Harnessing the Power of Vocal Signals in COVID-19 Detection Utilizing Machine Learning

Conference Paper · December 2024

DOI: 10.1109/BigData62323.2024.10825050

4 authors, including:

Ajinkya Jadhav
Gannon University

Richard Matovu
Gannon University

Harnessing the Power of Vocal Signals in COVID-19 Detection Utilizing Machine Learning

Aleesa Mann**, Ajinkya P. Jadhav†*, Richard Matovu†, Vibhuti Gupta*
*Department of Computer Science and Data Science, School of Applied Computational Sciences,
Meharry Medical College, Nashville, TN, USA
Email: amann22@mmc.edu, vgupta@mmc.edu
†Department of Computing and Information Science, Gannon University, Erie, PA, USA
Email: jadhav003@gannon.edu, matovu001@gannon.edu

***Abstract*—The global COVID-19 pandemic has strained healthcare systems and highlighted the need for accessible and efficient diagnostic methods. Traditional diagnostic tools, such as nasal swabs and biosensors, while accurate, pose significant logistical challenges and high costs, limiting their scalability. This paper explores an alternative, non-invasive approach to COVID-19 detection using machine learning algorithms to analyze vocal patterns, particularly cough and breathing sounds. Leveraging a publicly available dataset, we developed machine learning models capable of classifying audio samples as COVID-19 positive or negative. Our models achieve an AUC of up to 85% and an F1-score of 81%, demonstrating the potential of machine learning in enabling rapid, cost-effective COVID-19 diagnosis. These findings suggest that audio-based diagnostics could be a practical and scalable solution, particularly in resource-limited settings where traditional methods are less feasible.**

***Index Terms*—Coronavirus, COVID-19, machine learning, coughs, breathing, vocal signal analysis, COVID-19 detection**

I. Introduction

Coronavirus disease 2019 (COVID-19) is an infectious illness caused by the Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) [1]. Since its emergence in late 2019, the virus has triggered one of the most devastating pandemics in modern history, leading to a global health crisis. With over 700 million confirmed cases and nearly 7 million deaths worldwide, the pandemic has overwhelmed healthcare systems, particularly in regions with limited resources [2] [3], [4] [5] [6]. The rapid spread of the virus and its high transmission rates have made it difficult for healthcare providers to meet the growing demand for timely testing and treatment. This has driven an urgent need for more efficient and accessible diagnostic solutions [7], [8].

While current diagnostic methods, such as nasal swabs, biosensors, and blood tests, are highly reliable and accurate, these methods are not without limitations. These methods are often associated with high costs, making them inaccessible to a large portion of the global population, especially in low-income regions [9] [10] [11] [12]. Additionally, the logistical processes involved in collecting, handling, and transporting biological samples are time-consuming, further delaying the diagnosis. These challenges underscore the necessity for alternative diagnostic tools that are both scalable and non-invasive, enabling more widespread testing to be conducted more quickly and affordably.

Machine learning (ML) offers a promising solution to these challenges by providing a non-invasive approach to COVID-19 detection through the analysis of vocal signals. As COVID-19 primarily affects the respiratory system [13], it alters vocal and respiratory sound patterns, including coughing, which can be captured and analyzed for diagnostic purposes [12], [14]. Coughs, one of the most common symptoms of COVID-19, present a valuable and accessible source of data for machine learning models. By analyzing the frequency, amplitude, and other acoustic features of these cough sounds, ML-driven tools can detect subtle changes indicative of infection, providing a more rapid and cost-effective alternative to traditional methods [11], [12].

The development of machine learning algorithms capable of analyzing both respiratory and non-respiratory sounds introduces an innovative approach to diagnosing COVID-19. These algorithms, when trained on cough and other audio datasets, have demonstrated high accuracy in classifying audio samples as either COVID-19 positive or negative [14], [15]. This technology's scalability and non-invasive nature make it particularly valuable for large-scale screening, especially in low-resource settings where traditional testing methods may not be feasible.

This paper explores the design and evaluation of machine learning algorithms for detecting COVID-19 statuses using sound patterns. We make the following contributions:

1) *Design and Evaluation of machine learning models for COVID-19 detection:* Utilizing publicly available datasets, we have developed various machine learning models that non-invasively classify audio samples as either COVID-19 positive or negative. Our models have achieved an AUC of up to 85% and an F1-score of 81%. These findings show the potential of machine learning to provide rapid, cost-effective COVID-19 diagnosis, offering a practical alternative to traditional, resource-intensive testing methods like nasal swabs and biosensors.
2) *Comprehensive evaluation of developed models across various audio scenarios:* We have assessed the performance of our machine learning models under two

* The first and second authors contributed equally to this work.

practical scenarios: (1) using entire audio files, which includes background noise and silences, and (2) using segmented data after segmenting the input audio file that focuses specifically on certain cough or breathing patterns. This dual approach evaluates the models' effectiveness in both real-time, noisy environments and more controlled settings. The results provide valuable insights into potential adaptations of the models for telemedicine and large-scale health monitoring systems.

