Investigating into segmentation methods for diagnosis of respiratory diseases using adventitious respiratory sounds

Liqun Wu[1] and Ling Li[2]

Abstract— **Respiratory condition has received a great amount of attention nowadays since respiratory diseases recently become the globally leading causes of death. Traditionally, stethoscope is applied in early diagnosis but it requires clinician with extensive training experience to provide accurate diagnosis. Accordingly, a subjective and fast diagnosing solution of respiratory diseases is highly demanded. Adventitious respiratory sounds (ARSs), such as crackle, are mainly concerned during diagnosis since they are indication of various respiratory diseases. Therefore, the characteristics of crackle are informative and valuable regarding to develop a computerised approach for pathology-based diagnosis. In this work, we propose a framework combining random forest classifier and Empirical Mode Decomposition (EMD) method focusing on a multi-classification task of identifying subjects in 6 respiratory conditions (healthy, bronchiectasis, bronchiolitis, COPD, pneumonia and URTI). Specifically, 14 combinations of respiratory sound segments were compared and we found segmentation plays an important role in classifying different respiratory conditions. The classifier with best performance (accuracy = 0.88, precision = 0.91, recall = 0.87, specificity = 0.91, F1-score = 0.81) was trained with features extracted from the combination of early inspiratory phase and entire inspiratory phase. To our best knowledge, we are the first to address the challenging multi-classification problem.**

I. Introduction

There are over 400 million people being affected or dying from various respiratory diseases every year including COPD, asthma, lung cancer and etc [1]–[3]. Diagnosing of respiratory sounds is crucial for early prevision of respiratory diseases. Respiratory sounds reflect the condition of respiratory system as they are primarily generated by the turbulent and vorticose airflow through the trachea-bronchial tree. ARSs are usually expected from respiratory diseases affected subjects due to changes happened within the lung tissue and position of secretions within the trachea-bronchial tree [4]. Crackle, one of the, is discontinuous, explosive and non-musical sound superimposed in normal respiratory sound [5]. Crackle can be further categorised into two groups, early and late inspiratory crackle, based on the timing of start and the period carried on. A example of early and late inspiratory crackle were illustrated using phonopneumography as shown in Figure 1 [6]. It was revealed that an early inspiratory crackle takes place shortly after the beginning of inspiration and does not carry on beyond the first half of inspiration phase. In contrast, a late inspiratory crackles continues into second half of inhalation but may start at any time. Moreover, the timing of start of a crackle is highly correlated to subjects' respiratory condition. Specifically, early inspiratory crackle usually happens in patients with airway obstruction including chronic bronchitis, asthma and emphdema. While patients with lung defection, such as pneumonia, asbestosis and fibrosing alveolitis, are often encounter late inspiratory crackle [6].

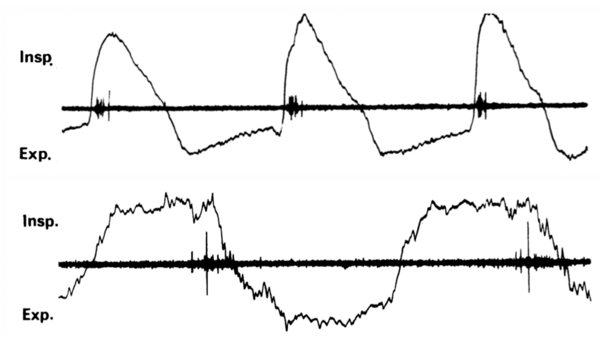

Fig. 1. Phonopneumogram of (top) RC containing early inspiratory crackles and (bottom) RC containing late inspiratory crackles [6]

Most previous research focuses on the detection of abnormal respiratory sounds including cough, crackle and wheeze [7]–[9]. However, due to the artifacts caused by heart beat, ambient and body motion involved respiratory sound, it is still challenging to develop computerised solutions to recognise adventitious sounds. Research focus has recently shifted to the development of automatic diagnosis in pathology level. Some recent work suggested approaches using various kinds of deep neural network for pathology based classification task. A wavelet transformation based approach was developed to identify the severity of asthma using fully connected neural network [10]. It has been demonstrated that by taking advantage of dominating features in sound analysis, Mel Frequency Cepstral Coefficient (MFCC) as inputs in convolutional neural network and recurrent neural can classify subjects with three respiratory conditions [11], [12]. [13] not only focuses on differentiating adventitious sounds but also proposed an approach to classify subjects into healthy and sick class.

