TussisWatch: A Smart-phone System to Identify Cough Episodes as Early Symptoms of Chronic Obstructive Pulmonary Disease and Congestive Heart Failure

Anthony Windmon, *Student Member, IEEE,* Mona Minakshi, Pratool Bharti, Sriram Chellappan, *Senior Member, IEEE,* Marcia Johansson, Bradlee A. Jenkins and Ponrathi R. Athilingam

***Abstract*—Chronic Obstructive Pulmonary Disease (COPD) and Congestive Heart Failure (CHF) are leading chronic health concerns among the aging population today. They are both typically characterized by episodes of cough that share similarities. In this paper, we design *TussisWatch*, a smart-phone based system to record and process cough episodes for early identification of COPD or CHF. In our technique, for each cough episode, we do the following: (1) filter noise; (2) use domain expertise to partition each cough episode into multiple segments, indicative of disease or otherwise; (3) identify a limited number of audio features for each cough segment; (4) remove inherent biases as a result of sample size differences; and finally, (5) design a two-level classification scheme, based on the idea of Random Forests, to process a recorded cough segment. Our classifier, at the first-level, identifies whether or not a given cough segment indicates a disease. If yes, the second level classifier identifies the cough segment as symptomatic of COPD or CHF. Testing with a cohort of 9 COPD, 9 CHF and 18 CONTROLS subjects spread across both genders, races and ages, our system achieves good performance in terms of Sensitivity, Specificity, Accuracy and Area under ROC curve. The proposed system has the potential to aid early access to healthcare, and may be also used to educate patients on self-care at home.**

***Index Terms*—Congestive Heart Failure, Chronic Obstructive Pulmonary Disease, Healthcare, Cough, Audio, Machine Learning, Smart-phones, Aging.**

I. Introduction

CHRONIC Obstructive Pulmonary Disease (COPD) and Congestive Heart Failure (CHF) are progressive disorders, and often the terminal stage of pulmonary and cardiac disease leading to death. It is estimated today that 15 million people have COPD [1] and 6.5 million people have CHF in the US [2]. Both COPD and CHF are systemic disorders with overlapping causes and pathophysiological processes. The most common cause of COPD is smoking, which accounts for 85% of cases, while the rest are due to factors like occupational smoke/dust and genetics [3]. The most common conditions that lead to CHF are coronary artery disease, high blood pressure and previous heart attack(s).

Cough is often regarded as a critical symptom of COPD and CHF, and listening to cough is still an important mechanism for physicians to gauge disease onset and severity. In this paper, we design *TussisWatch*, a smart-phone based system that is user-friendly and low cost to enable self-diagnosis of COPD and CHF by patients.

Specifically, our system consists of a) a simple and user-friendly mobile application to record cough; b) noise cancellation techniques to filter out ambient noise; c) careful extraction of a small number of audio features that provide discriminatory power among classes; and d) a two-level Random Forest based classification technique, where the first-level identifies the recorded cough as symptomatic of DISEASE (COPD or CHF) or otherwise (CONTROLS); followed by a second-level classification of the recorded cough as symptomatic of COPD or CHF, based on classification at the first-level. With a cohort of 9 COPD, 9 CHF and 18 CONTROLS subjects, spread across both genders, races and ages, we extensively evaluate our proposed system. We see good performance across Sensitivity, Specificity, Accuracy and Area under ROC curve, across multiple testing strategies, which demonstrates practical utility of our proposed system.

The paper is organized as follows. We survey related work in Section II. We present details on our Data Collection in Section III, and the Technical Approach in Section IV. Results of performance evaluations are presented in Section V, and the paper is concluded with important discussions in Section VI.

II. Related Work

We now present important related work in the space of designing algorithms to process cough audio for healthcare.

The first class of related work we survey relates to identification of basic cough from audio. In [7], a smartphone application has been developed to detect respiratory events like sneezing, coughing, sniffling and clearing the throat. Using a number of time and frequency domain features, followed by SVM based algorithms, the authors design a multi-level classifier (similar to our design in this paper) for classification. The accuracy achieved is 82% for respiratory events, and 99.1% for non respiratory events. Similar results to classify

A. Windmon, M. Minakshi, P. Bharti and S. Chellappan are with the Computer Science and Engineering Department, University of South Florida, 4202 E. Fowler Avenue, ENB 118, Tampa, FL 33620, USA (e-mail: awindmon@mail.usf.edu, mona6@mail.usf.edu, pratool@mail.usf.edu, sriramc@usf.edu). P. Bharti was a student while working on this paper.

P. Athilingam, M. Johansson and B. Jenkins are with the College of Nursing, University of South Florida, 12901 Bruce B. Downs Blvd MDC Box 22, Tampa, FL 33612, USA (email: pathilin@health.usf.edu, mjohans1@health.usf.edu, bradleej@health.usf.edu).

only cough from other noises are presented in [14]. In related works, like [12], [18] and [13], the problem is to classify wet cough from dry cough. The sources of data were external recording devices in [12] and [18], and a high fidelity data acquisition system in [13]. Features extracted include 1^{st}, 2^{nd} and 3^{rd} order formant frequencies, mel-cepstrum, non-Gaussianity, bispectrum, pitch, zero crossing rates, peaks of cough spectrum envelopes, and power ratios of frequency bands. Subsequently, using standard machine learning algorithms, good classification accuracies are achieved in these works. Another work in [19], specifically focuses on classifying cough in pediatric settings. Using specialized instruments to record cough, features extracted include MFCC, formant frequency, ZCR, non-Gaussian score and Shannon entropy. Then, using Neural Net models, Sensitivity, Specificity, and Cohen's Kappa of 93%, 98%, and 0.65, respectively, were achieved during classification.