II. Related Work

In the wake of the COVID-19 pandemic, several studies have been conducted to explore the potential of artificial intelligence to detect COVID-19 by analyzing cough audio [16]–[26]. Many studies collected the data through mobile phone apps and developed a unique dataset for analysis through machine learning algorithms. Authors in [16] and [17] developed a mobile application to collect breath, voice, and cough data and apply ML algorithms to detect COVID-19 with symptoms and COVID-19 without symptoms. Many studies [18]–[20] used crowdsourcing to collect data for COVID-19 however the analytics methods differ. The authors in [18] employed both traditional and deep learning techniques to detect COVID-19 using voice, cough, and breathing sounds. In contrast, [19] focused solely on cough and breathing sounds to differentiate COVID-19 from other respiratory conditions. [20] developed a generalized AI model to detect COVID-19 using cough samples only which predicts accurately when applied on Latin America and South Asia clinical samples. An interpretable AI model is developed using cough sound in [21]. Overall, these studies demonstrate the potential of AI in COVID-19 detection.

Common techniques in this area of research include extracting features such as Mel-frequency cepstral coefficients (MFCCs) and utilizing neural network classifiers like Convolutional Neural Networks (CNNs) [27]. Other approaches, such as transfer learning [23] [17], are also frequently employed. Challenges of working with poor audio quality [28], unbalanced datasets and insufficient evaluation strategies can lead to an over-optimistic assessment of model performance [24], but preprocessing techniques such as segmenting cough clusters [29], oversampling from minority classes [30] and frequency filtering techniques [26] can improve model performance. The results from these experiments are promising, indicating that COVID-19 exhibits a distinct pathophysiology that can aid in virus detection [17]. Additionally, in certain instances, AI-assisted diagnostic tools have been shown to enhance the testing capacity of healthcare systems by up to 43% [22]. However, many of these papers emphasize the need for clinical validation of their proposed methods [17]. In summary, the analysis of audio data from voice, cough, and breathing is challenging and requires rigorous clinical validation and interpretability to understand the features and outcomes. This paper analyzes the publicly available crowdsourced dataset for analysis of respiratory and non-respiratory audio signals using conventional machine learning approaches, along with rigorous performance evaluation across various audio scenarios.

III. Methodology

A. Dataset Description

Our study uses the Coswara dataset [2], a publicly available crowd-sourced dataset designed for the analysis of respiratory and non-respiratory audio signals, primarily aimed at detecting COVID-19 through sound. The dataset contains audio samples such as coughs, breathing patterns (both deep and shallow), and sustained vowel sounds, as well as speech recordings from 2,746 participants. These participants self-reported their health status, including whether they tested positive or negative for COVID-19, along with symptoms like cough, fever, or sore throat.

The Coswara dataset [2] includes several health status classifications. These classifications consist of healthy, no respiratory illness exposed, respiratory illness not identified, and recovered in full, which we grouped as COVID-19 negative, representing 1,984 individuals. Additionally, 81 participants were classified as under validation and were excluded from the study. The remaining categories — positive mild, positive moderate, and positive asymptomatic — were classified as COVID-19 positive, totaling 681 individuals.

B. Tools Used

We used the Python programming language for this work, with Jupyter Notebook as our coding environment. The Python library *librosa* was utilized for audio pre-processing and feature extraction, while Scikit-Learn was employed for building the machine learning models. Additionally, Pandas and Numpy were used for statistical analysis and data manipulation.

C. Dataset Preparation

To prepare our dataset for this study, we downloaded the dataset from the Coswara GitHub repository [31]. The dataset is organized by collection date folders, and within each collection date, there are subfolders corresponding to individual participants. For each participant, we accessed their CSV file located in their subfolder to examine the "COVID Status" column. Based on this status, we reassigned the original class labels: participants marked as "Healthy," "No Respiratory Illness Exposed," "Respiratory Illness Not Identified," and "Recovered in Full" were relabeled as 'Negative,' while those marked as "Positive Mild," "Positive Moderate," and "Positive Asymptomatic" were relabeled as 'Positive.' Entries labeled "Under Validation" were excluded from our dataset.