The major contribution of this work is to suggest a machine learning framework to firstly address the multi-classification problem of identifying subjects in 6 respiratory conditions. The diagram of the proposed framework is shown as Figure 2. It also presents a domain knowledge informed approach for feature extraction. In section II, we describe the dataset selected for the development and demonstrate the

[1]Liqun Wu is with School of Computing, University of Kent, Medway, United Kingdom `l.wu@kent.ac.uk`

[1]Ling Li is with School of Computing, University of Kent, Medway, United Kingdom `c.li@kent.ac.uk`

Authorized licensed use limited to: UNIVERSITY OF KENT. Downloaded on September 19,2020 at 13:34:22 UTC from IEEE Xplore. Restrictions apply.

steps of preprocessing and features extraction. In section III we discussed the experimental results and findings. Finally, conclusions of the work and future plan are included in section IV.

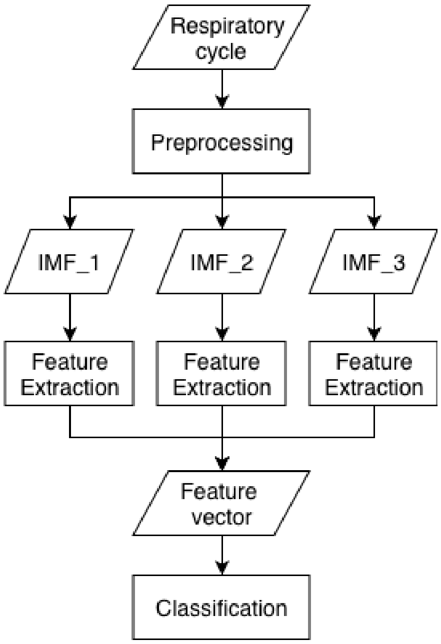

Fig. 2. The structure of classification system

II. Materials and methods

A. Data preparation

The dataset employed in this study is a respiratory sound collection originally developed for the competition of developing an automatic solution to detect ARSs held at International Conference on Biomedical and Health Information (ICBHI) 2017 [14]. The dataset contains a total of 5.5 hours of audio recordings including 6898 cycles collected from 126 participants with 8 respiratory conditions. There are 7 body locations and 3 recording devices being utilised in data acquisition. Annotations of respiratory cycles (RCs), respiratory condition of subject and label of respiratory sound are also provided by domain experts as ground truth. The original dataset is tailored to only preserve the respiratory cycles containing crackle to match the focus of this work. Table I presents the demographic insight of the original and modified dataset.

B. Preprocessing

1) Resampling: To reduce the size of signal processed and overcome the difference of sampling frequency among the acquisition equipment, each RC was firstly resampled at 4kHz.

2) Segmentation: A balanced inspiratory phase (IP) expiratory phase (EP) was assumed within a RC. Therefore, IP and EP are defined as the first and second half of a RC, respectively. Similarly, early inspiratory phase (EIP) and late inspiratory phase (LIP) are the first half and second half of an IP, respectively. Eventually, four segments are produced from a RC including EIP, LIP, IP and RC for the use of next steps.

