OTHER PEOPLE'S COUGHS: CHALLENGES AND PERSPECTIVES IN REMOTE COUGH MONITORING USING WEARABLE DEVICES

Hyfe Inc.

ABSTRACT

Accurate detection and classification of cough sounds is a key factor in monitoring respiratory health. While personal cough detection via wearables has been studied, detecting other people's coughs in proximity remains underexplored. This paper discusses the problem of detecting other people's coughs from multiple points of view, and then presents a novel dataset designed to address this challenge. The latter contains coughs recorded from both male and female subjects at near and far distances, providing a balanced representation of environmental noise conditions and room acoustics. Cough events are modeled using Mel-frequency energies to capture the spectral properties of the sounds, and these representations are fed into a variety of machine learning models, ranging from simple and interpretable ones, such as Logistic Regression, up to more complex ones, such as 2D convolutional neural networks (CNN), for classification. While standard ML models achieve a noteworthy performance, the proposed CNN models can effectively distinguish between near and far coughs, as demonstrated in a 5-fold stratified grouped cross validation, achieving an average Receiver Operating Characteristic Area Under the Curve (ROCAUC) score of 0.94. This research underscores the importance of uncontaminated cough monitoring, particularly in settings where distinguishing between personal and external coughs can enhance the reliability of health assessments and prevent data misinterpretation.

Index Terms— Cough classification, acoustic signal, neural network, wearables, near cough, far cough

1. INTRODUCTION

Cough is a fundamental physiological reflex that serves to clear the respiratory tract of irritants, pathogens, and secretions. It is an important clinical symptom across a range of diseases, including respiratory infections, asthma, chronic obstructive pulmonary disease (COPD), and pulmonary fibrosis. In healthcare, monitoring the frequency and severity of cough can provide valuable diagnostic and prognostic information. Persistent or abnormal cough patterns often indicate underlying pathology, and in certain diseases, cough is a critical marker for disease progression and treatment efficacy. Therefore, accurately tracking cough can enhance clinical decision-making and patient care.

Automatic cough detection and counting using wearable audio devices [1–14] have emerged as promising technologies in health monitoring. These devices leverage machine learning algorithms and acoustic analysis [4, 5, 7, 14–23] to detect cough events in real-time or off-line, offering a continuous, objective measure of cough frequency. Such systems are non-invasive and convenient, allowing patients to wear them during daily activities. The ability to automatically detect and count coughs can be particularly beneficial for diseases where cough is a key symptom, such as COVID-19 [24–27], providing healthcare professionals with detailed data for patient monitoring and early intervention [28–31].

Cough monitoring through wearable devices [32–35] holds significant potential in the realm of *remote* healthcare. With the increasing demand for telemedicine and digital health solutions, the ability to track cough patterns remotely can provide valuable insights into a patient's respiratory health. For individuals with chronic respiratory conditions, such as COPD or asthma, continuous monitoring of cough can offer an objective measure of disease exacerbations or response to treatment.

Furthermore, remote cough monitoring can enhance the management of patients with chronic illnesses who require frequent medical oversight but are unable to visit healthcare facilities regularly [36–39]. It can reduce the burden on healthcare systems by enabling early identification of complications or worsening symptoms, allowing for prompt intervention and reducing hospital admissions. In resource-limited settings, where access to healthcare professionals may be restricted, such systems can play a crucial role in ensuring continuous patient monitoring and improving health outcomes.

The integration of remote cough monitoring into healthcare also allows for better patient engagement and self-management. Patients can receive real-time feedback on their cough patterns, which can motivate adherence to treatment plans and promote greater awareness of their respiratory health. By empowering patients with insights into their symptoms, this technology can facilitate shared decision-making between patients and healthcare providers, leading to more personalized and effective care strategies.

Additionally, data from remote cough monitoring systems can contribute to population-level health monitoring, particularly during outbreaks of respiratory diseases. Large-scale data collection on cough patterns can help identify trends in disease spread, aiding public health efforts to contain infectious diseases. This type of surveillance can be particularly important in managing pandemics, where early detection of symptoms like cough can inform strategies for intervention and mitigation [40–44].

However, one significant challenge in cough monitoring systems is the problem of ”other people's coughs” (OPCs), where the system detects coughs from individuals other than the intended user (target user). This contamination factor can lead to inaccurate data, reducing the reliability of the cough monitoring system. In environments with multiple individuals, such as households or public spaces, OPCs can pose a considerable problem, as it becomes difficult to attribute cough events to a specific user. To mitigate this issue, advanced machine learning techniques and context-aware algorithms are required to distinguish between different users' coughs, ensuring the accuracy and utility of the monitoring system.

1.1. Audio-based Approaches

”Cougher verification” is an approach that draws inspiration from speaker verification techniques [45–47] in order to address the problem of attributing coughs to the correct individual. Similar to how speaker verification systems authenticate a user based on their unique vocal characteristics, cougher verification would use the acoustic signature of a person's cough to distinguish it from other people's coughs. Each individual's cough has subtle variations in sound frequency distribution, duration, and intensity, influenced by the anatomy of their respiratory system and the nature of their health condition. By employing machine learning models trained on these distinct cough features, a cougher verification system could accurately attribute cough events to a specific individual, even in noisy environments or shared spaces. This approach would involve creating a “cough profile” for each user based on a series of recorded coughs, which would serve as a reference for identifying future cough events.

Adapting speaker verification algorithms for cougher verification would involve the use of advanced signal processing techniques and deep learning architectures like convolutional neural networks (CNNs) [48], multitask learning with residual encoders [49], metric learning with CNNs [50], recurrent neural networks (RNNs) and their variants [51,52], and cough embeddings from DNNs [53]. These models would be trained to learn the unique patterns in a person's cough sound and differentiate them from others. Additionally, such systems could incorporate context-aware mechanisms, such as proximity detection and microphone arrays, to further isolate the source of the cough sound. A combination of individual cough profiling and acoustic verification could significantly reduce the issue of OPCs contaminating the data, making cough monitoring systems more reliable for remote healthcare applications. This methodology would not only improve the accuracy of cough detection but also enhance patient monitoring by ensuring that the data collected is truly reflective of the target user's health condition.