We then created two main directories, "COVID Positive" and "COVID Negative," to organize the cough and breathing audio files according to the relabeled COVID status. Within each category, we further divided the files into subfolders for "Cough Samples" and "Breathing Samples." Finally, a verification process was performed to ensure the correct classification and organization of all audio files within the dataset.

After the data reorganization, we performed a preprocessing step prior to feature extraction. During this step, we resampled the dataset to 22 kHz, normalized the audio data to reduce amplitude variability and pre-emphasized the audio signals using a first-order differencing filter. Audio files that were too short, too quiet, and potentially empty were dropped during this step. We then extracted features as described in the next section.

D. Exploratory Analysis of Audio Features for COVID-19 Status Classification

To explore the potential of utilizing audio recordings of coughing and breathing in our study, we examined both the time and frequency domain features of these sounds to determine if there are any discernible differences between positive and negative cases. Figure 1 showcases two examples of these audio recordings. The figure provides a comparative analysis of breathing patterns from individuals tested for COVID-19, with one testing positive and the other negative. Each subject's data is represented in two segments: the upper section shows the time-domain waveform, while the lower section displays the corresponding spectrogram of the audio.

For the COVID-positive subject, shown in Figure 1a, the waveform features notable spikes and variations in amplitude which are indicative of irregular breathing patterns — commonly associated with COVID-19. These irregularities are also reflected in the spectrogram as variations in intensity and color, particularly pronounced in the lower frequency bands, suggesting episodes of breathing difficulty. Conversely, the audio from the COVID-negative subject, depicted in Figure 1b, shows a more consistent and rhythmic breathing pattern, with minimal fluctuations in both the waveform and spectrogram.

This exploratory analysis not only facilitates the identification of potential differences in breathing sounds associated with COVID-19 but also suggests potential for a powerful, non-invasive tool for early COVID-19 screening. These preliminary findings are in line with earlier works, for example, in [2] [32], which demonstrate that spectrograms of breath sounds captured via smartphone can effectively distinguish between asthmatic conditions and those of healthy individuals through distinct patterns. Leveraging these insights, along with other time and frequency-domain features, we extracted features as described in the next section for our machine learning models.

E. Feature Extraction

Multiple audio handcrafted features were extracted from each audio sample to capture various characteristics of the sound signal, similar to the approach used in [19]. These features capture the temporal, spectral, and harmonic properties of the audio, enabling accurate classification of respiratory sounds.

#	Feature	Count
1.	Average, Standard Deviation, Zero Crossing Rate	3
2.	Spectral centroid, rolloff, contract, chroma stft	4
3.	The first 14 MFCC coefficients	14
4.	The first 14 Delta-MFCC coefficients	14
5.	The first 14 Delta-Delta MFCC coefficients	14

TABLE I: Features extracted from the cough sounds.

Table I summarizes the features extracted from our filtered cough sounds, using the Python audio and music processing package Librosa [33]. The features included statistical, spectral, and MFCC features that are popularly used for audio/sound classification. The feature extraction phase gave us a total of 49 features that were used as inputs to our models. This feature set was further standardized using z-scale to ensure features with different magnitudes don't disproportionately influence the models' classifications.

