Australasian Physical & Engineering Sciences in Medicine (2019) 42:1011–1024
https://doi.org/10.1007/s13246-019-00800-2

SCIENTIFIC PAPER

A novel approach towards non-obstructive detection and classification of COPD using ECG derived respiration

Surita Sarkar[1] · Parthasarathi Bhattacharyya[2] · Madhuchhanda Mitra[1] · Saurabh Pal[1]

Received: 12 February 2019 / Revised: 29 August 2019 / Accepted: 7 September 2019 / Published online: 10 October 2019

Abstract
The alarming rate of mortality and disability due to Chronic Obstructive Pulmonary Disease (COPD) has become a serious health concern worldwide. The progressive nature of this disease makes it inevitable to detect this disease in its early stages, leads to a greater demand for developing non-obstructive and reliable technology for COPD detection. The use of highly patient-effort dependent, time-consuming, and expensive methods are some major inherent limitations of previous techniques. Lack of knowledge about the disease and inadequacy of proper diagnostic tool for early detection of COPD is another reason behind the 3rd leading cause of death worldwide. For this reason, this study aims to explore the utility of ECG Derived Respiration (EDR) for classification between COPD patients and normal healthy subjects as EDR can be easily extracted from ECG. ECG and respiration signals collected from 30 normal and 30 COPD subjects were analysed. Error calculation and statistical analysis were performed to observe the similarity between original respiration and EDR signal. The morphological pattern changes of respiration and EDR signals were analysed and three different features were extracted from those. Classification was performed by different classifiers employing Decision Tree, Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). Apart from obtaining comparable classification performance it was seen that EDR has better potential than the original respiration signal for classification of COPD from normal population.

Keywords Chronic Obstructive Pulmonary Disease · Electrocardiogram · ECG derived respiration · Respiration · Classification

Introduction

Chronic Obstructive Pulmonary Disease, generally known as COPD, becomes a global health epidemic that causes millions of deaths every year. As reported by the World Health Organization (WHO), COPD is the 3rd leading cause of mortality and one of the major causes of morbidity worldwide [1]. Because of its progressive nature and lack of knowledge about the disease, COPD often ends up with early disability and death among patients, causing a huge socio-economic burden [2]. For this reason, the necessity of early detection of COPD is immense as it becomes rampant if remains undiagnosed and undertreated [3]. Only early intervention and possible prevention of COPD progression can improve the quality of life of the patient and thus reduce untimely death [4].

Till date spirometry is the only global standard method for COPD diagnosis [5]. But the scope of using spirometry is often restricted due to its limited availability outside the urban areas especially in developing countries [6]. However, it is observed that spirometry often performs poorly due to its high dependency on patient effort during the test [7] and poor guidance by unskilled technician [6]. The use of spirometry is also restricted for non-ambulatory and ICU patients. Hence, there is a need for developing non-obstructive and reliable methods for better diagnostic purpose.

Some other methods, developed for the detection of COPD, include lung imaging techniques like X-ray, CT scan etc. [8, 9], analysis of different volatile organic compounds (VOCs) presented in the expired air of COPD patients [10], measurement of flow obstruction of exhaled breath in COPD

✉ Saurabh Pal
spal76@gmail.com

1 Department of Applied Physics, University of Calcutta, Kolkata 700009, India

2 Institute of Pulmocare & Research, DG-8, Action Area 1, Newtown, Kolkata, India

patients using forced oscillation method [11], estimation of airflow limitation in COPD using impulse oscillometry system [12], quantitative analysis of capnogram shape [13], electronic nose [14]. Risk of radiation exposure, higher cost of test, greater time consumption and commercial unavailability are some of the inherent limitations of earlier reported techniques.

ECG is a well-known diagnostic tool practiced for long time and previous studies showed that COPD produces significant changes on cardiac function which are electrophysiologically expressed in ECG [15, 16]. Also, the breathing pattern changes in patients of chronic obstructive pulmonary diseases [17]. Previous researchers showed that respiratory information can be obtained from ECG in terms of EDR signal [18, 19]. There are some studies where heart rate variability method was used to investigate the activity and condition of COPD patients [20–22]. As per our knowledge there is no reported method on classification of COPD population from normal healthy subjects using peak amplitude variation method for EDR derivation. The novelty of this proposed study stands by this reason.

The main objective of this study is to develop an EDR based algorithm for detection of COPD using only single-lead ECG. The difference in underlying mechanism causes the respiratory pattern of COPD to vary from that of normal subject. This morphological pattern change in respiration also affects the ECG signal which leads to change in EDR in COPD patients. Based on these physiological changes, features were derived from both respiration and EDR signal.

Section "Materials and methods" of this paper describes the data collection and methodology of the study. Derivation of EDR, feature extraction and classification methods are discussed within the methodology section. In section "Result", value of extracted features and performance results are described. A comparative study with other literatures is also provided in this section. Finally, section "Conclusion" followed by the discussion of this working section "Discussion". The block diagram of the proposed study is shown in Fig. 1.

Hence, the main objective of this proposed work is to classify COPD patients from normal healthy population using EDR signal using different classifiers and their classification performance has been compared along with the results obtained from original respiration signal.

Materials and methods

Study population

In this study, a total number of 60 subjects were recruited for data collection purpose. The study protocol followed the Declaration of Helsinki and was approved by the Institutional Ethics Committee of Institute of Pulmocare and Research, Kolkata, India. Informed consent was obtained from all individual participants included in the study. Prior to the data collection, the subjects underwent a clinical assessment including previous history and physical examination done by expert physicians and their demographic details were recorded (shown in Table 2). Evaluation of the pulmonary function was done by spirometry as per American Thoracic Society (ATS) guideline [23]. 30 COPD patients, diagnosed with clinico-cardiological and spirometry test, constituted 'COPD group', whereas, thirty healthy subjects with normal spirometry and no clinical clue of any lung disease or significant systemic illness were included as 'Normal group'. Only subjects, within the age range of 18–75 years, were included in the study and those having arrhythmia, any severe cardiovascular or pulmonary diseases, sepsis, pregnancy, artificial pacemaker or any other chronic disease were excluded from the study.