Despite its potential, the implementation of cougher verification in real-life settings presents several limitations. One significant challenge is the variability of cough sounds, both within an individual and across different contexts. Factors such as the severity of illness, environmental noise, and changes in a person's respiratory condition over time can alter the acoustic characteristics of a cough. For instance, a person recovering from an illness might have a milder or less frequent cough compared to when they are acutely ill, making it more difficult for the system to consistently verify coughs against a static profile. This intra-person variability could reduce the accuracy of cougher verification models, leading to false positives or negatives, particularly in dynamic real-world environments where individuals may move between quiet and noisy spaces. More importantly, in monitoring approaches that involve a smartwatch or any other microphone-equipped device that allows variable distance between the cougher and the recording sensor, potential solutions face severe performance degradation: although the verification algorithm can be trained on clear, noise-free, close-proximity coughs, in real-life scenarios, target user's coughs can significantly vary due to the distance factor alone. Another limitation lies in the practicality of deploying such systems in multi-user or densely populated environments. In settings like hospitals, shared homes, or workplaces, there may be significant overlap in cough events, making it challenging for the system to isolate and verify coughs accurately. Additionally, while machine learning models can be trained to recognize subtle differences in cough sounds, the presence of other background noises, such as speech, traffic, or ambient sounds, can interfere with the system's ability to distinguish between different users' coughs. The technology also requires sufficient high-quality training data to build robust cough profiles, which may not always be available, particularly for individuals with rare respiratory conditions. Lastly, privacy concerns may arise, as audio recording required for user profiling and training may raise ethical questions regarding surveillance and data security in public and private spaces. Moreover, from a technical perspective, an algorithm designed to discriminate between OPCs and target user's coughs must achieve a high level of accuracy to ensure the effectiveness of the overall cough detection system. If the algorithm performs poorly, it could lead to two critical issues: false negatives, where the target user's coughs are incorrectly classified as OPCs and removed, thus underestimating the user's cough frequency; and false positives, where OPCs are mistakenly attributed to the target user, inflating cough counts. Both scenarios could significantly compromise the reliability of the data, potentially leading to incorrect clinical assessments or delayed interventions. Furthermore, issues with trustworthiness of remote health monitoring would arise in such cases, ultimately reducing the clinical utility of the system. Therefore, ensuring high precision and recall in cough discrimination algorithms is crucial for maintaining the accuracy and integrity of the entire cough detection system pipeline. On top of that, the OPCs problem is relatively rare in everyday life, making it an over-engineered solution in most contexts. In typical day-to-day scenarios, most people are not in close proximity to others who cough frequently enough to cause significant confusion in cough monitoring systems. The issue generally only arises in specific situations, such as when multiple people, like family members or roommates, share confined spaces like apartments during flu seasons or respiratory outbreaks. In these cases, distinguishing coughs becomes more relevant, but for the average user, the risk of picking up others' coughs in most environments is minimal. This rarity suggests that creating sophisticated systems to handle this problem may be unnecessary for the general population, where simpler, more user-friendly solutions could suffice without adding unnecessary complexity.

1.2. Heart Rate and Accelerometry based Approaches

A potential solution to the issue of OPCs could involve integrating additional data streams commonly available in wearable devices, such as heart rate and accelerometry, to enhance cough detection accuracy or as a sole datastream for cough detection [54–59]. By cross-referencing these physiological and motion data with cough detection events, the system could more reliably determine whether the detected cough originates from the target user. For example, a sudden increase in heart rate or specific movement patterns (such as the rapid chest or abdominal movements associated with coughing) could be used to confirm that a cough sound is coming from the user. These physiological changes tend to occur simultaneously with coughing, and incorporating them into the detection algorithm could help filter out OPCs by ensuring the cough is accompanied by physical signs from the user.

Heart rate data, in particular, provides a valuable biometric signal that could be used to confirm whether a detected cough sound corresponds with the user's physiological state. During coughing episodes, the body often experiences slight fluctuations in heart rate due to the exertion involved in expelling air from the lungs. By analyzing these patterns, the algorithm could validate whether a cough event is occurring in tandem with a physiological response from the user, improving the specificity of cough detection. This multi-modal approach would not only help in identifying genuine coughs from the target user but also allow the system to reject cough sounds that occur without corresponding changes in heart rate, which could be more likely to originate from other individuals nearby.

Accelerometry, which tracks body movement, could further refine cough detection by identifying the distinct motions associated with coughing, such as the rapid contraction of muscles in the chest or upper body. Combining accelerometry data with cough audio detection could allow the system to detect a user's characteristic coughing motion and verify the source of the sound. This additional layer of verification would be especially useful in environments where multiple individuals are present, reducing the risk of OPC contamination. By creating a multi-sensor system that leverages heart rate and accelerometry alongside audio-based cough detection, wearable devices could improve the precision of cough monitoring, ensuring that only the target user's coughs are recorded and analyzed for health monitoring purposes.

However, wearables such as belts and patches, which utilize accelerometry and heart rate detection for cough monitoring, face significant limitations when used for extended periods. One primary challenge is the discomfort and impracticality of wearing these devices for long durations, especially for continuous monitoring. Belts, for instance, may restrict natural body movements and cause discomfort during daily activities or sleep, while adhesive patches can cause skin irritation, making them unsuitable for sensitive skin. Moreover, the bulkiness or intrusive nature of these devices compromises user-friendliness and might engage social discomfort if placed in a visible body part, discouraging consistent usage over time. These wearables also face difficulties in differentiating the user's cough from nearby sounds, such as OPCs, due to their reliance on bodily signals rather than more targeted acoustic or spatial data. Thus, while these devices offer valuable health insights, their design limits long-term usability and general acceptance among users due to issues of comfort and practicality.