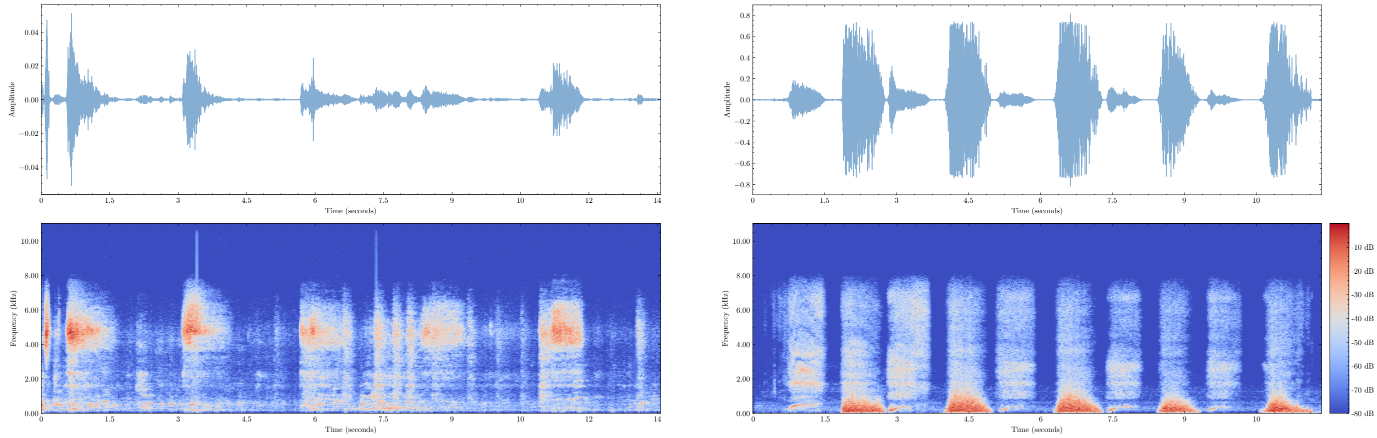
F. Model Training and Evaluation

During the initial exploration phase, we evaluated several machine learning algorithms, including k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Random Forest, and XGBoost. To efficiently search for optimal hyperparameters, we employed *RandomizedSearchCV*, which explored a range of hyperparameter combinations across the models. This method allowed for an efficient and comprehensive hyperparameter search by testing different configurations without exhaustive grid searches. A validation dataset comprising 40% of the original data was reserved for this search to support hyperparameter tuning and model comparison. In addition, K-fold cross-validation was applied to each model, splitting the dataset into multiple subsets for training and testing. This process reduced the likelihood of overfitting and provided a more reliable estimate of each model's generalization performance on unseen data.

Following this initial exploration, the two top-performing algorithms—Random Forest and SVM—were selected for final model training and evaluation. Random Forest performed best with the hyperparameters: *class_weight="balanced"*, *min_samples_leaf=2*, *min_samples_split=5*, and *n_estimators=200*, while SVM performed best with a radial basis function (RBF) kernel, degree=3, and class_weight="balanced". These parameters were specifically chosen to balance the class distribution and improve the robustness of the models, especially for imbalanced datasets, ensuring that both classifiers could effectively handle minority and majority classes in the COVID-19 detection task.

After selecting the classifiers, the models were trained and evaluated on the preprocessed dataset. The features were first standardized to have a mean of zero and a standard deviation of one, facilitating more consistent and efficient learning. Principal Component Analysis (PCA) was then applied to reduce the dataset's dimensionality while preserving 99% of its variance, removing noise and reducing computational complexity. To further address the issue of class imbalance, Synthetic Minority Oversampling Technique (SMOTE) was utilized, oversampling the minority class in the training set only. This process ensured that both classifiers learned from balanced data and avoided biases toward the majority class, leading to more equitable and accurate predictions.

To evaluate the performance of our models, we employed several widely-used metrics for imbalanced datasets, including area under the curve (AUC), precision, recall, and F1-score, given by the equations below:



(a) Breathing, COVID-positive.

(b) Breathing, COVID-negative.

Fig. 1: Time-domain waveform and spectrogram of breathing audio recordings from (a) a COVID-positive subject and (b) a COVID-negative subject.

$$\text{AUC} = \int_0^1 TPR(FPR)\, d(FPR) \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

where *TP* represents True Positives, the number of correctly predicted positive cases; *TN* represents True Negatives, the number of correctly predicted negative cases; *FP* represents False Positives, the number of negative cases incorrectly predicted as positive; and *FN* represents False Negatives, the number of positive cases incorrectly predicted as negative. For equation 1, *TPR* is the True Positive Rate while *FPR* is the False Positive Rate.

G. Evaluation Scenarios

To comprehensively evaluate the performance of our models, we consider two scenarios: (1) *using the whole audio file*, and (2) *using segmented data* as input to our models. These two approaches allow us to explore different aspects of the model's ability to process and classify cough sounds effectively in various practical situations.

In the *whole audio file* scenario, the entire recording — including coughs, pauses, and any background noise — is fed into the model for classification. This approach mirrors real-world conditions where users may submit continuous, unprocessed recordings. Evaluating the model's performance in this scenario is important because it tests the model's robustness to various environmental factors, such as noise and variability in the length and quality of the recordings.

In contrast, the *segmented data* scenario focuses on classifying specific portions of the audio, such as individual coughs, in real-time scenarios. This approach allows for quick, low-latency processing, which is crucial for applications requiring immediate responses, like telemedicine or remote health monitoring. It also optimizes resource usage, making the system more scalable and efficient, especially for low-power devices or large-scale implementations.

Under each of the above scenarios, we consider different audio sounds from the dataset to evaluate our models (1) breathing (deep and shallow) sounds only, (2) cough (deep and shallow) sounds only, and finally when both are merged, i.e., breathing and cough sounds. These further classifications of our dataset provide a more comprehensive evaluation.

