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

Automatic and Robust Identification of Spontaneous Coughs from COVID-19 Patients

Michael J. Pettinati*, Xiyu Zhang, Ali Jalali, Kuldeep Singh Rajput, and Nandakumar Selvaraj

Abstract— **Cough is one of the most common symptoms of COVID-19. It is easily recorded using a smartphone for further analysis. This makes it a great way to track and possibly identify patients with COVID. In this paper, we present a deep learning-based algorithm to identify whether a patient's audio recording contains a cough for subsequent COVID screening. More generally, cough identification is valuable for the remote monitoring and tracking of infections and chronic conditions. Our algorithm is validated on our novel dataset in which COVID-19 patients were instructed to volunteer natural coughs. The validation dataset consists of real patient cough and no cough audio. It was supplemented by files without cough from publicly available datasets that had cough-like sounds including: throat clearing, snoring, etc. Our algorithm had an area under receiver operating characteristic curve statistic of 0.977 on a validation set when making a cough/no cough determination. The specificity and sensitivity of the model on a reserved test set, at a threshold set by the validation set, was 0.845 and 0.976. This algorithm serves as a fundamental step in a larger cascading process to monitor, extract, and analyze COVID-19 patient coughs to detect the patient's health status, symptoms, and potential for deterioration.**

I. INTRODUCTION

Coronavirus 2019 (COVID-19) has fundamentally altered our world. The way in which this highly infectious virus manifests is variable from patient to patient; however, one of the most common symptoms, present in over half of patients, is cough [1]. The commonness of cough, and the ease with which cough is recorded has prompted researchers to collect audio data in the hopes that it can be used to construct models to rapidly identify the presence of COVID-19 in patients. Such models would allow for patients to be more quickly tested and more easily remotely monitored as the society at large begins the process of reopening.

Models have already been developed from crowd-sourced data in which study participants provided forced coughs and self-reported their COVID-19 status via smartphone applications and web-based browsers (e.g. [2]–[4]). These papers achieve high area under receiver operating characteristic (ROC) curve statistics, particularly for people who indicate that they are not making a self-diagnosis but received a COVID-19 test. Certain models show a high true positive rate even in cases of asymptotic COVID-19 patients [3].

Researchers have also used these data to construct and validate models that identify if participants provided a cough or no cough audio [2]. There is an extensive body of literature on automated cough extraction from audio that is recorded using widely available smartphone microphones in various environments that shows high accuracy [5].

These apps differ from what is discussed below for a couple of reasons. Previous solutions for COVID-19 patients are crowd sourced and little can be verified about contributing participants (e.g. [2]). In some other cases, coughs are not from COVID-19 patients at all (e.g. [5]). Finally, the coughs collected from COVID-19 patients were often instructed to be forced rather than natural. Different types of coughs may encode different information about the patient.

This paper is the first step in a larger effort to extract cough from audio and use these data, potentially in conjunction with other data, to detect a patient's health status as well as monitor for and predict patient deterioration (see also: [6]). This work is valuable beyond tracking COVID-19 patients. Identifying cough using audio allows for more precise extraction and additional analysis that can help track infections and remotely monitor chronic conditions.

This particular undertaking is spurred by the use of Biofourmis' Biovitals Sentinel platform and patient application for the remote monitoring of COVID-19 patients [7]. An array of health parameters from COVID-19 patients have been gathered at sites around the world using this multifaceted platform. This work focuses on one piece of analyzing the volunteered coughs provided by these patients using the patient's smartphone application. This algorithm decides whether or not the patient has provided a cough in a recorded audio file and could prompt the patient to provide a more distinguishable cough if no cough was provided.

The following section describes the full data gathering process along with how a cough classification algorithm that is robust to background noise was developed and validated. The subsequent section presents the results of the validation work. Finally, the discussion details the importance of and need for this algorithm in remote COVID-patient monitoring as well as potential future directions for how this algorithm can be used different contexts.

II. METHODOLOGY

The objective of our cough classification algorithm is to determine whether or not the patient has provided at least a single cough in an audio file recorded by their smartphone to allow for additional analysis to take place in the case a cough has been provided. The first step is to gather data for the training and validation of this model. The following section describes the collection of a novel dataset containing coughs from COVID-19 patients [7].