1.3. A Potential Audio-based Single-Stream Solution

A potential and implicit solution to address the challenge of distinguishing between the target user's coughs and OPCs involves creating a large dataset of cough sounds recorded at varying distances and in different environmental contexts. This dataset would be used to train machine learning algorithms for a binary classification task, where cough sounds are labeled as either "near" (close to the wearable and thus likely the target user's cough) or "far" (further from the microphone, indicating another person's cough). By capturing cough sounds in a range of real-world scenarios - such as quiet indoor spaces, noisy urban settings, and different proximities from the microphone - would allow the system to learn the acoustic features associated with distance and environmental interference. Features like spectral attenuation, spectral tilt, and echo patterns could be key in differentiating between "near" and "far" coughs. However, recent deep learning advancements rely on raw or close-to-raw data streams, such as spectrograms, to learn a highly non-linear mapping function between the input sound and the sound label, making feature engineering obsolete.

Training a binary classification model on this labeled dataset could significantly enhance the accuracy of cough detection systems in distinguishing between the target user's cough and OPCs. For instance, the system could learn to detect specific changes in the acoustic signal that occur as a result of distance, such as sound intensity drop-off, attenuated frequencies due to distance, and the presence of environmental noise. This classification model could then be integrated into wearable devices, allowing for real-time filtering of cough sounds based on their proximity to the user. By focusing on distance-based classification, the system would not only improve the precision of cough monitoring but also reduce the risk of data contamination from nearby individuals, making remote health monitoring more reliable and actionable in real-world settings.

1.4. Contributions

In this work we demonstrate that

- far (further than 1m) and near (less than or equal to 1m) coughs are distinguishable from a single smartwatch microphone
- mel-frequency energies (MFEs) have the capacity to represent near and far coughs adequately
- convolutional neural networks are a suitable approach for a binary classification task of detecting near from far coughs.

2. METHODOLOGY

In this section, we will present some information about our current technology, the approach taken, and the dataset we compiled for our task.

2.1. Hyfe's ID206 technology

Hyfe's current tech for Remote Patient Monitoring involves the use of a smartwatch named ID206. The smartwatch is equipped with a MEMS microphone capable of sampling continuous audio at different sampling rates. Due to complexity and performance issues, Hyfe's current sampling rate is 8000 Hz, with a floating point precision of 32-bit. Continuous audio is analyzed in a frame-wise manner and onset times are identified via spectral and temporal criteria, assuming that the explosive phase of a cough can be detected similarly to the attack of a percussive musical instrument sound or a voiceless stop sound in speech. A detected onset defines a 0.5-second segment that may, or may not, be a cough sound. The raw 0.5-second candidate waveform is transformed into Mel-Frequency Energies (MFEs), a two-dimensional spectrotemporal representation that reveals the acoustic energy distribution among time and mel-scaled frequency (a perceptually relevant transformation of the well-known linear frequency measured in Hertz). This representation is successively fed into a classifier trained on millions of 0.5-second sounds, labeled as either "cough" or "non cough".

The parametrization of the raw audio signal, as well as the architecture of the classifier, are carefully selected to optimize latency and memory constraints of ID206. In other words, the MFE representation should be compact but "sufficient", including all relevant information of the input audio, and the classifier (a neural network, in Hyfe's case) should be accordingly designed to extract useful information for the task, without being too complex or over-parameterized.

2.2. Dataset

Hyfe used a modified ID206 able to record and store continuous audio on device. Cough and ambient sound recordings were made in several different acoustic environments by several different coughers (22 in total), both male and female (14 and 8, respectively), aged from 20 to 67 years old. All coughs were recorded in specific distances from the microphone, starting from 5 cm and up to 300 cm. Each cough sound was labeled as either "near" or "far", with 100 cm being the separation threshold: coughs recorded at a distance less than or equal to 100 cm are considered as "near", while all others are labeled as "far". It should also be noted that coughers were asked to cough "silently, as they would in an crowded place", "normally, as they would at home", and "loudly, as if they were choking". Hence, the dataset contains a variety of cough sounds, ranging from loud and far coughs to faint and close-to-microphone coughs. In total, 829 near and 378 far coughs were collected, all sampled at 8000 Hz. Figure 1 shows two coughs from the same person, at the same environment, but at different distances.

2.3. Feature Extraction and Classification

In this section, we discuss the feature extraction pipeline and the classification approaches suitable for the task in hand.

2.3.1. Feature Extraction

The computation of Mel-Frequency Energies is a critical step in numerous audio processing applications, particularly in speech and music recognition systems. The process begins by converting a raw audio signal $x[n]$ into the frequency domain using the Short-Time Fourier Transform (STFT). The STFT is applied to overlapping frames of the signal, where each frame $x_m[n]$ (with m indicating the frame index) is multiplied by a window function $w[n]$, such as a Hamming window. The Fast Fourier Transform (FFT) is then used to

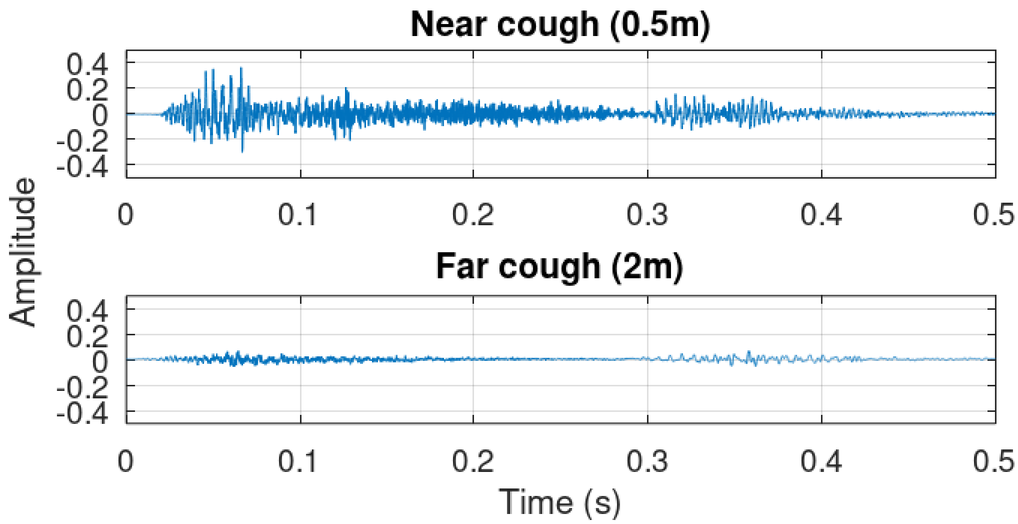

Fig. 1. *Top panel: a near cough. Bottom panel: a far cough. Both coughs are from the same person.*

