

Automatic Identification of Cough Events from Acoustic Signals

Renard Xaviero Adhi Pramono, Syed Anas Imtiaz, and Esther Rodriguez-Villegas

Abstract—Cough is a common symptom of numerous respiratory diseases. In certain cases, such as asthma and COPD, early identification of coughs is useful for the management of these diseases. This paper presents an algorithm for automatic identification of cough events from acoustic signals. The algorithm is based on only four features of the acoustic signals including LPC coefficient, tonality index, spectral flatness and spectral centroid with a logistic regression model to label sound segments into cough and non-cough events. The algorithm achieves sensitivity of 86.78%, specificity of 99.42%, and F1-score of 88.74%. Its high performance despite its small size of feature-space demonstrate its potential for use in remote patient monitoring systems for automatic cough detection using acoustic signals.

Index Terms—Cough sound, automatic detection, logistic regression

I. INTRODUCTION

Cough is a common symptom that manifests in many respiratory diseases such as croup, pertussis, and asthma [1]. It is a mechanism of the body to expel unwanted materials from the respiratory system, including excessive amount of mucus and foreign particles [2]. The mechanical events of a cough can be divided into three phases [3]. These include the expulsive act as characterized consecutively by inspiratory effort, expiratory effort against a closed glottis, and the opening of the glottis where rapid expiratory flow occurs [4]. In Chronic Respiratory Diseases (CRDs), such as asthma and Chronic Obstructive Pulmonary Disease (COPD), monitoring of coughs is useful for the management of the disease. In these cases, detecting the occurrence of coughs and the onset of their spells, and counting the number of cough events help to understand the status of the disease progression [5], [6]. Understanding the pattern of cough occurrence over a longer period of time also helps to provide appropriate and timely care to patients. This can be achieved by recording sound from the patients to determine these cough-related parameters.

Manual counting of cough events in recorded sounds, however, is a very laborious and error-prone task and suffers from inter-expert subjectivity in classification. To avoid these problems, there is a growing interest in automating the process of detecting and counting cough events. Several commercial solutions and research studies have presented methods to achieve partially or fully automated cough detection.

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK / grant agreement no. EP/P009794/1. Renard X. A. Pramono gratefully thanks the Indonesian Endowment Fund for Education (LPDP) for its financial support for his study.

R. X. A. Pramono, S. A. Imtiaz, and E. Rodriguez-Villegas are with the Wearable Technologies Lab, Department of Electrical and Electronic Engineering, Imperial College London, SW7 2BT, United Kingdom. Email: {renard.pramono14, anas.imtiaz, e.rodriiguez}@imperial.ac.uk

For example, VitaloJAK [7] removed silent parts of audio recordings to help experts manually annotate the remaining sound events, with Sensitivity (SE) and Specificity (SP) up to 98%. The combination of Leicester cough monitor and Hull automated cough counter performed automatic detection of cough, but needed to be re-trained for new patients. This cough counter achieved sensitivity and specificity of 57.9% and 98.2% respectively [8].

Artificial Neural Networks (ANN) have been used by several studies to classify cough sounds using a variety of features. Liu *et al.* [9] used a Deep Neural Network (DNN) for classification with Mel-frequency Cepstral Coefficient (MFCC) feature inputs and achieved performance of 90.1% SE and 85% SP. DNN was also used in [10], achieving 86.8% SE and 92.7% SP. Drugman *et al.* [11] employed a large set of features including MFCC, spectral flatness, and spectral centroid. These were fed into two different ANNs. After feature selection based on mutual information, 50 features were used for each ANN resulting in classification performance of 89.7% SE and 89.85% SP. In another study [12], ANN was used with MFCC, formant frequency, zero crossing rate, non-gaussianity score, and Shannon entropy were used as the input features achieving an average performance of 92.8% SE and 97.5% SP.

Other machine learning methods have also been used in order to detect cough sounds, where mostly cepstral-based features were used. A study in [13] used local Hu moment and k-Nearest Neighbour (KNN) and achieved 88.51% sensitivity with specificity of 99.72%. You *et al.* [14] employed Support Vector Machine (SVM) for the detection of cough with Non-Negative Matrix Factorisation (NMF), and achieved an average performance were SE of 80.1% and SP of 83.1%. In [15], XGBoost algorithm was used with MFCC features for cough detection, achieving average Area Under Curve of 0.916 with Leave-One-Out Cross-Validation (LOOCV). Matos *et al.* [16] used MFCC with Hidden Markov Model (HMM) achieving 71% sensitivity. The study in [17] used Sequential Minimal Optimisation (SMO) with MFCC achieving 75.5% SE with 99.6% SP. Our previous method in [18] used Logistic Regression Model (LRM) with MFCC and spectral features on a 50%-50% random data split, where the focus of the study was to diagnose pertussis. Part of the study was detecting cough events which achieved a performance of 85.20% SE and 98.32% SP.