The second class of work we survey attempts a finer grained classification of cough. The work in [15] attempts to differentiate pneumonia from asthma using a sample of 18 child subjects and cough data recorded from a low noise microphone. Using features like MFCC, non- gaussianity score and Shannon entropy, the authors design Artificial Neural Net classifiers to achieve a Sensitivity, Specificity and Kappa of 89%, 100%, and 0.89, respectively in classification. In [16], the problem is to evaluate the airflow and sound characteristics of a voluntary cough to classify lung diseases, wherein specialized instruments were designed to record signals. Using a sample of around 100 subjects, and a relatively large number of features (more than 100 of them), the authors design a Principal Component Analysis model, wherein the classification accuracies were in the range of 94% and 97% for female and male subjects, respectively. In another paper [17], the problem is to detect pertussis in children. Classification was performed using publicly available audio sources from 38 children patients to achieve an accuracy of 92%. In [20], work has been done to diagnose and screen pulmonary disease via cough sounds, using a sample of 33 healthy subjects and 54 patients having COPD, asthma, and allergic rhinitis. The source of data was a stethoscope to collect lung sounds data. Using 7 audio features (including kurtosis, variance, zero crossing rate, and rate of decay), and a logistic regression algorithm, classification accuracy of 80% was achieved.

Summary: To summarize, the problem we address in this paper follows the overall flavor of most related works. However, our problem, namely classification of cough symptomatic of COPD/CHF is unique and not explored yet, but very important. Naturally, the features we extract in this paper, explanations of their relevance, and design of classification algorithms are unique in this paper. Note that in a recent paper [11], we present results on detecting only COPD symptoms using cough recorded from smart-phones. However, the system in [11] does not address the issue of classifying both CHF and COPD cough, which is a much more difficult problem. This necessitates new features and methods for classification, which are new contributions of the system proposed in this paper.

III. Data Collection

We now present details on data collection.

A. Custom Mobile Application for Cough Recording

All cough episodes were recorded using a custom voice recording Android application, called $VoiceRecorder$, developed by the authors. This application was installed on a Samsung Galaxy S5 smart-phone, which uses Android Operating System $5.1.1$ Lollipop, and whose microphone has a sampling rate of 44100Hz. The recording applications works as follows. When the application is opened, it immediately initiates a 30 second timer, and the subject's cough will be recorded. A 'Stop' button is pressed to stop recording. Otherwise, the application will automatically close after 30 seconds. The recording is saved in the phone as 3GP file, later converted to a *.wav* format for feature extraction. We kept the recording time to 30 seconds, since for patients with COPD/ CHF, symptoms of the disease from cough are highly likely to manifest within this duration. Allowing a subject to cough beyond 30 seconds was unnecessary, and can be strenuous for elders.

B. Subjects Recruitment and Cough Recording

With the help of nurses at Tampa General Hospital, located in Downtown Tampa, we identified patients who were clinically diagnosed, by a physician, with early stage COPD and/or CHF. We also identified subjects of a similar age group who did not have COPD or CHF, and they served as CONTROLS. All subjects who gave us cough data consented to do so.

In our study, a registered nurse asked each subject to cough close to the microphone of the Samsung Galaxy S5 smartphone, and would then turn on our app. The duration of each cough ranged from 3 seconds to 17 seconds per subject. In this manner, cough data was collected from 9 COPD, 9 CHF and 18 CONTROLS subjects. After recording a cough episode, the corresponding audio file in the phone was renamed with a unique subject identifier appended with the subject type (i,e., "COPD" or "CHF" or "CONTROLS"). In our experiment, 40% of subjects were female, and 60% were male. The average age of subjects was 55 years, with a standard deviation of 7.

IV. Technical Approach

Fig. 1 presents the work-flow of *TussisWatch*. It is a two-level classification system, where at the first level, a cough segment is identified as either DISEASE or CONTROLS. In the case of the former, the second level of classification identifies the cough segment as symptomatic of COPD or CHF.

A. Removal Of Pauses and Noise

The first step is to remove pauses and noise. Occasionally, during a cough recording, there were few instances of audio data containing long pauses before, after, and in between coughs. Such pauses were discarded using an online audio cutting application. Afterwards, we applied a band-pass filter to remove additional noises. The cut-off frequencies are 300Hz and 1200Hz, since cough related sounds are primarily in this range [21]. This design choice is also applied in other related works for extracting extract cough signals from noise [12].

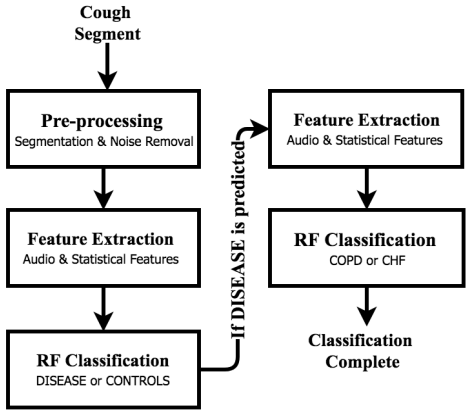


Fig. 1: Work-flow of our Two-Level Cough Classification Scheme

B. Segmentation Algorithm based on Physiology of Cough

We now present information related to the physiology of cough that necessitates us to partition a cough episode into multiple segments before analysis. This is because, when a patient with COPD/CHF suffers from an episode of cough, only selected segments within that episode indicate either condition, and the rest of the cough can appear as normal.

