This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Wearable stethoscope for lung disease diagnosis

Chundong Qiu, a,b Wenru Zeng, b Wei Tian, a Jingyi Xu, a Yingnan Tian, a Chao Zhao *a and Hong Liu *a

Lung disease is one of the most popular diseases especially in the era of COVID-19. Its diagnosis is of great importance, as different types have diverse treatments and prognosis. The most popular methods are Computed Tomography scan, ultrasonogram or by bioimpedance sensors, while they are not suitable for wearable applications. Here we developed a wearable stethoscope with accompanying algorithm for lung disease diagnosis. It was demonstrated on 18 patients in hospital with three types of lung disease. After lung sound signals collecting and pre-processing, several machine learning methods with optimized features were applied and achieved high classification metrics. The features of low frequency wavelets decomposed from the lung sound signals were found to be important, serving as potential biomarkers for different lung disease types. Overall, it was proven that our wearable stethoscope could provide a more user-friendly method and find more application scenarios for lung disease diagnoses.

Introduction

Lung diseases are one of the most common diseases in the world. As lung is a complex system, its disease may be divided into many types with totally different causes and symptoms, affecting its diagnosis, treatment, and prognosis. In general, lung diseases can be divided by its affecting regions, such as airways, air sacs, interstitium and others. For lung diseases affecting the airways, chronic obstructive pulmonary disease (COPD) is one of the most common. COPD causes obstructed air flow from the lungs, resulting in symptoms such as breathing difficulty, and is often triggered by long-term exposure to cigarette smoke. However, it’s treatable by bronchodilators using inhalers after diagnosis. For lung diseases affecting the air sacs, pneumonia is the most popular, especially in the era of COVID-19. Pneumonia may cause the air sacs filling with fluid or pus, affecting mostly on children younger than 2 years old and people older than 65 years old. After diagnosis, its symptoms may ease in a few days, while the feeling of tiredness can stay for a longer time. It is generally treated by antibiotics. For lung diseases affecting the interstitium, interstitial lung disease (ILD) is prevalent, which causes progressive scarring of lung tissue caused by long-term exposure to hazardous materials, such as asbestos. ILD is generally irreversible, so an early diagnosis is critical. Many ILD people are initially treated with a corticosteroid and other drugs to suppress the immune system.

In general, as the lung diseases happen in different parts of the respiratory system, its early and accurate diagnosis is critical for different treatment strategies with various prognosis. Currently, it is performed by pulmonary function test, which tests the amount of the air inhaling and exhaling from the lungs, artery blood gas analysis, pulse oximetry, and sputum test. Furthermore, chest X-ray, Computed Tomography (CT) scan, and echocardiogram could be utilized to find the severity and location of lung diseases. However, for diagnosis on rather healthy people, portable or even wearable biomedical devices may find more application scenarios than conventional instruments in the hospital, and provide a key factor for early diagnosis of lung diseases. Also, for people in hospital, wearable device provides a more friendly method than harassing the patient moving from the ward to the department with instruments, especially for patients in Intensive Care Unit (ICU).

Current wearable technology is based on sensing physical vital sign or biochemical signals of the subject. Optical method like photoplethysmography (PPG) used visible or IR light to sense the signal, but could not reach the lungs for only penetrating a few millimetres of the skin and is mostly used to detect the artery blood vessel. Electrophysiological method such as electrocardiography (ECG), electroencephalogram (EEG) or electromyography (EMG) could not be applied to the lungs, which generate too weak electrical signals. Alternatively, bioimpedance device was studied to monitor tidal volume and respiratory rate, and used to classify respiration disorders, like apnea and hypopnea. However, bioimpedance measurement requires the injection of a weak current into the body, which can be noticeable to the subject, affecting the user experience and making it unsuitable for long-term monitoring. Respiration caused vibration or micromotion on the chest could be directly sensed by strain or inertial sensors, while its information was not as rich as lung sounds. As a result, wearable lung sensing is focused on its generated sounds by ultrasonogram (USG). Yet, ultrasonic technique is harmful for the tissue, and not recommended for usage longer than 30 min, which may miss the abnormal lung sound. Also, its complicated
piezoelectric sensor array makes it not widely affordable. Nevertheless, stethoscope, which also detects sound signal, is the mostly widely used tool for doctors to detect lung diseases. It’s passive on receiving lung sounds which could be worn for hours or days, and experienced doctors could make an initial diagnosis on the type of lung disease directly using the stethoscope alone. Besides the sensing method, current research is focused on the lung sound type, such as crackling, rhonchi, wheezing and stridor20-27. However, the lung sound type is not directly related to the diagnosis of the type of lung disease, and one type is sometimes related with several abnormal lung sound.