Authors are with Biofourmis Inc, Boston, MA 02110, USA.
(*correspondence e-mail: michael.p@biofourmis.com)

A. Biovitals® Sentinel Platform - Data Collection

Biovitals® Sentinel (Biofourmis Inc, Boston, USA), a remote monitoring solution, has been deployed in countries across the world including: Singapore, Hong Kong, the United States, Australia, and the United Kingdom to help monitor patients who caregivers suspected to be suffering from COVID-19 or had confirmed COVID-19 [8]. This solution compiles both continuous and episodic data streams from these patients. There is continuous data available from the Everion device[1], which monitors numerous biomarkers including: typical vital signs, patient movement, e.g. steps, and derived measures of health and activity. In addition, there were episodic measurements taken by an attending caregiver, including, weight, blood pressure and temperature. Finally, there were episodic measures provided by a patient-facing smartphone application. During onboarding patients were given an introduction about how to use the application.

Using the patient-facing smartphone application, patients were able to provide their present symptoms, such as sore throat, cough, etc. They completed quality of life measures and recorded their coughs using the smartphone application. The application explicitly instructed patients to provide their coughs only if they were natural and not to force these coughs. This warning was given before every recording.

B. Datasets' split and annotations

The gathered cough data to validate our algorithm was manually scored by three researchers using the process enumerated below. Each patient provided audio file was coded as having a cough, no cough or exclude.

1) Two researchers independently indicated whether there was a cough in the file, there was not a cough in the file or they were not sure.
2) A third independent researcher listened to each of the files on which the first two scorers disagreed; the third scorer provided a cough, no cough, unsure rating.
3) If all three scorers disagreed, then the file was marked to be excluded. If at least two coders agreed, then that became the rating for the file (cough, no cough or unsure).
4) The original scorers made a forced cough, no cough choice on the files for which they were unsure.
5) In this case, if there was disagreement, the file was excluded. In the case of agreement, the file was given the agreed upon designation (cough or no cough)

The result of the data gathering and generation process was 181 files provided by 53 COVID-19 patients:

- 168 Cough Files
- 13 No Cough Files

The COVID patient cohort files are anticipated to have cough in most recordings due to the inherent study design and instructions for use targeted to capture natural spontaneous coughs in COVID patients that serve as true positives. This cough skewed in-house data, however, would not be sufficient enough to train algorithms to automatically recognize the coughs. This is a nontrivial task as shown by Fig. 1, which compares the waveforms and mel-frequency spectrograms of a COVID patients coughing to just a handful of non-cough sounds. Therefore, publicly available audio recordings have been included for the algorithm training and evaluation. The additional data included: ESC-50 [9], DCASE2016 [10], and Coswara [11], as well as public domain audio sources.

The researchers extracted and manually verified a subset of examples of cough and no cough files from these datasets and certain public domain audio websites for training, validation and testing purposes. The training set consisted of cough and no cough audio from numerous audio sources to ensure that the algorithm was learning features to differentiate cough and no cough, not differentiating audio sources.

The coughs in many of the publicly available datasets are very clean (minimal background noise) which could be far from real life recordings. Patients provided coughs in very different environments, sometimes these environments were very noisy. In order to make our algorithm more robust to noise, we up sampled the coughs in our training set by mixing the clean cough audio with audio of different background conversations, different music types, etc.

This up sampling procedure consisted of randomly selecting a subset of clean coughs from the publicly available cough audio in the training set. Segments of no cough files containing conversations and music of the length of the selected cough file were selected. The overlay function from Python's AudioSegment module[2] mixed the cough with no background noise audio with the audio from the files that did not have cough at one of three specified gains to simulate low, medium, and high noise environments.

Forty cough and forty no cough files were reserved to serve as a small validation set. This validation set was used to set the model's decision boundary before testing on reserved test set. The rest of the data was used to train the model. The number of files from each source in the training, validation and testing sets are given in Table 1.