In the case of a patient with COPD, whistling noises and/or sounds indicating a large mucus buildup due to infection are *heard in the middle of a cough episode*. In the case of a patient with CHF, crackling sounds from fluid buildup in the lower lungs are *heard in the breath drawn right after the cough*. In either case, the remainder of cough during that episode for a patient typically sounds as regular cough. In the case of a patient with neither disease (i.e., CONTROLS), the cough stems from the throat, and naturally sounds like regular cough.

As we can see, in any cough episode, only selected portions of the episode are indicative of COPD or CHF, and not the entire episode. The issue for us is how to partition a cough episode to only retrieve only those segments symptomatic of disease to enable learning. Fortunately, three authors of this paper (i.e., the fifth, sixth and seventh authors) have decades of combined experience in identifying COPD and/or CHF cough by listening, and they indicated that segmenting a cough episode into one second windows was optimum to catch the corresponding cough sounds of interest. A window size smaller than one second is too hard for the ear to process, and a window size larger than one second may contain data from more than one class, both of which are problematic.

As such, we partitioned each cough episode into multiple segments, each of one second duration. That is, for a cough episode of 10 seconds, we extract 10 segments each of one second duration. Then, our three COPD/CHF expert co-authors jointly listened to each (one second) segment of each recorded cough episode carefully, to agree and tag that segment as symptomatic of COPD or CHF or otherwise. At the conclusion of segmentation, we derived a total of 82 segments of cough that were symptomatic of COPD and 47 segments of cough symptomatic of CHF from all the patients. The rest of cough segments from these cohorts, and those from the CONTROLS cohort were labeled as CONTROLS cough. In this manner, we labeled 81 cough segments as CONTROLS. Note that the duration of each cough segment across all classes was one second. Also, recall that since the sampling rate of the smart-phone in our case was 44100Hz, each cough segment in our data set had that many data samples within. This dataset enabled subsequent model development.

C. Feature Extraction to Classify a Cough Segment as DISEASE or CONTROLS at the First Level

Any learning algorithm is sensitive to the features on which it is trained. At the first level of our scheme, we identify whether a given segment of cough is indicative of DISEASE (either COPD or CHF) or CONTROLS. To do so, we extract 6 features that we carefully identified for our problem. Each feature is computed for each (one second) segment of a cough episode in either class. The features are presented below.

Zero Crossing Rate (ZCR): Let $sgn(x)$ return $+1$ when x is positive, or when $x = 0$; and return -1 when x is negative. For a cough segment f, with L samples, ZCR is

$$ZCR(f) = \sum_{i=2}^{L} \frac{|sgn(s_i) - sgn(s_{i-1})|}{2(L-1)}. \quad (1)$$

ZCR [7] is a measure of the number of times the amplitude of sample points s_i in a segment f of a given cough episode passes through a value of zero. Fig. 2 demonstrates an instance of the significance of this feature for classification, where the DISEASE cough (COPD & CHF) has much higher ZCR's compared to CONTROLS cough.

Sound Pressure Level (SPL): SPL is a logarithmic measure of the actual sound pressure of a cough segment, with respect to a fixed reference pressure. For a given cough segment f, the SPL, measured in decibels dB, is defined as

$$SPL(f) = 20log_{10}\frac{r}{r_{ref}}dB. \quad (2)$$

Here, r denotes the average sound pressure of a cough segment and r_{ref} denotes a reference value of $20\mu Pa$, which is lowest hearing threshold of a healthy ear. The DISEASE class, when compared to the CONTROLS class, showed consistently higher SPL values. This is depicted in Fig. 3.

Interquartile Range (IQR): To derive IQR, we divide the frequencies within each cough segment into quartiles, the 75^{th} percentiles (upper quartile) and the 25^{th} percentile (lower quartile), and determine their difference. Cough segments corresponding to DISEASE class have consistently lower frequencies, due to lower pitches. CONTROLS cough segments have frequencies fluctuating more, which usually start high (upper quartile) and end much lower (lower quartile). Hence, the differences between the two quartiles were higher for CONTROLS cough in comparison to DISEASE cough.

Percentiles (PER): To calculate PER, we extract all frequencies available in a cough segment, and sort them in ascending order. The PER value for this cough segment is that frequency, below which 40% (which is a tunable parameter) of all frequencies contained in that segment are present. For our dataset, the PER values computed as above were consistently higher for CONTROLS cough compared to DISEASE cough.