compute the discrete Fourier transform $X_m[k]$ for each frame, where k represents the frequency bin. The resulting spectrum is represented as:

$$X_m[k] = \sum_{n=0}^{N-1} x_m[n]w[n]e^{-j\frac{2\pi kn}{N}},$$

where N is the number of points in the FFT, and $X_m[k]$ provides the magnitude and phase information in the linear frequency domain. However, the human auditory system perceives frequencies non-linearly, particularly more sensitively at lower frequencies and less sensitively at higher frequencies. To account for this, the frequency scale is transformed to the Mel scale, which is approximately linear below 1 kHz and logarithmic above it. The relationship between a frequency f in Hertz and its Mel-scaled counterpart f_{Mel} is given by:

$$f_{\text{Mel}} = 2595 \cdot \log_{10}\left(1 + \frac{f}{700}\right).$$

This transformation reflects the psychophysical properties of human hearing. The next step involves applying a bank of triangular band-pass filters uniformly spaced in the Mel domain. Each filter $H_i[k]$ corresponds to a range of frequencies in the linear domain but is placed according to Mel frequencies. The filters are constructed such that they overlap, with the response of each filter increasing linearly on one side and decreasing linearly on the other. Figure 3 shows 26 such filters for a frequency range of 4000 Hz, which corresponds to the Nyquist frequency of cough audio signals.

The Mel-frequency energies are computed by passing the linear magnitude spectrum $|X_m[k]|$ through this filter bank. For each filter i, the energy E_i is calculated as the weighted sum of the squared magnitudes of the Fourier coefficients:

$$E_i = \sum_{k=1}^{K} H_i[k]|X_m[k]|^2,$$

where K is the number of frequency bins, and $H_i[k]$ is the triangular filter response for the i-th filter. This process yields a set of energies, each corresponding to a different range of frequencies in the Mel scale, providing a perceptually relevant representation of the sound signal. These Mel-frequency energies are typically subjected to a logarithmic compression to further emulate the non-linear sensitivity of human hearing to intensity changes:

$$E_i^{\text{log}} = \log(E_i).$$

This final log-Mel representation is commonly used as input features in many audio-related machine learning applications, providing a robust and perceptually meaningful feature set. We utilize MFEs in this work and we peak-normalize all sounds before processing. Figure 2 shows MFEs for the two coughs in Fig. 1.

2.3.2. Classification

MFEs are understood as two-dimensional (2D), spectrotemporal representations of cough sounds. Conventional machine learning (ML) algorithms [60] can handle 2D data, such as Mel-Frequency Energies (MFEs), by transforming or adapting the data to fit the expected input formats of these algorithms. Most ML algorithms, such as Support Vector Machines (SVMs), Random Forests (RFs), and ensemble methods such as eXtreme Gradient Boosting (XGB), typically operate on one-dimensional (1D) feature vectors. Therefore, when dealing with 2D data, the key challenge is to reshape

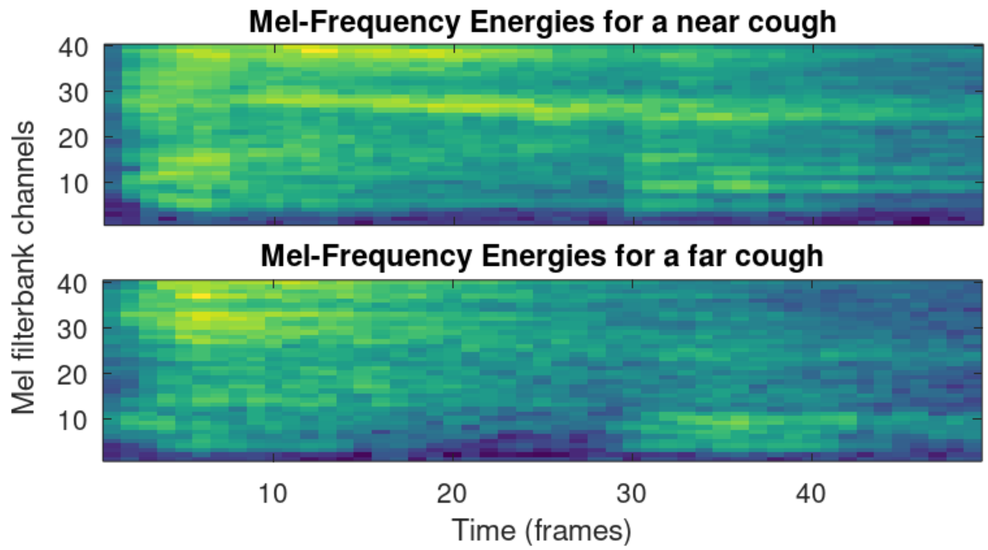

Fig. 2. *Top panel: MFE for a near cough. Bottom panel: MFE for a far cough. Both coughs are from the same person.*

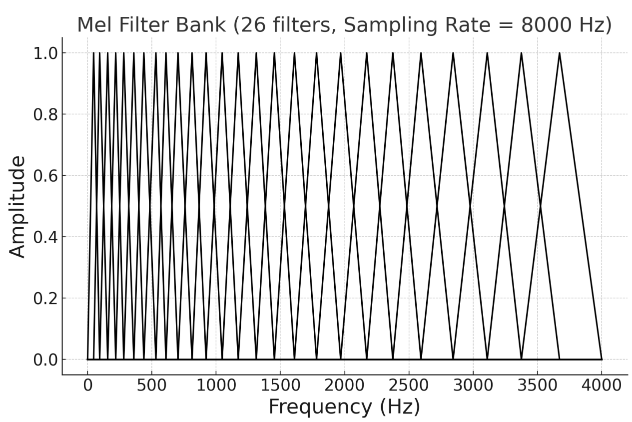

Fig. 3. 26 *filters of a filterbank covering* 4000 *Hz.*

or transform the data into a form that these algorithms can process effectively. We choose the most common approach, which is to "flatten" the data into a 1D vector. In the case of MFEs, which are typically represented as matrices where the rows correspond to time frames and the columns to frequency bands, each 2D matrix can be reshaped into a single long 1D vector by concatenating all rows or columns.