In order to take the advantage of stethoscope, make it digital, and solve the dilemma for lung disease diagnosis, we optimized a wearable small-scale electronic stethoscope (WSES) system developed by us before17, and applied it to 18 patients with lung diseases in hospital. The system has a microphone IC chip with an integrated microelectromechanical system (MEMS). Its small size, low power consumption, high signal to noise ratio (SNR) makes the wearable stethoscope applicable for patients in hospital. A user-friendly mobile application was developed to collect signals from the device for further diagnosis, making our device cable free. In order to diagnosis the lung disease type, several machine learning methods were compared to get the best performance. Our results showed that the system could reach a diagnosis accuracy higher than 90%, and could be applied in hospital to assist the doctors for lung diseases diagnosis. Fig. 1 showed the overall flow chart of our study design. Lung sounds from patients with different lung disease were collected, and pre-processed to remove noise and artifacts. Then, features in the time domain, from wavelet decomposition and demography were extracted. Next, different features sets were compared, and nine machine learning algorithms were demonstrated to find the optimized algorithms with highest performance metrics.

### Methods

#### Subjects and Ethical Considerations

<table>
<thead>
<tr>
<th>No</th>
<th>Sex</th>
<th>Age</th>
<th>HR</th>
<th>SpO₂</th>
<th>RR</th>
<th>SBP</th>
<th>DBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>male</td>
<td>91</td>
<td>60</td>
<td>99</td>
<td>30</td>
<td>137</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>female</td>
<td>84</td>
<td>114</td>
<td>95</td>
<td>50</td>
<td>110</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>male</td>
<td>87</td>
<td>75</td>
<td>98</td>
<td>17</td>
<td>130</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>female</td>
<td>88</td>
<td>114</td>
<td>97</td>
<td>18</td>
<td>128</td>
<td>79</td>
</tr>
<tr>
<td>5</td>
<td>female</td>
<td>75</td>
<td>96</td>
<td>98</td>
<td>36</td>
<td>105</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>male</td>
<td>67</td>
<td>70</td>
<td>99</td>
<td>25</td>
<td>105</td>
<td>67</td>
</tr>
<tr>
<td>7</td>
<td>female</td>
<td>70</td>
<td>89</td>
<td>100</td>
<td>13</td>
<td>143</td>
<td>78</td>
</tr>
<tr>
<td>8</td>
<td>female</td>
<td>72</td>
<td>84</td>
<td>99</td>
<td>26</td>
<td>101</td>
<td>51</td>
</tr>
<tr>
<td>9</td>
<td>female</td>
<td>82</td>
<td>76</td>
<td>99</td>
<td>18</td>
<td>138</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>male</td>
<td>75</td>
<td>65</td>
<td>98</td>
<td>24</td>
<td>113</td>
<td>62</td>
</tr>
<tr>
<td>11</td>
<td>female</td>
<td>55</td>
<td>53</td>
<td>95</td>
<td>22</td>
<td>110</td>
<td>66</td>
</tr>
<tr>
<td>12</td>
<td>male</td>
<td>79</td>
<td>70</td>
<td>99</td>
<td>14</td>
<td>111</td>
<td>N/A</td>
</tr>
<tr>
<td>13</td>
<td>female</td>
<td>80</td>
<td>96</td>
<td>98</td>
<td>21</td>
<td>159</td>
<td>71</td>
</tr>
<tr>
<td>14</td>
<td>male</td>
<td>69</td>
<td>71</td>
<td>83</td>
<td>30</td>
<td>152</td>
<td>91</td>
</tr>
<tr>
<td>15</td>
<td>male</td>
<td>63</td>
<td>106</td>
<td>88</td>
<td>15</td>
<td>120</td>
<td>72</td>
</tr>
<tr>
<td>16</td>
<td>male</td>
<td>55</td>
<td>62</td>
<td>96</td>
<td>17</td>
<td>98</td>
<td>64</td>
</tr>
<tr>
<td>17</td>
<td>female</td>
<td>74</td>
<td>72</td>
<td>99</td>
<td>24</td>
<td>126</td>
<td>73</td>
</tr>
<tr>
<td>18</td>
<td>male</td>
<td>69</td>
<td>87</td>
<td>89</td>
<td>24</td>
<td>125</td>
<td>83</td>
</tr>
</tbody>
</table>