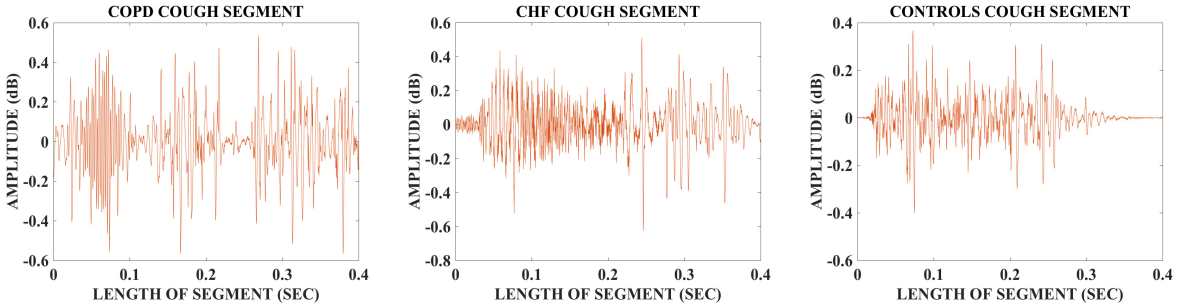


Fig. 2: Zero Crossing Rate of COPD, CHF and CONTROLS cough segments over time

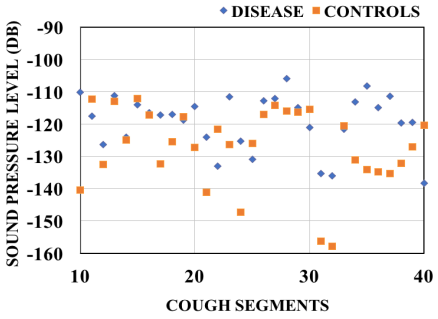


Fig. 3: Sound Pressure Level of DISEASE and CONTROLS cough segments

Mean Absolute Deviation (MAD): This parameter is the mean absolute deviation of the set of all frequencies contained within a cough segment. In our dataset, DISEASE cough had consistently higher MAD values in comparison to CONTROLS cough.

Standard Deviation (STD): This parameter is the standard deviation of the set of all frequencies contained within a cough segment. In our datasets, DISEASE cough had consistently higher STD values in comparison to CONTROLS cough.

D. Feature Extraction to Classify a Cough Segment as COPD or CHF at the Second Level

At the second-level of classification, recall that we want to classify a cough segment as symptomatic of COPD or CHF, if the first-level classifier identified the segment as belonging to the DISEASE class. To do so, we identify 11 features which we compute for each segment of cough in either class.

Spectral Centroid (SC): Let p_i (i = $1, 2, \ldots n$) represent the normalized magnitude of the i^{th} frequency bin of a cough segment f computed using Fast Fourier Transform (FFT). The Spectral Centroid is calculated as

$$SC(f) = \frac{\sum_{i=1}^{n}(i)(p_i)}{\sum_{i=1}^{n}(p_i^2)}. \tag{3}$$

Here, SC represents the "brightness", or loudness, of a cough segment. As shown in Fig. 4 (a) for a few instances, the SC is higher for the CHF class compared to COPD class. This is because CHF coughs contain crackles and fluids, which have higher sounds (i.e., volume) compared to COPD coughs that contain mucus, which produce lower, muffled sounds.

Spectral Roll-Off (SR): Consider the Total Energy of a cough segment, which is computed as $\sum_{i=0}^{n}(p_i)$, where n and p_i are defined above. In this case, SR is that frequency below which 85% of the energy of the cough segment is contained. The SR for CHF cough was consistently higher compared to COPD cough as seen in Fig. 4 (b).

Spectral Flatness (SF): SF characterizes the audio spectrum of each cough segment by determining how "noise-like" a cough is versus how "tone-like" it is. It is determined as

$$SF(f) = \frac{\sqrt[n]{\Pi_{i=1}^{n}(p_i)}}{\frac{\sum_{i=1}^{n-1}(p_i)}{n}} dB. \tag{4}$$

COPD cough segments, considering their common mucus sounds, are more "noise-like" compared to CHF segments. Thus, SF for COPD segments were higher than that of CHF.

Mel Frequency Cepstral Coefficients (MFCC): Let $\overline{C}$ denote the mean of the frequencies in a cough segment. Let $x = 1, 2 \ldots K$, where K = 44100 (the sampling rate of the smart-phone) and S_k represent the Discrete Cosine Transform (mel cepstrum) coefficients. The MFCC is calculated via