IV. Results and Discussion

In this section, we present and discuss the performance evaluation of our models in classifying the different types of respiratory sounds across various scenarios. We have organized our results first by the type of sound analyzed, and then further categorized them based on whether segmented data or the entire audio file was utilized. This structure allows us to thoroughly analyze how each model performs in different conditions, providing insights into the model's capabilities in handling diverse types of respiratory sounds and input formats. Our results are based on the two best-performing machine learning algorithms for the scenario and dataset used for training and testing. For reference, the parameters set for these models are detailed in the section above.

A. Classification Results Using Breathing Sounds Only

In this subsection, we present the performance evaluation of our models trained and tested using only breathing sounds (both deep and shallow) from the dataset. This analysis focuses on assessing how effectively the models classify COVID-19 statuses based solely on breathing sound patterns. The top two best-performing classifiers in this scenario were *XGBoost* and *SVM*, thus reporting results from these two classifiers in this subsection.

Classifier	Score Type	AUC	Precision	Recall	F1-Score
SVM	Macro	77%	64%	71%	65%
	Weighted		84%	77%	80%
XGBoost	Macro	77%	68%	66%	67%
	Weighted		83%	84%	84%

TABLE II: Model performance for COVID-19 status classification using breathing sounds only from the entire audio file.

1) Classification Results Using Entire Audio Files of Breathing Sounds Only: In this subsection, we present the results when the models are trained and tested using the entire audio files, without segmentation. The performance of our top two best-performing classifiers, *XGBoost* and *SVM*, in classifying COVID-19 statuses using the complete audio files for breathing sounds only, is shown in Table II. The performance metrics are calculated for both macro and weighted averages. Macro averages treat each class equally, whereas weighted averages account for class imbalances present in the dataset, ensuring that classes with more instances are given greater weight. Our dataset is largely imbalanced with more negative samples than positive.

For the SVM, the AUC score was 77%. The macro average precision, recall, and F1-Score were 64%, 71%, and 65%, respectively. When considering weighted averages, SVM's precision increased to 84%, with a recall of 77% and an F1-Score of 80%. XGBoost also achieved an AUC score of 77%, with a macro precision of 68%, a recall of 66%, and an F1-Score of 67%. The weighted average for XGBoost showed an improved performance with a precision of 83%, a recall of 84%, and an F1-Score of 84%. This indicates that XGBoost was slightly more effective in handling class imbalances and maintaining a balance between precision and recall in the weighted evaluation. Overall, XGBoost outperformed SVM slightly, particularly in the weighted averages, though the difference in macro scores was minimal.

Classifier	Score Type	AUC	Precision	Recall	F1-Score
SVM	Macro	84%	69%	77%	71%
	Weighted		84%	79%	81%
XGBoost	Macro	85%	76%	70%	72%
	Weighted		84%	85%	84%

TABLE III: Performance of our classifiers for COVID-19 status classification using segmented breathing sound audio.

2) Classification Results Using Segmented Audio Data of Breathing Sounds Only: In this subsection, we present the performance results when the models were trained and tested on segmented breathing sound only. As detailed in Section **??**, this approach focuses on analyzing individual breathing segments rather than the entire audio file. Table III shows the performance of our top two best-performing classifiers, XGBoost and SVM, in classifying COVID-19 statuses using segmented breathing sounds. SVM achieved an AUC score of 84%, with a macro precision of 69%, a recall of 77%, and an F1-Score of 71%. When considering weighted averages, SVM's performance further improved, reaching 84% in precision, 79% in recall, and an 81% F1-Score.

On the other hand, XGBoost performed slightly better overall, achieving an AUC score of 85%. It recorded a macro precision of 76%, a recall of 70%, and a macro F1-Score of 72%. When considering the weighted averages, XGBoost achieved a precision of 84%, a recall of 85%, and an F1-Score of 84%. Overall, XGBoost outperformed SVM while using the segmented audio.

Compared to classification using the entire audio file, both classifiers showed improved performance for both macro and weighted averages while using the segmented audio. This improvement could be attributed to the segmentation process, which created more samples for training, allowing the models to capture relevant features more effectively. Additionally, segmentation may have helped by removing portions of the audio that contained silence, noise, or other irrelevant content, which can interfere with feature extraction and degrade performance in the case of long, unsegmented audio.

B. Performance Evaluation using Cough Sounds Only

This subsection presents the performance evaluation of our models trained and tested using only *cough sounds* (both deep and shallow) from the dataset. It assesses how effectively the models classify COVID-19 statuses based solely on cough sound patterns. The two best-performing classifiers in this scenario were *Random Forest (RF)* and *XGBoost*, and we report the results from these two classifiers in this subsection. For the remainder of the results, we report only the weighted averages, as the cough sound data is largely imbalanced, with more negative than positive cases.

Classifier	AUC	Precision	Recall	F1-Score
Random Forest	77%	83%	85%	84%
XGBoost	73%	83%	83%	83%

TABLE IV: Model performance for COVID-19 status classification using cough sounds only from the entire audio file.

1) Classification Results Using Entire Audio Files of Cough Sounds Only: Table IV shows the performance of our two best classifiers, Random Forest and XGBoost, in classifying COVID-19 status using cough sounds from the entire audio file. Random Forest achieved slightly better overall results with an AUC of 77%, indicating stronger discriminative ability compared to XGBoost's AUC of 73%. Additionally, Random Forest had a higher Recall (85%) compared to XGBoost (83%), meaning it was better at correctly identifying COVID-19 positive cases. Both models demonstrated equally strong Precision at 83%, indicating that the models were similarly effective at minimizing false positives.

In terms of the F1-Score, Random Forest slightly outperformed XGBoost with a score of 84% compared to XGBoost's 83%. This indicates that while both models are well-balanced in identifying positive cases, Random Forest is marginally more effective, especially in handling true positives. Overall, both classifiers performed well, but Random Forest showed a slight edge in overall classification performance.

Classifier	AUC	Precision	Recall	F1-Score
Random Forest	76%	81%	81%	81%
XGBoost	75%	81%	80%	80%

TABLE V: Model performance for COVID-19 status classification using segmented cough sounds only.

2) Classification Results Using Segmented Audio Data of Cough Sounds Only: Table V shows the performance of our two best classifiers, Random Forest and XGBoost, in classifying COVID-19 status using segmented cough sounds. Random Forest achieved an AUC of 76%, slightly outperforming XGBoost, which had an AUC of 75%. Both models demonstrated equal Precision (81%), indicating that they were similarly effective in minimizing false positives. However, Random Forest had a marginally higher Recall (81%) compared to XGBoost (80%), meaning it was slightly better at identifying true positives. This resulted in F1-Scores of 81% for Random Forest and 80% for XGBoost, with Random Forest showing a slight overall performance advantage in handling segmented audio data.

C. Performance Evaluation Using Both Breathing and Cough Sounds

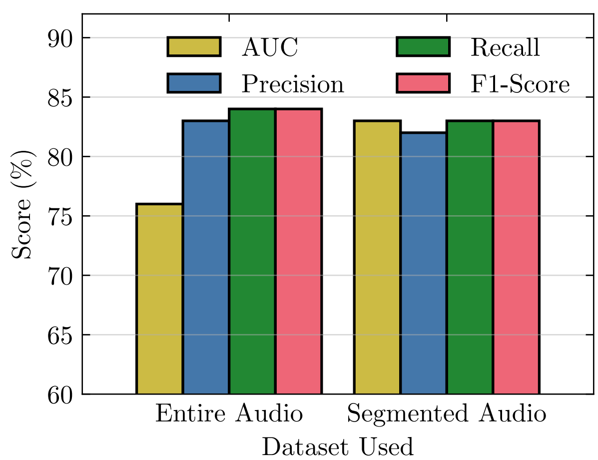


Fig. 2: Classifier performance using combined breathing and cough sounds for COVID-19 status classification

Figure 2 shows the performance of our best classifier, *Random Forest*, using combined breathing and cough sounds for COVID-19 status classification. The figure displays a comparison of classifier performance metrics across two different approaches: using entire audio files versus segmented audio sounds. For "Entire Audio," the classifier achieved slightly lower scores across all metrics compared to "Segmented Audio." Specifically, the AUC and F1-Score for "Entire Audio" are around 75% and 84% respectively, while for "Segmented Audio," these values are closer to 83% and 83%. Precision and Recall also show improvement in the segmented approach, emphasizing that processing audio in segments may help in enhancing the classifier's ability to accurately predict COVID-19 status based on both cough and breathing sounds.

V. Conclusion and Future Work

This study explored a non-invasive approach to COVID-19 detection using machine learning to analyze vocal patterns, specifically cough sounds, addressing the limitations of traditional diagnostic methods like nasal swabs and biosensors. By leveraging a publicly available dataset, we developed machine learning models that achieved an AUC of up to 85% and an F1-score of 81%, demonstrating the potential for rapid, cost-effective COVID-19 diagnosis. These findings highlight the promise of audio-based diagnostics, particularly in resource-limited settings where traditional testing is not feasible. While the results are promising, further research is needed to validate these models in real-world clinical environments and across diverse populations. Future work should focus on expanding datasets to enhance model robustness, addressing biases such as age and gender, and integrating these tools into telemedicine platforms for scalable, remote health monitoring. This could pave the way for the use of vocal signals not only for COVID-19 but also for broader applications in diagnosing respiratory diseases.

VI. Acknowledgment

This project was supported, in part, by grants from the U.S. National Science Foundation (IIS2334391), NIH RCMI Supplement award (3U54MD007586-37S3), and NLM Fisk-Meharry-Vanderbilt Summer research program (5R25LM014216). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of NSF or other funding agencies.

References

[1] T. A. Slater, S. Straw, M. Drozd, S. Kamalathasan, A. Cowley, and K. K. Witte, "Dying 'due to'or 'with'covid-19: a cause of death analysis in hospitalised patients," *Clinical medicine*, vol. 20, no. 5, pp. e189–e190, 2020.

[2] D. Bhattacharya, N. K. Sharma, D. Dutta, S. R. Chetupalli, P. Mote, S. Ganapathy, C. Chandrakiran, S. Nori, K. Suhail, S. Gonuguntla *et al.*, "Coswara: A respiratory sounds and symptoms dataset for remote screening of sars-cov-2 infection," *Scientific Data*, vol. 10, no. 1, p. 397, 2023.

[3] L. Verde, G. De Pietro, A. Ghoneim, M. Alrashoud, K. N. Al-Mutib, and G. Sannino, "Exploring the use of artificial intelligence techniques to detect the presence of coronavirus covid-19 through speech and voice analysis," *IEEE Access*, vol. 9, pp. 65 750–65 757, 2021.

[4] J. M. Sanders, M. L. Monogue, T. Z. Jodlowski, and J. B. Cutrell, "Pharmacologic treatments for coronavirus disease 2019 (covid-19): a review," *Jama*, vol. 323, no. 18, pp. 1824–1836, 2020.

[5] R. Filip, R. Gheorghita Puscaselu, L. Anchidin-Norocel, M. Dimian, and W. K. Savage, "Global challenges to public health care systems during the covid-19 pandemic: a review of pandemic measures and problems," *Journal of personalized medicine*, vol. 12, no. 8, p. 1295, 2022.

[6] S. A. Meo, A. A. Abukhalaf, A. A. Alomar, F. J. AlMutairi, A. M. Usmani, and D. C. Klonoff, "Impact of lockdown on covid-19 prevalence and mortality during 2020 pandemic: observational analysis of 27 countries," *European journal of medical research*, vol. 25, pp. 1–7, 2020.

[7] M. L. Robinson, A. Mirza, N. Gallagher, A. Boudreau, L. Garcia Jacinto, T. Yu, J. Norton, C. H. Luo, A. Conte, R. Zhou *et al.*, "Limitations of molecular and antigen test performance for sars-cov-2 in symptomatic and asymptomatic covid-19 contacts," *Journal of clinical microbiology*, vol. 60, no. 7, pp. e00 187–22, 2022.

[8] R. Tao, J. Downs, T. M. Beckie, Y. Chen, and W. McNelley, "Examining spatial accessibility to covid-19 testing sites in florida," *Annals of GIS*, vol. 26, no. 4, pp. 319–327, 2020.

[9] J. Hopman, B. Allegranzi, and S. Mehtar, "Managing covid-19 in low- and middle-income countries," *Jama*, vol. 323, no. 16, pp. 1549–1550, 2020.

[10] P. Lloyd-Sherlock, S. Ebrahim, L. Geffen, and M. McKee, "Bearing the brunt of covid-19: older people in low and middle income countries," 2020.

[11] S. S. Nayak, A. D. Darji, and P. K. Shah, "Machine learning approach for detecting covid-19 from speech signal using mel frequency magnitude coefficient," *Signal, image and video processing*, vol. 17, no. 6, pp. 3155–3162, 2023.

[12] R. Islam, E. Abdel-Raheem, and M. Tarique, "Early detection of covid-19 patients using chromagram features of cough sound recordings with machine learning algorithms," in *2021 International Conference on Microelectronics (ICM)*. IEEE, 2021, pp. 82–85.

[13] B. Hu, H. Guo, P. Zhou, and Z.-L. Shi, "Characteristics of sars-cov-2 and covid-19," *Nature reviews microbiology*, vol. 19, no. 3, pp. 141–154, 2021.

[14] K. Feng, F. He, J. Steinmann, and I. Demirkiran, "Deep-learning based approach to identify covid-19," in *SoutheastCon 2021*. IEEE, 2021, pp. 1–4.

[15] I. Aytekin, O. Dalmaz, K. Gonc, H. Ankishan, E. U. Saritas, U. Bagci, H. Celik, and T. Çukur, "Covid-19 detection from respiratory sounds with hierarchical spectrogram transformers," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[16] A. Ponomarchuk, I. Burenko, E. Malkin, I. Nazarov, V. Kokh, M. Avetisian, and L. Zhukov, "Project achoo: a practical model and application for covid-19 detection from recordings of breath, voice, and cough," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 2, pp. 175–187, 2022.

[17] A. Imran, I. Posokhova, H. N. Qureshi, U. Masood, M. S. Riaz, K. Ali, C. N. John, M. I. Hussain, and M. Nabeel, "Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app," *Informatics in medicine unlocked*, vol. 20, p. 100378, 2020.

[18] V. Despotovic, M. Ismael, M. Cornil, R. Mc Call, and G. Fagherazzi, "Detection of covid-19 from voice, cough and breathing patterns: Dataset and preliminary results," *Computers in Biology and Medicine*, vol. 138, p. 104944, 2021.

[19] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, and C. Mascolo, "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*, 2020, pp. 3474–3484.

[20] G. Chaudhari, X. Jiang, A. Fakhry, A. Han, J. Xiao, S. Shen, and A. Khanzada, "Virufy: Global applicability of crowdsourced and clinical datasets for ai detection of covid-19 from cough," *arXiv preprint arXiv:2011.13320*, 2020.

[21] A. Pal and M. Sankarasubbu, "Pay attention to the cough: Early diagnosis of covid-19 using interpretable symptoms embeddings with cough sound signal processing," in *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, 2021, pp. 620–628.

[22] P. Bagad, A. Dalmia, J. Doshi, A. Nagrani, P. Bhamare, A. Mahale, S. Rane, N. Agarwal, and R. Panicker, "Cough against covid: Evidence of covid-19 signature in cough sounds," *arXiv preprint arXiv:2009.08790*, 2020.

[23] M. Loey and S. Mirjalili, "Covid-19 cough sound symptoms classification from scalogram image representation using deep learning models," *Computers in biology and medicine*, vol. 139, p. 105020, 2021.

[24] J. Han, T. Xia, D. Spathis, E. Bondareva, C. Brown, J. Chauhan, T. Dang, A. Grammenos, A. Hasthanasombat, A. Floto *et al.*, "Sounds of covid-19: exploring realistic performance of audio-based digital testing," *NPJ digital medicine*, vol. 5, no. 1, p. 16, 2022.

[25] D. T. Pizzo and S. Esteban, "Iatos: Ai-powered pre-screening tool for covid-19 from cough audio samples," *arXiv preprint arXiv:2104.13247*, 2021.

[26] L. Orlandic, T. Teijeiro, and D. Atienza, "The coughvid crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms," *Scientific Data*, vol. 8, no. 1, p. 156, 2021.

[27] E. D. Haritaoglu, N. Rasmussen, D. C. Tan, J. Xiao, G. Chaudhari, A. Rajput, P. Govindan, C. Canham, W. Chen, M. Yamaura *et al.*, "Using deep learning with large aggregated datasets for covid-19 classification from cough," *arXiv preprint arXiv:2201.01669*, 2022.

[28] A. Khanzada, S. Hegde, S. Sreeram, G. Bower, W. Wang, R. P. Mediratta, K. D. Meister, and A. Rameau, "Challenges and opportunities in deploying covid-19 cough ai systems," *Journal of Voice*, vol. 35, no. 6, pp. 811–812, 2021.

[29] A. E. Ashby, J. A. Meister, K. A. Nguyen, Z. Luo, and W. Gentzke, "Cough-based covid-19 detection with audio quality clustering and confidence measure based learning," in *Conformal and Probabilistic Prediction with Applications*. PMLR, 2022, pp. 129–148.

[30] R. V. Sharan, "Productive and non-productive cough classification using biologically inspired techniques," *IEEE Access*, vol. 10, pp. 133 958–133 968, 2022.

[31] "Coswara-data," https://github.com/iiscleap/Coswara-Data, [Online; accessed September-2024].

[32] K. Anderson, Y. Qiu, A. R. Whittaker, and M. Lucas, "Breath sounds, asthma, and the mobile phone," *The Lancet*, vol. 358, no. 9290, pp. 1343–1344, 2001.

[33] "Librosa : audio and music processing in python," https://librosa.org/, [Online; accessed September-2024].