Table 1 Demographic and pathological characteristics of 18 subjects. HR: heart rate/BPM, RR: respiration rate/RPM, SBP: systolic blood pressure/mmHg, DBP: diastolic blood pressure/mmHg.

---

**Figure 1 Flow chart of diagnosis of lung disease types by our wearable stethoscope system.**

---

*This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.*
We collected around 3 minutes of lung sound for each patient. The lying flat, kept the body posture unchanged as much as possible. 3.7. demonstrated on MATLAB R2021a and Scikit-Learn 0.24.2 in Python based on NordicSemi nRF SDK 15.3. The algorithm was the general public. The firmware was developed with uVision 5.28 mAh. The device cost is tens of US dollars, and could be afforded by patient skin for sanitary purpose, and the battery used was 250 (Huaxi Sanitary Materials) was inserted between the device and which is enough for lung sounds collecting. A blue medical tape mobile phone. The device works at a sampling rate of 2.4 kHz, nRF52840, which has a Bluetooth module to send the data to the sense the sound from the body and transmit the digital signal to (Nordic Semiconductor) with 3.5 mm by 3.6 mm size. The ICS-43432 by 3.5 mm size, and the microcontroller unit (MCU) is nRF52840 wearing it. The sensing chip is ICS-43432 (InvenSense) with 3.5 mm Hardware and Experimental Procedure

This study was conducted in 18 patients with lung disease recruited from Nanjing Drum Tower Hospital. 9 males and 9 females were with a mean age of 74.2 years (standard deviation ± 10.4 years), with specific demographic and pathological characteristics detailed in Table 1. All subjects had normal hearing, normal vision, and normal speech function. The subjects were conscious and able to communicate with doctors, and the gold standard for lung disease type was evaluated by the doctor before the experiment. Prior to the experiment, subjects were informed of the procedure and related precautions, and that the experiment was not harmful to humans. All participants signed an informed consent form, and the study was approved by the Ethics Committee of Southeast University. Data were obtained in accordance with the guidelines of the University Ethics Committee and the ethical principles of the Declaration of Helsinki for medical research involving human beings.

Hardware and Experimental Procedure

Fig. 2 exhibits the design of the device and photo of a patient wearing our device. The sensing chip is ICS-43432 (InvenSense) with 3.5 mm by 3.5 mm size, and the microcontroller unit (MCU) is nRF52840 (Nordic Semiconductor) with 3.5 mm by 3.6 mm size. The ICS-43432 sense the sound from the body and transmit the digital signal to nRF52840, which has a Bluetooth module to send the data to the mobile phone. The device works at a sampling rate of 2.4 kHz, which is enough for lung sounds collecting. A blue medical tape (Huaxi Sanitary Materials) was inserted between the device and patient skin for sanitary purpose, and the battery used was 250 mAh. The device cost is tens of US dollars, and could be afforded by the general public. The firmware was developed with uVision 5.28 based on NordicSemi rnf SDK 15.3. The algorithm was demonstrated on MATLAB R2021a and Scikit-Learn 0.24.2 in Python 3.7.