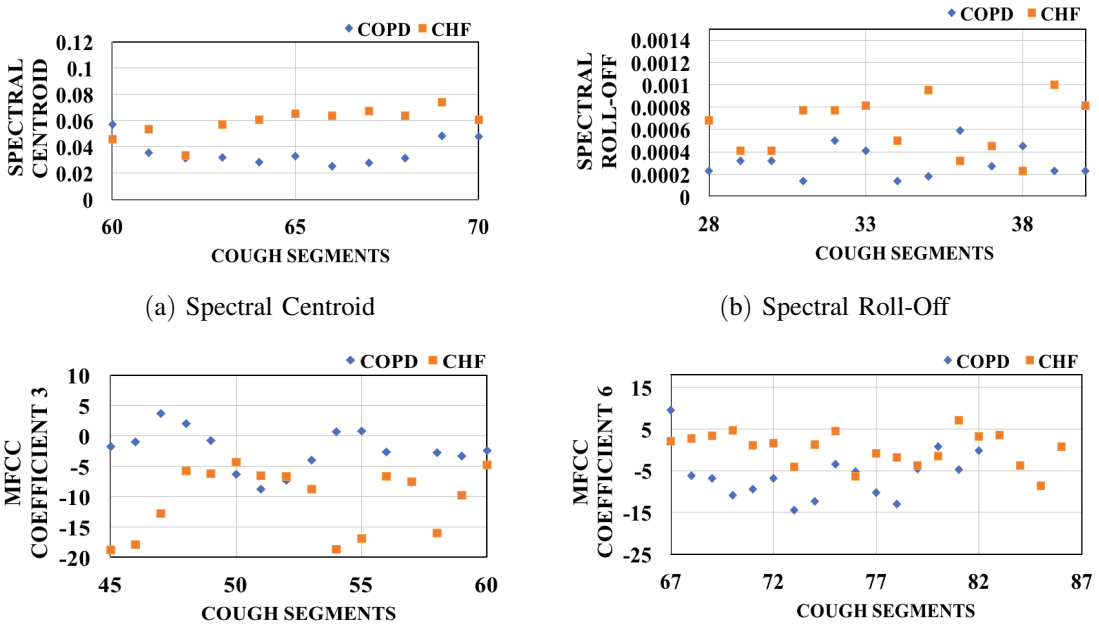
$$\overline{C_x} = \sum_{k=1}^{K}(logS_k)[x(k - \frac{1}{2})\frac{\pi}{K}]. \tag{5}$$

The MFCC represents the spectral envelope of a given cough segment, and its computation requires a series of complex steps [5] [9] [10], which are not elaborated here due to space limitations. For best performance of our cough classification scheme, only 2 of the 13 cepstrum coefficients were selected after analysis, which were the third and sixth coefficients. As seen in Figs. 4 (c) and (d) for a representative case, the third coefficient reflected a consistently higher MFCC for COPD cough, while the sixth coefficient reflected a consistently higher MFCC for CHF cough.

Short Time Energy (STE): Let y_i denote the amplitude of the i^{th} sample of a cough segment, and $h(w)$ denote impulse response of a linear filter of signal w. For a cough segment f,

$$STE(f) = \sum_{i=-\infty}^{\infty} y_i^2 * h(w - i). \tag{6}$$

The STE measures (in increments) energy increase of a cough segment. In our case, we set $w = 10$ms. We find that COPD cough segments had consistently higher values for STE as compared to CHF cough segments.



(a) Spectral Centroid

(b) Spectral Roll-Off

(c) MFCC Coefficient 3^{rd}

(d) MFCC Coefficient 6^{th}

Fig. 4: Spectral Centroid (a) & Roll-Off (b), and MFCC Coefficients 3^{rd} (c) & 6^{th} (d) for COPD and CHF cough segments.

Root Mean Square (RMS): Let L denote the number of samples in a cough segment f, and n_i denote the normalized amplitude value of the i^{th} sample in f. Then,

$$RMS(f) = \sqrt{\frac{1}{L}\sum_{i=1}^{L} n_i^2}. \quad (7)$$

RMS is used to characterize the energy contained in the sound waves of a given cough segment. We see a consistently higher RMS in COPD, compared to CHF cough segments, due to its higher averaged sound pressure.

Maximum Value (MAX), Variance (VAR), Median (MED) and Mean (AVG): These features denote the maximum value, variance, median and mean among frequencies contained in a cough segment. For each feature, values were higher for CHF cough compared to COPD cough.

E. Data Balancing via SMOTE

Recall that our dataset is imbalanced, containing 129 seconds of cough in DISEASE class (i.e., 82 seconds of COPD , and 47 seconds of CHF cough); and 81 seconds of CONTROLS cough. Classification on unbalanced datasets can create biased results. To alleviate this problem, we balance our datasets by oversampling the deficient classes, following the idea of Synthetic Minority Oversampling Technique (SMOTE) [6]. SMOTE is a widely used data balancing method in which feature values in the minority classes are oversampled by creating synthetic examples, rather than by replacement or creating copies [1]. We explain using Fig. 5.

[1]Undersampling the majority class is not viable due to our limited amount of data. Also, balancing techniques, like Cost Matrix are better for larger datasets, and are shown to not work as well as SMOTE when datasets are small [22] [23], as is the case with our dataset.

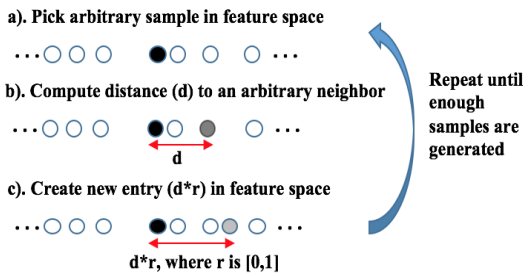


Fig. 5: Three step implementation of SMOTE

Let us consider an arbitrary feature in the minority class. In the SMOTE technique, for this feature, an arbitrary feature point from the feature vector is picked as seen in Fig. 5 (a). Then, the distance to a randomly chosen neighboring feature point among 5 closest ones is computed [2]. This is shown in Fig. 5 (b). The difference (i.e., distance) is multiplied by a random number in the range $[0, 1]$ and this value is added to the initial arbitrary data point picked, and this resulting data point becomes a new entry in the corresponding feature vector for the minority class, as shown in Fig. 5 (c). The process repeats until the desired number of feature points are computed and added. The process naturally repeats for each feature. The SMOTE technique is widely used, and has been identified as a robust technique for over-sampling (by sometimes up to 200% of the original data) to overcome class imbalance effects [6].