Most of the works mentioned above either require expert intervention or use large number of features with complex machine learning methods for classification. These may be acceptable in certain applications but for long-term remote

monitoring systems for patients with CRDs, there is a need to develop simple fully automatic cough detection methods. Hence, this paper proposes a new method of automatic cough detection based on only four features with logistic regression for classification. The features used in this work are based on the characteristics of cough sounds. These sounds have a wide spread spectrum, low tone prominence, and represent a sudden burst of energy in a short time. To detect such sounds, Linear Predictive Coding (LPC) coefficient is used as a time domain feature, while tonality index, spectral flatness, and spectral centroid are used as frequency domain features. Section II describes the database used, preprocessing steps, and the different features that are extracted from the signals for the algorithm proposed in this paper. Section III introduces the cough detection algorithm and reports its performance. Finally, Section IV discusses these results and compares its performance against those in other studies.

II. MATERIAL AND METHODS

A. Database

For this study, a total of 43 recordings from different subjects were obtained from multiple repositories, as described in [18]. The data collected contained cough sounds from various diseases including pertussis, croup, common cold, bronchitis, bronchiolitis, asthma, and COPD. As the recordings were collected from multiple public repositories, there were variations in the recording conditions, sensors, sampling frequencies, and noise level associated with each recording. In these recordings, cough sound events were manually annotated to be used as reference to compare the performance of the proposed cough detection method. As a result, a total of 980 events were labeled as cough sounds whilst more than 1000 separate non-cough sound events were also identified. The non-cough sound events present in the recordings included speech, laughter, sneeze, throat clearing, wheezing sound, whooping sound, machine noise, and other types of background noise.

B. Preprocessing

Before being used for analysis, all the recordings were first preprocessed to minimize the variations as a result of different recording conditions. The first preprocessing step resampled all the recordings to 8000 Hz since the cough-related information was observed to be contained below 4000 Hz. Subsequently, the amplitude of the recordings was scaled into a range between -1 and 1 without changing the distribution, such that important statistical parameters of the sounds were preserved. Signals were then segmented into frames of 200 ms width with 50% overlap between consecutive segments.

Prior to detection of cough sounds, it can be useful to remove silent parts of the recordings to ensure that all further processing is performed only on parts of the signals containing a sound event. However, this requires the silence removal approach to be much simpler than the further processing stages. To achieve this, in this study, silence and low noise parts of the recordings were detected and

removed using a simple threshold based on the mean of deviation in the previous samples. Thus, further processing was performed only when average energy of a frame was above this threshold.

C. Feature extraction

The different features extracted from the processed signals to detect the presence of coughs are explained below. Time domain features were extracted directly from the sound segments, while the frequency domain features were extracted from the power spectrum which was obtained for each segment by using the Welch's method, with Blackman window of size 256 and 50% overlap.

1) *Linear predictive coding coefficients*: Linear Predictive Coding (LPC) coefficient is a time domain feature that can be used to predict the next signal value from the previous samples. For a p^{th} order LPC, the sample $x[n]$ is predicted by the weighted sum of the previous p samples as shown in Eq. 1. The equation is then solved by using Levinson-Durbin algorithm. An 8th order LPC was used in this study to represent the formant frequency peaks to differentiate between cough and other sounds. It was found that the second coefficient, a_2 , which weighs the first previous sample has higher discriminatory ability compared to other coefficients. This coefficient was then kept as the only feature for the cough detection task.

$$\hat{x}[n] = \sum_{k=1}^p a_k x[n-k] \quad (1)$$

2) *Tonality index*: Tonality index was part of MPEG-7 standard to estimate short-term tonality measure of a signal [19]. It was recently used to detect wheezes in respiratory sounds in [20], where more detail on its implementation is outlined. In the context of cough detection it can be used to separate speech sounds, that are part of the recordings, from the cough events. Compared to speech, cough appears to be more like noise that has wide spectrum and low tone prominence and hence lower tonality index.