C. DNN Cough Classifier Architecture

We used a sound event detection model to automatically extract features as input for our model (see [12]). The sound event detection model is precompiled and automatically extracts 1024 features that have been shown to meaningfully separate hundreds of different sounds, including cough, using a convolution neural network architecture. This architecture divides an audio file into segments. Mel-frequency spectrograms, visual representations of the frequency composition of these segments, are derived for each segment. These images are then passed through an extensive series of convolutional and max pooling neural network layers. The output of the final convolutional layer is 1024 output channels for each segment. These segment-level outputs are then combined into an audio file level output via averaging or taking the

[1] support.biofourmis.com/hc/en-us/categories/201377109-Everion-Device

[2] https://pypi.org/project/audiosegment/

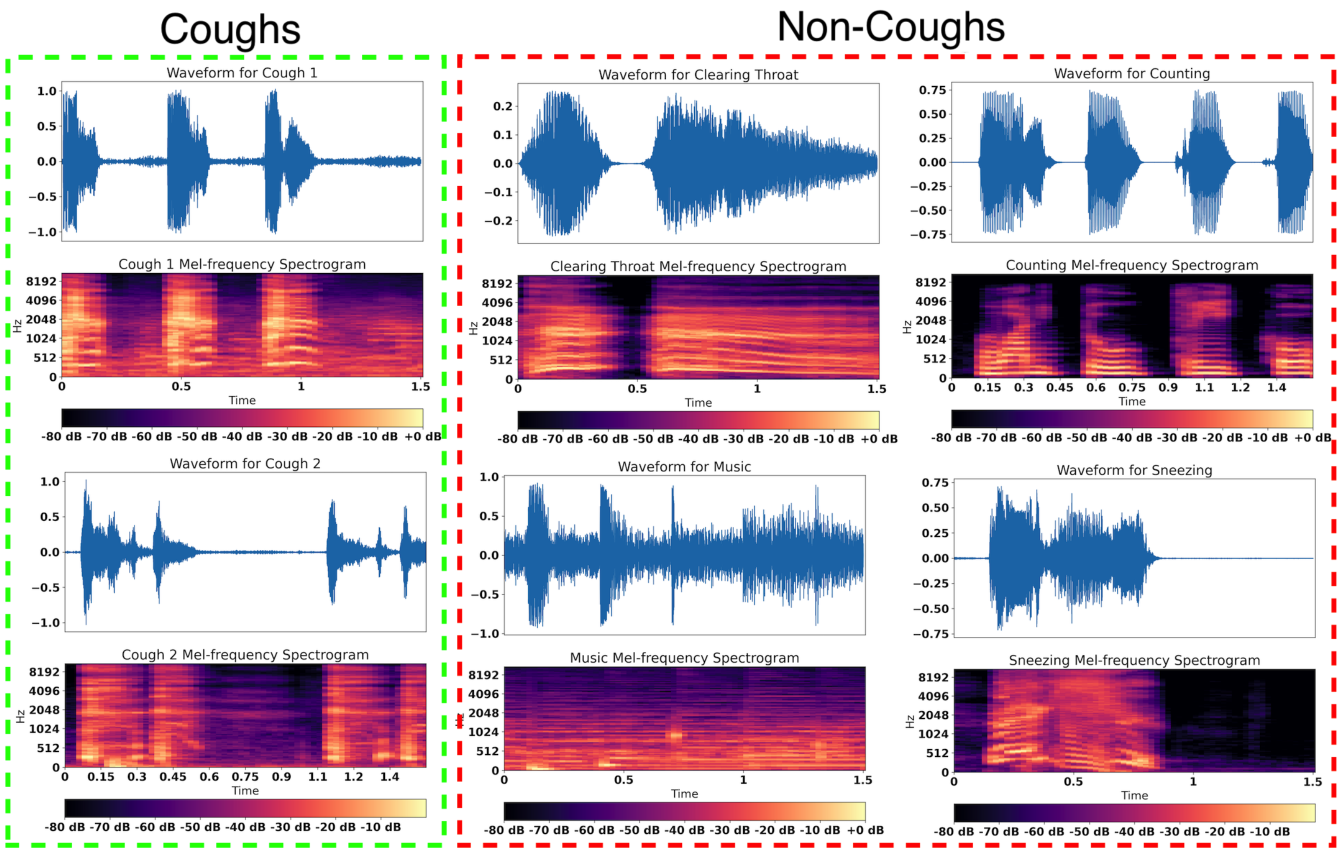

Fig. 1. Representative examples of cough (1st column) and non-cough (2nd and 3rd columns) audio recordings and their respective mel-frequency spectrograms illustrate unique time-frequency compositions and time-varying signal dynamics. Though the current datasets involve variety of non-cough sound recordings, a few non-cough sounds such as clearing throat, counting, sneezing and music are illustrated to showcase certain degree of resemblance and striking background noise that presents the challenge to the classification problem.