TABLE I

Summary of ICBHI 2017 challenge dataset. (Respiratory condition: 1-asthma, 2-bronchiectasis, 3-bronchiolitis, 4-chronic obstructive pulmonary disease (COPD), 5-healthy, 6-lower respiratory tract infection (LRTI), 7-pneumonia and 8-upper respiratory tract infection (URTI))

Respiratory cycle	Respiratory condition							
	1	2	3	4	5	6	7	8
Normal	2	76	51	2725	303	31	240	214
Crackles	0	12	15	1779	16	0	21	21
Wheezes	4	31	64	752	2	1	24	8
Cs&Ws	0	10	5	490	1	0	0	0
Crackles (train)	0	8	10	37	13	0	10	18
Crackles (test)	0	4	5	16	3	0	11	3

3) Band-pass filter: Each segment was filtered by a 12th order band-pass filter with a targeted frequency range of 120-1800Hz. It removed artifacts caused by heart beat, ambient and body motion and only preserved the primary frequency components of crackles.

4) Hamming window: Hamming window was applied to each segment to prevent frequency leakage.

5) Empirical Mode Decomposition: EMD is highly recommended for the analysis of non-linear and non-stationary signal, such as respiratory sound, in previous studies [9], [15], [16]. EMD decomposes a signal into a set of sub-band signal components, also known as intrinsic mode functions (IMF), based on an iterative process and a set of evaluation criteria. Only the first three IMFs of each segment are retained for feature extraction in next stage. The results of applying above steps to a RC containing crackles collected from a subject with URTI is shown as Figure 3.

C. Feature extraction and classification

Each of the three IMFs from a segment was considered for feature extraction. A set of time domain and time-frequency domain features were taken into account including root mean square, maximum upper envelop amplitude, mean instantaneous frequency. Additionally, a number of widely used spectral features including centroid, bandwidth, flatness and rolloff were considered and their statistical variations including standard deviation, mean, median, maximum and minimum were also computed. Lastly, the mean of each of the first 13 MFCCs were derived. Totally, there were 36 features generated from an IMF. The features extracted from 4 respiratory phases (EIP, LIP, IP and RC) were then assembled as multiple feature vectors based on various experiment settings which are discussed in next section. The feature vectors were eventually be utilised to train a number of random forest classifiers.

D. Performance criteria

The modified dataset was next divided into 70% for training and 30% for test as shown in Table I. The performance

Authorized licensed use limited to: UNIVERSITY OF KENT. Downloaded on September 19,2020 at 13:34:22 UTC from IEEE Xplore. Restrictions apply.

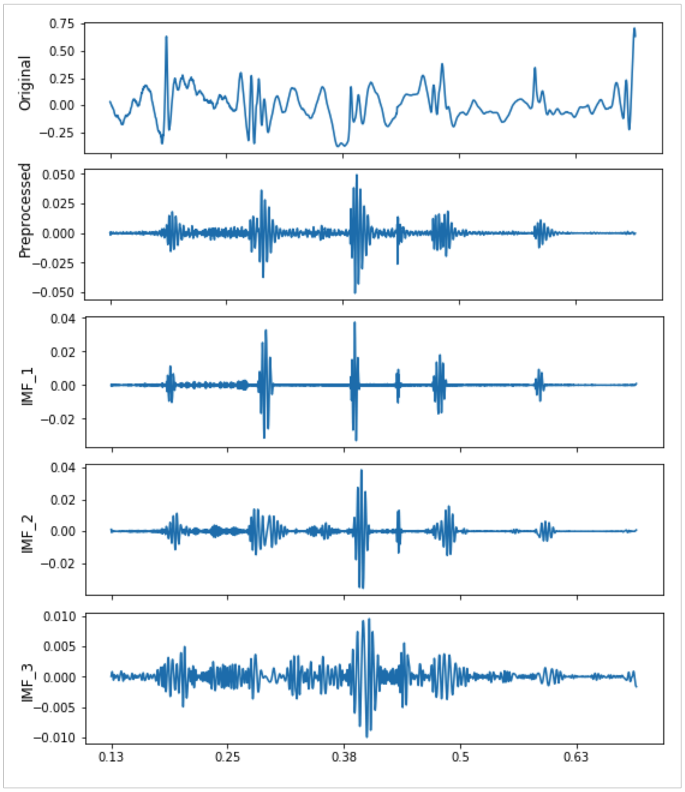

Fig. 3. (a) A RC signal containing crackles collected from a subject with URTI. (b) The results of applying resampling, band-pass filter and hamming window to the RC signal. (c), (d) and (e) illustrate the first 3 IMFs

of each of the classifiers was measured against the test set based on the following criteria:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad (5)$$

where TP is the number of true positive events, FP is the number of false positive events, TN is the number of true negative event and FN is the false negative events. Each of the classifier was evaluated and compared on macro averaged accuracy, precision, recall, specificity and F1-score.

III. Experimental Results

The primary goal of this study was to explore the combination of respiratory phases for feature extraction which provided optimal overall classification performance in pathology level. Totally, 15 classifier were trained using features established from a set of combinations of respiratory phases. The detailed settings and results of each of the classifiers are presented in Table II.

TABLE II
Performance of the 15 classifiers trained

Classifier	Phase(s)	Acc.	Pre.	Rec.	Spe.	F1.
1	EIP	0.69	0.71	0.64	0.93	0.66
2	LIP	0.48	0.45	0.35	0.89	0.37
3	IP	0.62	0.65	0.6	0.92	0.59
4	RC	0.57	0.56	0.6	0.92	0.54
5	EIP, LIP	0.67	0.61	0.58	0.92	0.54
6	**EIP, IP**	**0.88**	**0.91**	**0.87**	**0.97**	**0.87**
7	EIP, RC	0.67	0.62	0.68	0.94	0.6
8	LIP, IP	0.55	0.62	0.52	0.9	0.49
9	LIP, RC	0.62	0.61	0.63	0.92	0.61
10	IP, RC	0.69	0.63	0.69	0.94	0.63
11	EIP, LIP, IP	0.76	0.85	0.67	0.94	0.7
12	EIP, LIP, RC	0.62	0.56	0.61	0.93	0.53
13	EIP, IP, RC	0.67	0.61	0.68	0.93	0.61
14	LIP. IP, RC	0.67	0.64	0.63	0.94	0.6
15	EIP, LIP, IP, RC	0.69	0.66	0.68	0.94	0.63

A. Comparison of the 4 classifiers trained using features extracted from only one respiratory phase

It is widely applied that only using features extracted from RC in developing classification solutions. The characteristics contained in any particular respiratory phase within a RC were overlooked. This part of experiment was designed to explore if the classifiers trained with any specific respiratory phase have better performance than with RC. The results indicated that using EIP (0.66 in F1-score) and IP (0.59 in F1-score) in training classifiers provided better classification performance than using RC (0.54 in F1-score). However, it is interesting that the classifier performed the worst (0.37 in F1-score) was trained using features extracted from LIP.

B. Comparison of the 8 classifiers trained using features extracted from RC combining additional phase(s)

Since using either EIP or IP in training a classifier has improved performance than using RC, it is valuable to consider the impact of utilising additional phases(s) in conjunction with RC. According to the performance of classifier 4, 7, 9 and 10, EIP, LIP and IP had relatively equivalent improvement in term of F1-score (F1-score of classifier 4, 7, 9, 10 are 0.54, 0.6, 0.61 and 0.63). Furthermore, similar impact was also obeserved when comparing classifier trained with features extracted from two or three additional segments (F1-score of classifier 4, 13, 14 and 15 are 0.54, 0.61, 0.6 and 0.63). Therefore, it can be concluded that utilising additional phase(s) can result in approximately 6-9% improvement in F1-score.

C. Discuss of the classifier with overall best performance

Classifier 6 was trained using features extracted from EIP and IP which obtained the best performance (0.88 in accuracy, 0.91 in precision, 0.87 in recall, 0.97 in specificity and 0.87 in F1-score). It outperformed classifier 11 which was the second best classifier with a significant improvement (0.12 in accuracy, 0.06 in precision, 0.2 in recall, 0.03 in specificity and 0.17 in F1-score). The prediction results for each class of classifier 6 are presented in Table III. As presented in Table IV, it is shown that all RC samples collected

Authorized licensed use limited to: UNIVERSITY OF KENT. Downloaded on September 19,2020 at 13:34:22 UTC from IEEE Xplore. Restrictions apply.

from subjects with bronchiectasis, COPD and URTI were recognised. However, URTI class had the relatively most false positive events resulting only 0.6 precision. Importantly, the non-healthy class RC samples did not happen to be misclassified into healthy class.

TABLE III
DETAIL PERFORMANCE OF CLASSIFIER 6 WHICH WAS TRAINED USING FEATURES EXTRACTED FROM EIP AND IP

Respiratory condition	Precision	Recall	Specificity	F1-score
Bronchiectasis	**1**	**1**	**1**	**1**
Bronchiolitis	1	0.8	1	0.89
COPD	0.84	1	0.88	0.91
Healthy	1	0.67	1	0.8
Pneumonia	1	0.73	1	0.84
URTI	0.6	1	0.95	0.75

TABLE IV
CONFUSION MATRIX OF CLASSIFIER 6 WHICH WAS TRAINED USING FEATURES EXTRACTED FROM EIP AND IP

Actual class	Predicted class					
	Bronchie.	Bronchio.	COPD.	H.	P.	URTI.
Bronchie.	**4**	**0**	**0**	**0**	**0**	**0**
Bronchio.	0	4	0	0	0	1
COPD	0	0	16	0	0	0
H.	0	0	1	2	0	0
P.	0	0	2	0	8	1
URTI	0	0	0	0	0	3

IV. CONCLUSION AND FUTURE WORK

This work suggests a domain knowledge informed framework to tackle the challenging problem of identifying subjects in 6 respiratory conditions. A significant improvement of classification performance was achieved by utilising a combination of EIP and IP in feature extraction. It shows the importance of considering proved clinic findings in developing computerised solution. However, it was assumed that every RC has an even proportion of IP and EP. Thus, the most crucial work considered for the future is to develop an algorithm to objectively segment IP and EP from a RC and finally achieve phase dependent feature extraction. Importantly, it is worth to apply dimension reduction approach to reduce the total number of features used. Lastly, future work will also focus on exploring a wide range of classifiers, such as support vector machine and various kind of deep neural networks.

REFERENCES

[1] P. G. Burney, J. Patel, R. Newson, C. Minelli, and M. Naghavi, "Global and regional trends in COPD mortality, 1990-2010," *European Respiratory Journal*, vol. 45, no. 5, pp. 1239–1247, 2015. [Online]. Available: http://dx.doi.org/10.1183/09031936.00142414

[2] Global Asthma Network, "The Global Asthma Report," Auckland, 2018. [Online]. Available: www.globalasthmanetwork.org

[3] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," *CA: A Cancer Journal for Clinicians*, vol. 65, no. 2, pp. 87–108, 3 2015. [Online]. Available: http://doi.wiley.com/10.3322/caac.21262

[4] P. Forgacs, A. R. Nathoo, and H. D. Richardson, "Breath sounds," *Thorax*, vol. 26, no. 3, pp. 288–295, 5 1971. [Online]. Available: http://thorax.bmj.com/cgi/doi/10.1136/thx.33.6.681http://thorax.bmj.com/cgi/doi/10.1136/thx.26.3.288

[5] M. Sarkar, I. Madabhavi, N. Niranjan, and M. Dogra, "Auscultation of the respiratory system," *Annals of Thoracic Medicine*, vol. 10, no. 3, p. 158, 2015. [Online]. Available: http://www.thoracicmedicine.org/text.asp?2015/10/3/158/160831