In a quiet environment in the ward, the subjects wore our device lying flat, kept the body posture unchanged as much as possible. We collected around 3 minutes of lung sound for each patient. The

| Table 2 Features of time domain, frequency domain and nonlinear domain |

<table>
<thead>
<tr>
<th>Feature Type</th>
<th>Max</th>
<th>Min</th>
<th>Max-min</th>
<th>Standard deviation</th>
<th>Mean of absolute of coefficients</th>
<th>Mean of power of coefficients</th>
<th>Standard deviation of coefficients</th>
<th>Kurtosis</th>
<th>Mean of absolute of derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time domain features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelet decomposition features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Pre-processing and Feature Extraction

The APP saved lung sound data to the memory of the smartphone as a text file. The name of the file contains time information, such as "Tue Nov 08 10:09:47 GMT+08:00 2023-282". The first six digits of the last string are time stamps and the seventh digit is the order of collected data. The text file is read into MATLAB, and the sound data of the corresponding period was searched and intercepted, according to the time stamp recorded during the experiment. The data is processed with a high pass filter of 100 Hz and a wavelet denoising algorithm.

After the pre-processing, the features in the time domain, frequency domain, and nonlinear domain of the signal were extracted, listed in Table 2. Other information such as HR, SpO2, RR, SBP and DBP on Table 1 were collected but not included as features, because during collecting period they were not continuously monitored. They were only utilized to confirm that the patient status is stable. Here, 9 features of the maximum value, minimum value, maximum minus minimum value, standard deviation, mean of absolute, median of absolute, mean, kurtosis, mean of absolute of derivative in the time domain were extracted. Next, mean of absolute of coefficients, mean of power of coefficients, standard deviation of coefficients of five wavelets were extracted, resulting in 15 features. Also, the ratio between mean of absolute of coefficients were derived as features. Adding sex and age as features, in total there are 30 features involved.

Machine Learning Models

Supervised learning is the most applicable method for lung disease classification, where input data with labels are propagated by an algorithm that then learns the patterns associated with each label. A supervised machine learning model is trained with extracted features as the dataset and pre-obtained lung disease types as labels. This study utilizes the base classification models ExtraTreesClassifier, DecisionTreeClassifier, LinearDiscriminantAnalysis, LogisticRegression,
QuadraticDiscriminantAnalysis, KNeighborsClassifier, AdaBoostClassifier, SVC, MLPClassifier from the python machine learning library scikit-learn. Parameter optimization is performed by RandomizedSearchCV. The performance of the above machine learning models is judged by Accuracy, Precision, Recall, and F1 score. Accuracy, recall, precision, F1 scores were used to evaluate our algorithms. They are related to True Positives (TP), True Negative (TN), False Positives (FP), and False Negatives (FN). Accuracy is calculated by \( \frac{TP+TN}{TP+TN+FP+FN} \), recall is quantified by \( \frac{TP}{TP+FN} \) and precision is \( \frac{TP}{TP+FP} \). F1 metric is the harmonic mean of precision and recall for overall performance evaluation. Also, confusion matrix is studied.

The features were firstly standardized by removing the mean and scaling to unit variance. Dataset was divided into training and test subsets at a ratio of 9:1. The classifier was trained by using the training subsets, and then utilized to predict on the test subsets. Algorithm’s capability and performance to tell the difference between difference lung disease was studied. K-fold was adopted for cross-validation. The dataset was split randomly into K parts, with one part as test subsets and the remaining as training subsets.

**Results and Discussion**

**Feature Extraction and Selection**

The acquired lung sound signals were filtered using 100 Hz high pass filter as lung sound frequency is typically higher than 100 Hz, and then decomposed using 'wavedec' at level 5 with the wavelet 'coif4' in MATLAB. Then, 'wthresh' and 'waverec' were utilized to denoise and reconstruct the lung sounds signals. Fig. 3 showed the raw data, and signals denoised using wavelet thresholds.
After feature extraction, the ranking of feature importance is calculated and shown in Fig. 4a. Two sets of features (28 features vs. 30 features) were tried. The classification using 28 features in the time domain and wavelet decomposition results in accuracy at 0.85 ± 0.04, recall at 0.84 ± 0.04, precision at 0.84 ± 0.03, F1 at 0.83 ± 0.03, which is rather good already. Interestingly, Fig. 4b showed that after adding the demographic information, all performance metrics boosted with accuracy at 0.99 ± 0.01, recall at 0.98 ± 0.02, precision at 0.99 ± 0.01, F1 at 0.99 ± 0.01. Indeed, age and gender are the most important features, which indicates the relationships between lung disease type and demographic information. Thus, the features were decided including all 30 features for all machine learning methods. It should be noted that obviously only age and sex could not decide the lung disease type, as all three lung disease patients are with wide range of ages and random sex. Besides age and sex, the features related to low frequency wavelets ranked higher, such as coefficients from the decomposition at level 5 and level 4, that indicate the lower frequency lung sounds are biomarkers for diagnosis. Also, it could be observed that the higher frequency wavelet features, the lower its rankings in the frequency importance.

**Classification Results**

After feature selection, nine machine learning methods were compared and the model performance is calculated. The classification results of the top four algorithms were shown in Fig. 5. K-fold validation was utilized on accuracy, recall, precision and F1 to test the robustness of the model. ExtraTreesClassifier’s optimized hyperparameters were `n_estimators = 125, min_samples_split = 3, min_samples_leaf = 1, criterion = entropy`, with the mean cross validation F1 score at 0.98 ± 0.016. MLPClassifier’s optimized hyperparameters were `alpha = 0.0001, hidden_layer_sizes = (200, 200, 200), max_iter = 1000, learning_rate_init = 0.001, tol = 0.0001, beta_1 = 0.99, beta_2 = 0.99, epsilon = 1e-7` with the mean cross validation F1 score of 0.96 ± 0.015. KNeighborsClassifier’s optimized hyperparameters were `algorithm = ball_tree, leaf_size = 40, n_neighbors = 30, p = 1, weights = distance` with the mean cross validation F1 score at 0.94 ± 0.022. svm.SVC’s optimized hyperparameters were `C = 2, kernel = rbf, gamma = scale, tol = 0.0001` with the mean cross validation F1 score at 0.95 ± 0.018. ETC performed the best, followed by MLPC and KNN. Fig. 5b demonstrated the confusion matrix of ETC especially with high metric for PI and IPF. The only missed label is wrongly classified others as PI. With our wearable device with corresponding optimized diagnosis algorithm, it is proven that our system could accurately diagnosis lung disease, and could find further applications in clinics.

Nevertheless, several insights and limitations were identified. Firstly, different biocompatible adhesives could be applied to attach our device to the patient skin to minimize the environmental noise and match the sound impedance, to get the highest SNR. Secondly, different types of sound could be categorized in detail, to find the relationships between the sound type and lung disease types. Thirdly, most of the patients were older than 60 years old, and our system’s robustness on lung disease diagnosis should be proven with other range of ages.

**Conclusions**

Lung disease types classification is critical for deciding the treatment strategy and prognosis monitoring. Currently most lung disease study with wearable technology are focused on sound types classification which is not clinically significant. Here, we introduced a wearable stethoscope and applied on 18 lung disease patients with three most popular disease types. Machine learning methods were used to achieve high performance metrics, proving its potential for application in clinics.

**Author Contributions**

...
C.Z. and H.L. conceived the research idea. C.Q. and W.Z. performed experiments. W.T. analyzed the data. C.Q. and C.Z. wrote the paper. All authors discussed the results and reviewed the manuscript.

Conflicts of Interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Key Research and Development Program of Jiangsu Province (BE2021700), Science and Technology Development Program of Suzhou (SYG202117), the National Natural Science Foundation of China (62001104, 62271136), Natural Science Foundation of Jiangsu Province (BK20200357), Key Project and Open Research Fund of State Key Laboratory of Bioelectronics, the Fundamental Research Funds for the Central Universities and Zhishan Young Scholars of Southeast University. We thank all doctors, nurses and patients related to this paper.

Notes and References


21. N. S. Haider, Biomedical Signal Processing and Control, 2021, 64.