We trained four different ML algorithms: Logistic Regression (LR), Support Vector Machines (SVMs), Random Forests (RFs), and eXtreme Gradient Boosting (XGBoost), all trained on flattened vectors of the input MFEs. Logistic Regression is a linear model used for binary classification, estimating the probability of a class by applying a logistic function to a linear combination of input features. It is simple, interpretable, and works well when the classes are linearly separable. SVMs are powerful classification algorithms that find the optimal hyperplane to separate classes in the feature space, maximizing the margin between them. They are effective in high-dimensional spaces and are robust to overfitting, especially with a proper kernel choice. Random Forests are ensemble methods that construct multiple decision trees during training and aggregate their outputs for improved accuracy and robustness. They handle high-dimensional data well, are less prone to overfitting, and provide feature importance estimates. XGBoost is an optimized gradient boosting algorithm that builds an ensemble of decision trees sequentially, improving on previous trees by minimizing errors. It is highly efficient, scalable, and often achieves state-of-the-art performance in many classification and regression tasks.

However, more modern machine learning algorithms, especially deep learning models like Convolutional Neural Networks (CNNs) [61], are highly effective at handling 2D data, such as images or spectrogram-like data (such as MFEs). CNNs excel in recognizing spatial patterns by applying convolutions that preserve local structure and hierarchical features. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks can also process 2D data with a temporal dimension, capturing dependencies over time. These models automatically extract relevant features from raw 2D data, making them superior to traditional ML algorithms. We trained two CNNs, a deeper, more resource-demanding network, and a shallow one, more computationally efficient, suitable for embedded devices. Both operate on the raw 2D MFE data, with the only difference being in the parametrization of the MFEs and in their depth (model complexity). CNN_1 is a deep, multi-layer network, consisting of convolutional layers, followed by max-pooling layers and a dropout layer. A ReLU activation function is used in all convolutional layers. Furthermore, a flattening layer transforms the resulting feature map into a 1D representation which is fed to a series of dense-plus-dropout layers, also activated by a ReLU. The final layer utilizes a sigmoid activation function that outputs a score for each sound. Similarly, CNN_2 follows the same basic architecture but it has significantly less convolutional layers and only one dense-plus-dropout layer. A sigmoid also drives the final prediction of any audio input.

3. RESULTS

A nested, stratified, grouped cross-validation scheme was used for hyperparameter optimization and performance metric computation. Care was taken to implement a cougher-independent classification, with zero information leak between train, validation, and test sets. A variety of performance metrics were selected, namely sensitivity, specificity, F1-score, precision, and ROC-AUC. Sensitivity, also known as the true positive rate, is defined as the ratio of true positives (TP) to the sum of true positives and false negatives (FN), given by

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

Specificity, or the true negative rate, measures the proportion of actual negatives correctly identified, calculated as

$$\text{Specificity} = \frac{TN}{TN + FP}$$

where TN represents true negatives and FP denotes false positives. Precision, or positive predictive value, indicates the accuracy of positive identifications and is expressed as

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

The F1-score, which balances precision and sensitivity, is calculated using the formula

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$$

Finally, the ROC-AUC (Receiver Operating Characteristic - Area Under Curve) evaluates the model's discrimination ability by plotting sensitivity against (1 - specificity) across various thresholds, with an AUC of 1 indicating perfect discrimination and 0.5 indicating no discrimination. Together, these metrics provide a comprehensive assessment of classification model performance.

Table 1 shows performance metrics for each conventional ML model. Support Vector Machines (SVMs) consistently performs well, achieving the highest ROC-AUC (0.885) and strong Precision (0.862) and F1-score (0.779). This is not a surprising outcome since SVMs are well-known for their ability to handle high-dimensional data and small datasets. Random Forest achieves the highest Sensitivity (0.878), indicating its strength in correctly identifying positive cases, but it shows weakness in Specificity (0.571), which is significantly lower than other models. Logistic Regression provides a balanced performance, with strong Precision (0.860) and ROC-AUC (0.856), while Extreme Gradient Boosting demonstrates competitive results across all metrics, particularly in Sensitivity (0.863) and ROC-AUC (0.846), though it has lower Specificity (0.584). Overall, SVMs and XGBoost offer the most consistent and competitive performances across multiple metrics.

Complementary to conventional ML models, Table 2 illustrates a comparison between our two convolutional neural networks. Specifically, CNN_1 consistently outperforms CNN_2 in all metrics, achieving an impressive ROC-AUC of 0.943, indicating a strong ability to discriminate between classes. Also, it achieves noteworthy Precision (0.922), F1-score (0.912), and Sensitivity (0.904), showing its overall robust classification performance. Specificity for CNN_1 is also high at 0.816, reflecting its ability to correctly identify negative cases. CNN_2, while still performing reasonably well, exhibits lower values across the board, particularly in Specificity (0.697) and ROC-AUC (0.900), though it still maintains solid results in Precision (0.848) and Sensitivity (0.884). Overall, CNN_1 demonstrates superior performance, especially in its ability to balance sensitivity and specificity. However, it should be noted that our shallow model (CNN_2) performs adequately in such a small audio dataset. It is evident that CNNs are more promising candidate models for solving the OPC problem.

