Journal of Medical Systems (2019) 43: 255
https://doi.org/10.1007/s10916-019-1388-0

IMAGE & SIGNAL PROCESSING

Respiratory Sound Based Classification of Chronic Obstructive Pulmonary Disease: a Risk Stratification Approach in Machine Learning Paradigm

Nishi Shahnaj Haider [1] · Bikesh Kumar Singh [1] · R. Periyasamy [2] · Ajoy K. Behera [3]

Received: 12 March 2019 / Accepted: 14 June 2019 / Published online: 28 June 2019

Abstract

This article investigates the classification of normal and COPD subjects on the basis of respiratory sound analysis using machine learning techniques. Thirty COPD and 25 healthy subject data are recorded. Total of 39 lung sound features and 3 spirometry features are extracted and evaluated. Various parametric and nonparametric tests are conducted to evaluate the relevance of extracted features. Classifiers such as support vector machine (SVM), k-nearest neighbor (KNN), logistic regression (LR), decision tree and discriminant analysis (DA) are used to categorize normal and COPD breath sounds. Classification based on spirometry parameters as well as respiratory sound parameters are assessed. Maximum classification accuracy of 83.6% is achieved by the SVM classifier while using the most relevant lung sound parameters i.e. median frequency and linear predictive coefficients. Further, SVM classifier and LR classifier achieved classification accuracy of 100% when relevant lung sound parameters, i.e. median frequency and linear predictive coefficient are combined with the spirometry parameters, i.e. forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV_1). It is concluded that combining lung sound based features with spirometry data can improve the accuracy of COPD diagnosis and hence the clinician's performance in routine clinical practice. The proposed approach is of great significance in a clinical scenario wherein it can be used to assist clinicians for automated COPD diagnosis. A complete handheld medical system can be developed in the future incorporating lung sounds for COPD diagnosis using machine learning techniques.

Keywords Chronic obstructive pulmonary disease diagnosis · Lung sound · Feature extraction · Spirometry · Machine learning · Risk stratification

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is an obstructive pulmonary disease which is progressive by nature. This disease ranked at the 3rd position for causing world-wide health crisis [1–3]. It is a wide umbrella term consisting of diseases like emphysema, chronic bronchitis, and refractory asthma [1]. As per the global health organization, around 3.17 million people have lost their lives because of such life-threatening disease. COPD is though not treatable, the progress of it can be controlled with proper medication [2, 4].

Chronic Obstructive Pulmonary Disease is a chronic respiratory disorder which affects the respiratory process, thereby deteriorating the exhalation capability of the lungs, subjecting to poor health quality. Cough, sputum production, breathing difficulty and wheezing are its symptoms. Symptoms at its very severe stage are weight loss, coronary heart disease, depression, anorexia, cognitive dysfunction and lung cancer [5]. The air sacs and airways are elastic by nature. Lungs are having stretching property. They expand and compress during air passages through it while respiration process. This elastic property deteriorates slowly in people with COPD [6].

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Bikesh Kumar Singh
bsingh.bme@nitrr.ac.in

1 Biomedical, NIT Raipur, G E Road, Raipur 492010, India

2 Department of Instrumentation and Control Engineering, NIT Tiruchirappall, Tamil Nadu 620015, India

3 Department of Pulmonary Medicine, AIIMS Raipur, Tatibandh, G E Road, Raipur 492010, India

Stretching property of the lungs is affected, airways are obstructed and mucus formation in the lungs increases during COPD [7]. Tobacco smoke, biomass fuel smoke, pulmonary tuberculosis, frequent respiratory infections from childhood, indoor air pollution, outdoor air pollution, aging, alpha-antitrypsin deficiency, and long term asthma are the risk factors of developing COPD. People exposed to particulate matter in a poorly ventilated house are more liable to develop COPD [8, 9]. Outdoor air pollution comprises vehicular, emissions from chemical plants, poisonous gas emissions from chimneys. Aging could be another significant fact behind COPD [10].

There are clinical methods to classify COPD and healthy cases. Monitoring the symptoms, blood gas testing, pulmonary function testing and radiographic methods, including X-Ray, CT, etc. It is difficult to detect COPD if the patient has a long history of asthma. Patient's symptoms and spirometry outcomes are not enough parameter to diagnose COPD correctly [10]. Moreover, spirometry performance entirely depends on the patient's efforts. Spirometry results are affected if the patient is aged, weak or noncooperating. In addition, aging causes morphological changes in the respiratory organs which therefore affects the respiratory functions and hence the spirometry results [10, 11]. In such situations, it becomes difficult to diagnose COPD [12]. In such cases, the diagnosis is done using radiographic methods. Such techniques are costly as well as harmful for the patient's health due to irradiation [13, 14]. COPD mostly prevails in low and middle-income countries [3]. It is thus required to develop an economical, non-invasive and harmless method to diagnose COPD.

Normal lung sound is the sound produced by the lungs during respiration. The characteristics of lung sound are affected by respiratory diseases [15]. Adventitious sounds are the added sounds like wheezes; crackles, etc. which are common in people with COPD [16, 17]. Keen observation of variations in lung sound behavior can aid COPD diagnosis. There are few studies available on lung sound based COPD diagnosis. Malmberg et al. [18] assessed respiratory sound features of Asthma, COPD and normal lungs. A non-paired t-test was used to differentiate lung sound features available among the study groups. Quantitative lung data (QLD) were examined in the lower and upper lung. A lower median frequency of mean lung sound spectra was observed in COPD than in asthma. The study was insufficient to distinguish the spectral behavior of asthma and COPD. Morillo et al. [19] introduced the detection of acute exacerbation stage of COPD using computerized breath sound analysis. Frequency domain features were evaluated. Cluster analysis was done to investigate the association of variable groups. The principal component analysis was used to decrease the high degree of association between the features. Student t-test and Chi-square (χ^2) test were employed for bivariate comparisons. The significant association was obtained between forced expiratory volume in 1 s (FEV1) and χ^2 value. This study could aid the remote analysis of COPD exacerbations.

Mineshita et al. [20] proposed the relationship between respiratory sound distribution and pulmonary function parameters in COPD patients. Lung sound intensity was found higher in the lower lung compared to the upper lung in COPD patients. Mann-Whitney U-test was used to differentiate the lower, middle and upper QLD of COPD and the healthy subject. The correlation between pulmonary function parameters and the ratio of lower QLD to upper QLD was determined using Spearman's rank correlation coefficient. The obtained correlation was moderate with $r = 0.45$ and $p < 0.005$. The ratio of lower QLD to upper QLD was found to decrease with respect to disease severity. Bennette et al. [21] introduced a study to identify the relationship between the crackle features and HRCT variables in people with COPD. Lung sound feature extraction was done. One-way analysis of variance (ANOVA) and post hoc comparison tests was used along with Bonferroni correction to contrast the mean of the three subject groups mean. The linear relationship between independent and dependent variables was tested using the Pearson correlation coefficient. A positive correlation was obtained between conductive airway measurements and crackle characteristics.