Data acquisition

Both electrocardiogram and respiration signals were acquired simultaneously in a normal, non-soundproof room using Biopac MP-45 data acquisition device by Biopac Systems Inc.[24]. At the time of signal collection, the subjects were resting in supine condition. Recording in other conditions like standing, sitting, and while exercising was avoided in this case. To minimize the presence of noise in signal subjects were advised to take regular breaths and not to talk and/or move. For ECG measurement, three disposable Ag/

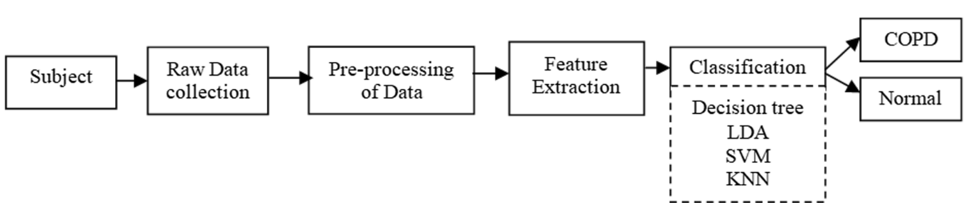

Fig. 1 Block diagram of proposed method for COPD-Normal classification. Note: If the collected data was ECG, EDR signal was extracted from pre-processed ECG data as discussed in section "Signal pre-processing", prior to feature extraction. However, for the collected respiration data, the path shown in the block diagram was followed.

AgCl electrodes were placed on the right anterior forearm and medial surface of each leg constituting Lead II configuration. A respiratory effort transducer was tied across the chest of each subject for respiration measurement. The recording was done for 300 s time duration with 1 KHz sampling frequency.

Signal pre-processing

The signals were pre-processed before further utilization. In raw ECG signal, 50 Hz powerline interference and baseline wander were predominant. For removal of noise, second order Bandpass Butterworth filter was designed separately for denoising of ECG and respiration signal. Filtered signals were normalized between zero to unity. The frequency range of the recorded signals with their pass bands at the time of filtering are given below in Table 1.

Derivation of ECG Derived Respiration

Previous researchers found that there is a direct effect of respiration on cardiovascular activity [19, 25]. The inflation and deflation of lungs during breath in and breath out result in apparent cardiac axis variation which leads to the modulation of the QRS amplitude [18, 26]. This modulation in R peak amplitude can be used to produce the surrogate respiration signal. In order to derive the EDR signal using R peak amplitude variation method, R peaks were detected using the following steps.

Step 1 Sliding window with a window length of 2 s. was applied on the normalized ECG signal for detecting all peaks. To determine the window length, intervals between two consecutive cycles were measured and the maximum interval was identified. After that a duration greater than the maximum interval time was selected as window in order to detect a minimum of two successive peaks.

Step 2 The maximum amplitude was found out within the specific range and a threshold of 60% of the maximum amplitude was applied to exclude points other than possible R-peaks.

Step 3 Potential R peaks were identified from another 200 ms window (100 ms each side) surrounding the possible peaks. The minimum distance between possible R-peak to the adjacent highest peak on each side was calculated and the window length was taken less than that.

Step 4 Finally the peak with highest co-ordinate value within the duration was plotted as the actual R peak as shown in Fig. 2d.

Table 1 Frequency range of signals and their pass bands

Signal	Frequency range of recorded signal (Hz)	Pass band (Hz)
ECG	0.05–100	1–47
Respiration	0.1–0.7	0.2–0.7

During respiration signal extraction the precision of respiration frequency and the discrimination between inspiration and expiration should be observed from reconstructing EDR. Considering the frequency range of respiration (0.2–0.7 Hz in this study), a surrogate signal was reconstructed using cubic-spline interpolation with 1 kHz sampling rate. Assume the original respiration be $u_j = f\left(v_j\right)$, where $j = 0, 1, 2, \ldots, n-1$. Then within the span of $[v_j, v_{j+1}]$, the cubic-spline interpolation can be written as shown in Eq. (1),

$$u = Cu_j + Du_j + 1 + Eu_j'' + Fu_j'' + 1 \tag{1}$$

where C, D, E, F are coefficients and v should be $v_0 \le v \le v_{n-1}$.

The small phase difference occurred between the original respiration and EDR was neglected in further calculations. In our previous study we observed that the calculated ECG derived respiration rate was almost equal for all subjects with original respiration rate acquired from thoracic belt [27]. To validate the EDR signal with the original respiration signal, respiration rate was calculated from both the signals and error calculations were done. Respiration rate (ReR) was calculated using the formula as shown in Eq. 2.

$$\mathrm{Re}R = \frac{60}{Time\ difference\ between\ two\ con\sec utive\ peaks} \tag{2}$$

Using the average of the respiration rates calculated from original and derived respiration signals, percentage error (PE), mean absolute error (MAE) and root mean square error (RMSE) were calculated as shown in Eq. 3 a, b, c.

$$PE = \frac{1}{N}\sum_{n=1}^{N} \frac{\left|\mathrm{Re}R_D(n) - \mathrm{Re}R_R(n)\right|}{\mathrm{Re}R_R(n)} \times 100\% \tag{3a}$$

$$MAE = \frac{1}{N}\sum_{n=1}^{N} \left|\mathrm{Re}R_D(n) - \mathrm{Re}R_R(n)\right| \tag{3b}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N} \left[\mathrm{Re}R_D(n) - \mathrm{Re}R_R(n)\right]^2} \tag{3c}$$