In our proposed implementation for class balancing, each feature in the CHF class was increased from 47 to 82 (75%); and each feature in the CONTROLS class was increased from

[2]Choosing 5 neighboring feature points gave us best results, and is also recommended in [6].

81 to 162 (100%). As a result, we now have 164 data points for each feature in the DISEASE class (i.e., 82 for COPD cough, and 82 for CHF cough); and 162 data points for each feature in CONTROLS class resulting in a balanced dataset.

F. Random Forest Classifier

Finally, we design a Random Forests (RF) based classification algorithm at both levels. The RF algorithm creates a random subset of training samples from the cough datasets, for both classification levels by assembling a congregation of decision trees. Each decision tree predicts a class, based on a majority vote made by each individual tree. Then, the decision tree utilizes the majority vote to determine the final predicted class. The parameters of our RF model, that gave us best results for both classification levels are as follows: 121 decision trees, information gain as splitting criteria, 6 as the maximum depth of each decision tree, 5 as minimum number of samples to split the internal node, and bag size percent of 100. We found these parametric values optimal by applying grid search on wide range of parameters.

V. Results and Related Discussions

A. Results and Interpretations

1) First Level Results: At this level, the classification of a cough segment was between DISEASE and CONTROLS. Employing 10-Fold Cross Validation, we achieved a Sensitivity of 80%, Specificity of 82% and an Accuracy of 80.67%. Furthermore, the ROC Area is 83%. Employing Leave-One-Out Validation, we achieved a Sensitivity of 77%, Specificity of 79% and an Accuracy of 75.45%. The ROC Area was 78%.

2) Second Level Results: Here, the classification of a cough segment was between COPD and CHF. Employing 10-Fold Cross Validation, we achieved a Sensitivity of 82%, Specificity of 75% and an Accuracy of 78.05%. Furthermore, the ROC Area is 80%. Employing Leave-One-Out Validation, we achieved a Sensitivity of 75%, Specificity of 73% and an Accuracy of 74.63%. Furthermore, the ROC Area was 76%.

B. Justification for our Two-level Classification Scheme

Note that both COPD and CHF related cough sound similar, even for experts. In our system, when we attempted a single level classification scheme to differentiate between COPD, CHF and CONTROLS cough, our system learns to recognize CONTROLS cough from the other two classes better, but there was significant confusion between COPD and CHF cough. This is because the features used to differentiate CONTROLS cough from DISEASE (i.e., either COPD or CHF) do not work well enough to isolate COPD cough from CHF cough. This is reasonable, since the differences between physiology of CONTROLS cough and DISEASE cough are more pronounced. Once we identify that the cough belongs to the DISEASE class, the separate set of features we designed to separate COPD cough from CHF cough perform much better, hence explaining our design choice.

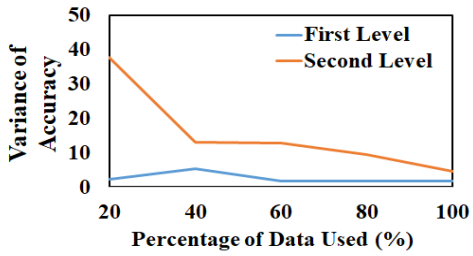


Fig. 6: Variance trends for different subsets of our dataset

C. A Note on Representativeness of our Dataset

It is vital for any machine learning system to train on sufficient sized datasets to avoid overfitting. We present insights on the representativeness of our datasets, by employing a variance based approach that is used in the literature [24] [25]. In employing this approach, we randomized our entire dataset by dividing it into 10 non-overlapping subsets initially. We then classified cough using our technique for the first 10% subset of the dataset to get 10 numbers for classification accuracy for the case of 10-fold cross validation. Next, we quantify the variance among accuracies derived. Then, we do the same for 20% of datasets, and again we compute the variance in classification accuracies among the 10 values for 10-fold cross validation. We do the same for 30%, 40%, 50%, 60%, 70%, 80%, 90% and the entire 100% of our dataset. Fig. 6 plots the resulting variances. As we see, for small sized datasets, the variances are higher, but they start to decrease with more datasets. Beyond 70% of datasets, the variances are stable and very low as well. This, we believe, gives us confidence on representativeness of our datasets for our system at both levels.

D. Complexity of Execution

We designed a simple Android app to record cough, remove noise, segment cough, extract features identified above, and execute our Random Forest Algorithm for classification. On a Samsung Galaxy S5 smart-phone, our app consumed a memory of less than 5MB. Processing a 10 second cough episode consumed 40Joules (a very small amount), with a processing time of less than 5 seconds on the average. These numbers are highly encouraging for practical deployment and potential extensions of our *TussisWatch* system for better care.

E. Comparing Performance of Classification Algorithms

In Table I, we show classification results from implementing different learning algorithms for our problem, and found that Random Forests performs the best. This is because Random Forests Classifiers are one of the most accurate learners available, and better reduce the likelihood of over-fitting [8].

VI. Conclusion and Future Work

In this paper, we designed *TussisWatch*, a smart-phone based system to record and detect cough patterns indicative of COPD, CHF or no disease, using a two-level classification scheme. We believe our system is the first to demonstrate the

TABLE I: Comparing Accuracy (%) of Different Machine Learning Algorithms

Algorithm	First Level	Second Level
Random Forests	80.67%	78.05%
Support Vector Machine	76.77%	75.60%
k-Nearest Neighbors	69.75%	67.13%
Naive Bayes	62.07%	72.03%

feasibility of a smart-phone based system for self-detection of cough symptomatic of COPD/CHF.