Poreva et al. [22] presented diagnostic signs of COPD based on breath sound. An algorithm was developed to separate the inspiration and expiration phases of the signal. The values of bicoherence coefficient and skewness coefficient values were evaluated. Bicoherence was lying in between 20 and 50 with bifrequency pair $f1 \neq f2$. The frequency value at which maximum bicoherence was achieved is the bifrequeny pair. Skewness was found less than 0.15 in healthy subjects, whereas greater than 0.15 in COPD. Ishimatsu et al. [23] demonstrated a study on lung sound intensity during tidal breathing of COPD subjects. Inspiratory and expiratory phases were identified using computer programs. Noisy sections were removed by observing the spectrogram of the breath sound signal. Power is computed in low, middle and high frequency bands with the help of fast frequency transform. Octave power band values were averaged during deep inspiration, deep expiration, resting inspiration and resting expiration for each subject. These values were contrasted using unpaired t-test between the control and COPD group. Similarly, unpaired t-test was used to compare airflow between the two groups. During resting breathing, high lung sound intensity was observed during both inspiration and expiration phases. Whereas during deep inspiration, the breath sound intensity was found to be decreased in COPD compared to controls.

Recently, a study on the diagnosis of COPD based on the computerized lung sound analysis conducted by Jacome and Marques et al. [24]. Reliability and variability of lung sound were assessed. Intraclass correlation coefficient (ICC) method to determine relative reliability and Bland-Altman method were used to compute the absolute reliability. The coefficient of variation was used to compute intersubject variability. Breath sound could be a significant biomarker at an airflow of 0.4–0.6 L/s was concluded in this study. Recently, Jacome et al. [25] presented the breath sound based discrimination of acute exacerbation COPD (AECOPD) and stable COPD. Butterworth band-pass of order 8 was used to denoise the recorded lung sound signal. Re-sampling of the signal was done before computing short time Fourier transform. Box filtering was used for signal smoothening. Sample characteristics were determined using descriptive statistics. Statistical evaluation was done by independent t-test, Mann-Whitney U test, and chi-square test. The wheezes and crackles occurrence rate was found lesser in stable COPD than in AECOPD.

After reviewing the aforementioned studies it is found that lung sound based automated COPD diagnosis using machine learning techniques is a little explored area. Most of the existing studies attempt to find a correlation between COPD and lung sound features using statistical significance analysis methods, parametric and non-parametric tests. Techniques such as t-test, Mann Whitney U-test, correlation analysis, Spearman's rank correlation coefficient, ANOVA, etc. have been employed by researchers for the same. Thus, the development of a model for risk stratification of COPD using machine learning techniques is of immense significance in the field of medical systems.

There are certain shortcomings in the existing studies. Most of the studies are conducted on a limited data set. They have presented a moderate association of lung sounds with respect to pulmonary function parameters. No studies are available on the classification of COPD and healthy subjects using lung sound examination. Thus, there is a need to explore the classification of COPD subjects based on lung sound. The hypothesis states that machine learning techniques can assist clinicians in deriving unforeseen relationships within lung sound features thereby significantly improving the time complexity, observational human errors, and prediction accuracy of COPD. This article presents an automated, convenient, economic, harmless and non-invasive way of classification of COPD. This work includes the classification of COPD using respiratory sound analysis. The contributions of this work are:

(a) The lung sound data used in this study are original. Overall, 42 features are computed which includes 39 features extracted from lung sound and 3 spirometry based features for COPD diagnosis. Using such a large number of features and combined lung sound features and spirometry data for COPD classification is a novel aspect of this study.
(b) Statistical approaches, as well as machine learning protocols, are implemented and evaluated to determine the most significant features to diagnose COPD. The proposed method can aid an automated risk stratification system for COPD diagnosis.

The organization of the rest of the article is as follows: second section discusses the material and methods. It involved the data collection, lung sound signal denoising, feature calculation, and statistical feature evaluation and performance evaluation. Third section demonstrates the result and discussion followed by conclusions and future work in fourth section.

Material and methods

Figure 1 represents the strategy for the development of a risk stratification system for automated COPD diagnosis. It includes two phases: the online phase and the offline phase. The right section of Fig. 1 is the online phase, whereas the left side of it is the offline phase. The offline phase represents the indirect way of determination of COPD biomarker. The offline phase includes various steps like feature extraction, statistical significance analysis and development of machine learning model using ground truth data. The machine learning model thus developed is then used for automated diagnosis of COPD in online mode. In the online phase only relevant i. e. statistically significant biomarkers are extracted for COPD diagnosis to save computational time.

Study design and participants

The data included in the study is collected from the Department of Pulmonary Medicine & Tuberculosis, All India Institute of Medical Sciences, Raipur, Chhattisgarh. Approval of Institutional Ethical Committee was taken before conducting the study. Thirty stable COPD and 25 normal subjects aged between 20 and 60 years were enrolled in the study. Most of the participants were from Chhattisgarh state. Informed consent was obtained from all the participants before enrolling them into the study. Twenty-six male and four female COPD participants were recruited into this study. Fourteen male and 11 female healthy subjects participated in the data recording of the control group. Those COPD patients diagnosed strictly by the standard criteria of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) were included in this study. Only those participants were included in the study that performed spirometry successfully. Patients with a history of respiratory, neurological and musculoskeletal disorders were excluded from the study. Both smoker and non-smoker subjects were included. A spirometer is used to perform pulmonary function testing. Subjects with normal spirometry and no past history of respiratory disease were enrolled in the control group. Breath sounds were recorded

