Evaluating the Effectiveness of Inhaler Use Among COPD Patients via Recording and Processing Cough and Breath Sounds from Smartphones

Anthony Windmon[1(✉)], Sriram Chellappan[1], and Ponrathi R. Athilingam[2]

[1] Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620, USA
{awindmon,sriramc}@usf.edu
[2] College of Nursing, University of South Florida, Tampa, FL 33612, USA
pathilin@usf.edu

Abstract. Chronic Obstructive Pulmonary Disease (COPD) is a major health concern for elders today. Chronic cough and wheezing, which occur in the lungs as a result of mucus buildup are the main symptoms of COPD. COPD patients are advised to regularly medicate themselves via an inhaler, which delivers medicine to the lungs to break down mucus and relieve wheezing. Unfortunately, many patients do not use their inhaler devices correctly, resulting in no improvement of COPD symptoms, and worsened health. In this paper, we design machine learning (Support Vector Machine) algorithms operating on Mel-frequency Cepstral Coefficients of cough and breath sounds of patients (recorded via smartphones before and after inhaler usage) to detect the effectiveness of inhaler usage. Using a cohort of 55 clinically diagnosed COPD patients, spread across both genders, we evaluate our system from multiple metrics, including Precision, Recall, Sensitivity and Specificity. Our system achieved accuracies close to 80% in detecting effectiveness of inhaler usage. Our proposed system can aid COPD patients in improved selfcare routines, and also reduce the rate of re-hospitalizations caused by exacerbated symptoms.

Keywords: COPD · Lungs · Health · Cough · Breath · Machine learning · Smartphones · Aging

1 Introduction

Chronic Obstructive Pulmonary Disease (COPD) is the fourth leading cause of death worldwide, and is estimated to become the third by 2020 [1]. The prevalence of COPD in the US alone is 24 million patients today. However it is estimated that the number is actually much higher, due to millions not diagnosed yet, but who are living with impaired lung functions [2]. The most common cause of COPD is smoking, accounting for 85%

J. Liu et al. (Eds.): MobiCASE 2020, LNICST 341, pp. 102–120, 2020.
https://doi.org/10.1007/978-3-030-64214-3_7

of cases, while occupational smoke/dust and genetic factors are responsible for COPD in the remaining 15% of the population today [2].

Chronic cough and wheezing from the lungs are main symptoms of COPD, due to excess mucus production. Pharmacological therapy for COPD includes regular self-use of an inhaler (to deliver medicine directly to the lungs to breakdown mucus), and is validated in several clinical trials [3, 4]. However, it is a fact that a significant percentage of patients engage in sub-optimal inhaler techniques during self-care [5], which, as a consequence does not breakdown mucus enough, leading to worsened symptoms/health, and sometimes re-hospitalizations.

Our Contributions: In this paper, we propose an in-home smartphone based system to enable a COPD patient to determine the effectiveness of inhaler use via processing cough and breath sounds. To do so, we recruited a cohort of 55 clinically diagnosed COPD patients, spread across both genders[1]. Each subject was asked to cough and take deep breaths before inhaler use (to detect presence of mucus) and after correct inhaler use (to detect break-up of mucus and symptoms improvement). All data was recorded via a smartphone. After removing noise, three experts (one of them, our third co-author) listened to each audio segment recorded, to classify the cough and breath sounds as symptomatic of COPD (i.e., excess mucus build-up) or otherwise (i.e., mucus breakup due to correct inhaler use). After appropriate pre-processing, a total of 430 s of cough audio, and 1161 s of breath audio were obtained, evenly spread before and after inhaler use.

From this audio dataset, we then extracted Mel-frequency Cepstral Coefficients (MFCC) for post-processing. Very briefly, the mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of an audio signal, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. The Mel-frequency cepstral coefficients (MFCC) are those that together make up an MFC. MFCC provides a robust feature set for our classification problem (i.e., assess improvement in cough and breath due to inhaler use), as this feature performs best at capturing the spectral envelope of cough and breath sounds. As we present later in the paper, the spectral envelope is a critical component in audio signal processing that best captures features unique to sounds like cough and breath.

We then designed a spectrum of machine learning algorithms to process the MFCC extracted in order to classify cough and breath of COPD patients. Specifically, we want to discern those cough and breath sounds that indicate absence of mucus (with correct inhaler use) compared to sounds that indicate presence of mucus. It is easy to see that if we are successful in achieving our goal, quick feedback can be given to patients indicating either a) they are correctly using the inhaler; or b) they are incorrectly using the inhaler (while also directing them to tutorials on correct inhaler use).

To address our goal, we found that a Support Vector Machine (SVM) algorithm performed the best among k-Nearest Neighbors, Random Forests, and Logistic Regression in terms of standard metrics like Precision, Recall, Sensitivity and Specificity. Our overall accuracies were close to 80% for both cough and breath. We believe that our

[1] The study was approved by the University of South Florida's Institutional Review Board (IRB). **IRB Reference Number: Pro00035013.**

paper is the first to actually design an in-home smartphone based system to record and process cough & breath to evaluate the effectiveness of inhaler use. We expect our system to have significant value to educate patients, improve health outcomes and reduce re-hospitalization rates in relation to COPD.

The remainder of this paper is organized as follows: Sect. 2 covers related work. Section 3 explains our data collection procedures. Section 4 elaborates on data processing, and Sect. 5 details our feature extraction and classification techniques. Section 6 presents evaluation results, and, finally, we conclude and discuss future works in Sect. 7.

2 Related Work

In this section, we elaborate upon important related work within the scope of this paper.

Analyzing Cough for Healthcare: There are a number of studies in the recent past that process cough for healthcare. Researchers collected cough sounds from 38 subjects, 17 with tuberculosis and 21 without, to develop an algorithm capable of detecting symptoms of Tuberculosis [6]. Cough samples were recorded, voluntarily, from both infected and uninfected patients, using a Tascam DR44WL hand-held audio recording device with a 44100 Hz sampling rate. The features processed were log spectral energies and MFCC (similar to our system in this paper), and the classifier combines decision trees and logistic regression methods. For this problem, the authors achieved an 82% accuracy, 95% sensitivity, 72% specificity, and an area under the curve (AUC) score of 0.95 [6].