4. CONCLUSIONS AND FUTURE WORK

In conclusion, this work demonstrates that it is indeed possible to predict (in a binary setting) the distance of a cough sound from the target individual with high accuracy using machine learning models, as evidenced by the strong performance metrics, including a ROCAUC score of 0.94 attained by convolutional neural networks. The development of a novel dataset that includes both near and far coughs, under varied environmental conditions, played a crucial role in achieving these results. However, while these findings highlight the potential for accurate detection of external coughs, further validation with larger and more diverse datasets is necessary to ensure robustness and generalizability across different environments and populations. Future work should focus on expanding the dataset and exploring additional factors that may influence cough sound detection and classification, ultimately contributing to more reliable respiratory health monitoring.

Metric	Logistic Regression	Support Vector Machines	Random Forest	eXtreme Gradient Boosting
ROC-AUC	0.856 (0.050)	**0.885 (0.053)**	0.844 (0.040)	0.846 (0.043)
Precision	0.860 (0.085)	**0.862 (0.086)**	0.809 (0.104)	0.810 (0.105)
F1-score	0.746 (0.058)	**0.779 (0.058)**	0.759 (0.063)	0.754 (0.067)
Sensitivity	0.750 (0.120)	0.802 (0.101)	**0.878 (0.056)**	0.863 (0.080)
Specificity	**0.757 (0.121)**	0.749 (0.122)	0.571 (0.197)	0.584 (0.190)

Table 1. *Performance metrics for Logistic Regression, Support Vector Machines, Random Forest, and eXtreme Gradient Boosting over a speaker-independent 5-fold stratified, grouped cross-validation. Best performance in bold.*

Metric	$\mathbf{CNN_1}$	$\mathbf{CNN_2}$
ROC-AUC	**0.943 (0.029)**	0.900 (0.068)
Precision	**0.922 (0.047)**	0.848 (0.092)
F1-score	**0.912 (0.038)**	0.860 (0.077)
Sensitivity	**0.904 (0.043)**	0.884 (0.107)
Specificity	**0.816 (0.106)**	0.697 (0.157)

Table 2. *Performance metrics for CNN_1 (deep model) and CNN_2 (shallow model) over a speaker-independent 5-fold stratified, grouped cross-validation. Best performance in bold.*

5. REFERENCES

[1] SS Birring, T Fleming, S Matos, AA Raj, DH Evans, and ID Pavord, "The leicester cough monitor: preliminary validation of an automated cough detection system in chronic cough," *European Respiratory Journal*, vol. 31, no. 5, pp. 1013–1018, 2008.

[2] Filipe Barata, Kevin Kipfer, Maurice Weber, Peter Tinschert, Elgar Fleisch, and Tobias Kowatsch, "Towards device-agnostic mobile cough detection with convolutional neural networks," in *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, 2019, pp. 1–11.

[3] Thomas Drugman, Jerome Urbain, and Thierry Dutoit, "Assessment of audio features for automatic cough detection," in *2011 19th European Signal Processing Conference*. IEEE, 2011, pp. 1289–1293.

[4] Renard Xaviero Adhi Pramono, Syed Anas Imtiaz, and Esther Rodriguez-Villegas, "Automatic cough detection in acoustic signal using spectral features," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 7153–7156.

[5] Jesús Monge-Álvarez, Carlos Hoyos-Barceló, Paul Lesso, and Pablo Casaseca-De-La-Higuera, "Robust detection of audio-cough events using local hu moments," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 184–196, 2018.

[6] Yan Shi, He Liu, Yixuan Wang, Maolin Cai, and Weiqing Xu, "Theory and application of audio-based assessment of cough," *Journal of Sensors*, vol. 2018, 2018.

[7] Brian H Tracey, Germán Comina, Sandra Larson, Marjory Bravard, José W López, and Robert H Gilman, "Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis," in *2011 Annual international conference of the IEEE engineering in medicine and biology society*. IEEE, 2011, pp. 6017–6020.

[8] Matthijs D Kruizinga, Ahnjili Zhuparris, Eva Dessing, Fas J Krol, Arwen J Sprij, Robert-Jan Doll, Frederik E Stuurman, Vasileios Exadaktylos, Gertjan JA Driessen, and Adam F Cohen, "Development and technical validation of a smartphone-based pediatric cough detection algorithm," *Pediatric Pulmonology*, vol. 57, no. 3, pp. 761–767, 2022.

[9] Jaclyn A Smith, Kimberley Holt, Rachel Dockry, Shilpi Sen, Kitty Sheppard, Philip Turner, Paul Czyzyk, and Kevin McGuinness, "Performance of a digital signal processing algorithm for the accurate quantification of cough frequency," *European Respiratory Journal*, vol. 58, no. 2, 2021.

[10] Mingyu You, Huihui Wang, Zeqin Liu, Chong Chen, Jiaming Liu, Xiang-Huai Xu, and Zhong-Min Qiu, "Novel feature extraction method for cough detection using nmf," *IET Signal Processing*, vol. 11, no. 5, pp. 515–520, 2017.

[11] YH Hiew, JA Smith, JE Earis, Barry MG Cheetham, and AA Woodcock, "Dsp algorithm for cough identification and counting," in *2002 IEEE International Con-*

ference on Acoustics, Speech, and Signal Processing. IEEE, 2002, vol. 4, pp. IV–3888.

[12] Mingyu You, Zeqin Liu, Chong Chen, Jiaming Liu, Xiang-Huai Xu, and Zhong-Min Qiu, "Cough detection by ensembling multiple frequency subband features," *Biomedical Signal Processing and Control*, vol. 33, pp. 132–140, 2017.

[13] Thomas Drugman, "Using mutual information in supervised temporal event detection: Application to cough detection," *Biomedical Signal Processing and Control*, vol. 10, pp. 50–57, 2014.

[14] Samantha J Barry, Adrie D Dane, Alyn H Morice, and Anthony D Walmsley, "The automatic recognition and counting of cough," *Cough*, vol. 2, no. 1, pp. 1–9, 2006.

[15] Igor DS Miranda, Andreas H Diacon, and Thomas R Niesler, "A comparative study of features for acoustic cough detection using deep architectures," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 2601–2605.