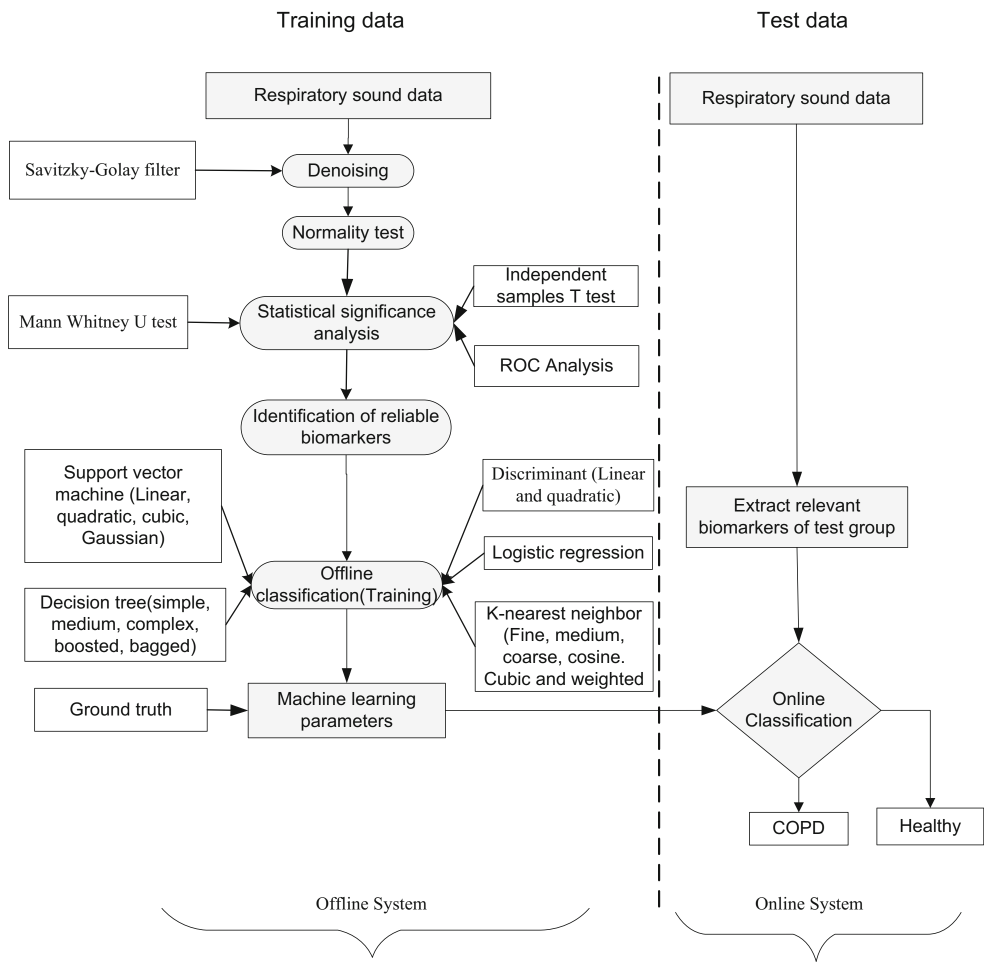

Fig. 1 The proposed risk stratification system

with the help of electronic stethoscope (3 M Littmann Electronic Stethoscope model 3200). Breath sounds were recorded at deep breathing for a period of 20 s at a sampling rate of 4000 Hz from the left posterior chest location.

Signal processing and feature extraction

Noise often degrades the quality of lung sound signal. Signal processing can be flexibly integrated to improve the quality of lung sound signal [26]. The recorded signal is first allowed to pass through a bandpass filter of order 6; with the passband frequency chosen as 20–2000 Hz. The breath sound signal is then filtered using a Savitzky-Golay filter wherein the main objective is to adjust a specific polynomial to a frame of the signal by using the least squares method recently applied in [27]. Compare to the other existing approaches, Savitzky-Golay filter achieves improved signal smoothening ability with retained characteristics of the original lung sound signal. The details of experiments conducted for lung sound preprocessing can be found in [27]. Signal processing is done in MATLAB® environment. Total of 39 lung sound features are computed namely, 13 Mel-frequency cepstral coefficients (MFCC1, MFCC2… MFCC13), Linear predictive coefficients (LPC1 and LPC2), first three formant frequencies, median frequency, spectral centroid, dominant frequency, standard deviation, mean, root mean square value (RMS), variance, minimum value, maximum value, peak value, dynamic range, crest factor, autocorrelation time, sample entropy, zero crossing rate, spectral roll-off, skewness, maximum frequency, Renyi entropy, kurtosis and Katz's fractal dimension (Katz fd). Three spirometry features namely forced expiratory volume (FEV_1), forced vital capacity (FVC) and a ratio of FEV_1 and FVC are also utilized.

Statistical significant analysis

Not all the extracted features are equally important to discriminate the sample between groups. The large pool of features only enhances computational complexity and processing time. Hence, the removal of redundant and unimportant features is done by statistical significance analysis to reduce the feature space. In this work independent t-test and Mann-Whitney U test is carried out to check the significance level of an individual feature. As the COPD, lung sound data are highly varied in

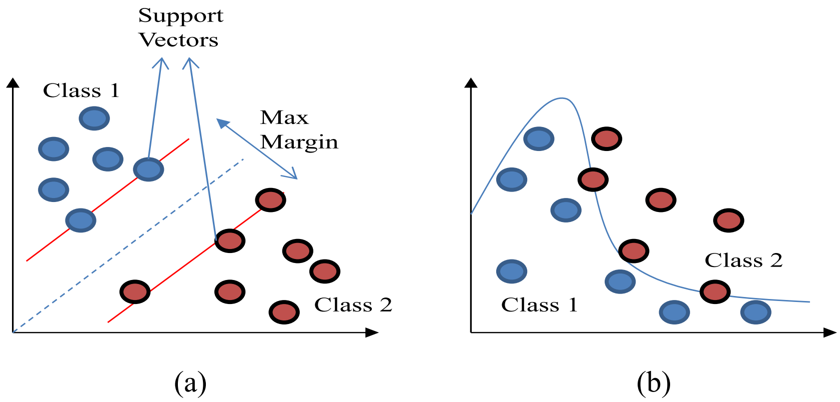

Fig. 2 **a** Linear Support Vector Machine Classifiers, **b** Non-Linear Support Vector Machine Classifiers

nature; it's not easy to predict the form of distribution. Therefore, the application of non-parametric test alongside t-test gives a superior view for reasonable characterization of information. All tests are done by using the SPSS (Statistical Package for Social Sciences) software and the significance level was chosen to be less than 0.05.

ROC analysis of individual features

Receiver operating characteristics (ROC) analysis is a widely used approach to evaluate machine learning models. The area under the ROC curve (AUC) is a well-known parameter which can be derived from ROC and used as a performance measure. In the present work, ROC analysis is used as an additional tool to evaluate the most relevant features which had the highest discrimination capabilities for the diseased and control groups. The ROC analysis for individual features is done using MedCalc software following the principles and method given by DeLong et al. [28] for the calculation of standard errors of AUC.