As future work, first, we want to improve our classification accuracies further. To do so, we are carefully planning new data collection experiments. This time, instead of going to clinics, we will release our app to a larger sample of patients and control groups, and request them to use our app to record and store cough data for (say) a month. All cough data locally stored on the phone will be collected and processed subsequently by us. In this manner, our data sets will be bigger. The data will be more realistic since it is obtained in natural settings of patients, and not in clinics where the cough is more voluntary. There will also be better diversity in terms of locations commonly visited by patients as well (e.g., clinics, homes, gyms, shopping centers etc.). With such kinds of data, our team plans to investigate (1) deep learning approaches for classification that do away with feature extraction; (2) design personalized models to enhance classification accuracies; (3) attempt more complex problems like evaluate the impact of an intervention (e.g., an inhaler [4]) via processing cough before and after administering the intervention. Interviewing patients to understand their security and privacy expectations during design and deployment of a comprehensive self-care system are additional directions of our future work. All these are exciting avenues for which results from this paper serve as a strong foundation.

ACKNOWLEDGMENTS

This work was supported in part by US National Science Foundation under grants HRD 1612347, IIS 1559588 and CNS 1718071. Any results, opinions, findings and conclusions are those of the authors alone and do not necessarily reflect the views of the funding agency.

REFERENCES

[1] Center for Disease Control and Prevention, “Chronic Obstructive Pulmonary Disease (COPD): Data and Statistics”, Aug. 15, 2016. [Online]. Available: https://www.cdc.gov/copd/data.html.

[2] E.J. Benjamin, M.J. Blaha, S.E. Chiuve, et al, “Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association”, pp. e146-e603, Mar. 2017, DOI: 10.1161/CIR.0000000000000485.

[3] GOLD, “Pocket Guide to COPD Diagnosis, Management, and Prevention: A Guide for Health Care Professionals Global Initiative for Chronic Obstructive Lung Disease”, 2017. [Online] Available http://goldcopd.org/wp-content/uploads/dlm_uploads/2016/12/wms-GOLD-2017-Pocket-Guide-Final.pdf

[4] A. N. Amin, V. Ganapathy, A. Roughley, and M. Small, “Confidence in correct inhaler technique and its association with treatment adherence and health status among US patients with chronic obstructive pulmonary disease”, *Patient Prefer Adherence*, 11: 1205-1212, Jul. 2017, DOI:10.2147/PPA.S140139.

[5] H. Rashidul, M. Jamil, G. Rabbani, and S. Rahman, “Speaker Identification Using Mel Frequency Cepstral Coefficients”, 3^{rd} *International Conference on Electrical and Computer Engineering ICECE 2004*, pp. 565-568, Dec. 2004.

[6] N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, ”Smote: Synthetic minority over-sampling technique”. *Journal of Artificial Intelligence Research*, pp. 321-357, June 2002.

[7] X. Sun, Z. Lu, W. Hu, and G. Cao, ”SymDetector: detecting sound-related respiratory symptoms using smartphones”, *UbiComp ’15*, pp. 97–108, Sep. 2015, DOI: 10.1145/2750858.2805826.

[8] L. Breiman, “Random Forests”, *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, DOI: 10.1023/A:1010933404324.

[9] J. Lyons, “Mel Frequency Cepstral Coefficient (MFCC) tutorial”, 2012. [Online] Available: http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/

[10] H. Fayek, “Speech Processing for Machine Learning: Filter banks, Mel-Frequency Cepstral Coefficients (MFCCs) and What’s In-Between”, 2016. [Online] Available: http://haythamfayek.com/2016/04/21/speech-processing-for-machine-learning.html

[11] A. Windmon, M. Minakshi, S. Chellappan, P.R. Athilingam, M. Johansson, and B.A. Jenkins, “On Detecting Chronic Obstructive Pulmonary Disease (COPD) Cough using Audio Signals Recorded from Smart-Phones”, *Proc. of 11th Intl. Conf. on Health Informatics (HealthInf)*, pp. 329-338, Jan. 2018, DOI: 10.5220/0006549603290338

[12] H. Chatrzarrin, A. Arcelus, R. Goubran, ”Feature Extraction for the Differentiation of Dry and Wet Cough Sounds”, *2011 IEEE International Symposium on Medical Measurements and Applications*, July 2011, DOI: 10.1109/MeMeA.2011.5966670.

[13] V. Swarnkar, U.R. Abeyratne, Y.A. Amrulloh, A. Chang, ”Automated algorithm for Wet/Dry cough sounds classification”, *34th Annual International Conference of the IEEE EMBS*, Nov. 2012, DOI: 10.1109/EMBC.2012.6346632

[14] V. Swarnkar, U.R. Abeyratne, Y. Amrulloh, C. Hukins, R. Triasih, A. Setyati, ”Neural network based algorithm for automatic identification of cough sounds”, *35th Annual International Conference of the IEEE EMBS*, July 2013, DOI: 10.1109/EMBC.2013.6609862.