[16] Ebrahim Nemati, Md Mahbubur Rahman, Viswam Nathan, and Jilong Kuang, "Private audio-based cough sensing for in-home pulmonary assessment using mobile devices," in *13th EAI International Conference on Body Area Networks 13*. Springer, 2020, pp. 221–232.

[17] Cuong Pham, "Mobicough: real-time cough detection and monitoring using low-cost mobile devices," in *Intelligent Information and Database Systems: 8th Asian Conference, ACIIDS 2016, Da Nang, Vietnam, March 14–16, 2016, Proceedings, Part I 8*. Springer, 2016, pp. 300–309.

[18] Leonardo Di Perna, Gabriele Spina, Susannah Thackray-Nocera, Michael G Crooks, Alyn H Morice, Paolo Soda, and Albertus C den Brinker, "An automated and unobtrusive system for cough detection," in *2017 IEEE Life Sciences Conference (LSC)*. IEEE, 2017, pp. 190–193.

[19] Sergio Matos, Surinder S Birring, Ian D Pavord, and H Evans, "Detection of cough signals in continuous audio recordings using hidden markov models," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1078–1083, 2006.

[20] Jia-Ming Liu, Mingyu You, Zheng Wang, Guo-Zheng Li, Xianghuai Xu, and Zhongmin Qiu, "Cough detection using deep neural networks," in *2014 IEEE international conference on bioinformatics and biomedicine (BIBM)*. IEEE, 2014, pp. 560–563.

[21] Justice Amoh and Kofi Odame, "Deep neural networks for identifying cough sounds," *IEEE transactions on biomedical circuits and systems*, vol. 10, no. 5, pp. 1003–1011, 2016.

[22] Justice Amoh and Kofi Odame, "Deepcough: A deep convolutional neural network in a wearable cough detection system," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2015, pp. 1–4.

[23] Hui-Hui Wang, Jia-Ming Liu, Mingyu You, and Guo-Zheng Li, "Audio signals encoding for cough classification using convolutional neural networks: A comparative study," in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2015, pp. 442–445.

[24] Yunus Emre Erdoğan and Ali Narin, "Covid-19 detection with traditional and deep features on cough acoustic signals," *Computers in Biology and Medicine*, vol. 136, pp. 104765, 2021.

[25] Gauri Deshpande and Björn Schuller, "An overview on audio, signal, speech, and language processing for covid-19," *arXiv preprint arXiv:2005.08579*, 2020.

[26] Abdelfatah Hassan, Ismail Shahin, and Mohamed Bader Alsabek, "Covid-19 detection system using recurrent neural networks," in *2020 International conference on communications, computing, cybersecurity, and informatics (CCCI)*. IEEE, 2020, pp. 1–5.

[27] Harry Coppock, Alex Gaskell, Panagiotis Tzirakis, Alice Baird, Lyn Jones, and Björn Schuller, "End-to-end convolutional neural network enables covid-19 detection from breath and cough audio: a pilot study," *BMJ innovations*, vol. 7, no. 2, 2021.

[28] Juan Carlos Gabaldon-Figueira, Joe Brew, Dominique Hélène Doré, Nita Umashankar, Juliane Chaccour, Virginia Orrillo, Lai Yu Tsang, Isabel Blavia, Alejandro Fernández-Montero, Javier Bartolomé, Simon Grandjean Lapierre, and C Chaccour, "Digital acoustic surveillance for early detection of respiratory disease outbreaks in spain: a protocol for an observational study," *BMJ Open*, vol. 11, no. 1, 2021, Accessed: 2024-09-19.

[29] Maximilian Boesch, Frank Rassouli, Florent Baty, Anja Schwärzler, Sandra Widmer, Peter Tinschert, Iris Shih, David Cleres, Filipe Barata, Elgar Fleisch, and Martin H. Brutsche, "Smartphone-based cough monitoring as a near real-time digital pneumonia biomarker," *European Respiratory Society Open Research*, vol. 9, no. 1, 2023, Accessed: 2024-09-19.

[30] NHS Transformation Directorate, "Remote monitoring for early detection of respiratory conditions," *NHS AI Lab*, 2020, Accessed: 2024-09-19.

[31] Juan C. Gabaldón-Figueira, Eric Keen, Gerard Giménez, Virginia Orrillo, Isabel Blavia, Dominique Hélène Doré, Nuria Armendáriz, Juliane Chaccour, Alejandro Fernandez-Montero, Javier Bartolomé, Nita Umashankar, Peter Small, Simon Grandjean Lapierre, and Carlos Chaccour, "Acoustic surveillance of cough for detecting respiratory disease using artificial intelligence," *European Respiratory Society Open Research*, vol. 8, no. 1, 2021, Accessed: 2024-09-19.

[32] Michael Crooks, Yvette Hayman, Albertus den Brinker, Peter Hill, and Alyn Morice, "Daily remote cough monitoring in COPD," *European Respiratory Journal*, vol. 44, no. Suppl 58, pp. P4008, 2014.

[33] "Precision medicine and ai monitoring in home-based COPD management," *Medical Economics*, 2023, Accessed: 2023-09-19.

[34] Diego Sales-Lerida, Blanca Priego-Torres, and Antonio León-Jiménez, "Cough detection using acceleration signals and deep learning techniques," *Electronics*, vol. 13, no. 12, pp. 2410, 2024.

[35] "Remote cough monitoring in COPD," *Hyfe AI*, 2023, Accessed: 2023-09-19.

[36] B. Ruyobeza, S.S. Grobbelaar, and A. Botha, "Hurdles to developing and scaling remote patients' health management tools and systems: a scoping review," *Systematic Reviews*, vol. 10, 2021, Accessed: 2024-09-19.

[37] McKinsey and Company, "Digital tools could boost efficiency in african health systems," *McKinsey and Company*, 2021, Accessed: 2024-09-19.

[38] Maram Almufareh, Fatima Al-Quayed, Sulaiman Alateyah, and Mohammed Alatiyyah, "Improving healthcare facilities in remote areas using cutting-edge technologies," *Applied Sciences*, vol. 13, no. 11, pp. 6479, 2023.