Classification

Classification is the last stage of the proposed risk stratification system. The job of a classifier is to map the input variables to the correct class label. In this paper supervised learning

Table 1 Significant features and its performance

Name of feature	Normality test (*p* value)	Independent sample t-test (*p* value)	Mann-Whitney U test (*p* value)	ROC analysis
FEV1/FVC	0.003	0.000	0.000	AUC =0.977
FEV1	0.005	0.000	0.000	AUC =0.977
FVC	0.013	0.000	0.000	AUC =0.895
Standard deviation	0.200	0.012	0.008	AUC = 0.711
Variance	0.007	0.020	0.008	AUC = 0.711
Zero crossing rate	0.200	0.003	0.003	AUC = 0.733
Spectral roll off	0.001	0.001	0.001	AUC = 0.765
Median frequency	0.005	0.000	0.000	AUC = 0.855
Spectral centroid	0.200	0.003	0.001	AUC = 0.752
Dominant Frequency	0.000	0.012	0.001	AUC = 0.799
LPC2	0.001	0.000	0.000	AUC = 0.837
Kurtosis	0.000	0.049	0.020	AUC = 0.684
Sample entropy	0.047	0.014	0.014	AUC = 0.693
Formant frequency 2	0.000	0.004	0.035	AUC = 0.667
Formant frequency 3	0.000	0.001	0.003	AUC = 0.636
MFCC2	0.200	0.047	0.035	AUC = 0.667
MFCC3	0.090	0.000	0.000	AUC = 0.805
MFCC5	0.168	0.035	0.002	AUC = 0.747
MFCC6	0.200	0.000	0.000	AUC = 0.816
MFCC7	0.200	0.001	0.000	AUC = 0.784
MFCC8	0.200	0.039	0.039	AUC = 0.663

Table 2 Classifier performance using spirometry data at 33% holdout

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Fine Gaussian SVM	100	100	100	1
Linear SVM	94.4	88.9	100	1
Quadratic SVM	94.4	88.9	100	1
Cubic SVM	94.4	88.9	100	1
Linear discriminant	94.4	88.9	100	0.98
Quadratic Discriminant	88.9	88.9	88.9	0.98
Logistic Regression	88.9	100	77.78	0.94

based classifiers namely support vector machines, linear discriminant classifier, quadratic discriminant classifier, logistic regression, decision trees and K-NN classifiers as depicted in Fig. 1 are implemented and evaluated.

Support vector machine

Support vector machines (SVMs) are the famous and widely utilized supervised classifiers developed by Vapnik et al. [29]. SVM is a binary classifier which separates the two distinct groups based on their element values. The separation is done in the form of a hyperplane between the classes based on support vector theory. Figure 2 displays an example of classification by SVM. Figure 2a illustrates an example of a linear SVM classifier. Here the two classes are being classified by a separating hyperplane given by the dashed blue line. There are several hyperplanes that can be constructed to separate both the classes, but the optimum one with maximum margin criterion between support vectors (SVs) of two classes is chosen [30, 31].

Practical classification problems are hardly distinguishable by any linear classifier as the data points are distributed in a random manner. In that case, a non-linear approach is employed as in Fig. 2b. This is done by SVM with kernel trick [32]. Here, the input feature space is transformed into some higher dimensional space where the data points are perfectly separable. The transformation is denoted by $\psi : M \rightarrow \psi(M)$. The kernel function performs inner product of variables in higher dimensional feature space as $k(m_i, m_j) = \Psi(m_i).\ \Psi(m)$. Examples of kernel functions investigated in this study are the linear kernel, polynomial kernel, Gaussian kernel (Radial Basis Function), etc. Further, the subtypes of polynomial kernel namely quadratic SVM (polynomial kernel with degree = 2), cubic SVM (polynomial kernel with degree = 3) and subtypes of Gaussian SVM kernels (coarse (standard deviation = 2), medium (standard deviation = 0.7) and fine (standard deviation = 0.25)) are also evaluated.

Discriminant classifiers

Discriminant classifiers are based on Fisher's discriminant theory which primary purpose is to have a projection of feature vectors of lower dimensional feature space to increase the separation between the classes. This is done by maximizing the ratio of the variance of between-class to the variance of within-class [33]. This concept is known as linear discriminant analysis (LDA). Between-class variance is given by computing distance between the mean of two classes and within-class variance is defined by calculating the distance between the mean and individual samples of that particular class.

The quadratic discriminant analysis (QDA) works in close principle as LDA but does not follow the assumption of the identical covariance matrix. So it is more suitable when the two classes possess different variance structures. Here, the distance calculation for each class is done through the sample variance-covariance matrix of each class instead of considering the overall matrix for all samples [34]. As the name, QDA draws a boundary in the form of a quadratic curve, which in turn can also classify the non-linearly distributed samples.

Logistic regression

Logistic regression (LR) is a statistic based method to predict an outcome or label with the help of some independent variables of the data set. LR is widely used for the binary level problem. Here, regression refers to the process of mapping the

Table 3 Classifier performance using spirometry data at 5-fold

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Logistic Regression	98.2	96.7	100	0.98
Fine Gaussian SVM	96.4	96.7	96	0.99
Cubic SVM	96.4	93.3	100	0.96
Quadratic SVM	94.5	90	100	0.98
Medium Gaussian SVM	94.5	90	100	0.98

Table 4 Classifier performance using the spirometry data at 10-fold

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Logistic Regression	96.4	93.3	100	1
Fine Gaussian SVM	96.4	96.7	96	0.99
Medium Gaussian SVM	96.4	93.3	100	0.99
Quadratic SVM	94.5	90	100	0.97
Cubic SVM	94.5	90	100	0.96
Linear discriminant	92.7	86.7	100	0.95
Quadratic Discriminant	92.7	90	96	0.98

input explanatory features which serve as independent variables to a dichotomous characteristic of interest which is nothing but dependent variables or outcomes consist of two values: 1 (yes) or 0 (no). The LR is modeled to compute the probability, what is its value corresponding to a particular outcome when subjected to input variables. It is not a classifier by nature, but by deciding an appropriate threshold value and comparing it with the probability computed for input variables could make it work as a binary classifier.

The name LR is because of the function used at the core of the method, the logistic function. It is in the form of the sigmoid function. According to Brownlee, et al. [35], "The logistic function was developed to describe properties of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment". The logistic function is a continuous function between 0 and 1, whose shape resembles the letter 'S'. Mathematically,

$$P_i = \frac{1}{1 + e^{\left(-\left(1, m_i^T\right)K\right)}} \qquad (1)$$

Where 'm_i' is the sample input features, K is the parameter vector and P_i is the probability of an outcome.

The expression for 'P_i' to attain the value 1 can be given as:

$$P_i = P(y = 1 | S_x, m_i) \qquad (2)$$

The values of parameters or coefficient vectors K is defined by maximum-likelihood estimation. The best values of K tend to minimize error in probabilities of predicting the correct class by model according to the input variables.

K-nearest neighbor

K-nearest neighbor (KNN) is a popular classification technique which classifies a sample by determining its distance from the closest sample (K refer to the number of neighbors). Some common distance measures used are Euclidean, Chebyshev, and Manhattan. In this study, different types of KNN are used depending on the value of K and distance metric used that is, fine KNN (K = 1, Euclidean), medium KNN (K = 10, Euclidean), coarse KNN (K = 100, Euclidean), cosine KNN (K = 10, cosine), cubic KNN (K = 10, Minkowski (cubic)) and weighted KNN (K = 10, Euclidean with squared inverse distance weight).

Decision trees

Decision trees are supervised techniques used in classification and regression problems. It includes branches and nodes. Two types of nodes specifically, root node and leaf node are present. The classification of an unknown sample is done based on the split criterion in the root node and finally reaching the leaf node through branches. In the present work, different types of decision trees are investigated namely simple, medium, complex, boosted, bagged.

Table 5 Classifier performance using lung sound parameters (median frequency and LPCC2) at 33% holdout

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Fine Gaussian SVM	72.2	77.8	66.7	0.73
Medium Gaussian SVM	72.2	77.8	66.7	0.79
Linear SVM	66.7	66.7	66.7	0.80
Quadratic SVM	72.2	88.9	55.6	0.84
Cubic SVM	66.7	88.9	44.4	0.78
Linear Discriminant	72.2	88.9	55.6	0.85
Quadratic Discriminant	72.2	88.9	55.6	0.84
Logistic Regression	66.7	66.7	66.7	0.78
Simple tree	77.8	77.8	77.8	0.78

Table 6 Classifier performance using significant lung sound parameters (median frequency and LPCC2) at 5-fold

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Fine Gaussian SVM	83.6	80	88	0.83
Quadratic SVM	80	83	76	0.85
Medium KNN	80	80	80	0.83
Cubic KNN	80	80	80	0.84
Bagged trees	78.2	76.67	80	0.79

Performance evaluation

Several experiments are conducted to obtain the best classification performance. Different feature combinations are applied to classifier input to evaluate features as well as classifiers. Firstly, the classification performance is evaluated by taking spirometry parameters alone. Then two most relevant lung sound parameters, i.e. median frequency and linear predictive coefficient are used at the input of classifier for COPD diagnosis. Further, classification is performed by taking the two most significant lung sound parameter, i.e. median frequency and linear predictive coefficient along with the spirometry parameters. The classifier performance is evaluated to understand the system's capability to classify COPD and normal lung sound correctly. The performance of the classifier is evaluated in terms of accuracy, sensitivity, and specificity. Various data division techniques such as holdout and k-fold cross validation are used to evaluate classifiers on separate training and test dataset. In the holdout approach, 33% of the samples are used for testing the classifiers while remaining samples are used for training. In the cross-validation technique, 5-fold and 10-fold schemes are used. In each case, one fold is used for testing the classifier model while the remaining folds are used for its training. Further, the area under the ROC curve (AUC) is also used for evaluation of classifiers. The ROC curve demonstrates the true positive rate versus false positive rate for the presently chosen trained classifier. It utilizes various thresholds which are used to calculate every point in the ROC curve. The individual point of the ROC curve (i.e. threshold) represents the particular values of sensitivity and specificity. The AUC is an abstract parameter of concern that signifies if on average a true positive is graded higher the above the false positives. The optimal performance at a specific threshold is chosen. To calculate sensitivity and specificity we have used confusion matrix which gives true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). On the other hand, AUC reported in the paper is estimated for optimal performance at a specific threshold using classification learner application of MATLAB software.

Results and discussions

Development of an automated approach for diagnosing COPD using lung sound features is important predominantly in settings with low resources. The existing studies suffer from various shortcomings due to restrictions imposed by the small size of data used. Most of the studies concentrated on evaluating the correlation between lung sound parameters and COPD using statistical significance analysis. Application of machine learning strategies for COPD diagnosis for risk stratification is little explored area. No benchmark data is available to compare different machine learning approaches for COPD diagnosis. Thus, evaluating different machine learning models for risk stratification of COPD using lung sound features is important for clinical acceptability of such systems.

Total of 42 features are extracted, of which 39 are lung sound parameters and 3 are spirometry parameters. The temporal, spectral and spectro-temporal lung sound features were computed. The temporal features include mean, standard deviation, variance, skewness, kurtosis, number of zero crossings, entropy, linear predictive coefficients (LPCs), etc. Spectral features include spectral signatures like dominant frequency, spectral roll-off, median frequency, maximum frequency, and spectral centroid. Mel frequency cepstral coefficients (MFCCs) are the spectro-temporal features calculated in this study. The statistical analysis can be flexibly included to see the statistical significance of the feature. Table 1 represents the features which are found statistically significant along with their AUC performance.

Table 7 Classifier performance using lung sound parameter (median frequency and LPCC2) at 10-fold

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Medium KNN	81.8	80	84	0.84
Linear discriminant	78.2	86.7	68	0.84
Fine Gaussian SVM	78.2	76.7	80	0.81
Quadratic SVM	76.4	83.3	68	0.85

Table 8 Classifier performance using significant lung sound parameters (median frequency and LPCC2) and spirometry data at 33% holdout

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Coarse Gaussian SVM	94.4	88.9	100	1
Fine Gaussian SVM	94.4	88.9	100	1
Medium Gaussian SVM	94.4	88.9	100	1
Linear SVM	100	100	100	1
Quadratic SVM	100	100	100	1
Cubic SVM	100	100	100	1
Linear discriminant	94.4	88.9	100	1
Quadratic Discriminant	94.4	88.9	100	1
Logistic Regression	100	100	100	1
Cubic KNN	88.9	77.78	100	1

Referring the second column of Table 1, it is observed that some of the features enlisted in the first column are normally distributed ($P > 0.05$) while others are not following a normal distribution ($P < 0.05$). For normally distributed features independent samples t-test is conducted while for other Mann-Whitney U test is conducted at 95% confidence interval. Out of 42 features extracted, only 21 features are found statistically significant i. e. with p value <0.05. The efficacy of these features in diagnosing COPD cases are further evaluated using ROC analysis. Among the lung sound parameters extracted in this study, it is found that the highest AUC value of 0.855 is achieved by median frequency followed by a linear predictive coefficient (LPC2) which achieves AUC of 0.837. Hence, it is concluded that median frequency and linear predictive coefficient (LPC2) of lung sound features are the two best performing parameters to diagnose the COPD cases in terms of AUC. Spirometry features namely, FEV1/FVC, FEV1, and FVC attain highest AUC which is expected as these values are current GOLD standards used for COPD diagnosis. The following sections present the results of classifier techniques using spirometry data, two most relevant lung sound features, and combined lung sound features and spirometry features.