[15] Y. Amrulloh, U. Abeyratne, V. Swarnkar, R. Triasih, ”Cough Sound Analysis for Pneumonia and Asthma Classification in Pediatric Population”, *2015 6th International Conference on Intelligent Systems, Modelling and Simulation*, Oct. 2015, DOI: 10.1109/ISMS.2015.41.

[16] A.A. Abaza, J.B. Day, J.S. Reynolds, A.M. Mahmoud, W.T. Goldsmith, W.G. McKinney, E.L. Petsonk, D.G. Frazer, ”Classification of voluntary cough sound and airflow patterns for detecting abnormal pulmonary function”, Nov. 2009, DOI: 10.1186/1745-9974-5-8.

[17] R.X.A Pramono, S.A. Imtiaz, E. Rodriguez-Villegas, ” A Cough-Based Algorithm for Automatic Diagnosis of Pertussis”, Sept. 2016, DOI: 10.1371/journal.pone.0162128.

[18] V. Swarnkar, U.R. Abeyratne, A.B. Chang, Y.A. Amrulloh, A. Setyati, R. Triasih, ”Automatic Identification of Wet and Dry Cough in Pediatric Patients with Respiratory Diseases”, *Annals of Biomedical Engineering*, pp. 1016-1028, May 2013, DOI: 10.1007/s10439-013-0741-6.

[19] Y.A. Amrulloh, U.R. Abeyratne, V. Swarnkar, R. Triasih, ”Automatic Cough Segmentation from Non-contact Sound Recordings in Pediatric Wards”, *Biomedical Signal Processing and Control*, pp. 126-136, May 2015.

[20] C. Infante, D. Chamberlain, R. Fletcher, Y. Thorat, R. Kodgule, ”Use of cough sounds for diagnosis and screening of pulmonary disease”, *2017 IEEE Global Humanitarian Technology Conference (GHTC)*, Dec. 2017, DOI: 10.1109/GHTC.2017.8239338.

[21] J. Korpas, J. Sadlonova, M. Vrabec, ”Analysis of the Cough Sound: an Overview”, *Pulmonary Pharmacology*, Vol. 9, pp. 261-268, Oct. 1996, DOI: 10.1006/pulp.1996.0034.

[22] G.M. Weiss, K. McCarthy, B. Zabar, ”Cost-Sensitive Learning vs. Sampling: Which is Best for Handling Unbalanced Classes with Unequal Error Costs?”, *Proc. Int. Conf. Data Mining*, pp. 35-41, 2007.

[23] M.M. Rahman, D.N. Davis, ”Addressing the Class Imbalance Problem in Medical Datasets”, *International Journal of Machine Learning and Computing*, Vol. 3, NO. 2, pp. 224-228, Apr. 2013.

[24] R. Kohavi, ”A study of cross-validation and bootstrap for accuracy estimation and model selection”, *Proc. of the 14th Intl. Joint Conf. on Artificial Intelligence*, Vol. 2, pp. 1137-1143, 1995.

[25] Y. Bengio and Y. Grandvalet, ”No Unbiased Estimator of the Variance of K-Fold Cross-Validation”, *Journal of Machine Learning Research 5*, pp. 1089-1105, Dec. 2004.

Anthony Windmon received his BS degree in Computer Engineering, with a minor in Mathematics, from Bethune-Cookman University. He is currently a graduate student, studying Computer Science and Engineering, at the University of South Florida. He is a NSF LSAMP Bridge to the Doctorate and McKnight Doctoral Fellow. His research interests are in smart healthcare and machine learning.

Mona Minakshi is a PhD student in Computer Science and Engineering at the University of South Florida. Her research interest is in healthcare, computer vision and deep learning.

Pratool Bharti is a research and development manager. His research interest is in smart healthcare, computer vision and deep learning. He has a PhD in Computer Science and Engineering from the University of South Florida.

Sriram Chellappan is an associate professor in the Department of Computer Science and Engineering at the University of South Florida. Previously, he was a faculty member in the Dept. of Computer Science at Missouri University of Science and Technology. His current research interests include social and technological aspects in smart healthcare, privacy and cybersecurity. Chellappan has a PhD in Computer Science and Engineering from The Ohio State University.

Marcia Johansson , DNP is an assistant professor in the College of Nursing. She is the Adult-Gerontology Acute Care Concentration Director and received her Doctor of Nursing Practice Degree from the University of South Florida. She has done extensive research in the use of mobile technology to assist patients with heart failure.

Bradlee A. Jenkins is currently a PhD student at University of South Florida in the College of Nursing. He is a Jonas Scholar and was the Entrepreneurial Lead for NSF Innovation Corps- grant, PI Dr. Athilingam. He has obtained his Masters of Nursing from the University of South Florida in the College of Nursing. His research interests are in smart healthcare and biomedical informatics within the COPD population.

Ponrathi R. Athilingam is an Associate Professor at University of South Florida College of Nursing and Assistant Professor in the College of Medicine Cardiology. Her research is focused on technology-based interventions to improve selfcare of patients and thus reduce readmissions. She has received funding from National Science Foundation and NINR for the design, development and testing of a mobile phone application to improve self-care of patients with heart failure. She has received the Nurse Investigator Award for the year 2008 from Heart Failure Society of America. She was named Fellow of the American Academy of Nurse Practitioners (FAANP) and Fellow of the Heart failure Society of America (FHFSA).