There are also systems that process audio signals to *only* detect cough. Related works in this space are [7] and [8], where data is collected using smartphones and close to 1000 s of cough served as the training dataset. Using machine learning techniques like k-NN and SVM, accuracies close to 90% were achieved in classification. In another paper, a slightly more complex problem - namely detection of respiratory events (cough, sneezing, throat-clearing and sniffling) is addressed [9]. Data pertaining to these events was recorded from a cohort of 16 subjects using a smartphone, unobtrusively for six weeks. This technique is a multi-layered SVM approach that processes several time and frequency domain features. In this paper, researchers achieved an accuracy of 82% for detection of respiratory events, and 99.1% for detection of non-respiratory events [9]. More recently, and with advances in deep learning via neural networks, there are some works that design convolutional neural network (CNN) based techniques in the domain of cough detection. Work in this domain include [10, 11] and [12]. In these papers though, the number of subjects recruited was relatively small (ranging from only 9 to 14), and accuracies close to 95% were achieved.

While these are all important related work, we point out that the mere detection of cough was the problem of interest here, and not finer grained classification of cough for a specific health condition (i.e. COPD) as we do in this paper.

Breath Analysis for Healthcare: Processing breath sounds is important for healthcare. Machine Learning algorithms have been designed and implemented to analyze breathing techniques to differentiate patients with lung cancer from healthy patients, and from a mixed group of patients with other lung diseases (i.e., COPD, asthma, pneumonia, etc.)

in works like [13]. Also, systems to detect various phases on breathing (without a specific disease context) have been developed in works like [14]. Other related work in the space of breath detection is [15], where algorithms are devised to measure lung function, including exacerbation, by processing breath sounds via a spirometer connected to a smart-phone. There are also many papers related to processing breath sounds for sleep and exercise detection, which we do not elaborate here due to space limitations.

Our Prior Work: We have done prior work in the space of designing AI techniques to process cough. In our first paper [16], we designed a classification system to discriminate cough indicative of COPD symptoms, from cough collected from healthy subjects that do not have COPD, wherein the cough was recorded via smartphones. In our second paper [17], we expanded our system by enabling it detect COPD and CHF (Congestive Heart Failure) from normal cough. In this paper, we have similar, but still orthogonal goals, in that we are now attempting to detect specifically the effectiveness of inhaler use by monitoring and analyzing cough and breath sounds before and after its use. To the best of our knowledge, this is a unique problem not yet attempted in the literature, but one that has big impact.

A Note on Sample Sizes: We wish to point out that the process of collecting data (elaborated in the next section) was difficult. We could only recruit 55 patients with COPD that met our criteria for recruitment (including age, gender, mental health conditions, approved for inhaler use etc.), and it was a nine month effort. Many patients did not consent to our study, and it is normal to do so. As such, in this paper, we do not attempt deep learning (i.e., featureless) techniques, due to non-availability of truly big data. But we are confident that our machine learning techniques in this paper are rigorous. Attempting deep learning techniques is part of our future work with much more data collected.

3 Data Collection

3.1 Recruitment of Subjects with COPD

During Spring and Summer 2019, we collaborated with respiratory therapists at Tampa General Hospital in Downtown Tampa, FL. With their assistance, we identified 55 (34 Female and 21 Male) clinically diagnosed COPD patients. Each subject was asked to sign an Institutional Review Board (IRB) approved consent form, indicating their willingness to participate in our study. Additionally, subjects were asked to provide their demographic information (i.e. age, gender, martial status, etc.), documented in Table 1, and to complete the COPD ABC and Leicester Cough Questionnaires. The COPD ABC Questionnaire measures the burden of COPD [22], and the Leicester Cough Questionnaire assesses the impact of cough on various aspects of life (i.e., personal, professional, etc.) [23].

3.2 Our Procedure for Recording Cough and Breath Sounds

Cough and breath sounds were recorded using a custom application developed by the authors. This application was installed onto a Motorola Moto E SmartPhone device, containing Android version 4.4.4 KitKat, recording at a sampling rate of 44100 Hz and bit rate of 16 (per second). This sample and bit rate are standard [34] and recommended [55].

Table 1. Demographic information of subjects.

Description	Category		Data
		(N=55)	62.09 ±
	Mean Std. Deviation	12.54	
	Female	34 (61%)	
	Male	21 (39%)	
	Married	16 (29.6%)	
	Never Married	11 (18.5%)	
	Windowed, Divorced or Separated		28 (51.9%)
Education:	<High School		11) (18.5%
	High School or GED	16 (29.6%)	
	Some College	12 (22.2%)	
Mean of COPD	Degree or Professional	16 (29.6%)	
	Current Smoker	9 (16.7%)	29.5
Score: Burden	Score on COPD ABC Questionnaire		5.11 ±
Mean COPD Severity Score:	Score on Leicester Cough Questionnaire		8.57 ± 3.59

We collected *four* samples from each subject: (1) cough and (2) breath sounds before inhaler use, and (3) cough and (4) breath sounds after inhaler use. To collect cough, each subject would simply cough into the phone's microphone via our app. However, to collect breath sounds (i.e., wheezing), we developed a custom recorder using the diaphragm of an actual stethoscope. Connected it to is an Audio-Technica ATR-3350IS Omnidirectional Condenser Lavalier microphone, which records breath sound waves as audio files. The microphone, shown in Fig. 1, connects to our smart-phone via wire and is available in the market.

The process of collecting data required care. Initially, *before* each subject was administered their inhaler medication, we recorded a sample of their cough. To do so, each subject was asked to cough directly into the microphone of our smartphone in which the sound was recorded. Recording breath required a series of steps. First, we started off using an actual stethoscope to manually listen to each subject's wheezing sounds in their lungs, which indicate quality of breath sounds. The objective was to locate the clearest wheezing sounds being projected. Most often, wheezing sounds are best heard on four different areas of the subject's front, mid-to-lower chest, where the lungs are located. Additionally, wheezing sounds can be heard on eight different areas of the subject's mid-to-lower back. The best areas of auscultation, as depicted in Fig. 2, vary by subject. Once we identify the best location where wheezing was heard, we used our

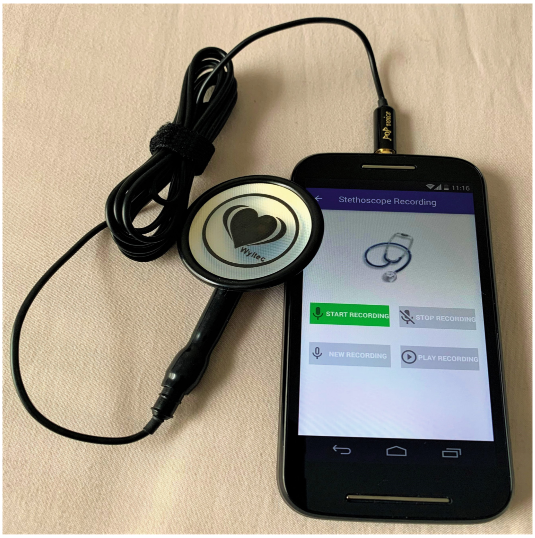

Fig. 1. The Audio-Technica ATR-3350IS omnidirectional condenser lavalier microphone (with heart shape on surface) connected to the Motorola Moto E smart-phone device to collect breath sounds from COPD patients. The custom mobile application is shown on the screen of the smart-phone.

stethoscope, Audio-Technica microphone and mobile application to record and store the clearest wheezing sound from that location for that subject. All recorded data was appropriately labeled (without any identifiable information).

In the second round of recording, the subject was required to take their inhaler medication, which was correctly administered by a respiratory therapist. About five minutes *after* inhaler use, the exact same cough and breath recording process was executed, and all data was labeled. This process was repeated for all subjects. Since all patients used the inhaler correctly, their cough and breath sounds sounded differently due to mucus break down after inhaler use (which can be gleaned by a trained ear). Automating the classification via this ground-truth data is our problem.

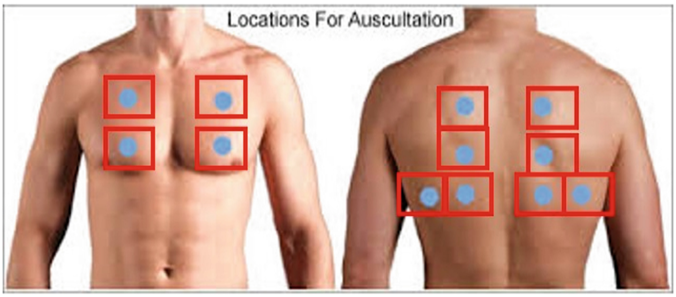

Fig. 2. Locations for auscultation, where wheezing sounds are heard best. Front, mid-to-lower chest (left), mid-to-lower back (right).

Table 2. Number of seconds (sec.) for cough and breath classes, before (BP) and after (AP) data pre-processing.

Class	Sec. (BP)	Sec. (AP)
Cough before inhaler use	289 s	219 s
Breath before inhaler use	627 s	579 s
Cough after inhaler use	261 s	211 s
Breath after inhaler use	632 s	582 s

4 Data Pre-processing

4.1 Cough and Breath Data Before Pre-processing

In our data collection procedures, cough sounds, on an average, lasted about 7 s each. Breath sounds, on the contrary, were collected for approximately one minute each, consisting of 5–7 deep breaths. Using this data, we developed four different classes: cough before inhaler use, breath before inhaler use, cough after inhaler use and breath after inhaler use. As shown in Table 2, the second column contains the duration of raw data before any pre-processing (denoted as BP). Data in the third column is after pre-processing (denoted as AP), and is discussed next.

4.2 Pause and Noise Removal

First, we identified unintentional pauses in our audio files post recording, and cut them using an audio cutting application [17]. Next, we applied a band-pass filter to lower background ambient noise caught in our cough data (i.e., noise due to surrounding conversations, medical equipment, etc.), with cut-off frequencies as 300 Hz and 1200 Hz because cough lies in-between those frequencies [20, 21]. There was little to no noise in our breath data, because the stethoscope was placed tightly on the subject's skin for recording.

4.3 One Second Windowing Algorithm

The next issue is accurate ground-truthing of data collected. This is a little tricky. Even if a patient is clinically diagnosed with COPD and chronic cough, it is still not the case that the entire cough episode indicates symptoms of COPD. It often happens that only a certain subset of their entire cough episode indicates COPD symptoms. Naturally, with proper inhaler use, subsequent cough segments will be completely devoid of COPD symptoms. Fortunately, the third co-author of this paper has decades of experience working with COPD patients, and she indicated that a one second segment of cough can indicate presence of COPD symptoms to a trained human ear. The same is true for breath also. As such, we split our entire dataset of cough and breath (after pause and noise removal) from all patients into one second segments and our third co-author, joined by other experienced nurses, listened to each cough and breath segment to tag it as indicative of concerning COPD symptoms (that necessitated inhaler intervention for that episode of cough/breath) or very mild/no discernible COPD symptoms (that does not necessitate any further inhaler use for that episode).

As a result of these procedures, we generated a new dataset containing 219 s of cough and 579 s of breath audio that indicated COPD symptoms necessitating inhaler intervention; and 211 s of cough and 582 s of breath audio that demonstrated improved COPD conditions not warranting any further inhaler use for that episode. This data is shown in the third column in Table 2. This dataset is what we train and validate our machine learning algorithms on.

5 Feature Extraction & Classification Algorithms

We now elaborate on extraction of Mel Frequency Cepstral Coefficients (MFCC) from our dataset as features, and our Support Vector Machine based algorithm for classification.

5.1 Mel Frequency Cepstral Coefficients

For sound recognition systems, a primary goal is to classify a sound (i.e., speech, singing, breath, cough, etc.) as produced by a human. Human sounds are produced via the larynx (voice box) and vibrations of vocal cords. This sound is then filtered by their vocal tract, which determines how the sound produced is, *both*, shaped and ejected from the mouth. The vocal tracts for a human consists of the lips, nose, tongue, teeth and throat areas [43], as depicted in Fig. 3. The corresponding shape of the sound is defined within the envelope of the short time power spectrum, which estimates loudness and timbre, also shown in Fig. 3. The MFCC is the strongest audio feature capable of accurately defining that envelope [26, 35], which inturn serves as robust features to classify sounds like cough and breath, since the shape of the vocal tract defines how these sounds emanate [39–41]. The MFCC accomplishes this by generating Cepstral Coefficients. The calculation to generate the Cepstral Coefficients, depicted for each of our four classes in Fig. 4, is explained below and mapped out in Fig. 5.

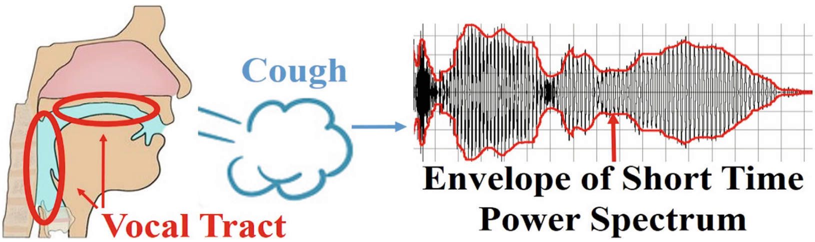

Fig. 3. Sound type (i.e. sound of cough or breath) is produced by the shape of the vocal tract. This shape is defined within the envelope of the Short Time Power Spectrum, which is best characterized by cepstral coefficients of the MFCC.

Computing MFCC Features: First, each audio (i.e., cough or breath) signal h is split into a small number of frames of duration 20 ms (ms). If the frame is shorter we do not have enough samples to get a reliable spectral estimate. If it is longer, the signal changes too much throughout the frame, making it highly non-stationary [26]. Since, our sampling rate to record the cough and breath audio signals is 44100 Hz, the frame length s of each audio signal is now 0.020 * 44100 = 882 samples. Next, from each frame, we extract one set of 13 MFCC coefficients.

To do so, we denote our pre-framed audio signal as $h(s)$, and our framed audio signal as $h_i(s)$. Then, we calculate the Discrete Fourier Transform (DFT) $D_i(k)$ of each i^{th} frame as,

$$D_i(k) = \sum_{s=0}^{S-1} h_i(s) w(s) x(s) e^{\frac{2\pi ks}{S}}. \quad (1)$$

Here, $w(s)$ represents the length of the hamming window function, where $w(s) = 0.54 - 0.46\cos(\pi s/S)$. S is the length of the discrete-time signal $x(s)$, which represents the quantity of signals in s. Also, k is the sampling frequency of the DFT, where $k = 0,1 \ldots S - 1$.

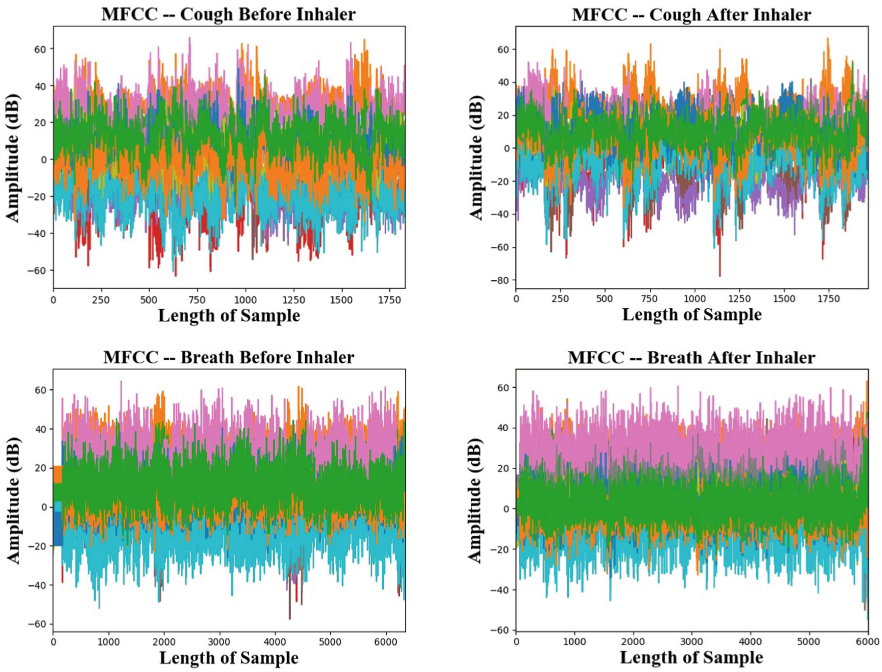

Fig. 4. MFCC features (cepstral coefficients) of before & after cough, and before & after breath samples, which are represented in the shape of the envelope of the short time power spectrum in Fig. 3. These images represent the lower 13 MFCC coefficients, which contain the highest quantity of information about the overall spectral shape produced by cough and breath sounds. For cough, there's a difference in amplitude due to reduction of mucus build-up after inhaler use. For breath, there's a difference in amplitude consistency, also due to reduction in mucus build-up after inhaler use. This figure is best viewed in color.

We then compute the Periodogram Estimate $M_i(k)$ as,

$$M_i(k) = \frac{1}{S}|D_i(k)|^2, \tag{2}$$

to identify which frequencies are present in each frame, and decipher cough and breath sound frequencies analogous to the human ear [26].

Next, to produce the Mel-Filter Bank, we applied a Triangular Filter, depicted in Fig. 6 [27]. The Triangular Filter, roughly, captures energy within the spectral envelope of a frequency bin. In other words, the filter provides an estimate of the given audio sample's spectral envelope shape. The Triangular Filter is applied on a Mel-Scale to the power spectrum. The Mel-Scale imitates the linear frequency of the human ear's perception to sound.

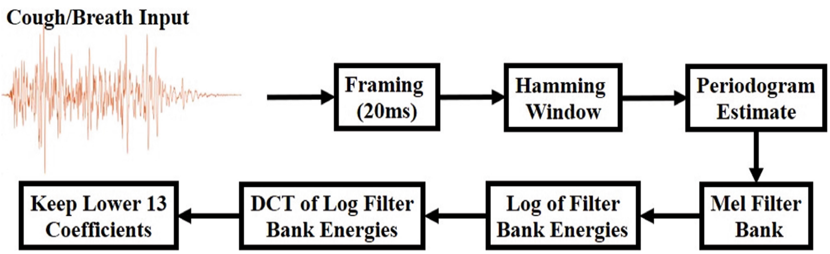

Fig. 5. The multitude of steps required to calculate the mel frequency cepstral coefficients (MFCC) for both cough and breath sounds.

The Mel-Filter Bank contains 26 vectors, and 257 coefficients. We multiple each Filter Bank by the power spectrum, calculated using Eq. 2, then add the coefficients. This results in 26 numbers, representing the amount of energy in each Filter Bank. The logarithm of these 26 numbers is calculated, which imitates what is heard by a human ear.

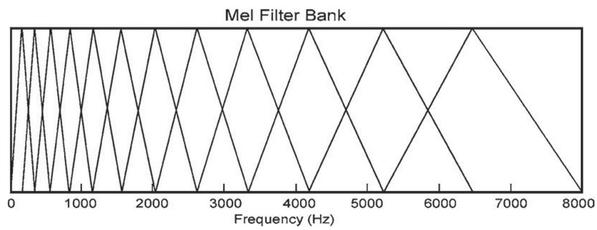

Fig. 6. Triangular filter applied to produce the mel filter bank [27].

Finally, we take the Discrete Cosine Transform (DCT) of the 26 log numbers. This results in 26 cepstral coefficients. We only keep the lower 13 coefficient, as these coefficients contain the strongest quality of information about the spectral envelope's shape [26]. We discard the higher coefficients because they represent fast changes in the

Mel-Filter Bank energies, which decrease cough and breath recognition performance. We see a small, but noticeable, increase in performance by dropping the higher coefficients.

After applying the DCT, the MFCC can be expressed as,

$$C_m = \sum\nolimits_{k=1}^{K} (logD_i(k))\left[m\left(k - \frac{1}{2}\right)\frac{\pi}{K}\right], \tag{3}$$

where C_m denotes the MFCC, $m = 1{,}2\ldots13$, which equates to the 13 MFCC coefficients, and $K = 44100$ which is the sampling rate of our recording device.