It is important to mention here that among the various classifiers evaluated i. e. support vector machines (fine Gaussian SVM, medium Gaussian SVM, course Gaussian SVM, linear SVM, cubic SVM, quadratic SVM), linear discriminant classifier, quadratic discriminant classifier, and logistic regression, the best three performing classifiers in terms of classification accuracy are shown in Tables 2, 3, 4, 5, 6, 7, 8, 9 and 10.

Results of classification using only spirometry data

In the present study, classification of normal and COPD subjects are conducted using spirometry parameters, lung sound parameters and including both spirometry and lung sound parameters as well. Tables 2, 3 and 4 presents the classifier performances at 33% holdout, 5-fold and 10-fold respectively in distinguishing the two categories using spirometry data alone. The fine Gaussian SVM classifier (hold out) outperforms other techniques achieving a classification accuracy of 100% followed by logistic regression (5-fold) with an accuracy of 98.2%. Though spirometry parameters achieve good performance in diagnosing COPD cases, its major drawback is that its outcome mainly depends on the technician skill, patient's efforts and cooperation. Hence, spirometry results alone are not enough parameter to diagnose COPD. Thus, this study investigates appropriate lung sound based parameters which can be used in conjunction with spirometry data for COPD diagnosis.

Results of classification using most significant lung sound parameters (median frequency and linear predictive coefficients)

Median frequency and LPC2 are found to be the most relevant feature with their significantly better AUC performance as observed in Table 1. Tables 5, 6 and 7 presents the classifier outcome at 33% holdout, 5-fold and 10-fold respectively, in classifying COPD and healthy cases using these two lung sound parameters. It is found that the classifier performance

Table 9 Classifier performance using significant lung sound parameters (median frequency and LPCC2) and spirometry data at 5-fold

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Linear SVM	98.2	96.7	100	1
Quadratic SVM	98.2	96.7	100	1
Cubic SVM	98.2	96.7	100	1
Logistic Regression	98.4	100	100	1
Fine KNN	96.4	93.3	100	1

Table 10 Classifier performance using significant lung sound parameters (median frequency and LPCC2) and spirometry data at 10-fold

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Linear SVM	98.2	96.7	100	0.97
Quadratic SVM	98.2	96.7	100	0.98
Cubic SVM	98.2	96.7	100	0.99
Linear Discriminant	96.4	90.3	100	0.97
Quadratic Discriminant	96.4	90.3	100	0.97
Logistic Regression	100	100	100	1

is moderately good when the experimentation is done using only two statistically relevant and best-performing lung sound parameters. The maximum classification accuracy achieved is only 83.6% using the fine Gaussian SVM classifier (5-fold). The classifier performance using only the best lung sound feature was lesser than previous experiments on taking only the spirometry data. It is thus interesting to investigate the classifier performance by combining both spirometry data and lung sound features.

Results of classification by including significant lung sound parameters, i.e. median frequency and linear predictive coefficient along with spirometry data

The classification is conducted, including median frequency, linear predictive coefficients along with the spirometry parameters. Tables 8, 9 and 10 represent the classifier outcome at 33% holdout, 5-fold and 10-fold respectively. It is found that the SVM classifier (linear, quadratic and cubic, 33% holdout) and logistic regression (33% holdout and 10 fold) outperformed other classifiers with 100% classification accuracy. Therefore, it can be concluded that the classifier performance is significantly improved by combining reliable lung sound parameters with spirometry parameters. Figure 3 shows the ROC of two best performing classifiers achieving a classification accuracy of 100% and AUC of 1.

Comparison of the proposed approach with existing studies

To the best of the knowledge of authors, there are very few studies in the area of automated classification/risk stratification of COPD using lung sound parameters. Existing studies are limited to statistical analysis and determining the correlation between lung sound parameters and COPD. The data size used in the existing studies is also limited. The study proposed by Mineshita et al. [20] has correlated its outcomes with spirometry parameters with an only moderate correlation coefficient of r = 0.45 was obtained. Another study presented by Bennette et al. [21] has compared the lung sound feature statistical outcomes with respect to that of High-resolution computed tomography (HRCT) variables. The study is insufficient to differentiate the two different categories of lung sound. Table 11 present the comparison between the proposed study with the existing respiratory sound based approaches to diagnose COPD.

To diagnose COPD, all existing literature has evaluated the lung sound based parameter and then they have correlated their studies with respect to some standard parameters like spirometry results or imaging outcomes. The present study is a unique study as it involved both statistical feature evaluation and feature classification. This study is able to categorize the two different classes of respiratory sound, i.e. normal and COPD, with the highest classification of 100% accuracy, 100% sensitivity, and 100% specificity. The study

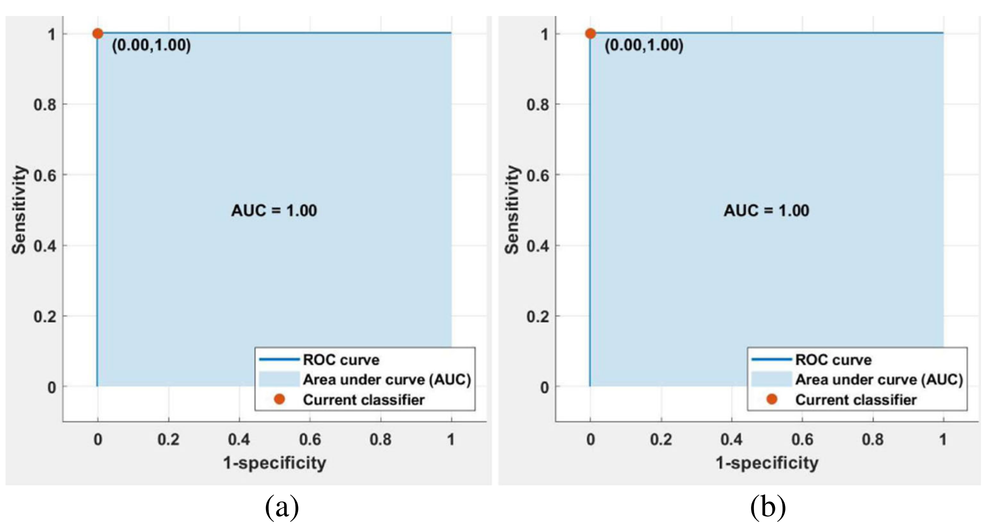

Fig. 3 Receiver operating characteristics of **a** Logistic regression classifier, **b** SVM classifier