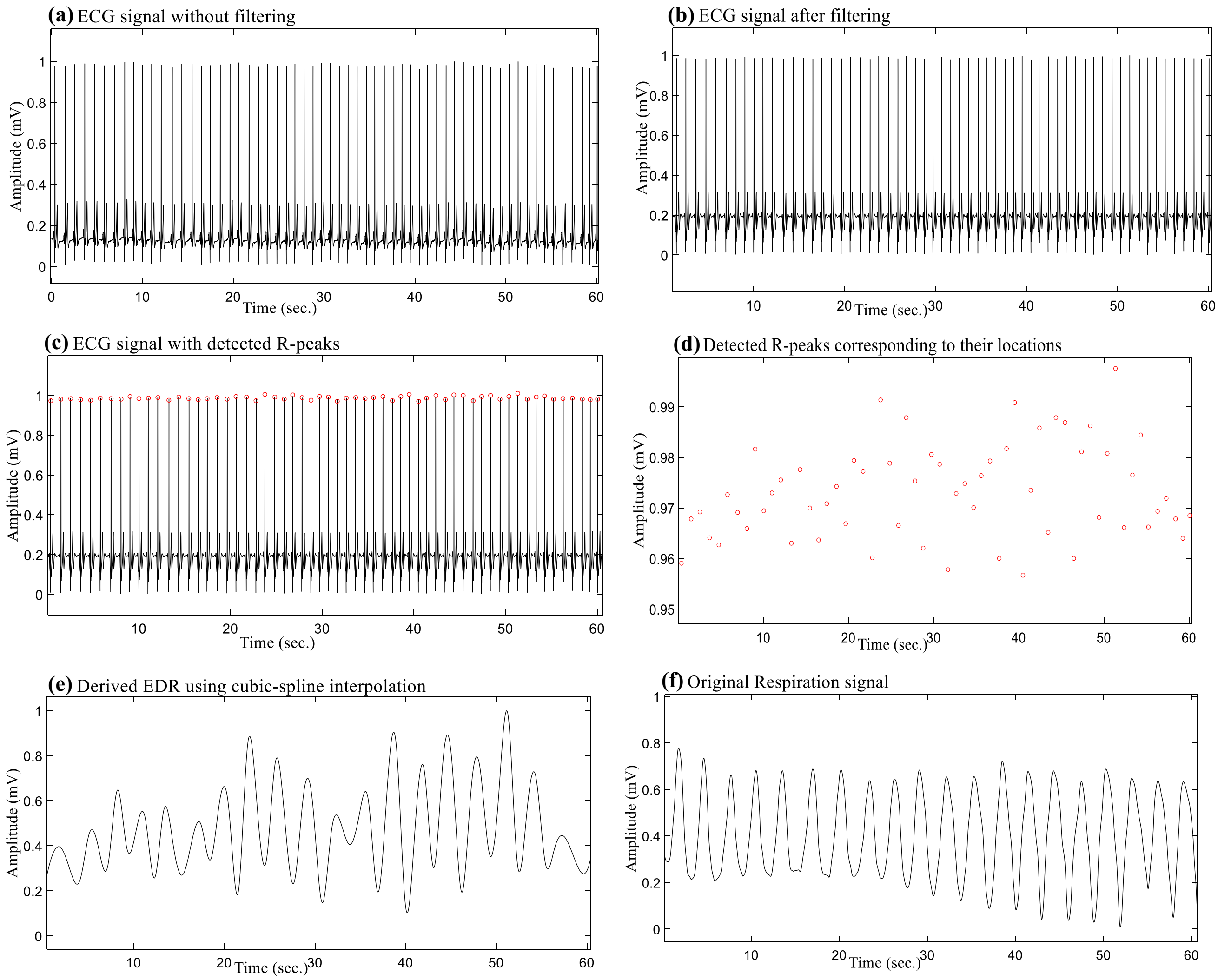

Fig. 2 The steps of EDR extraction are- **a** the original ECG signal, **b** ECG signal after denoising, **c** ECG signal with detected R peaks, **d** Detected R peaks corresponding to their locations against baseline (few R peaks are shown here in a magnified scale), **e** Derived EDR signal using cubic-spline interpolation. The respiration signal extracted from the thoracic belt is shown in (**f**)

where ReR_D is the respiration rate derived from EDR signal, ReR_R is the respiration rate derived from original respiration signal, n is the no. of subjects, and N is the total number of subjects.

Feature extraction

To utilize the morphological pattern changes in EDR and respiration signal, the discriminating parameters related to those changes were identified. As because COPD causes airway obstruction and airflow limitation, the respiration waveform of COPD patients differs from that of normal subjects [28]. The pattern changes can be seen in both original respiration signal and EDR signal of COPD patients as shown in Fig. 3.

For feature extraction, both the signals were taken for a time period of eighty seconds. Set of five randomly selected cycles from the window were picked out then for individual feature extraction. Based on the pattern differences, three features- area ratio, time ratio and skewness ratio were selected from both original and derived respiration signal for each individual subject. After denoising of respiration and EDR, those signals were normalized between 0 to unity. For one complete cycle, the co-ordinates of start point, peak and end points (Fig. 4) were detected using peak finding algorithm. The portion from starting point to peak was the inspiration, whereas, the portion between the peak and the ending point represented the expiration part.

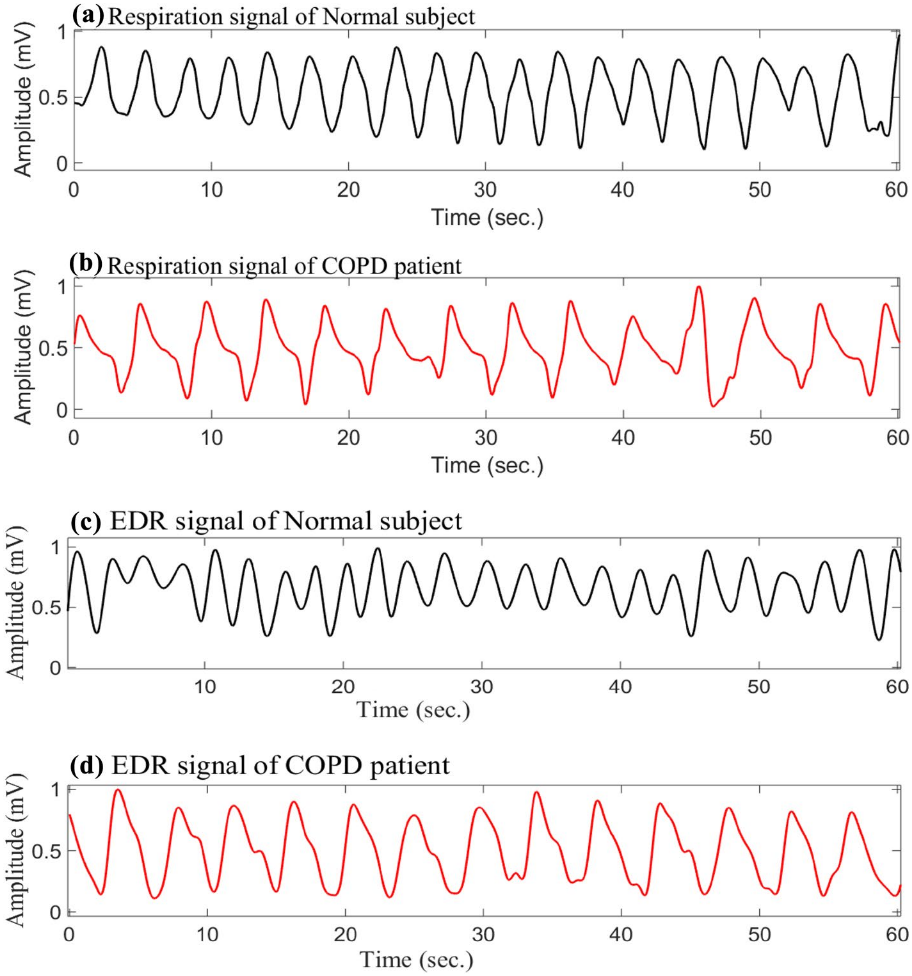

Fig. 3 Pattern difference between normal and COPD- (**a**) and (**b**) for original respiration signal (**c**) and (**d**) for EDR signal. The black solid lines represent the respiration and EDR signal for normal subject, and the red solid lines represent the respiration and EDR signal for COPD patient