[6] A. R. Nath and L. H. Capel, "Inspiratory crackles–early and late," *Thorax*, vol. 29, no. 2, pp. 223–227, 3 1974. [Online]. Available: http://thorax.bmj.com/cgi/doi/10.1136/thx.29.2.223

[7] B. M. Rocha, L. Mendes, I. Chouvarda, P. Carvalho, and R. P. Paiva, "Detection of Cough and Adventitious Respiratory Sounds in Audio Recordings by Internal Sound Analysis," ser. IFMBE Proceedings. Singapore: Springer Singapore, 2018, vol. 66, no. November, pp. 51–55. [Online]. Available: http://link.springer.com/10.1007/978-981-10-7419-6http://link.springer.com/10.1007/978-981-10-7419-6_9

[8] G. Serbes, S. Ulukaya, and Y. Kahya, "An Automated Lung Sound Preprocessing and Classification System Based OnSpectral Analysis Methods," ser. IFMBE Proceedings, N. Maglaveras, I. Chouvarda, and P. de Carvalho, Eds. Singapore: Springer Singapore, 2018, vol. 66, no. January, pp. 45–49. [Online]. Available: http://link.springer.com/10.1007/978-981-10-7419-6http://link.springer.com/10.1007/978-981-10-7419-6_8

[9] M. Lozano, J. A. Fiz, and R. Jane, "Automatic Differentiation of Normal and Continuous Adventitious Respiratory Sounds Using Ensemble Empirical Mode Decomposition and Instantaneous Frequency," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 2, pp. 486–497, 3 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7021891/

[10] F. Z. Göğüş, B. Karlık, and G. Harman, "Classification of Asthmatic Breath Sounds by Using Wavelet Transforms and Neural Networks," *International Journal of Signal Processing Systems*, vol. 3, no. 2, pp. 106–111, 2014. [Online]. Available: http://www.ijsps.com/index.php?m=content&c=index&a=show&catid=36&id=125

[11] D. Perna, "Convolutional Neural Networks Learning from Respiratory data," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 12 2018, pp. 2109–2113. [Online]. Available: https://ieeexplore.ieee.org/document/8621273/

[12] D. Perna and A. Tagarelli, "Deep Auscultation: Predicting Respiratory Anomalies and Diseases via Recurrent Neural Networks," in *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 6 2019, pp. 50–55. [Online]. Available: https://ieeexplore.ieee.org/document/8787435/

[13] G. Chambres, P. Hanna, and M. Desainte-Catherine, "Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis," in *2018 International Conference on Content-Based Multimedia Indexing (CBMI)*, vol. 2018-Septe. IEEE, 9 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8516489/

[14] B. Rocha, D. Filos, L. Mendes, G. Serbes, S. Ulukaya, Y. P. Kahya, N. Jakovljevic, T. L. Turukalo, I. M. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, N. Maglaveras, R. Pedro Paiva, I. Chouvarda, and P. de Carvalho, "An open access database for the evaluation of respiratory sound classification algorithms," *Physiological Measurement*, vol. 40, no. 3, p. 035001, 3 2019. [Online]. Available: http://stacks.iop.org/0967-3334/40/i=3/a=035001?key=crossref.e1fee2d24f8b96697b4f52483002604c

[15] M. Lozano, J. A. Fiz, and R. Jane, "Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 7 2013, pp. 981–984. [Online]. Available: http://ieeexplore.ieee.org/document/6609667/

[16] C. Zong and M. Chetouani, "Hilbert-Huang transform based physiological signals analysis for emotion recognition," in *2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 12 2009, pp. 334–339. [Online]. Available: http://ieeexplore.ieee.org/document/5407547/

Authorized licensed use limited to: UNIVERSITY OF KENT. Downloaded on September 19,2020 at 13:34:22 UTC from IEEE Xplore. Restrictions apply.