Table 11 Comparison of the present work with the existing studies

S.N	Author and year	Sample size	Modality	Feature evaluation technique	Classification	Performance / Remark
1	Malmberg et al. (1995) [18]	17 COPD, 10 Asthma, 11 healthy	Mean, total spectral power, maximum frequency in upper-frequency limits for the second (F50) and the third (F75) quartile of power spectra at the inspiratory phase of breath sound.	Non-paired t-test, correlative analysis, Pearson's correlation coefficient	Nil	No correlation obtained in COPD with respect to FEV1.
2	Mineshita et al. (2014) [20]	47 male COPD	The ratio of lower QLD to upper QLD	Mann-Whitney U test, ROC analysis, Spearman's rank correlation Coefficient	Nil	$\rho = 0.45$, $p < 0.005$
3	Bennett et al. (2015) [21]	8 COPD, 9 control, 9 smokers with airway obstruction	Crackles count per breathing cycle and crackle 2-cycle duration	One-way analysis of variance, post hoc comparison tests	Nil	Correlated with respect to HRCT variables with $r = 0.48$, $p = 0.01$
4	Poreva et al. (2016) [22]	55 COPD, 37 controls	Skewness coefficient (c3) and kurtosis coefficients (c4)	Cross-correlation function (CCF)	Nil	In 84% COPD cases $c3 > 0.2$, c3 of the CCF is positive for COPD and negative for healthy cases.
5	Ishimatsu et al. (2014) [23]	20 COPD, 20 healthy	Octave power in the low, middle and higher frequency band, Breath sound intensity	unpaired t-test, multiple regression analysis	Nil	Diminished breath sound intensity during deep inspiration, increased breath sound intensity during resting breathing
6	Present work	30 COPD, 25 healthy	39 lung sound features in time, frequency and time-frequency domain	Normality test, independent t-test, Mann-Whitney U test, and ROC analysis	SVM, Logistic regression, Decision tree, discriminant analysis, K-NN	SVM-100%, Logistic regression - 100%

successfully classifies the COPD respiratory sound and normal respiratory sound.

From Table 11, it is also observed that very few studies have incorporated machine learning techniques for analyzing lung sound features of COPD subjects. Traditional approaches like t-test, Mann Whitney U-test, correlation analysis, Spearman's rank correlation coefficient, ANOVA, etc. are mostly used by researchers to determine the association between lung sound features and COPD. On the contrary, this study extracts multiple lung sound time and frequency features and evaluate different machine learning techniques for automated classification of lung sounds using these features. Further, statistical significance analysis methods are used to determine the relevance of extracted lung sound features. Other than spirometry features, median frequency and linear predictive coefficients were found to be the most relevant lung sound attributes which can be used for COPD detection. Median frequency is among the popular features used for describing breath sounds. It is the frequency at the core of the power spectrum. The other relevant feature LPCC is also a widely used metric in sound analysis applications. These coefficients are obtained by representing every sample by a linear combination of some of the previous samples.

A note on strengths, weakness, and scope of future work

This is a novel study, which involved the classification of COPD and normal respiratory sound. In existing studies, none of the studies have classified COPD respiratory sounds based on a combination of features extracted in this study. Combining spirometry data with lung sound features for improving COPD diagnosis is also a novel aspect of this study. However, this study has certain limitations. The sample size is small and data collection is single centered. However, as observed in Table 11, the size of the dataset used in this study is higher than many of the existing studies. Still, larger datasets are required in the future, to implement and compare the performance of the other advanced machine learning techniques like a deep neural network. Further, benchmark open datasets of COPD and non-COPD lung sounds are required to compare different machine learning techniques on a common platform. Differences in lung sound parameters among smokers and nonsmokers for COPD and non-COPD cases can be conducted in the future. Future work is also concentrated on multicentric and multi-modal studies for COPD diagnosis along with clinical validation of technologies developed.

Conclusion and future scope

Computerized lung sound analysis is a non-invasive technique which is attaining the wide popularity in terms of respiratory disease diagnosis and analysis. A present study is a useful approach for classifying the COPD and the healthy lung sound. It is one of the non-invasive, convenient, irradiation free, safe and economic methods. Normality test, independent t-test, Mann-Whitney test, ROC analysis is the statistical tests conducted. The median frequency and linear predictive coefficients are found to be the two best significant biomarkers for COPD. Lung sound parameter and spirometry data are accompanied to conduct the classification experiments. Superior classification results are obtained when both spirometry and lung sound parameters are accompanied. The SVM and LR are the two classifiers which successfully classifies COPD and healthy lung sound data with the accuracy of 100%. Future work will be to include multi-centered data as well as to recruit COPD subjects with comorbidities.

Compliance with ethical standards

Conflict of interest Author Nishi Shahnaj Haider, declares that she has no conflict of interest. Author Bikesh Kumar Singh declares that he has no conflict of interest. Author R. Periyasamy, declares that he has no conflict of interest. Author Ajoy K. Behera, declares that he has no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Laniado-Laborin, R., Smoking and chronic obstructive pulmonary disease (COPD). Parallel epidemics of the 21 century. *Int. J. Environ. Res. Public Health* 6(1):209–224, 2009.
2. The top 10 causes of death, World Health Organization. http://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. Accessed on 4 Oct. 2018.
3. May, S. M., and Li, J. T., Burden of chronic obstructive pulmonary disease: Healthcare costs and beyond. *Allergy Asthma Proc.* 36(1): 4–10, 2015.
4. Viniol, C., and Vogelmeier, C. F., Exacerbations of COPD. *Eur. Respir. Rev.* 27:170103, 2018.
5. Hui, S., How, C. H., and Tee, A., Does this patient really have chronic obstructive pulmonary disease? *Singap. Med. J.* 56(4): 194–196, 2015.
6. Shapiro, S. D., and Ingenito, E. P., The pathogenesis of chronic obstructive pulmonary disease. *Am. J. Respir. Cell Mol. Biol.* 32(5):367–372, 2005.
7. Williamson, J. P. et al., Elastic properties of the central airways in obstructive lung diseases measured using anatomical optical coherence tomography. *Am. J. Respir. Crit. Care Med.* 183(5):612–619, 2011.
8. Lee, S. J., Kim, S. W., Kong, K. A., Ryu, Y. J., Lee, J. H., and Chang, J. H., Risk factors for chronic obstructive pulmonary disease among never-smokers in Korea. *Int. J. Chron. Obstruct.*