5.2 Justification for MFCC Audio Features

As of today, health-care professionals listen to the sound of a patient's cough and wheezing to determine the presence and severity of COPD [50]. The point of our study is to automate this process using machine learning processed on cough and breath. To do so, we need to capture the spectral envelope of cough and wheezing sounds. In the literature, it has been demonstrated various times that the MFCC audio features are most capable of capturing the spectral envelope [26, 43]. The MFCC does so by converting sounds to a Mel-Scale which analyses those sounds at frequencies that humans speak, and are capable of hearing. The MFCC is an ideal feature for our problem for the following reasons: (1) MFCC uses the mel-scale to analyze sound in a manner similar to humans [26]; (2) It has been successfully used in several similar applications related to cough [16, 20, 45, 46], breath [47], wheezing [48, 49], music [51] and speech recognition systems [52]; (3) MFCC is considered a classic front-end algorithm capable of significant and accurate performance in sound and speech recognition systems [43, 44]. Lastly, there are several studies that suggest a significant relationship between the MFCC audio feature and Support Vector Machine classification combination (which were both utilized in this study), when used in the domains of cough [12, 52, 53] and breath [48, 54] analysis.

Other audio features, like Spectral Centroid, Spectral Flatness or Spectral Flux, could have been used to solve this problem. However, these features do not capture the spectral envelope's perception as the MFCC does. Thus, we did not incorporate those features into this study. Furthermore, using a large number of features on our classification models can cause overfitting problems and also increased overhead. Hence, we stick with MFCC features alone for our problem, and are confident about our decision to do so.

5.3 Support Vector Machine

Based on MFCC features presented above, we briefly present our Support Vector Machine (SVM) based algorithm for classification, which performed the best among other techniques. Broadly speaking, SVM classifiers aim to find the best hyperplane between two classes. A hyperplane is a line which linearly separates the data points between the classes. In SVM, a hyperplane is considered "best" when it produces the largest margin between two classes. SVM uses Support Vectors, which are the classes' data points closest to the hyperplane, to calculate margin maximization [28]. Recall again that our problem is to identify improvement in symptoms before and after inhaler usage using cough and breath data in Table 2. We design two separate SVM classifiers to do so – one to process cough and the other to process breath.

Our classifiers for cough and breath were developed using the scikit-learn machine learning software, built into python programming language. The parameters for our SVM classification models, which produced best results were as follows:

Cough: degree of the Radial Basis Function (RBF) kernel function is 7, the cache size is 120, the random state is 4 and the kernel is linear, regularization = 1.0, tolerance for stopping criterion = 1e−3, class weight = balance, decision function shape = one-vs-rest, maximum number of iterations = −1 (default), gamma = "scale" and shrinking = True.

Breath: degree of the Radial Basis Function (RBF) kernel function = 5, the cache size = 215, the random state = 4, the kernel is linear, regularization = 1.0, tolerance for stopping criterion = 1e−3, class weight = balance, decision function shape = one-vs-rest, maximum number of iterations = −1 (default), gamma = "scale" and shrinking = True.

The RBF kernel was selected and works best for our study because (1) parts of our data has overlap making it difficult for SVM to find the right hyperplane to separate the data; (2) RBF provides better discriminative ability in a much higher dimensional subspace [28]; (3) RBF provided a much faster classification time in comparison to the Polynomial kernel.

6 Results

Classification results, presented below were measured using the following cross validation methods: 10-Fold (10-FCV) and Leave-One-Out (LOOCV). Metrics are Specificity, Sensitivity, Precision, Recall and F1-Score. Results of our SVM classifier, as well as other machine learning algorithms are recorded in Tables 3, 4, 5 and 6.

Cough: For our cough classification scheme, testing the cough before inhaler and cough after inhaler classes, we achieved the following results. Employing 10-Fold Cross Validation, we averaged an accuracy of 79.00%, precision of 81.00, recall of 81.00%, sensitivity of 84.52%, specificity 77.61% and a F1-score of 81.00%. Employing Leave-One-Out Cross Validation, we averaged an accuracy of 80.69%, precision of 80.00, recall of 80.00%, sensitivity of 83.45%, specificity 78.02% and a F1-score of 80.00%.

Breath: For classification, testing the breath before inhaler and breath after inhaler classes, we achieved the following results. Employing 10-Fold Cross Validation, we averaged an accuracy of 84.49%, precision of 84.00, recall of 83.00%, sensitivity of 83.00%, specificity 93.30% and a F1-score of 82.00%. Employing Leave-One-Out Cross Validation, we averaged an accuracy of 84.32%, precision of 83.00, recall of 83.00%, sensitivity of 82.00%, specificity 91.49% and a F1-score of 80.00%.

6.1 Comparison of Results Using Other Algorithms

Tables 3, 4, 5 and 6 show classification results using several popular machine learning approaches (i.e. k-Nearest Neighbors, Random Forests, Logistic Regression and Multilayer Perceptron). As shown, Support Vector Machine (SVM) provided the best results for the majority of metrics. This is because SVM works best for binary classification problems, and also works well with linearly separable data [28], such as our data. Figure 7 illustrates the Receiver Operating Characteristic (ROC) curves for classification performance based on cough and breath data, with SVM performing the best with Area under the Curve (AUC) scores close to 94% for cough and 93% for breath, when applying 10-Fold Cross Validation. When applying Leave-One-Out Cross Validation, the AUC scores are 87.5% for cough and 88.4% for breath.

6.2 Complexity of Execution

Our Support Vector Machine model, was compiled on a Windows 10 Dell PC, containing an Intel(R) Core(TM) $i7$-5600U processor, 2.60 GHz with 16 GB RAM. All processing (i.e. audio cutting/filtering, feature extraction and classification) was done with this system. MATLAB R2017b software was used for feature extraction, audio filtering and