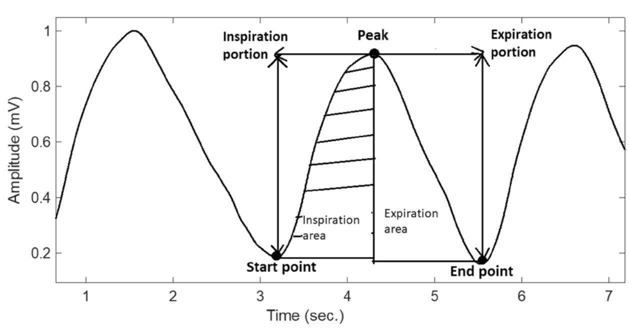

Fig. 4 EDR cycle with the starting, peak and ending point pointed and the portion from where the inspiration area and expiration area were calculated are shown. The inspiration and expiration portions are also shown in the figure

Area ratio

The inspiration area (I_A) and expiration area (E_A) were calculated by taking the area under the curve of the inspiration area portion and expiration area portion respectively as shown in Fig. 3. The area ratio (A_R) was determined by taking the ratio between expiration area to inspiration area as shown in Eq. 4.

$$A_R = \frac{1}{n}\sum_{i=1}^{n}\frac{E_{Ai}}{I_{Ai}} \tag{4}$$

where $i = 1,2,..n$ represents the ith number of randomly selected cycles.

Time ratio

The inspiration time (I_T) was the time duration between Start point to Peak, whereas, the expiration time (E_T) was the duration between Peak to End point as shown in Fig. 4. The time ratio (T_R) was computed using Eq. 5.

$$T_R = \frac{1}{n}\sum_{i=1}^{n}\frac{E_{Ti}}{I_{Ti}} \tag{5}$$

Skewness ratio

Skewness of inspiration portion (I_S) and expiration portion (E_S) were calculated separately and their ratio was taken as shown in Eq. 6.

$$S_R = \left| \frac{1}{n} \sum_{i=1}^{n} \frac{E_{Si}}{I_{Si}} \right| \tag{6}$$

Pearson correlation coefficient method was used to observe the similarity between the feature extracted from EDR and the same feature extracted from original respiration signal [29]. The correlation coefficient C_{mn} can be written as,

$$C_{mn} = \frac{\sum_{i=1}^{k} \left(m_i - \overline{m}\right)\left(n_i - \overline{n}\right)}{\sqrt{\sum_{i=1}^{k} \left(m_i - \overline{m}\right)^2} \sqrt{\sum_{i=1}^{k} \left(n_i - \overline{n}\right)^2}} \tag{7}$$

where k is the no. of data, m_i is the data points of original respiration signal, n_i is the data points of EDR signal,

$$\overline{m} = \frac{1}{k} \sum_{i=1}^{k} m_i,\ \overline{n} = \frac{1}{k} \sum_{i=1}^{k} n_i$$

Classification

Four different classifiers- Decision Tree, Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) and K Nearest Neighbor (KNN) were used in this study for a comparison between the classification performances. Classification was performed by tenfold cross validation method. In tenfold cross validation approach the data set was divided into 10 subsets, among which 9 subsets were selected to train the model and the remaining 1 subset was used as test data.

Decision Tree

Decision Tree algorithm is basically a supervised learning algorithm that can be used to solve both classification and regression problem. Decision Tree predicts class by learning the decision rule inferred from the training data set [30]. With training vector, $f_x \in K^n$, x = 1,2,..,r and a level vector $g \in K^r$, the Decision Tree repetitively split the dataset into samples with similar labels. The data P at node m, with a data split $\alpha = \left(u, v_m\right)$, [$u$ is a feature and v_m is the threshold], was partitioned into two subsets $P_{left}(\alpha)$ and $P_{right}(\alpha)$.

where $P_{left}(\alpha) = (f, g)|f_u \le v_m$, and $P_{right}(\alpha) = P/P_{left}(\alpha)$.

Linear Discriminant Analysis

LDA is a commonly used classifier that makes predictions by calculating the probability of a new set of data belongs to each class. The class with maximum probability value is the output and based on that a prediction is built. The probability is estimated using Bayes Theorem [30]. The model estimates the mean and variance of the data for each class. For the mean value M of each input i for each class c, can be calculated as, $M = 1/nc * \sum(i)$, where nc is the no. of instances for class c. The variance is, $\sigma^2 = 1/(n-c) * \sum\left((i-M)^2\right)$, where n is the number of instances. The discriminate function can be written as, $Dc(i) = i * \left(Mc/\sigma^2\right) - \left(Mc^2/\left(2 * \sigma^2\right)\right) + \ln(Pc)$, where P_c is the prior probability calculated from Bayes Theorem.

Support Vector Machine

SVM is a binary class classifier that maps the input dataset into higher dimension feature space [31–33]. SVM generally creates a hyperplane with maximum-margin, between different class-dataset, but it can be converted to non-linear classifier by applying the kernel to the hyperplane. Kernels are used to transform data to generate linear hyperplane in the higher dimensional feature space. Suppose, the training data belongs to binary class,

$$\xi = \left[\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots, \left(x_n, y_n\right)\right], x \in \psi^N, y \in (\pm 1) \tag{8}$$

where ψ is the Radius of the hyper-plane enclosing the complete input dataset, N is the total number of parameters. If the training dataset can be linearly discriminated by a hyperplane $\langle f, g\rangle + k = 0\ with\ f\psi^n,\ g \in \psi$, then the hard classifier can be written as shown in Eq. (9),

$$d(g) = \text{sgn}\left[\left\langle f', g\right\rangle + k\right] \tag{9}$$

Similarly, a soft classifier can be used for linearly interpolating the margin as shown in Eq. (10),

$$f(g) = x\left[\left\langle f', g\right\rangle + k\right] \tag{10}$$

where $x(m) = \left\{\begin{matrix} -1 : m < -1 \\ m : -1 < m < 1 \\ +1 : m > 1 \end{matrix}\right\}$.

K-Nearest Neighbor

KNN is a fundamental classifier that allocates the parametric values to a given class according to the decision which is made up by examining the input labels x on the K-Nearest Neighbors and taking the majority instances [34, 35]. Suppose x_i is one of the neighbor in the training set is $f\left(x_i, y_i\right) \in (0, 1)$, where i = 1,2,…,n, and $x_i$belongs to class

y_i [36]. Then for feature class C, the probability density function can be written as Eq. (11),

$$P(C, y_i) = \sum_{x_i \in kNN} S(C, x_i) \cdot f(x_i, y_i) \tag{11}$$

where $S(C,x_i)$ is the similarity function of C and x_i. In this study $S(C,x_i)$ was calculated using Euclidean distance metric where distance h can be written as,

$$h = \sqrt{\sum_{m=1}^{N} (p_m - q_m)} \tag{12}$$

where p_m is the input test data, q_m is the training data, and N is the total number of features. The label of K-Nearest Neighbor and the probability distribution of similarity function decide the class of input vector.