Pulmon. Dis. 10:497–506, 2015. https://doi.org/10.2147/COPD.S77662.
9. Koul, P. A., Chronic obstructive pulmonary disease: Indian guidelines and the road ahead. *Lung India* 30(3):175–177, 2013.
10. Andreeva, E., Pokhaznikova, M., Lebedev, A., Moiseeva, I., Kuznetsova, O., and Degryse, J. M., Spirometry is not enough to diagnose COPD in epidemiological studies: A follow-up study. *NPJ Prim. Care Respir. Med.* 27:62, 2017.
11. Lowery, E. M., Brubaker, A. L., Kuhlmann, E., and Kovacs, E. J., The aging lung. *Clin. Interv. Aging* 8:1489–1496, 2013.
12. Johns, D. P., Walters, J. A., and Walters, E. H., Diagnosis and early detection of COPD using spirometry. *J. Thorac. Dis.* 6(11):1557–1569, 2014.
13. Washko, G. R., Diagnostic imaging in COPD. *Semin. Respir. Crit. Care Med.* 31(3):276–285, 2010.
14. Guarascio, A. J., Ray, S. M., Finch, C. K., and Self, T. H., The clinical and economic burden of chronic obstructive pulmonary disease in the USA. *Clinicoecon Outcomes Res.* 5:235–245, 2013. https://doi.org/10.2147/CEOR.S34321.
15. Reichert, S., Gass, R., Brandt, C., and Andres, E., Analysis of respiratory sounds: State of the art. *Clin. Med. Circ. Respirat. Pulm. Med.* 2:45–58, 2008.
16. Melbye, H., Garcia-Marcos, L., Brand, P., Everard, M., Priftis, K., and Pasterkamp, H., Wheezes, crackles and rhonchi: Simplifying description of lung sounds increases the agreement on their classification: A study of 12 physicians' classification of lung sounds from video recordings. *BMJ Open Respir. Res.* 3(1):e000136, 2016. https://doi.org/10.1136/bmjresp-2016-000136.
17. Sarkar, M., Madabhavi, I., Niranjan, N., and Dogra, M., Auscultation of the respiratory system. *Ann Thorac. Med.* 10(3):158–168, 2015.
18. Malmberg, L. P., Pesu, L., and Sovijarvi, A. R. A., Significant differences in flow standardised breath sound spectra in patients with chronic obstructive pulmonary disease, stable asthma, and healthy lungs. *Thorax* 50:1285–1291, 1995.
19. Sanchez Morillo, D., Astorga Moreno, S., Fernandez Granero, M. A., and Leon Jimenez, A., Computerized analysis of respiratory sounds during COPD exacerbations. *Comput. Biol. Med.* 43:914–921, 2013.
20. Mineshita, M., Kida, H., Handa, H., Nishine, H., Furuya, N., Nobuyama, S., Inoue, T., Matsuoka, S., and Miyazawa, T., The correlation between lung sound distribution and pulmonary function in COPD patients. *PLoS ONE* 9(9):e107506, 2014.
21. Bennett, S., Bruton, A., Barney, A., Havelock, T., and Bennett, M., The relationship between crackle characteristics and airway morphology in COPD. *Respir. Care* 60(3):412–421, 2015.
22. Poreva, A. S., Karplyuk, Y. S., Makarenkova, A. A., and Makarenkov, A. P., Detection of COPD's auscultative symptoms using higher order statistics in the analysis of respiratory sounds. *Radioelectron. Commun. Syst.* 59(2):83–88, 2016.
23. Ishimatsu, A., Nakano, H., Nogami, H., Yoshida, M., Iwanaga, T., and Hoshino, T., Breath sound intensity during tidal breathing in COPD patients. *Intern. Med.* 54:1183–1191, 2015. https://doi.org/10.2169/internalmedicine.54.3161.
24. Jacome, C., and Marques, A., Computerized respiratory sounds are a reliable marker in subjects with COPD. *Respir. Care* 60(9):1264–1275, 2015.
25. Jacome, C., Oliveira, A., and Marques, A., Computerized respiratory sounds: A comparison between patients with stable and exacerbated COPD. *Clin. Respir. J.* 11(5):612–620, 2017. https://doi.org/10.1111/crj.12392.
26. Mondal, A., Bhattacharya, P., and Saha, G., Detection of lungs status using morphological complexities of respiratory sounds. *Sci. World J.* 182938:1–9, 2014.
27. Haider, N. S., Periyasamy, R., Joshi, D., and Singh, B. K., Savitzky-Golay filter for denoising lung sound. *Braz. Arch. Biol. Technol.* 61: e18180203, 2018.
28. DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L., Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44(3):837–845, 1988.
29. Cortes, C., and Vapnik, V., Support-vector networks. *Mach. Learn.* 20(3):273–297, 1995. https://doi.org/10.1007/BF00994018.
30. Duda, R. O., Hart, P. E., and Stork, D. G., *Pattern classification*. Wiley, New York, 2000. https://doi.org/10.1038/npp.2011.9.
31. Subashini, T. S., Ramalingam, V., and Palanivel, S., Breast mass classification based on cytological patterns using RBFNN and SVM. *Expert Syst. Appl.* 36(3):5284–5290, 2009. https://doi.org/10.1016/j.eswa.2008.06.127.
32. Boser, B. E., Guyon, I. M., and Vapnik, V. N., A training algorithm for optimal margin classifiers. *COLT '92. Proceedings of the Fifth Annual Workshop on Computational Learning Theory.* 144–152, 1992. https://doi.org/10.1145/130385.130401
33. Tharwat, A., Gaber, T., Ibrahim, A., and Hassanien, A. E., Linear discriminant analysis: A detailed tutorial. *AI Commun.* 30(2):169–190, 2017. https://doi.org/10.3233/AIC-170729.
34. Dixon, S. J., and Brereton, R. G., Comparison of performance of five common classifiers represented as boundary methods: Euclidean distance to centroids, linear discriminant analysis, quadratic discriminant analysis, learning vector quantization and support vector machines, as dependent on data structure. *Chemom. Intell. Lab. Syst.* 95(1):1–17, 2009.
35. Brownlee, J., Logistic regression for machine learning. 2016. https://machinelearningmastery.com/logistic-regression-for-machine-learning/. Accessed on 28 Feb 2019.