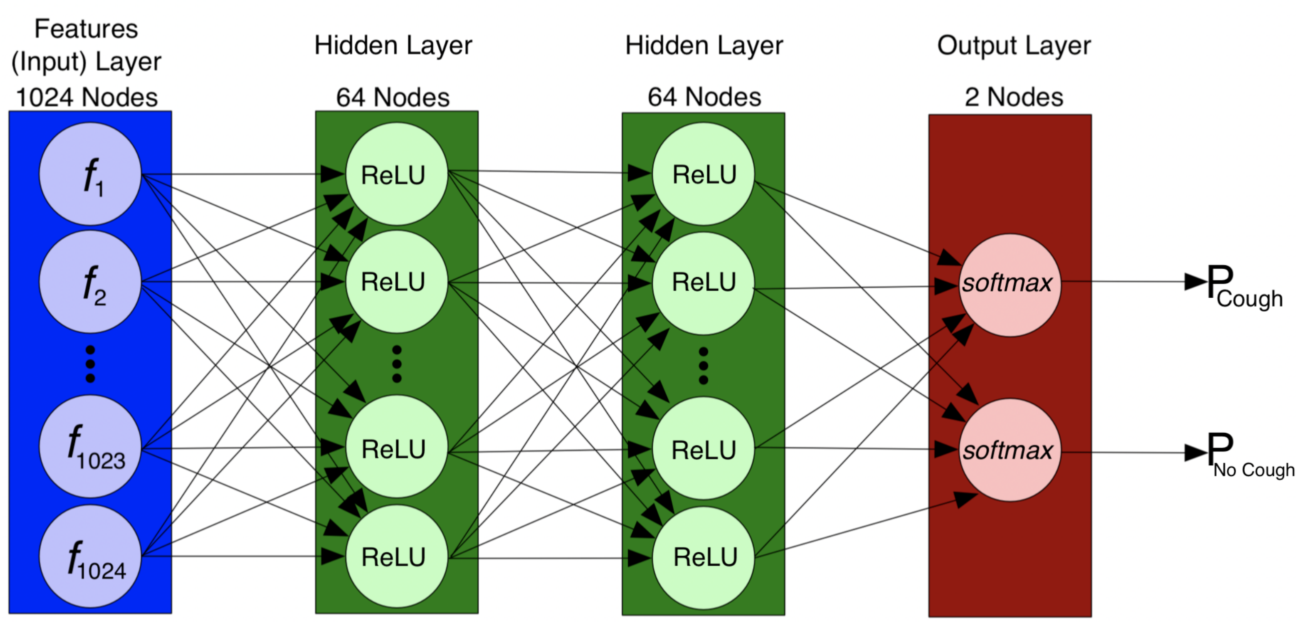

Fig. 2. The proposed deep neural network architecture consisting of an input layer of automatically extracted 1024 feature nodes, two hidden layers each of 64 ReLU nodes and an output layer for identifying cough versus no cough nodes, where all layers are fully connected.

TABLE I
TRAINING, VALIDATION AND TESTING SET SOURCES

File Type	# Cough Files	# No Cough Files
Training Set		
In-House Dataset	44	4
ESC-50	40	700
DCASE2016	20	140
Coswara	512	130
Public Domain	40	157
Engineered Cough+Noise	600	0
Total	1256	1131
Validation Set		
In-House Dataset	40	3
ESC-50	0	14
DCASE2016	0	7
Coswara	0	11
Public Domain	0	5
Total	40	40
Testing Set		
In-House Dataset	84	6
ESC-50	0	36
DCASE2016	0	18
Coswara	0	24
Total	84	84

maximum value. Convolutional neural network architectures have been shown to be superior to deep neural networks on certain tasks, including audio processing (see: [13], [14]).

Part of the work presented in the sound event detection model paper [12] showed how a classifier for a targeted task could be built on top of the automatically extracted 1024 features. Our model leverages this automated feature extraction and builds a deep neural network (DNN) classifier upon it. A deep neural network was employed because our 1024 features were automatically extracted, and this architecture exploits the tacit knowledge contained in these features to make the appropriate classification. Our targeted classification task is assessing whether or not a cough has been provided in recorded patient audio.