The classification performance was further analysed by determining the sensitivity, specificity and accuracy of the test result.

Result

Signals were acquired from a total 60 subjects (30 from each group) were recruited for this study. More than 80% of the COPD patients were smoker and rest of them have a history of either long term asthma or exposure to heavy indoor/outdoor pollution. Demographic details are shown in Table 2.

In the database, the age distribution of the normal subjects differs from that of the COPD patients as COPD generally attacks the older population. Normal group had a mean age of 47 years (range 23–72 years) while the mean age of COPD group was 58 years (range 45–71 years). However, previous researches showed changes in EDR with age, and presence of cardiovascular diseases [37–39]. For age be a physiological parameter, subjects from both groups with different age range were included in this study. However, subjects with cardiovascular diseases were excluded from this study as mentioned in section "Study population". The respiration rates and error calculation using EDR and original respiration signal for all the subjects are shown in Table 3.

Results show that the respiration rates calculated for both the subject groups using both the signals were almost same. Hence, the small phase difference between the derived and original respiration signal was ignored and both the signals were used for extraction of features. Three features, area ratio, time ratio and skewness ratio, were computed

Table 2 Demographic data of subject groups

Group	No. of subjects	Age (years)	Height (cm)	Weight (Kg)	BMI	Male: Female	No. of smoker/ non-smoker
Normal	30	47.57 ± 13.8	163 ± 10.12	61.43 ± 10.56	23.45 ± 2.78	3:2	6/24
COPD	30	58.82 ± 5.94	164.29 ± 8.33	60.78 ± 14.47	22.42 ± 4.64	7:3	25/5

Table 3 Respiration rate (RR) measurement from EDR and Respiration signal

Subject	RR from Respiration signal Average ± Standard deviation	RR from EDR signal Average ± Standard deviation	MAE	PE (%)	RMSE
Normal	16.5 ± 1.4	16.5 ± 1.3	0.23	1.43	0.48
COPD	23.6 ± 2.9	23.5 ± 3.1	0.37	1.57	0.61

Table 4 Statistical measurement on different extracted features for Normal and COPD group

Features extracted	COPD		Correlation coefficient	Normal		Correlation coefficient
	Respiration (mean ± SD)	EDR (mean ± SD)		Respiration (mean ± SD)	EDR (mean ± SD)	
Area ratio	1.66 ± 0.47	1.63 ± 0.36	0.762	0.94 ± 0.09	0.93 ± 0.11	0.76
Time ratio	1.71 ± 0.38	1.73 ± 0.39	0.764	0.95 ± 0.1	0.93 ± 0.09	0.783
Skewness ratio	4.42 ± 4.41	4.38 ± 4.29	0.987	0.69 ± 0.36	0.63 ± 0.37	0.855

from both derived and original respiration signal. The mean ± standard deviation (SD) value of the features for both normal and COPD group and their correlation co-efficient values are given in Table 4. Each feature was extracted five times from randomly selected waveforms and their mean value was taken for further calculation.

From Table 4, it can be observed that there is a big difference between the mean values of same feature extracted from two different groups. Though the correlation coefficient calculated between the feature derived from EDR to the same feature extracted from respiration signal for same subject group is high. This establishes the fact that the feature values extracted from EDR signal are almost similar to the features extracted from original respiration signal. Hence, EDR can be used as an alternative of respiration signal for showing respiration related information.

Two features at a time and a combination of all three features were used for classifying normal to COPD population. Figure 5 shows two-dimensional scatter plots, where each EDR-extracted feature set was classified using a classifier that achieved the highest classification accuracy.

The classification result of different classifier is shown in Table 5.

The result shows that for respiration-extracted features, the minimum accuracy of 85.00% (sensitivity 70% and specificity 100%) was achieved using LDA classifier on area ratio versus skewness ratio, while, the maximum accuracy level was 98.33% achieved using Decision Tree, SVM and KNN classifier on time ratio versus skewness ratio, and using Decision Tree classifier on the combination of three features. On the other side, the evaluation of classification performance on EDR based feature sets shows that the minimum accuracy level was 93.33% (sensitivity 86.67%, specificity 100%) by using LDA on area ratio versus skewness ratio and the maximum accuracy level was 100% in case of using Decision Tree and KNN on area ratio versus time ratio, using Decision Tree, SVM and KNN on area ratio versus skewness ratio, and using SVM and KNN on time ratio versus skewness ratio.

Figure 6 shows that the overall classification performance achieved using EDR signal was better than the performance achieved using respiration signal.

Outcome of some previously published researches along with the proposed one is shown in Table 6.

The comparable classification performance (showed in Table 6) demonstrated that the proposed method offers a

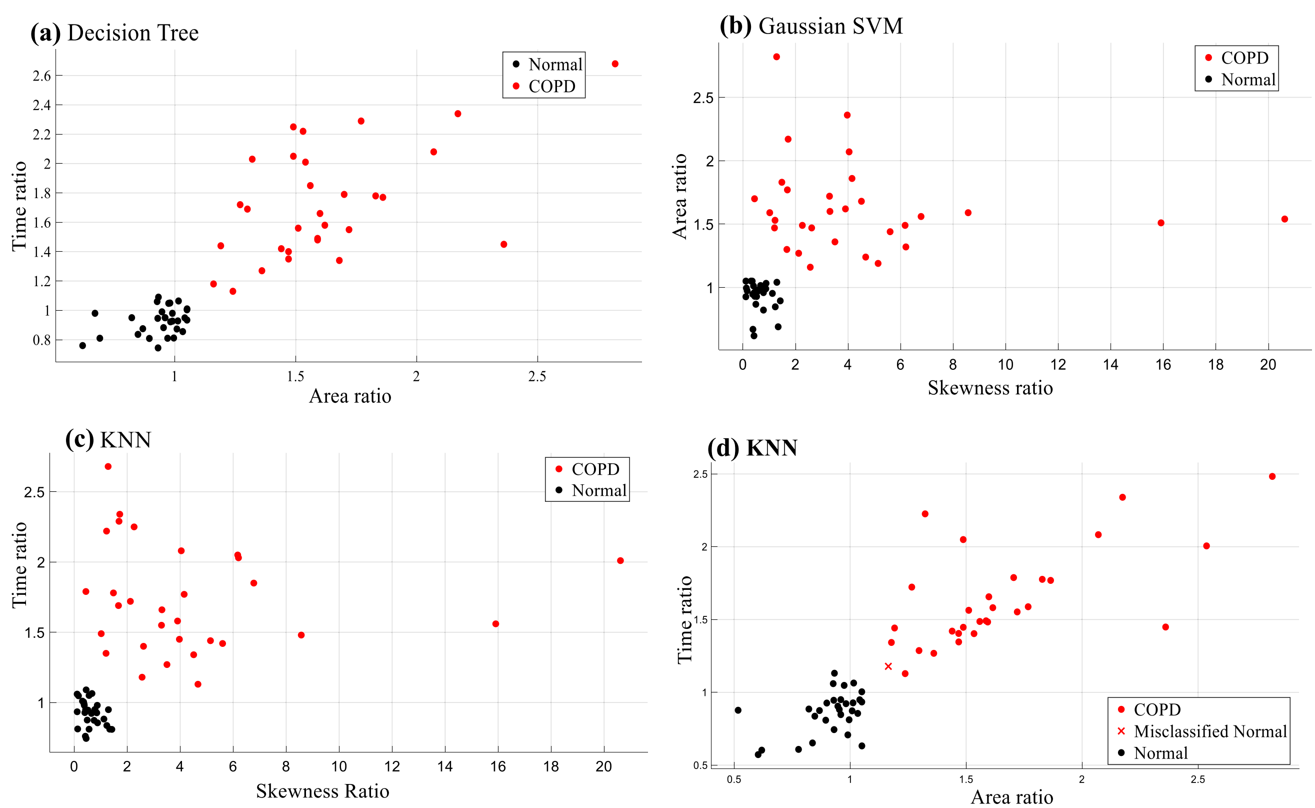

Fig. 5 Two dimensional scatter plot using features extracted from EDR, where, **a** shows area ratio versus time ratio plot using Decision Tree, **b** shows area ratio versus skewness ratio plot using Gaussian SVM, **c** shows time ratio versus skewness ratio plot using KNN. Figure **d** shows the combination of three features plotted against two dimensional axis (time ratio vs. area ratio) using KNN

Table 5 Classification performance result of different classifiers

Feature set used for classification	Classifier	Respiration			EDR		
		Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)
Area ratio vs. Time ratio	Decision Tree	96.67	96.67	96.67	100.00	100.00	100.00
	LDA	88.33	76.67	100.00	96.67	93.33	100.00
	Gaussian SVM	96.67	96.67	96.67	98.33	100.00	96.67
	KNN	95.00	93.33	96.67	100.00	100.00	100.00
Area ratio vs. Skewness ratio	Decision Tree	90.00	93.33	86.67	100.00	100.00	100.00
	LDA	85.00	70.00	100.00	93.33	86.67	100.00
	Gaussian SVM	95.00	93.33	96.67	100.00	100.00	100.00
	KNN	93.33	90.00	96.67	100.00	100.00	100.00
Time ratio vs. Skewness ratio	Decision Tree	98.33	96.67	100.00	95.00	93.33	96.67
	LDA	90.00	80.00	100.00	96.67	93.33	100.00
	Gaussian SVM	98.33	100.00	96.67	100.00	100.00	100.00
	KNN	98.33	96.67	100.00	100.00	100.00	100.00
Combination of three features	Decision Tree	98.33	100.00	96.67	96.67	100.00	93.33
	LDA	90.00	80.00	100.00	96.67	93.33	100.00
	Gaussian SVM	95.00	100.00	90.00	95.00	100.00	90.00
	KNN	95.00	90.00	100.00	98.33	96.67	100.00

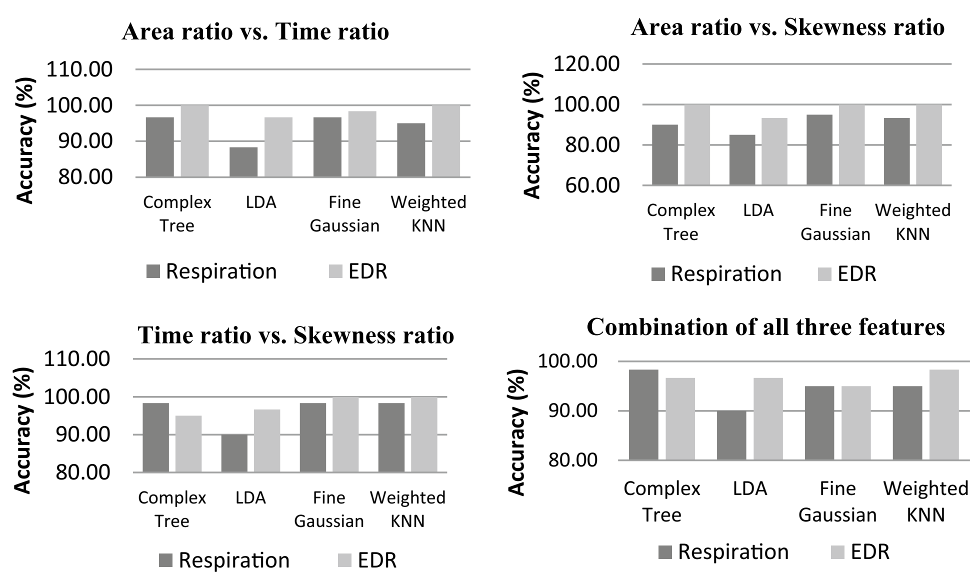

Fig. 6 Comparison of classification accuracies obtained from respiration and EDR signal by using Decision Tree, LDA, Gaussian SVM and KNN classifier

computationally simple and reduced feature set with good classification results.

As per Global initiative for chronic Obstructive Lung Disease (GOLD) guideline, diagnosis of COPD is mainly based on two parameters- the ratio of forced expiratory volume in 1st second to the forced vital capacity (FEV_1/ FVC) < 0.7; and $FEV_1 > 80\%$ for mild COPD, $80\% \geq FEV_1 > 50\%$ for moderate COPD, $50\% \geq FEV_1 > 30\%$ for severe COPD and $FEV_1 \leq 30\%$ for very severe COPD patients.

Statistical significance of these two spirometric parameters (FEV_1 and FEV_1/FVC)were compared with the features extracted from EDR signal (Table 7).

The significant p values (< 0.0001), show that the features extracted from EDR signals also possess significant difference between normal and COPD similar to those spirometric parameters.

Discussion

This study was accomplished to investigate the characteristics of original respiration signal as well as the EDR signal derived from the ECG of normal and COPD groups. The major findings of this study are that patients with

Table 6 Comparative study of some previous studies on COPD detection with the proposed one

Sl	Publication	Features used	Classifier used	Accuracy %	Specificity %	Sensitivity %
1	Van Berkel et al. [10]	13 VOCs from exhaled air (isoprene, C16 hydrocarbon, 4,7-dimethyl-undecane, etc.)	SVM	91%	81%	100%
2	Dellaca et al. [11]	Features extracted from within breath reactance during inspiration and expiration	Mann–Whitney	Not specified	100%	100%
3	Mieloszyk et al. [13]	Exhalation duration, maximum end-tidal $PeCO_2$ and its time duration, and end-exhalation slope	Quadratic discriminant analysis	93.9% (95% CI: 81.8-100%) for COPD-normal classification	88% for COPD-normal classification	88% for COPD-normal classification
4	Phillips et al. [40]	Level of isoprene and total VOCs (excluding isoprene)	J48 decision tree	74.13% for J48	Not specified	Not specified
5	Patel et al. [41]	Intensity, range of motion, orientation, modulation and signal complexity of the data taken from accelerometer, heart rate and respiration rate	Instance Based Learning (IBL), Naïve Bayes, J48, Multilayer Perceptron (MLP), Random Forest, SVM	Classification error (%) was 1.26 ± 0.56 for IBL, 2.92 ± 1.25 for Naïve Bayes, 5.92 ± 2.30 for J48, 1.24 ± 0.43 for MLP, 2.08 ± 0.44 for Random Forest and 1.31 ± 0.55 for SVM using 10 cross validation		
6	Proposed method	Area ratio, time ratio, skewness ratio and combination of all three features	Decision tree, LDA, SVM, KNN	100% for EDR using decision tree, SVM and KNN	100% for EDR using decision tree, SVM and KNN	100% for EDR using decision tree, SVM and KNN

Table 7 Statistical measurement of spirometric and EDR parameters

Features	COPD (Mean ± SD)	Normal (Mean ± SD)	p value
EDR area ratio	1.63 ± 0.36	0.93 ± 0.11	< 0.0001
EDR time ratio	1.73 ± 0.39	0.95 ± 0.1	< 0.0001
EDR skewness ratio	4.38 ± 4.29	0.69 ± 0.36	> 0.1
FEV1 (post bronchodilator % predicted)	0.45 ± 0.18	0.87 ± 0.07	< 0.0001
FEV1/FVC (post bronchodilator absolute value)	0.53 ± 0.12	0.84 ± 0.05	< 0.0001

COPD showed morphological pattern changes in case of both respiration and EDR signal and EDR performed better for discriminating one subject group from the other.

The number of COPD patients is escalating all over the world specially in developing countries where the presence of risk factors like air pollution and number of smokers are massive. The total number of undiagnosed and undertreated patients is still more than the treated one. The high rate of premature death occurs due to lack of knowledge about this disease and continuous adjustment to undesirable living conditions. Therefore, the necessity of early detection of this disease for reducing mortality and morbidity rate is gaining attention.

In some literatures, different devices were used for measuring respiration signal and to study its parameters. Even many signal-based feature extraction methods were studied for classifying COPD patients from healthy normal subjects [12, 15, 41]. Among these methods features like reactance change in expiration, exhalation duration, and end exhalation slope were also used for analysing respiration signal. The expiration pattern of COPD patient changes due to their insufficient expiration ability. Therefore, features extracted from the expiration signal of the patient can be helpful for diagnosing this disease.

Most of the studies for COPD detection were based on either spirometry, or plethysmography [7, 11, 16]. Besides, few independent techniques like nonlinear analysis of electrodermal activity signals [42] and lung sound analysis [43, 44] were also developed later.

The use of ECG derived respiration in disease detection is already an established method. The applicability of EDR for diagnosing various diseases like obstructive sleep apnoea [45] and chronic heart failure [46] were also reported. But very few studies related to EDR signal for diagnosis of COPD were reported [20, 21].

In present study, the respiration and ECG signals were recorded simultaneously from all the subjects of normal and COPD groups. EDR signals were derived later for investigating its characteristics in normal and abnormal conditions. COPD is a progressive lung disease with severe airway obstruction and airflow limitation and even parenchymal destruction. Due to change in internal lung structure, respiration signal also changes in patients with COPD. Even some previous literatures stated changes of heart rate variability in COPD patients. For example, heart rate variability changes in COPD patients [47], in patients with exacerbations [48], and also in patients during exercise [49]. In this study, the signal analysis showed visible respiration as well as EDR pattern change in COPD patient from the normal one. The previously reported investigations were mainly based on the heart rate variability analysis of chronic obstructive pulmonary disease patients. So, from that point of view, this is the first study which is based on the pattern analysis of EDR signal using amplitude variation method.

In this study, 'COPD group' was constituted using data from all COPD patients staring from early stage to very severe stage, whereas, 'Normal group' consisted healthy subjects. Features were extracted from both original respiration signal and from EDR signal. It was observed during our study that the value of each feature for COPD patients was increased than value of normal subjects. Also, the increased values of all the features denote the morphological expiration pattern change in COPD patients. Four classifiers- Decision Tree, Linear Discriminant Analysis, Support Vector Machine and K Nearest Neighbor were used for classifying the two groups. The classification results showed that the EDR-based features performed better than the respiration-based features. In one of our previous studies, we observed that EDR also performed better in classifying two normal subject groups- one without any symptoms, and other with symptoms present (cough, cold, etc.) [50]. This indicates that the respiration signal pattern may change for even mild physiological changes, whereas, EDR is not easily affected unless significant cardiopulmonary changes occur. This makes EDR more robust for classifying abnormalities. The result also showed that binary feature classification produced better classification performance than multi-feature classification. The comparative study with other literatures (Table 6) showed that the classification performance of this study outnumbered against most of the performance results reported earlier.

The EDR-based features were validated along with the standard spirometric parameters FEV1 and FEV1/FVC in case of COPD-Normal classification. The statistically significant p values (<0.0001) for area ratio and time ratio indicate that these two features can be used alike the standard parameters for classification purpose. The result also showed that mean ± SD value of area ratio and time ratio were closely spaced like the two spirometric parameters. This indicates that these two EDR-based features might be sufficient for discriminating COPD patients of any stage from the normal subjects. Therefore, the proposed method provides substantial benefits such as non-obstructive measurement, computational simplicity of feature extraction, compact feature dimension, small execution time, reduced lead set for patient comfort.

This study has confronted several inadequacies. Though the number of subjects in both groups was same, but the age, gender ratio, and number of smokers were different in COPD group compared to the normal one because COPD is predominant among smokers as well as in elderly group. In some previous works, researchers showed that these parameters had some influence on breathing patterns [51–53]. This might affect the classification performance. The result might be better and more reliable if the number of total study population, were more and the subject groups were age and sex-matched. COPD patients from all stage groups were included in this study. However, the total number of mild and moderate patients was much less than the number of severe and very severe patients. Classification of mild and moderate stage COPD patients with normal subjects might give more information about the possibility of early disease detection using EDR based method.

Conclusion

In this work, a non-obstructive technique for discriminating COPD from normal population using EDR signal was developed. The results obtained from the proposed method revealed that EDR can be incorporated to develop an automated COPD detection modality for a portable COPD analysis device since early stage treatment is the main way for reducing the higher mortality rate. Morphological variations have been observed in both respiration and EDR waveform. Extensive studies are planned to observe the effect of COPD staging on ECG derived respiration signal which might correlate with the standard GOLD staging. Study with higher number of samples is needed to improve accuracy and to validate the result for its applicability in clinical field.

Acknowledgements The authors would like to thank the medical technicians and research assistants of Institute of Pulmocare & Research, Kolkata, India for their valuable assistance in this study. The first author acknowledges the support of Council of Scientific & Industrial Research, Human Resource Development Group, India through the CSIR-SRF Fellowship.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

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

References

1. WHO (2017) Fact sheet. The top ten causes of death. Fact Sheet-310. World Health Organization, Geneva.
2. Vrbica Z, Labor M, Gudelj I, Labor S, Juric I, Plavec D (2017) Early detection of COPD patients in GOLD 0 population: an observational non-interventional cohort study- MARKO study. BMC Pulm Med 17(1):36–923
3. van Schayck OCP, D'Urzo A, Invernizzi G, Roman M, Stallberg B, UrbinaC, (2003) Early detection of chronic obstructive pulmonary disease (COPD): the role of spirometry as a diagnostic tool in primary care. Prim Care Respir J 12(3):90–93
4. Decramer M, Miravitlles M, Price D, Rodriguez MR, Llor C, Welte T, Buhl R, Dusser D, Samara K, Siafakas N (2011) New horizons in early stage COPD-improving knowledge, detection and treatment. Respir Med 105:1576–1587–1587
5. Rabe KF (2007) Global initiative for chronic obstructive lung disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 176:532–555
6. Petty TL (2005) Benefits of and barriers to the widespread use of spirometry. Curr Opin Pulm Med 11:115–120
7. Yamauchi Y, Kohyama T, Jo T, Nagase T (2012) Dynamic change in respiratory resistance during inspiratory and expiratory phases of tidal breathing in patients with chronic obstructive pulmonary disease. Int J Chronic Obstr Pulm Dis 7:259–269
8. Taplin GV, Tashkin DP, Chopra SK, Anselmi OE, Elam D, Calvarese B, Coulson A, Detels R, Rokav SN (1977) Early detection of chronic obstructive pulmonary disease using radionuclide lung-imaging procedures. Chest 71:567–575
9. Hosseini MP, Zadeh HS, Akhlaghpoor S (2012) Detection and severity scoring of chronic obstructive pulmonary disease using volumetric analysis of lung CT images. Iran J Radiol 9(1):22–27
10. Van Berkel JJBN, Dallinga JW, Moller GM, Godschalk RWL, Moonen EJ, Wouters EFM, van Schooten FJ (2010) A profile of volatile organic compounds in breath discriminates COPD patients from controls. Respir Med 104:557–563
11. Dellaca RL, SantusP Aliverti A, Stevenson N, Centanni S, Macklem PT, Pedotti A, Calverley PMA (2004) Detection of expiratory flow limitation in COPD using the forced oscillation technique. Eur Respir J 23:232–240
12. Frantz S, Nihlen U, Dencker M, Engstrom G, Lofdahl CG, Wollmer P (2012) Impulse oscillometry may be of value in detecting early manifestations of COPD. Respir Med 106:1116–1123
13. Mieloszyk RJ, Verghese GC, Deitch K, Cooney B, Khalid A, Mirre-González MA, Heldt T, Krauss BS (2014) Automated quantitative analysis of capnogram shape for COPD-normal

and COPD-CHF classification. IEEE Trans Biomed Eng 61(12):2882–2890
14. Velasquez A, Duran CM, Gualdron O, Rodriguez JC, Manjarres L (2009) Electronic nose to detect patients with COPD from exhaled breath. AIP Conf Proc 1137(1):452
15. Warnier MJ, Rutten FH, Numans ME, Kors JA, Tan HL, de Boer A, Hoes AW, De Bruin M (2013) Electrocardiographic characteristics of patients with chronic obstructive pulmonary disease COPD. J Chronic Obstr Pulm Dis 10:62–71
16. Rachaiah NM, Rachaiah JM, Krishnaswamy RB (2012) A correlative study of spirometric and ECG changes in patients with chronic obstructive pulmonary disease. Int J Biol Med Res 3(1):1322–1326
17. Loveridge B, West P, Anthonisen NR, Kryger MH (1984) Breathing patterns in patients with chronic obstructive pulmonary disease. Am Rev Respir Dis 130:730–733
18. Moody GB, Mark RG, Zoccola A, Mantero S (1985) Derivation of respiratory signals from multi-lead ECGs. Comput Cardiol 12(1985):113–116
19. Zhao L, Reisman S, Findley T (1994) Derivation of respiration from electrocardiogram during heart rate variability studies. Computers in Cardiology, Bethesda, Maryland, USA
20. Lado MJ, Méndez AJ, Vila XA, Rodríguez-Liñares L, Félix P (2012) HRV patterns and exacerbations of COPD patients following routine controls: a preliminary study. 7th Iberian Conference on Information Systems and Technologies, Madrid.
21. Pantoni CBF, Reis MS, Martins LEB, Catai AM, Costa D, Borghi-Silva A (2007) Study of heart rate autonomic modulation at rest in elderly patients with chronic obstructive pulmonary disease. RevistaBrasileira de Fisioterapia 11(1):33–38
22. Handa R, Paonta L, Rusu D, Albu A (2012) The role of heart rate variability in assessing the evolution of patients with chronic obstructive pulmonary disease. Rom J Intern Med 50:83–88
23. Celli BR, MacNee W (2004) Standards for the diagnosis and treatment of patients with COPD: asummary of the ATS/ERS position paper. Eur Respir J 23(6):932–946
24. Pflanzer R, McMullen W (2014) Electrocardiography (ECG) I Introduction Biopac Student Lab Manual. BIOPAC Systems Inc, Goleta
25. Sobron A, Romero I, Lopetegi T (2010) Evaluation of methods of respiratory frequency from the ECG. Computing in Cardiology, IEEE, New Jersey
26. Chan AM, Ferdosi N, Narasimhan R (2013) Ambulatory respiratory rate detector using ECG and a triaxial accelerometer. 35th Annual International Conference of the IEEE EMBS, Osaka.
27. Sarkar S, Bhattacherjee S, Pal S (2015) Extraction of respiration signal from ECG for respiratory rate estimation. Proceeding of Michael Faraday IET International Summit, Kolkata
28. Sarkar S, Bhattacherjee S, Bhattachrayya P, