Biomedical Signal Processing and Control 46 (2018) 94–103

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

A robust audio classification system for detecting pulmonary edema

K.J. Hong[a,b,*], S. Essid[a,b,**], W. Ser[a,b], D.C.-G. Foo[b,c]

[a] Nanyang Technological University, 50 Nanyang Ave, Singapore 639798, Singapore
[b] LTCI, Télécom ParisTech, Université Paris-Saclay, Paris 75013, France
[c] Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng, Singapore 308433, Singapore

ARTICLE INFO

Article history:
Received 13 April 2018
Received in revised form 11 June 2018
Accepted 10 July 2018

Keywords:
Non-negative matrix factorization
Robust testing
Pulmonary edema
Biomedical signal processing
Feature learning

ABSTRACT

In this paper we present a robust audio classification system to efficiently detect pulmonary edema. The system uses a feature learning technique based on (NMF), then classified with logistic regression. A study was done to compare feature engineering approaches with feature selection techniques against NMF. Different NMF schemes were investigated and also compared with Principal Component Analysis. NMF scored 95% F1 score, which was superior to feature engineering techniques that had scores from 83% to 93%. Background noise collected from hospitals and speech from a speech corpus database was used to simulate noisy data. The system was then tested using noisy data. The best NMF scheme scored 74%, while other feature engineering techniques scored lower; from 66% to 71%. NMF was also used as a signal enhancement tool. It improved the F1 score to 77%. Lastly, only inhalations from breath sounds were considered and this further improved classification results to 86%. The proposed robust classification system using NMF thus proved to be an effective method for audio-based detection of pulmonary edema. If implemented in real-time, the proposed system can be used as a screening tool.

1. Introduction

Pulmonary edema is a condition where water engorges the alveolar beds. The water is pushed into air spaces and thus reduces normal oxygen movement. It can cause hemoptysis, difficulty in breathing, shorted of breath, and gurgling and wheezing sounds during breathing. It can be caused by congestive heart failure, kidney failure, high altitude exposure, lung damage, and major injuries.

Currently, physicians will carry out physical examinations before performing X-rays, CT-scans and electrocardiograms. Physicians mainly use auscultations to check for abnormal heart sounds, crackles, increased heart rate and rapid breathing. Most patients with pulmonary edema must face the disease and its implications for the rest of their lives. They must visit the physician for a check-up whenever they show symptoms, and if left untreated, they could suffer suffocation. These constant trips to the physician can add stress, worry and inconvenience. Physicians also in turn face a greater workload.

Nomenclature

ADC	analog-to-digital converter
CV	cross-validation
DAS	discontinuous adventitious sounds
ELW	excessive lung water
EN	elastic net
IS	Itakura–Saito
k-NN	k-nearest neighbour
KL	Kullback–Leibler
LPC	Linear predictor coefficients
LR	logistic regression
MFCC	Mel-frequency cepstral coefficients
MU	multiplicative update
NMF	non-negative matrix factorization
PAF	“popular” audio features
PCA	principal component analysis
RFE	recursive feature elimination
SN	sensitivity
SP	specificity
SVM	Support vector machine

Thus, there exists a need for an automatic, quick, accessible, and simple screening solution for the detection of pulmonary edema.

* Corresponding author at: Nanyang Technological University, 50 Nanyang Ave, Singapore 639798, Singapore.
** Principal corresponding author.
E-mail addresses: khong008@e.ntu.edu.sg (K.J. Hong), slim.essid@telecom-paristech.fr (S. Essid), ewser@ntu.edu.sg (W. Ser).

https://doi.org/10.1016/j.bspc.2018.07.004
1746-8094/

With this, patients need not solely rely on check-ups by physicians to ascertain excessive accumulation of water but will be able to carry out screening with greater convenience.

The paper proposes a more thorough study on the use of audio classification techniques for the automatic detection of excessive lung water. It introduces an original system which proves highly effective for this task. The system is tested and validated by augmented samples collected from a local hospital. This paper presents results and analysis of a proposed general framework that extracts, transforms, and classifies features for the detection of excessive lung water in real world sound recordings. Moreover, special care is taken to assess the robustness of the system in noisy recording conditions.