The DNN classifier (architecture shown in Fig. 2) takes as input the 1024 features automatically extracted from the raw patient audio. The classifier has two fully connected hidden layers of 64 nodes each which use rectified linear unit (ReLu) activation. Finally, it has a dense output layer with two nodes with softmax activation, the probabilities of the file containing the cough or no cough. This classifier was implemented in Keras[3].

D. Cough Classifier Validation Process

When analyzing coughs, it is important to understand the targeted population and cough type because these can change the mechanisms that produce the cough and the cough sound in important ways. Our targeted population was COVID-19 patients who had been asked to self record spontaneous coughs. This is a unique dataset gathered in COVID-19 patients using Biovitals® Sentinel Platform. Therefore, this Sentinel dataset was equally split and reserved for testing and training purposes.

Each training file had a label of cough or no cough. The 1024 features described above were derived for each of the audio training files and were used to train our DNN classifier. The process for generating the labels, the features and the datasets are described above.

[3]https://keras.io

The algorithm to test the trained model made a single cough/no cough prediction for each audio file in our validation and reserved test sets. To do this, the algorithm segmented the files into audio clips. The audio clips length ranged from 1.5 seconds to the length of file. The audio clip start times began at the beginning of the file with a moving step size of one second. The algorithm computed a probability of an audio clip having a cough for every clip in the file using our trained model. The audio clips included every combination of starting position and clip length. The probability of the file containing cough was the maximum probability computed from the different audio clips extracted from that file. The maximum value was chosen because certain files had lengthy periods of silences in them, so average or median probabilities of cough in the file may not have been informative. Further, as noted above, this algorithm serves as input to additional algorithms that process the audio and have the opportunity to rule out noncough audio ([6]). Therefore, the design of this algorithm prioritized sensitivity.

III. RESULTS

Consider the problem of identifying coughs from audio alone that is visualized in Figure 1. There are many common sounds: human, environmental, and synthetic that have similarly shaped audio signals to cough. Even the frequency decomposition of these signals do not seem to clearly separate the sounds from cough. The deep neural network architecture described above had to use the examples in the training set to learn subtle differences between cough and no cough raw audio signals and Mel-frequency spectrograms to successfully stratify them.

This learned model was then applied to a balanced represented validation set that was described above. The AUROC for the model on the validation set was 0.977. The ROC curve is shown in Fig. 3. The area under the precision and recall curve for the validation set is 0.978. This curve is shown in Fig. 4. This validation set was used to assess the optimal threshold when testing on our reserved test set.

The results for our cough identification model on this test set are summarized in Table 2.

The threshold determined by the validation set was the threshold for which the product of the specificity and sensitivity was highest on our validation set. This threshold was then applied to the probabilities generated by the model for the test set. The prediction algorithm correctly identified 84.5% true negatives (no cough files) and 97.6% true positives (cough files).

We also computed the statistics for the optimal threshold for the test set (the threshold which maximized the product of specificity and sensitivity on the reserved test set). The sensitivity, specificity and precision were all at or above 0.929.

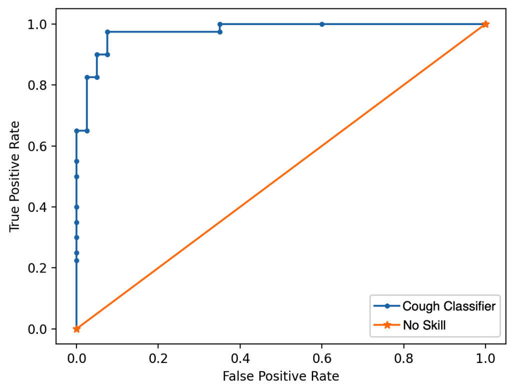

Fig. 3. ROC curve for the validation set. AUROC = 0.977

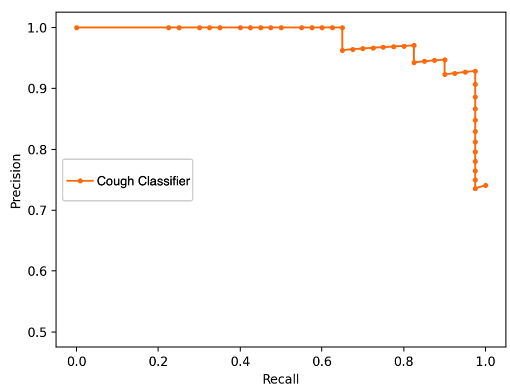

Fig. 4. Precision and recall curve for the validation set. AUPRC = 0.978

Our model's high precision and recall indicate a model that would perform well in real world settings. When a patient provides a cough in an audio recording, this model has a very high chance of recognizing an audio file with cough. The algorithm also can rule out a lot of audio in which the patient does not cough. Our test set consisted of many noises that could easily be confused as cough (e.g. throat clearing).

These results are important. This algorithm allows for additional processing to be done on the cough audio data under the assumption that the file being processed likely has a cough in it because of the low false positive rate. Subsequent algorithms analyzing the audio can be more precise in identifying conditions or physical qualities associated with cough when fed coughs only compared to algorithms fed a lot of confounding audio. The high true positive rate ensures that patient provided coughs are analyzed.

IV. DISCUSSION

We introduced an algorithm to identify whether or not COVID-19 patients provided a spontaneous cough using their smartphone application. The algorithm makes use of 1024 automatically derived features from a sound event detection CNN model that have been shown to successful separate hundreds of environmental sounds in previous work [12].

TABLE II
PERFORMANCE METRICS FOR TRAINED COUGH IDENTIFICATION ALGORITHM ON RESERVED TEST SET.

Threshold	Statistic Type	Value
Determined by Validation Set	Precision	0.863
	Specificity	0.845
	Sensitivity	0.976
Optimal	Precision	0.940
	Specificity	0.940
	Sensitivity	0.929

These features act as input to a new DNN classifier which was trained on our unique and carefully curated dataset.

The coughs in this test set were provided by patients who had confirmed COVID-19 patients or those suspected by professional caregivers. Patients were instructed to only provide natural coughs and not to force coughs. These factors make what was tested here unique from previous work; the previous work often was not done on confirmed COVID-19 patients and had patients forcing cough.

In a remote patient monitoring setting, this algorithm is important and necessary for a couple of reasons. First, it provides insights into whether patients are appropriately engaging with the patient-facing application and providing the data that are necessary to monitor their health statuses. If a patient is not providing cough, but providing lots of recordings, the patient may need assistance using the application. If the patient is not providing recordings, the patient may also need a caregiver to check in.

If the patient is not providing a cough, then additional computational resources should not be taken to perform additional processing, e.g. removing noise from the audio, extracting a cough, etc. This is the first step in a larger effort to use patient coughs as one of many potential ways to remotely monitor COVID-19 patients. This first step showed how in the presence of noise, when contending with cough-like no cough sounds, an automated algorithm could identify with high accuracy when a remotely monitored COVID-19 patient had provided cough audio.

Once a cough has been identified by this algorithm, noise from the audio must be removed and the cough must be extracted. The algorithm to carryout this extraction has been completed and validated (see [6]). Subsequently, the cough audio can be analyzed for the purposes of diagnosis, concurring symptom prediction, deterioration prediction, etc.

This paper presented an already widely distributed remote patient monitoring system that is gathering a new type of data. Patients' coughs were gathered along with numerous other self-reported and automatically monitored health parameters. Patients were explicitly told to provide natural coughs, which encode different information than forced coughs. This new data, which we have shown coughs can reliably be identified in, may reveal new biomarkers for symptoms or deterioration in COVID-19 patients.

Coughs can be differentiated by whether they are reflexive due to an environmental stimulus, natural due to an underlying infection, or forced; these different types of coughs activate different musculature and make different sounds [15].Beyond the cough type, studies have shown how cough sound, even reflexive cough sounds, can be changed through volitional control [16]. The method by which the cough is collected and the cough type must be considered in order to improve robustness of preceding algorithms and develop new algorithms. These factors change the model and could reveal different information about the patient.

As society begins to return to a new normal, caregivers must be vigilant about understanding who has COVID-19 or one of its many variants and limiting exposure to those patients while simultaneously providing them support. COVID-19 has highlighted the need for telemedicine in our healthcare systems. It affords a more cost effective and in many cases safer means for many patient-client interactions.