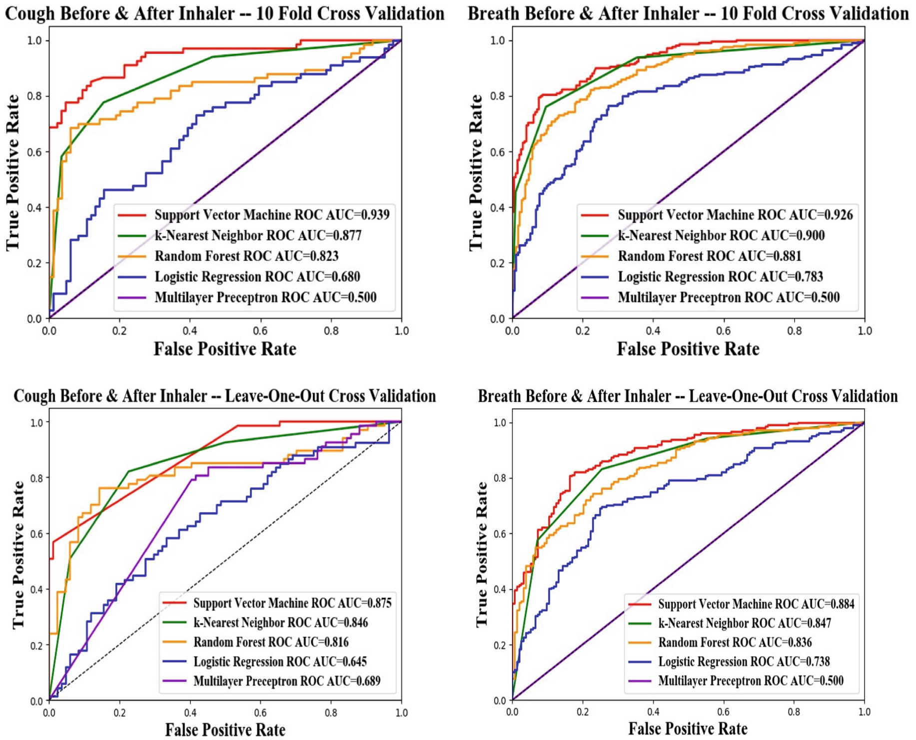

Fig. 7. Receiver Operating Characteristic (ROC) curves for cough classes and breath classes using 10-fold (top) and leave-one-out cross validation (bottom).

Table 3. Classification of cough based on 10-fold cross validation.

Algorithm	Accuracy precision recall sensitivity specificity F1-score			
Support vector machine 79.00%	81.00% 81.00%	84.52%	77.61%	81.00%
k-Nearest neighbors 77.00%	81.00% 81.00%	83.33%	77.61%	81.00%
Random forests 65.00%	76.00% 51.00%	92.00%	77.60%	48.00%
Logistic regression 66.00%	73.00% 73.00%	77.27%	69.04%	73.00%
Multilayer perceptron 62.00%	20.00% 44.00%	74.32%	54.66%	27.00%

Table 4. Classification of cough based on leave-one-out cross validation.

Algorithm	Accuracy precision recall sensitivity specificity F1-score			
Support vector machine 80.69%	80.00% 80.00%	83.45%	78.02%	80.00%
k-Nearest neighbors 79.06%	82.00% 82.00%	82.57%	79.00%	81.00%
Random forests 62.08%	74.00% 53.00%	89.00%	79.45%	46.00%
Logistic regression 67.45%	71.00% 71.00%	75.54%	68.65%	71.00%
Multilayer perceptron 58.85%	22.00% 43.00%	73.08%	52.27%	28.00%

Table 5. Classification of breath based on 10-fold cross validation.

Algorithm	Accuracy precision recall sensitivity specificity F1-score			
Support vector machine 84.49%	84.00% 83.00%	83.00%	93.30%	82.00%
k-Nearest neighbors 84.00%	84.00% 83.00%	76.00%	90.40%	83.00%
Random forests 80.00%	82.00% 81.00%	74.16%	87.87%	81.00%
Logistic regression 70.00%	72.00% 71.00%	77.27%	65.04%	71.00%
Multilayer perceptron 50.00%	23.00% 48.00%	46.56%	38.66%	31.00%

Table 6. Classification of breath based on leave-one-out cross validation.

Algorithm	Accuracy precision recall sensitivity specificity F1-score				
Support vector machine	84.32%	83.00% 83.00%	82.00%	91.49%	80.00%
k-Nearest neighbors	83.89%	82.00% 80.00%	73.00%	87.32%	81.00%
Random forests	82.25%	83.00% 80.00%	75.22%	86.32%	81.00%
Logistic regression	70.70%	73.00% 70.00%	78.01%	66.55%	72.00%
Multilayer perceptron	47.71%	25.00% 46.00%	44.52%	40.33%	34.00%

dataset construction. The custom mobile application used to record cough and breath samples, consumed a memory 4.55 MB on Motorola Moto E Smart-Phone containing 270 MB. Python 3.6 Sklearn libraries were used to develop all classifiers. The classification time during execution on the Windows 10 Dell PC was less than a minute. Ensuring all processing and classification to happen within the smartphone itself is part of future work.

7 Conclusions and Future Work

In this paper, we design a system capable of discerning improvements in cough and breath symptoms for COPD patients as a result of correct inhaler use (due to reduced mucus build-up). Data from 55 patients was used to developed our model. Our ultimate goal is to improve health of the individual and reduce costly re-hospitalizations, by detecting and recommending correct health procedures in-home. Our results are very favorable. We also believe that our work is the first to address the issue of designing algorithms to automatically evaluate the effectiveness of prescribed interventions (in this case, inhalers) for COPD.

In the future, we propose to integrate our algorithm as an AI based in-home care system for patients with COPD. With immediate feedback upon symptom monitoring, patients could be better notified of when to use an inhaler, and also offered tutorials on correct inhaler use which can definitely reduce re-hospitalizations. This system will also include the ability to inform patients when their cough symptoms have exacerbated enough for them to seek medical attention. Integrating these AI designs into current COPD care products in the market like [29–32, 37] is our strategy.

Also, we are aiming to conduct a longer study to design gender based algorithms, since clinical studies suggests that women experience significantly harsher COPD symptoms than men, throughout their lifespan [36]. While in this paper, we processed cough and breath separately, we are also looking into integrate both sounds and related features together to detect correctness of inhaler usage. Conducting longitudinal experiments to design personalized models for each patient is also part of our future work.

Acknowledgment. Support for Anthony Windmon was provided by the Florida Education Fund's McKnight Doctoral Fellowship Program. We thank registered nurses and scientists, respiratory therapists and participating patients at Tampa General Hospital for their support and cooperation.