2. Previous work

Conventionally, auscultations are used to help diagnose a vast number of other conditions, as well as pulmonary edema. The physician pays attention to abnormal respiratory sounds called adventitious sounds. In the case of pulmonary edema, the physician listens to (DAS) called crackles. Thus, there has been much research developed to detect DAS found in lung sounds. Recently, Pramono et al. [23] presented a systematic review of automatic adventitious respiratory sound analysis. Regarding the analysis of crackles, they cited 36 papers that involve detection or classification and listed their data sources, amount of data, validation method, features used, classification method and performance. They concluded that performance of recent studies showed a high agreement with conventional non-automatic identification and suggested that automated adventitious sound detection or classification was a promising solution to overcome the limitations of conventional auscultation and to assist in the monitoring of relevant diseases. However, depending on the severity of the disease, these DAS might not manifest frequently or at all. Furthermore, studies such as [22,28,33,21,27,32] present event-based classification, where the DAS was isolated manually before classification. This approach would perform poorly if crackles do not appear in a particular segment of audio data. Also, the same DAS appear in various diseases. Fine crackles appear in patients with bronchiectasis, pulmonary edema, asthma, chronic bronchitis, severe sepsis, pneumonia, congestive heart failure, pulmonary fibrosis, and acute bronchitis. Consequently, the classification of DAS does not systematically translate to the classification of diseases. Also, the aim of the proposed study is to detect pulmonary edema in any segment of the breath sound, and not just segments with crackles.

On the other hand, there have been a few works that involved classifying diseases using lung sounds. Hernandez et al. [15] presented a classification system for recordings of lung sounds with diffused interstitial pneumonia against healthy subjects. The authors used a multivariate auto-regressive model with a supervised neural network as a classifier. [24] presented a system for automated diagnosis of pertussis using audio signals by analyzing cough and whoop sounds. Yang and Wee first proposed a signal processing approach for detecting excessive lung water using sound-based sensing [31]. Different features and classifiers were explored, and Mel-frequency cepstral coefficients (MFCC) together with k-nearest neighbor (k-NN) produced the best result. In our previous work [7], we proposed a feature extraction method that segmented the magnitude spectrum of lung sound recordings into sub-bands. The sub-band spectral coefficients were used as features that were ranked using principle component analysis and support vector machine-based recursive feature elimination. classification was done using k-NN and Support Vector Machine (SVM). *F*-measure of up to 99% was reported. However, although the sample size was small, the use of appropriate testing and validation procedures was lacking, which puts into question the statistical validity of the results and the generalization ability of the previously proposed system. Also, the system hyperparameters were not optimized. Lastly, there was no study of the robustness to noise.

Secondly, all studies cited use different kinds of feature engineering techniques to obtain features. For example, the standard choice for speech and audio classification task, MFCC [10,4], was used by [25,18,29,19,17]. All proposed feature engineering methods as part of their solutions had success of various degrees. However, in recent years feature learning has become the dominant trend in machine learning problems. Similar to learning in classifiers, feature learning entails algorithms automatically learning features determined by data. It is an alternative to feature engineering (as shown in Fig. 1). Deep learning [11] has been a popular choice for many classification and regression problems. It also has been used in audio related problems [2,13]. However, deep learning methods require huge amounts of data, and thus for applications with small datasets, signal specific feature engineering is preferred. Nevertheless, due to the nature of this problem, the sample size involved in this work is small and despite the popularity of deep learning, it could not be implemented successfully. Still, research such as [1] has shown that another popular method called non-negative matrix factorization (which has been successful in audio source separation and enhancement [12,26,20]) had success similar to deep learning methods in the audio classification domain when there is a lack of training data. Also, NMF was used in [14] for blind source separation of heart and lung sounds. Thus, we adopt NMF as a feature learning method in this paper.

Thus, we propose a robust system for the detection of pulmonary edema using classical classification techniques as well as state-of-the-art non-negative matrix factorization for feature learning. This is an alternative to indirectly detecting crackles to screen for pulmonary edema. The system was also tested under the introduction of environmental noise found in hospitals. The system hyperparameters were optimized, and cross-validation was used to reliably assess its performance.

3. Methodology

Fig. 1 shows the flow of data processing as a chart. First, the breathing sounds of the subjects were recorded using a stethoscope sensor. Next, synthetic data containing recordings of excessive water were generated so that the number of ill patient samples and healthy subjects were approximately equal. On top of that, synthetic test data was generated to assess the generalization ability of our system and improve the reliability of the evaluation metrics. A skewed ratio between the two classes would result in a biased sensitivity, specificity and *F*-measure. More details are in Section 4. Data was segmented into 250 ms segments and normalized. Features of the data were then extracted. The system was tested with and without feature selection and feature transformation techniques. The logistic regression (LR) classifier was trained using training data and the class probability was obtained using testing data. Finally, the class of each test data point was predicted.

It is common to transform or select some of the initial features to increase final classification accuracies. In the next subsection, a brief introduction of the techniques used in the feature extraction, transformation and classification steps are given.

3.1. Feature extraction

First, we consider various of popular audio features to compare with our NMF-based approach. Table 1 shows the list of these