3) *Spectral flatness*: Spectral flatness is a measure of how noise-like a sound is and is computed as the ratio between the geometric mean and arithmetic mean of a power spectrum $X[n]$ with N frequency bins, as shown in Eq. 2. The value of spectral flatness varies in accordance with the existence of peaks in the spectrum. Hence, a spectrum with more peaks would have a lower spectral flatness value compared to a flatter spectrum. For cough detection, spectral flatness allows for the detection of peaks in the spectrum which represents tone and harmonics, which would be absent in cough sounds.

$$spectralFlatness = \frac{e^{(\frac{1}{N} \sum_n \log(X[n]))}}{\frac{1}{N} \sum_n X[n]} \quad (2)$$

4) *Spectral centroid*: Spectral centroid represents the equivalent of the center of mass in a spectrum and is computed as a weighted mean of the spectrum as shown in Eq. 3 where $F[n]$ is the frequency bin and $X[n]$ is the PSD

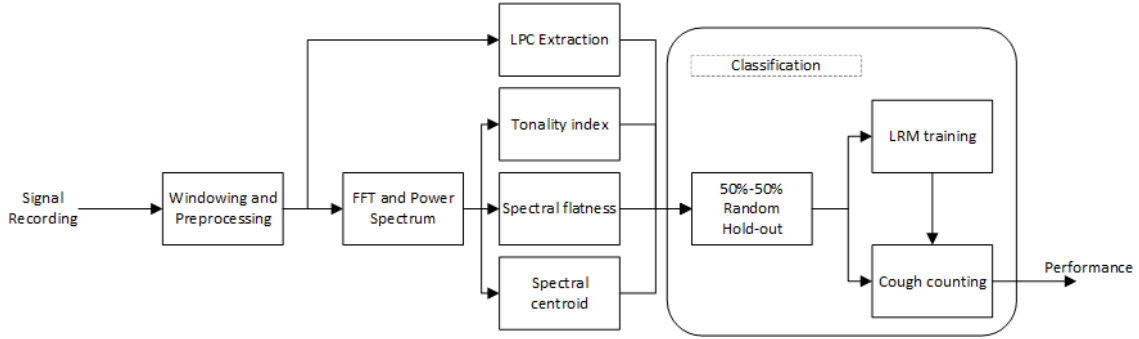


Fig. 1. Overview of the proposed cough detection algorithm.

estimate with N frequency bins. It can be used to identify cough sounds since it is expected that the wide spectrum cough sound would have higher spectral centroid compared to other sounds.

$$\text{spectralCentroid} = \frac{\sum_{n=0}^{N-1} F[n]X[n]}{\sum_{n=0}^{N-1} X[n]} \quad (3)$$

III. COUGH SOUND DETECTION ALGORITHM

A. Overview

Fig. 1 shows an overview of the proposed cough detection algorithm. The first step is preprocessing to normalize the sound recordings as well as to remove silent segments from them. The segments not removed in these stages were then used to extract the four features for the cough detection algorithm. The LPC coefficient was extracted from the preprocessed signal while the spectral features were extracted after obtaining the power spectrum of the preprocessed signal. Once the features were extracted, the segments were assessed for the presence of cough using a Logistic Regression Model (LRM) for classification.

To evaluate the classification performance of the algorithm, the recordings were separated into training and test set with 50%-50% random hold-out. The data split was performed on the recording level, such that the training and test set contained different recordings and that the test set consisted of entirely unseen data. The whole process was repeated 10 times in order to further randomize the split, where repeatability of the experiment was ensured by using the default randomisation `rng('default')` in MATLAB.

B. Performance Metrics

The performance of the cough detection algorithm was evaluated by counting the number of correctly detected cough events as True Positive (TP), incorrectly missed cough events as False Negative (FN), correctly rejected non-cough frames as True Negative (TN), and incorrectly labeled non-cough frames as False Positive (FP). From these, the Sensitivity (SE), Specificity (SP), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F1-score were calculated as in Eq. 4.

$$\begin{aligned} SE &= \frac{TP}{TP + FN}, \quad SP = \frac{TN}{TN + FP} \\ PPV &= \frac{TP}{TP + FP}, \quad NPV = \frac{TN}{TN + FN} \\ F1score &= \frac{2(SE)(PPV)}{SE + PPV} \end{aligned} \quad (4)$$

TABLE I
PERFORMANCE OF THE REPEATED 50%-50% RANDOM HOLD-OUT.