References

1. Global Initiative for Chronic Obstructive Lung Disease (GOLD): GOLD 2020 global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, 2020 report. https://goldcopd.org/wpcontent/uploads/2019/12/GOLD-2020
2. Center for Disease Control and Prevention (CDC): Chronic Obstructive Pulmonary Disease: Data and Statistics, May 2018. https://www.cdc.gov/copd/data.html
3. Barnes, N., Calverley, P.M., Kaplan, A., Rabe, K.F.: Chronic obstructive pulmonary disease and exacerbations: clinician insights from the global Hidden Depths of COPD survey. Curr. Med. Res. Opin. **304**, 667–684 (2014). https://doi.org/10.1185/03007995.2013.867842

4. Dolovich, M.B., et al.: Device selection and outcomes of aerosol therapy: evidence-based guidelines: American College of Chest Physicians/American College of Asthma, Allergy, and Immunology. Chest **127**(1), 335–371 (2005). https://doi.org/10.1378/chest.127.1.335
5. Melani, A.S., et al.: Inhaler mishandling remains common in real life and is associated with reduced disease control. Respir. Med. **105**(6), 930–938 (2011). https://doi.org/10.1016/j.rmed.2011.01.005
6. Botha, G.H.R., et al.: Detection of tuberculosis by automatic cough sound analysis. Physiol. Meas. **39**(4), 045005 (2018). https://doi.org/10.1088/13616579/aab6d0
7. Barcelo, C.H., Alvarez, J.M., Shakir, M.Z., Alcaraz-Calero, J.M., Casasecade-la-Higuera, P.: Efficient k-NN implementation of real-time detection of cough events in smartphones. IEEE J. Biomed. Health Inf. **22**(5), 1662–1671 (2018). https://doi.org/10.1109/jbhi.2017.2768162
8. Monge-Alvarez, J., Barcelo, C.H., San-Jose-Revuelta, L.M., Casaseca-de-laHiguera, P.: A machine hearing system for robust cough detection based on a highlevel representation of band-specific audio features. IEEE Trans. Biomed. Eng. **66**(8), 2319–2330 (2019). https://doi.org/10.1109/tbme.2018.2888998
9. Sun, X., Lu, Z., Hu, W., Cao, G.: SymDetector: detecting sound-related respiratory symptoms using smartphones. In: UbiComp 2015: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 97–108, September 2015. https://doi.org/10.1145/2750858.2805826
10. Amoh, J., Odame, K.: Deep neural networks for identifying cough sounds. IEEE Trans. Biomed. Circ. Syst. **10**(5), 1003–1011 (2016). https://doi.org/10.1109/tbcas.2016.2598794
11. Kadambi, P., et al.: Towards a wearable cough detector based on neural networks. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2018. https://doi.org/10.1109/icassp.2018.8461394
12. Amoh, J., Odame, K.: DeepCough: a deep convolutional neural network in a wearable cough detection system. In: 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS), Atlanta, GA, pp. 1–4, September 2015. arXiv:1509.02512
13. Tirzite, M., Bukovskis, M., Strazda, G., Jurka, N., Taivans, I.: Detection of lung cancer in exhaled breath with an electronic nose using support vector machine analysis. J. Breath Res. **11**(3), 036009 (2017). https://doi.org/10.1088/17527163/aa7799
14. Yahya, O., Faezipour, M.: Automatic detection and classification of acoustic breathing cycles. In: Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education, Bridgeport, CT, pp. 1–5, May 2014. https://doi.org/10.1109/aseezone1.2014.6820648
15. Larson, E.C., Goel, M., Boriello, G., Heltshe, S., Rosenfeld, M., Patel, S.N.: SpiroSmart: using a microphone to measure lung function on a mobile phone. In: UbiComp 2012, pp. 280–289, September 2012. https://doi.org/10.1145/2370216.2370261
16. Windmon, A., Minakshi, M., Chellappan, S., Athilingam, P.R., Johansson, M., Jenkins, B.A.: On detecting COPD cough using audio signals recorded from smart-phones. In: Proceedings of 11th International Joint Conference on Biomedical Engineering Systems and Technologies: HEALTHINF, vol. 5, pp. 329–338, February 2018. https://doi.org/10.5220/0006549603290338
17. Windmon, A., et al.: TussisWatch: a smart-phone system to identify cough episodes as early symptoms of chronic obstructive pulmonary disease and congestive heart failure. IEEE J. Biomed. Health Inf. **23**(4), 1566–1573 (2019). https://doi.org/10.1109/jbhi.2018.2872038
18. Ramalho, G.L.B., Filho, P.P.R., Sombra de Medeiros, F.N., Cortez, P.C.: Lungdisease detection using feature extraction and extreme learning machine. Rev. Bras. Eng. Biomed. 30(3), 207–214 (2014). https://doi.org/10.1590/rbeb.2014.019
19. Waltisberg, D., Amft, O., Brunner, D.P., Troster, G.: Detecting disordered breathing and limb movement using in-bed force sensors. IEEE J. Biomed. Health Inf. **21**(4), 930–938 (2017). https://doi.org/10.1109/jbhi.2016.2549938

20. Korpas, J., Sadlonova, J., Vrabec, M.: Analysis of cough sound: an overview. Pulm. Pharmacol. **9**, 261–268 (1996). https://doi.org/10.1006/pulp.1996.0034
21. Chatrzarrin, H., Arcelus, A., Goubran, R.: Feature extraction for the differentiation of dry and wet cough sounds. In: Proceedings of IEEE International Symposium on Medical Measurements and Applications, pp. 162–166, July 2011. https://doi.org/10.1109/memea.2011.5966670
22. Slok, A.H.M., et al.: Development of the assessment of burden of COPD tool: an integrated tool to measure the burden of COPD. NPJ Prim. Care Respir. Med. **24**(1–4), 14021 (2014). https://doi.org/10.1038/npjpcrm.2014.21
23. Birring, S.S., Spinou, A.: How best to measure cough clinically. Curr. Opin. Pharmacol. **22**, 37–40 (2015). https://doi.org/10.1016/j.coph.2015.03.003
24. Muda, L., Begam M., Elamvazuthi, I.: Voice recognition algorithms using melfrequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. J Comput. **2**(3), 138–143 (2010). arXiv:1003.4083
25. Logan, B.: Mel frequency cepstral coefficients for music modeling. In: International Symposium on Music Information Retrieval (ISMIR) (2000). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.9216
26. Lyons, J.: Mel frequency cepstral coefficient (MFCC) tutorial (2012). http://practicalcryptography.com/miscellaneous/machinelearning/guide-mel-frequency-cepstral-coefficients-mfccs/
27. Nisar, S., Shahzad, I., Khan, M.A.: Pashto spoken digits recognition using spectral and prosodic based feature extraction. In: 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI), Doha, pp. 64–78, February 2017. https://doi.org/10.1109/icaci.2017.7974488
28. Cortes, C., Vapnik, V.: Support-vector networks. J. Mach. Learn. **20**(3), 273–297 (1995). https://doi.org/10.1023/a:1022627411411
29. Leidy, N.K., et al.: Development of the exacerbations of chronic obstructive pulmonary disease tool (EXACT): a patient-reported outcome (PRO) measure. Value Health: J. Int. Soc. Pharmacoecon. Outcomes Res. **13**(8), 965–975 (2010). https://doi.org/10.1111/j.1524-4733.2010.00772.x
30. Leidy, N.K., Murray, L.T.: Patient-reported outcome (PRO) measures for clinical trials of COPD: the EXACT and E-RS. COPD: J. Chron. Obstruct. Pulm. Dis. **10**(3), 393–398 (2013). https://doi.org/10.3109/15412555.2013.795423
31. Wac, K., Hausheer, D.: COPD24: from future internet technologies to health telemonitoring and teletreatment applications. In: 12th IFIP/IEEE International Symposium on Integrated Network Management and Workshops, Dublin, pp. 812–826, August 2011. 978-1-4244-9221-3
32. Ding, H., Moodley, Y., Kanagasingam, Y., Karunanithi, M.: A mobile-health system to manage chronic obstructive pulmonary disease patients at home. In: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, pp. 2178–81, September 2012. https://doi.org/10.1109/embc.2012.6346393
33. Smith, J.A., Ashurst, H.L., Jack, S., Woodcock, A.A., Earis, J.E.: The description of cough sounds by healthcare professionals. Cough **2**(1) (2006). https://doi.org/10.1186/1745-9974-2-1
34. Self, D.: Audio Engineering Explained, 1st edn, pp. 200 & 446. Focal Press, Burlington (2012). ISBN 0240812735. E-Book
35. Huang, X., Acero, A., Hon, H.: Spoken Language Processing: A Guide to Theory, Algorithm, and System Development, 1st edn. Prentice Hall, Upper Saddle River (2001). ISBN-10: 0130226165
36. DeMeo, D.L., et al.: Women manifest more severe COPD symptoms across the life course. Int. J. Chron. Obstruct. Pulm. Dis. **13**, 3021–3029 (2018). https://doi.org/10.2147/copd.s160270

37. Xu, W., et al.: mCOPD: mobile phone based lung function diagnosis and exercise system for COPD. In: ACM International Conference on Pervasive Technologies Related to Assistive Environments (PETRA 2013) Proceeding Series, Rhodes, Greece, May 2013. https://doi.org/10.1145/2504335.2504383
38. Kopparapu, S.K., Laxminarayana, M.: Choice of mel filter bank in computing MFCC of a resampled speech. In: 10th International Conference on Information Sciences, Signal Processing and their Applications (ISSPA, 2010), Kaula Lumpur, pp. 121–124, May 2010. https://doi.org/10.1109/isspa.2010.5605491
39. French, P., Foulkes, P., Harrison, P., Hughes, V., Segundo, E.S., Stevens, L.: The vocal tract as a biometric: output measures, interrelationships, and efficacy. In: Proceedings of ICPhS, University of Glasgow, vol. 18 (2015). ISBN 978-0-85261-941-4
40. Serizel, R., Giuliani, D.: Vocal tract length normalisation approaches to DNN based children's and adults' speech recognition. In: 2014 IEEE Spoken Language Technology Workshop (SLT), South Lake Tahoe, NV, pp. 135–140, April 2015. https://doi.org/10.1109/slt.2014.7078563
41. Singh, N., Khan, R.A., Shree, R.: MFCC and prosodic feature extraction techniques: a comparative study. Int. J. Comput. Appl. **54**(1), 9–13 (2012). https://doi.org/10.5120/8529-2061
42. Cassidy, R.J.: Audio Speech Research Note. Stanford University, July 2003. https://ccrma.stanford.edu/rjc/pubs/audio-speech/audio-speech.pdf
43. Terasawa, H., Berger, J., Makino, S.: In search of a perceptual metric for timbre: dissimilarity judgments among synthetic sounds with MFCC-derived spectral envelopes. J. Audio Eng. Soc. **60**(9). 674–685 (2012). http://www.aes.org/elib/browse.cfm?elib=16372
44. Rabiner, L., Juang, B.H.: Fundamentals of Speech Recognition, 1st edn, pp. 183–190. Pearson, London (1993). ISBN-13: 978-0130151575
45. Liu, J.M., et al.: Cough signal recognition with gammatone cepstral coefficients. In: 2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, China, pp. 160–164, July 2013. https://doi.org/10.1109/chinasip.2013.6625319
46. Pramono, R.X.A., Imtiaz, S.A., Rodriguez-Villegas, E.: 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), Berlin, Germany, October 2019. https://doi.org/10.1109/embc.2019.8857792
47. Abushakra, A., Faezipour, M.: Acoustic signal classification of breathing movements to virtually aid breath regulation. IEEE J. Biomed. Health Inform. **17**(2), 493–500 (2013). https://doi.org/10.1109/jbhi.2013.2244901
48. Palaniappan, R., Sundaraj, K., Lam, C.K.: Reliable system for respiratory pathology classification from breath sound signals. In: 2016 International Conference on System Reliability and Science, Paris, pp. 152–156, January 2017. https://doi.org/10.1109/icsrs.2016.7815855
49. Lin, B.-S., Lin, B.-S.: Automatic wheezing detection using speech recognition technique. J. Med. Biol. Eng. **36**(4), 545–554 (2016). https://doi.org/10.1007/s40846-016-0161-9
50. Smith, J.A., Ashurst, H.L., Jack, S., Woodcock, A.A., Earis, J.E.: The descriptionof cough sounds by healthcare professionals. Cough **2**(1), 1–9 (2006). https://doi.org/10.1186/1745-9974-2-1
51. Jensen, J.H., Christensen, M.G., Murthi, M.N., Jensen, S.H.: Evaluation of MFCC estimation techniques for music similarity. In: 14th European Signal Processing Conference, Florence, Italy, September 2006. http://www2.imm.dtu.dk/pubdb/edoc/imm4413.pdf
52. Mohan, B.J., Babu, R.N.: Speech recognition using MFCC. In: International Conference on Advances in Electrical Engineering (ICAEE), Vellore, India, pp. 1–4, January 2014. https://doi.org/10.1109/icaee.2014.6838564

53. Shi, Y., Liu, H., Wang, Y., Cai, M., Xu, W.: Theory and application of audio based assessment of cough. Hindawi J. Sens. **2018**, 1–8 (2018). Article ID 9845321. https://doi.org/10.1155/2018/9845321
54. Boujelben, O., Bahoura, M.: Efficient FPGA-based architecture of an automatic wheeze detector using a combination of MFCC and SVM algorithms. J. Syst. Arch. **88**, 54–64 (2018). https://doi.org/10.1016/j.sysarc.2018.05.010
55. Audacity Manual: Sample Rates, November 2019. https://manual.audacityteam.org/man/sample-rates.html