A remote patient monitoring (RPM) platform affords this opportunity. Patients can record their spontaneous coughs, voice, symptoms; they also can be tracked by wearable devices. This will allow caregivers to understand and monitor the patient's health status without the need to be physically present with the patient risking the opportunity for infection. This paper introduced a unique RPM that provides this coverage and is unique within the COVID-19 research space.

The algorithm introduced as part of this paper provides the opportunity to recognize when the patient is spontaneously coughing. The algorithms described in our concurrent submission [6] provide the opportunity for the extraction and analysis of cough to detect the presence of COVID-19. This can help caregivers monitor the patient to ensure patient illness is caught as early as possible and monitored as closely as possible, so caregivers are able to provide the optimal care for the patient and patients are able to limit exposure to others. The proposed remote patient monitoring solution is an important opportunity to help in the world's return to the new normal.

REFERENCES

[1] J. Shuja, E. Alanazi, W. Alasmary, and A. Alashaikh, "Covid-19 open source data sets: A comprehensive survey," *Applied Intelligence*, pp. 1–30, 2020.

[2] A. Imran, I. Posokhova, H. N. Qureshi, U. Masood, S. Riaz, K. Ali, C. N. John, and M. Nabeel, "Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app," *arXiv preprint arXiv:2004.01275*, 2020.

[3] J. Laguarta, F. Hueto, and B. Subirana, "Covid-19 artificial intelligence diagnosis using only cough recordings," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 1, pp. 275–281, 2020.

[4] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "Covid-19 cough classification using machine learning and global smartphone recordings," *arXiv preprint arXiv:2012.01926*, 2020.

[5] E. C. Larson, T. Lee, S. Liu, M. Rosenfeld, and S. N. Patel, "Accurate and privacy preserving cough sensing using a low-cost microphone," in *Proceedings of the 13th international conference on Ubiquitous computing*, 2011, pp. 375–384.

[6] X. Zhang, M. J. Pettinati, A. Jalali, K. S. Rajput, and N. Selvaraj, "Novel covid-19 screening using cough recordings of a mobile patient monitoring system," *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2021, accepted.

[7] C. K. Wong, D. T. Y. Ho, A. R. Tam, M. Zhou, Y. M. Lau, M. O. Y. Tang, R. C. F. Tong, K. S. Rajput, G. Chen, S. C. Chan *et al.*, "Artificial intelligence mobile health platform for early detection of covid-19 in quarantine subjects using a wearable biosensor: protocol for a randomised controlled trial," *BMJ open*, vol. 10, no. 7, p. e038555, 2020.

[8] G. Lee, "Biofourmis deploys biovitals sentinel platform in singapore," *BioWorld RSS*, Jul 2020, https://www.bioworld.com/articles/496399-biofourmis-deploys-biovitals-sentinel-platform-in-singapore.

[9] K. J. Piczak, "Esc: Dataset for environmental sound classification," in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015, pp. 1015–1018.

[10] A. Mesaros, T. Heittola, E. Benetos, P. Foster, M. Lagrange, T. Virtanen, and M. D. Plumbley, "Detection and classification of acoustic scenes and events: Outcome of the dcase 2016 challenge," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 2, pp. 379–393, 2018.

[11] N. Sharma, P. Krishnan, R. Kumar, S. Ramoji, S. R. Chetupalli, P. K. Ghosh, S. Ganapathy *et al.*, "Coswara–a database of breathing, cough, and voice sounds for covid-19 diagnosis," *arXiv preprint arXiv:2005.10548*, 2020.

[12] A. Kumar, M. Khadkevich, and C. Fügen, "Knowledge transfer from weakly labeled audio using convolutional neural network for sound events and scenes," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 326–330.

[13] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*. Ieee, 2017, pp. 1–6.

[14] T. N. Sainath, A.-r. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for lvcsr," in *2013 IEEE international conference on acoustics, speech and signal processing*. IEEE, 2013, pp. 8614–8618.

[15] C. Magni, E. Chellini, F. Lavorini, G. A. Fontana, and J. Widdicombe, "Voluntary and reflex cough: similarities and differences," *Pulmonary pharmacology & therapeutics*, vol. 24, no. 3, pp. 308–311, 2011.

[16] K. W. Hegland, D. C. Bolser, and P. W. Davenport, "Volitional control of reflex cough," *Journal of Applied Physiology*, vol. 113, no. 1, pp. 39–46, 2012.