[39] Authors unknown, "Remote health solutions: Bridging gaps in rural healthcare," *Tenovi*, 2022, Accessed: 2024-09-19.

[40] AK Tripathy, AG Mohapatra, SP Mohanty, E Kougianos, AM Joshi, and G Das, "Easyband: a wearable for safety-aware mobility during pandemic outbreak," *IEEE Consumer Electronics Magazine*, vol. 9, pp. 57–61, 2020.

[41] IDE Morais, B Filho, G Aquino, R Malaquias, G Girão, and S Melo, "An iot-based healthcare platform for patients in icu beds during the covid-19 outbreak," *IEEE Access*, vol. 1, pp. 1–171, 2021.

[42] SK Sood and I Mahajan, "Wearable iot sensor-based healthcare system for identifying and controlling chikungunya virus," *Computers in Industry*, vol. 91, pp. 33–44, 2017.

[43] S Ketu and PK Mishra, "Enhanced gaussian process regression-based forecasting model for covid-19 outbreak and significance of iot for its detection," *Applied Intelligence*, vol. 51, pp. 149–160, 2021.

[44] Muhammad Tukur, Ghassan Saad, Fahad M AlShagathrh, Mowafa Househ, and Marco Agus, "Telehealth interventions during covid-19 pandemic: a scoping review of applications, challenges, privacy and security issues," *BMJ Health and Care Informatics*, 2022, Accessed: 2024-09-19.

[45] Frédéric Bimbot, Jean-François Bonastre, Corinne Fredouille, Guillaume Gravier, Ivan Magrin-Chagnolleau, Sylvain Meignier, Teva Merlin, Javier Ortega-García, Dijana Petrovska-Delacrétaz, and Douglas A Reynolds, "A tutorial on text-independent speaker verification," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, pp. 1–22, 2004.

[46] Aaron E Rosenberg, "Automatic speaker verification: A review," *Proceedings of the IEEE*, vol. 64, no. 4, pp. 475–487, 1976.

[47] Douglas A Reynolds, Thomas F Quatieri, and Robert B Dunn, "Speaker verification using adapted gaussian mixture models," *Digital signal processing*, vol. 10, no. 1-3, pp. 19–41, 2000.

[48] Michal Muszynski, Jeffrey Okyere, Ruchi Mahindru, and Thomas Brunschwiler, "Cough diary based on sound classification, source validation and event detection," in *2022 IEEE 10th International Conference on Healthcare Informatics (ICHI)*, 2022, pp. 143–150.

[49] Matt Whitehill, Jake Garrison, and Shwetak Patel, "Whosecough: In-the-wild cougher verification using multitask learning," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 896–900.

[50] Stefan Jokić, David Cleres, Frank Rassouli, Claudia Steurer-Stey, Milo A. Puhan, Martin Brutsche, Elgar Fleisch, and Filipe Barata, "Tripletcough: Cougher identification and verification from contact-free smartphone-based audio recordings using metric learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2746–2757, 2022.

[51] Mingyu You, Weihao Wang, You Li, Jiaming Liu, Xianghuai Xu, and Zhongmin Qiu, "Automatic cough detection from realistic audio recordings using c-bilstm with boundary regression," *Biomedical Signal Processing and Control*, vol. 72, pp. 103304, 2022.

[52] Madhurananda Pahar, Marisa Klopper, Byron Wp Reeve, Robin Warren, Grant Theron, Andreas H. Diacon, and Thomas R. Niesler, "Wake-cough: cough spotting and cougher identification for personalised long-term cough monitoring," *2022 30th European Signal Processing Conference (EUSIPCO)*, pp. 185–189, 2021.

[53] Korosh Vatanparvar, Ebrahim Nemati, Viswam Nathan, Md Mahbubur Rahman, and Jilong Kuang, "Coughmatch – subject verification using cough for personal passive health monitoring," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2020, pp. 5689–5695.

[54] Helia Mohammadi, Ali-Akbar Samadani, Catriona Steele, and Tom Chau, "Automatic discrimination between cough and non-cough accelerometry signal artefacts," *Biomedical Signal Processing and Control*, vol. 52, pp. 394–402, 2019.

[55] Madhurananda Pahar, Igor Miranda, Andreas Diacon, and Thomas Niesler, "Automatic non-invasive cough detection based on accelerometer and audio signals," *Journal of Signal Processing Systems*, vol. 94, no. 8, pp. 821–835, 2022.

[56] Takehiro Otoshi, Tatsuya Nagano, Shintaro Izumi, Daisuke Hazama, Naoko Katsurada, Masatsugu Yamamoto, Motoko Tachihara, Kazuyuki Kobayashi, and Yoshihiro Nishimura, "A novel automatic cough frequency monitoring system combining a triaxial accelerometer and a stretchable strain sensor," *Scientific Reports*, vol. 11, no. 1, pp. 9973, 2021.

[57] Shibo Zhang, Ebrahim Nemati, Minh Dinh, Nathan Folkman, Tousif Ahmed, Mahbubur Rahman, Jilong Kuang, Nabil Alshurafa, and Alex Gao, "Coughtrigger: Earbuds imu based cough detection activator using an energy-efficient sensitivity-prioritized time series classifier," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 1–5.

[58] Tamer Elfaramawy, Cheikh Latyr Fall, Martin Morissette, François Lellouche, and Benoit Gosselin, "Wireless respiratory monitoring and coughing detection using a wearable patch sensor network," in *2017 15th IEEE international new circuits and systems conference (NEWCAS)*. IEEE, 2017, pp. 197–200.

[59] Thomas Drugman, Jerome Urbain, Nathalie Bauwens, Ricardo Chessini, Carlos Valderrama, Patrick Lebecque, and Thierry Dutoit, "Objective study of sensor relevance for automatic cough detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 699–707, 2013.

[60] Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2, Springer, 2009.

[61] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al., "Recent advances in convolutional neural networks," *Pattern recognition*, vol. 77, pp. 354–377, 2018.