Rep #	SE (%)	SP (%)	PPV (%)	NPV (%)	F1 (%)
1	87.10	99.63	89.40	99.54	88.24
2	84.94	99.57	89.24	99.37	87.04
3	85.56	99.42	90.87	99.03	88.14
4	88.43	99.28	91.19	99.03	89.79
5	88.62	99.80	94.12	99.60	91.28
6	85.09	99.30	90.02	98.90	87.49
7	85.54	99.24	90.38	98.80	87.89
8	88.60	99.25	88.81	99.24	88.70
9	88.46	99.21	90.45	99.03	89.44
10	85.52	99.44	93.58	98.63	89.37
Average	86.78	99.42	90.81	99.12	88.74
SD	1.52	0.19	1.68	0.30	1.19

C. Results

Table I shows the performance of the cough detection method on the test data for each repetition of the random hold-out. In each repetition, the training set consisted of 22 recordings which were used to train the LRM while the other 21 recordings were then used as the test set to evaluate the model. It can be observed from Table I that the average SE achieved was 86.78% while the average PPV and F1-score was 90.81% and 88.74% respectively. The sensitivity range on all repetitions was between 84.94% and 88.62%, while the PPV range was between 88.81% and 94.12% showing that the method was able to achieve consistence performance across multiple iterations. The specificity and NPV values are also consistently very high showing the ability of the algorithm to reject most non-cough events correctly.

IV. DISCUSSION

In this paper, a new method for automatic detection of cough events is presented. It uses only four features (one time and three spectral) with a logistic regression model

TABLE II
COMPARISON OF PERFORMANCE WITH OTHER STUDIES.

Ref.	SE (%)	SP (%)	PPV (%)	F1 (%)	#
[9]	90.10	85.50	-	40.00	39
[10]	86.80	92.70	-	-	64
[11]	89.70	89.85	-	-	50
[12]	92.80	97.50	-	-	22
[13]	88.51	99.72	87.51	88.02	13
[14]	80.10	83.10	83.50	81.76	60
[15]	AUC 91.6		-	-	12
[16]	71.00	FAR 13/h		-	39
[17]	75.50	99.60	-	-	39
[18]	85.20	98.32	84.62	84.91	26
This work	86.78	99.42	90.81	88.74	4

indicates the number of features used in the study

for classification of acoustic sound segments as cough or non-cough events. The performance of the algorithm was evaluated on a database that consisted of 980 cough sounds from 43 different recordings. In 10 iterations using 50%-50% random hold-out data partition, consistent results were obtained with the range of SE between 84.94% and 88.62%, and the range of PPV between 88.81% to 94.12%. The small range of performance variation in each iteration demonstrates the capability of the proposed algorithm to work well with unseen data. It also shows that the features used in this work provide enough information that can be used for reliable classification of cough sounds events.

Table II shows the comparison of cough detection performance between the proposed algorithm and those in other studies. It can be seen that the result achieved by using the proposed features is comparable to other studies, despite having less number of features and simpler classification method. For example [13], uses kNN classification with local Hu moment as the main feature which, as reported by the authors, is very time-consuming to compute. The algorithm in [9] uses a deep neural network for classification with 13 MFCC features while [12] also uses a neural network with MFCC together with additional time and frequency features. In contrast, the algorithm proposed in this paper uses a simple classifier with only four features and achieves a comparable detection performance. It should be noted though that since the data used to report performances of algorithms is not the same, the differences in signal acquisition methods, recording conditions, and types of non-cough signals present in different data would have an impact on the achieved performance. However, since the data used by most studies is not publicly available, it is not possible to make a direct comparison between all the different methods. Nevertheless, the results reported in this study demonstrate the potential for the proposed algorithm to be used in remote patient monitoring systems for, potentially real-time, automatic cough detection.

