4.2. Results

Table 4 shows the sensitivity, specificity and accuracy of each row. In experiments (a) to (c), only MFCC was extracted from each database and analyzed with LSTM. Experiment (a) had a sensitivity and specificity of about 60% and 62%, respectively, and (b) had a sensitivity and specificity of about 77% and 65%, respectively. Experiment (c) had a sensitivity and specificity of about 71% and 76%, respectively, higher than (a) and (b). The lowest performance was found in (a), using relatively high-quality data. In addition, in (b) using the smallest number of data, the performance was relatively higher than that of (a) and (c). This fact suggests that more high-quality data was needed. Experiment (d) produced a diagnostic accuracy of about 88% to 90%, which was lower than (e) and (f). However, (d) showed that the ResNet-50 model was effective in diagnosing COVID-19 cough sounds. Experiment (e) was the model excluding the clinical feature in [13]. Here, the clinical features indicated whether the individual had fever symptoms or underlying respiratory illness. We extracted a new 36-dimensional feature set and spectrogram from [21] and trained the ResNet-50 + DNN model. It showed sensitivity and specificity of 93% and 94%, respectively, and had higher diagnostic accuracy than (e), even though the clinical features were not used.

Table 4. Performance comparison results of each study.

	Database	Features	Model	Performance		
				Accuracy	Sensitivity	Specificity
(a)	[21]	MFCC	LSTM	62%	60%	62%
(b)	[12]	MFCC	LSTM	71%	77%	65%
(c)	[20]	MFCC	LSTM	73%	71%	76%
(d)	[21]	Spectrogram	ResNet-50	88%	90%	88%
(e)	[21]	MFCC + spectrogram	ResNet-50 + DNN	89%	93%	86%
(f)	[21]	Proposed feature set + spectrogram	ResNet-50 + DNN	94%	93%	94%

For the experiment shown in Table 3, each ROC curve is as shown in Figure 8. The proposed method had an AUC value of 0.98 at the maximum area under the ROC curve as the red line (f). The blue line (e) had an AUC value of 0.96, which was lower than that of the proposed method. The orange line depicts method (d), and has an AUC value of about 0.95, lower than the red line (f) and the blue line (e). Below that, they show lower performance than the above three experimental results. From these results, it was found that the extracted feature set including the spectrograms and MFCC were important to distinguish the cough sound of COVID-19 positive people from the cough sound of healthy people. In addition, when the ResNet-50+DNN model was trained with this feature set, it showed higher diagnostic accuracy than [13].

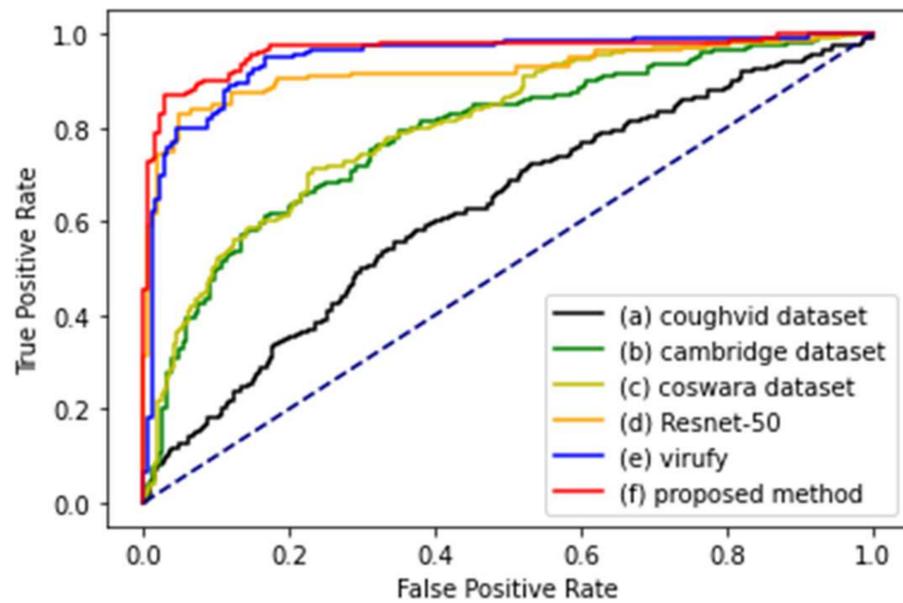


Figure 8. The ROC curve for each experiment.

5. Conclusions

COVID-19 affects the human respiratory system. Coughing is an audio signal that comes out of the body only after going through the respiratory system. Therefore, if the COVID-19 virus has affected the respiratory system, it will inevitably affect the audio signal generated. For this reason, COVID-19 diagnostic studies using coughing sound are being actively conducted.

In this study, we used only crowdsourced cough sound data to distinguish between the cough sound of COVID-19 positive people and that of healthy people. A new feature set based on features identified in previous studies was extracted to improve the performance of the analysis of cough sounds. If COVID-19 diagnosis is possible with only the cough sound, it will be possible to receive a faster diagnosis for prescreening purposes before undergoing the RT-PCR test. Depending on the diagnosis results, the user will be able to go to the hospital for an accurate diagnosis and minimize contact with the outside world. This would be able to help prevent infection of COVID-19.

We have created a model to diagnose the COVID-19 condition from cough sounds; however, before the implementation of machine learning algorithms, reliable data are needed for the generalization of the model [27]. Therefore, in order to ensure high performance in all locations and situations, ML solutions must be trained and tested with collected data from various people in various locations. Future research aims to find high-quality data collected from more diverse locations to compensate for this point. In addition, due to the inevitable lack of COVID-19-positive cough sound data, we will study the best way to combine the databases and how to combine models to compensate for this. Furthermore, the main challenge of clinical COVID-19 diagnosis is that the symptoms are similar to those of other common respiratory, lung and heart diseases [28]. Therefore, models should be tested to distinguish COVID-19 from other diseases, such as non-COVID-19 pneumonia, respiratory infections, asthma, and chronic lung disease exacerbations [29,30]. Therefore, we will conduct additional sub-analysis tests and conduct research on effective audio features. Using this approach, we hope to develop faster and larger-scale diagnostic technology that can be used by anyone with a smartphone or computer. In the future, it is expected that a technology for diagnosing other respiratory diseases using audio signals will be studied.

Author Contributions: Conceptualization, M.-J.S.; methodology, M.-J.S. and S.-P.L.; investigation, M.-J.S.; writing—original draft preparation, M.-J.S.; writing—review and editing, S.-P.L.; project administration, S.-P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Sangmyung University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Experiments used publicly available datasets.

Acknowledgments: This work was supported by the 2021 Research Grant from Sangmyung University.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. COVID-19 Dashboard. Available online: <https://coronaboard.kr/en/> (accessed on 2 February 2022).
2. Watson, J.; Whiting, P.F.; Brush, J.E. Interpreting a covid-19 test result. *BMJ* **2020**, *369*, m1808. [[CrossRef](#)] [[PubMed](#)]
3. Ondoa, P.; Kebede, Y.; Loembe, M.M.; Bhiman, J.N.; Tessema, S.K.; Sow, A.; Sall, A.A.; Nkengasong, J. COVID-19 testing in Africa: Lessons learnt. *Lancet Microbe* **2020**, *1*, e103–e104. [[CrossRef](#)]
4. Bae, S.; Lee, Y.W.; Lim, S.Y.; Lee, J.-H.; Lim, J.S.; Lee, S.; Park, S.; Kim, S.-K.; Lim, Y.-J.; Kim, E.O.; et al. Adverse Reactions Following the First Dose of ChAdOx1 nCoV-19 Vaccine and BNT162b2 Vaccine for Healthcare Workers in South Korea. *J. Korean Med. Sci.* **2021**, *36*, e115. [[CrossRef](#)] [[PubMed](#)]
5. Rajinikanth, V.; Dey, N.; Joseph, A.N.; Hassanien, R.A.E.; Santosh, K.C.; Raja, N. Harmony-search and otsu based system for coronavirus disease (COVID-19) detection using lung ct scan images. *arXiv* **2020**, arXiv:2004.03431v1.
6. Yildirim, M.; Cinar, A. A Deep Learning Based Hybrid Approach for COVID-19 Disease Detections. *Trait. Signal* **2020**, *37*, 461–468. [[CrossRef](#)]
7. Mouawad, P.; Dubnov, T.; Dubnov, S. Robust Detection of COVID-19 in Cough Sounds. *SN Comput. Sci.* **2021**, *2*, 34. [[CrossRef](#)]
8. Imran, A.; Posokhova, I.; Qureshi, H.N.; Masood, U.; Riaz, M.S.; Ali, K.; John, C.N.; Hussain, I.; Nabeel, M. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Inform. Med. Unlocked* **2020**, *20*, 100378. [[CrossRef](#)]
9. Hassan, A.; Shahin, I.; Alsabek, M.B. COVID-19 Detection System using Recurrent Neural Networks. In Proceedings of the 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), Sharjah, United Arab Emirates, 3–5 November 2020; pp. 1–5. [[CrossRef](#)]
10. Nessiem, M.A.; Mohamed, M.M.; Coppock, H.; Gaskell, A.; Schuller, B.W. Detecting COVID-19 from Breathing and Coughing Sounds using Deep Neural Networks. In Proceedings of the 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), Aveiro, Portugal, 7–9 June 2021; pp. 183–188. [[CrossRef](#)]
11. Laguarda, J.; Hueto, F.; Subirana, B. COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings. *IEEE Open J. Eng. Med. Biol.* **2020**, *1*, 275–281. [[CrossRef](#)]
12. Brown, C.; Chauhan, J.; Grammenos, A.; Han, J.; Hasthanasombat, A.; Spathis, D.; Xia, T.; Cicuta, P.; Mascolo, C. 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, Virtual Event, 6–10 July 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 3474–3484. [[CrossRef](#)]
13. Fakhry, A.; Jiang, X.; Xiao, J.; Chaudhari, G.; Han, A.; Khanzada, A. Virufy: A Multi-Branch Deep Learning Network for Automated Detection of COVID-19. *arXiv* **2021**, arXiv:2103.01806.
14. Bagad, P.; Dalmia, A.; Doshi, J.; Nagrani, A.; Bhamare, P.; Mahale, A.; Rane, S.; Agarwal, N.; Panicker, R. Cough Against COVID: Evidence of COVID-19 Signature in Cough Sounds. *arXiv* **2020**, arXiv:2009.08790v2. Available online: <https://arxiv.org/abs/2009.08790v2> (accessed on 19 April 2021).
15. Pahar, M.; Klopper, M.; Warren, R.; Niesler, T. COVID-19 Cough Classification using Machine Learning and Global Smartphone Recordings. *arXiv* **2020**, arXiv:2012.01926v2. Available online: <https://arxiv.org/abs/2012.01926v2> (accessed on 16 August 2021). [[CrossRef](#)] [[PubMed](#)]
16. Oletic, D.; Bilas, V. Energy-Efficient Respiratory Sounds Sensing for Personal Mobile Asthma Monitoring. *IEEE Sens. J.* **2016**, *16*, 8295–8303. [[CrossRef](#)]
17. Sotoudeh, H.; Tabatabaei, M.; Tasorian, B.; Tavakol, K.; Sotoudeh, E.; Moini, A. Artificial Intelligence Empowers Radiologists to Differentiate Pneumonia Induced by COVID-19 versus Influenza Viruses. *Acta Inform. Medica* **2020**, *28*, 190–195. [[CrossRef](#)] [[PubMed](#)]
18. Laguarda, J.; Hueto, F.; Rajasekaran, P.; Sarma, S.; Subirana, B. Longitudinal speech biomarkers for automated Alzheimer’s detection. *Front. Comput. Sci.* **2021**, *3*, 624694. [[CrossRef](#)]
19. Chatzarrin, H.; Arcelus, A.; Goubran, R.; Knoefel, F. Feature extraction for the differentiation of dry and wet cough sounds. In Proceedings of the 2011 IEEE International Symposium on Medical Measurements and Applications, Bari, Italy, 30–31 May 2011; pp. 162–166. [[CrossRef](#)]

20. Sharma, N.; Krishnan, P.; Kumar, R.; Ramoji, S.; Chetupalli, S.R.; Nirmala, R.; Ghosh, P.K.; Ganapathy, S. Coswara—A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis. *arXiv* **2020**, arXiv:2005.10548. Available online: <http://arxiv.org/abs/2005.10548> (accessed on 14 January 2021).
21. Orlandic, L.; Teijeiro, T.; Atienza, D. The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms. *Sci. Data* **2021**, *8*, 156. [[CrossRef](#)]
22. VGGish. Available online: <https://github.com/tensorflow/models/tree/master/research/audioset/vggish> (accessed on 14 September 2021).
23. Kaiming, H.; Xiangyu, Z.; Shaoqing, R.; Jian, S. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.0338.
24. Librosa Package. Available online: <https://librosa.github.io/librosa> (accessed on 30 August 2021).
25. Mermelstein, P. Distance measures for speech recognition, psychological and instrumental. In *Pattern Recognition and Artificial Intelligence*; Chen, C.H., Ed.; Academic Press: New York, NY, USA, 1976; pp. 374–388.
26. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
27. Najafabadi, M.M.; Villanustre, F.; Khoshgoftaar, T.M.; Seliya, N.; Wald, R.; Muharemagic, E. Deep learning applications and challenges in big data analytics. *J. Big Data* **2015**, *2*, 1. [[CrossRef](#)]
28. Sahu, K.K.; Mishra, A.K.; Martin, K.; Chastain, I. COVID-19 and clinical mimics. Correct diagnosis is the key to appropriate therapy. *Monaldi Arch. Chest Dis.* **2020**, *90*, 1327. [[CrossRef](#)]
29. Korpas, J.; Sadlonova, J.; Vrabec, M. Analysis of the Cough Sound: An Overview. *Pulm. Pharmacol.* **1996**, *9*, 261–268. [[CrossRef](#)] [[PubMed](#)]
30. Meister, J.; Nguyen, K.; Luo, Z. Audio feature ranking for sound-based COVID-19 patient detection. *arXiv* **2021**, arXiv:2104.07128v1. Available online: <https://arxiv.org/abs/2104.07128> (accessed on 5 February 2022).