REFERENCES

- [1] A. B. Goldsobel and B. E. Chipps, "Cough in the pediatric population," *The Journal of pediatrics*, vol. 156, no. 3, pp. 352–358, 2010.
- [2] K. F. Chung and I. D. Pavord, "Prevalence, pathogenesis, and causes of chronic cough," *The Lancet*, vol. 371, no. 9621, pp. 1364–1374, 2008.
- [3] F. D. McCool, "Global physiology and pathophysiology of cough: Accp evidence-based clinical practice guidelines," *Chest*, vol. 129, no. 1, pp. 48S–53S, 2006.
- [4] A. Morice, G. Fontana, M. Belvisi, S. Biring, K. Chung, P. Dicpinigaitis, J. Kastelik, L. McGarvey, J. Smith, M. Tatar, *et al.*, "Ers guidelines on the assessment of cough," *European respiratory journal*, vol. 29, no. 6, pp. 1256–1276, 2007.
- [5] British Thoracic Society and Scottish Intercollegiate Guidelines Network, "British guideline for the management of asthma; a national clinical guideline (sign 153)," September 2016, last accessed 07 February 2018. [Online]. Available: <https://www.brit-thoracic.org.uk/guidelines-and-quality-standards/asthma-guideline/>
- [6] —, "Chronic obstructive pulmonary disease in over 16s: diagnosis and management," June 2010, last accessed 07 February 2018. [Online]. Available: <https://www.nice.org.uk/guidance/CG101>
- [7] K. McGuinness, K. Holt, R. Dockry, and J. Smith, "P159 validation of the vitalojak 24 hour ambulatory cough monitor," *Thorax*, vol. 67, no. Suppl 2, pp. A131–A131, 2012.
- [8] M. G. Crooks, Y. Hayman, A. Innes, J. Williamson, C. E. Wright, and A. H. Morice, "Objective measurement of cough frequency during copd exacerbation convalescence," *Lung*, vol. 194, no. 1, pp. 117–120, 2016.
- [9] J.-M. Liu, M. You, Z. Wang, G.-Z. Li, X. Xu, and Z. Qiu, "Cough detection using deep neural networks," in *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Dept. of Control Sci. Eng., Tongji Univ., Shanghai, China. Piscataway, NJ, USA: IEEE, 2-5 Nov. 2014 2014, pp. 560–3.
- [10] J. Amoh and K. Odame, "Deep neural networks for identifying cough sounds," *IEEE transactions on biomedical circuits and systems*, vol. 10, no. 5, pp. 1003–1011, 2016.
- [11] T. Drugman, J. Urbain, N. Bauwens, R. Chessini, A.-S. Aubriot, P. Lebecque, and T. Dutoit, "Audio and contact microphones for cough detection," in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.
- [12] Y. A. Amrulloh, U. R. Abeyratne, V. Swarnkar, R. Triasih, and A. Setyati, "Automatic cough segmentation from non-contact sound recordings in pediatric wards," *Biomedical Signal Processing and Control*, vol. 21, pp. 126–136, 2015.
- [13] J. Monge-Alvarez, C. Hoyos-Barcelo, P. Lesso, and P. Casaseca-de-la Higuera, "Robust detection of audio-cough events using local hu moments," *IEEE Journal of Biomedical and Health Informatics*, 2018.
- [14] M. You, H. Wang, Z. Liu, C. Chen, J. Liu, X.-H. Xu, and Z.-M. Qiu, "Novel feature extraction method for cough detection using nmf," *IET Signal Processing*, vol. 11, no. 5, pp. 515–520, 2017.
- [15] L. Di Perna, G. Spina, S. Thackray-Nocera, M. G. Crooks, A. H. Morice, P. Soda, and A. C. den Brinker, "An automated and unobtrusive system for cough detection," in *Life Sciences Conference (LSC)*, 2017 IEEE. IEEE, 2017, pp. 190–193.
- [16] S. Matos, S. S. Biring, I. 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. 1078–83, 06 2006.
- [17] S. Larson, G. Comina, R. H. Gilman, B. H. Tracey, M. Bravard, and J. W. López, "Validation of an automated cough detection algorithm for tracking recovery of pulmonary tuberculosis patients," *PLoS one*, vol. 7, no. 10, p. e46229, 2012.
- [18] R. X. A. Pramono, S. A. Imtiaz, and E. Rodriguez-Villegas, "A cough-based algorithm for automatic diagnosis of pertussis," *PLoS one*, vol. 11, no. 9, p. e0162128, 2016.
- [19] J. D. Johnston, "Estimation of perceptual entropy using noise masking criteria," in *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on.* IEEE, 1988, pp. 2524–2527.
- [20] M. Wiśniewski and T. P. Zieliński, "Joint application of audio spectral envelope and tonality index in an e-asthma monitoring system," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 3, pp. 1009–1018, 2015.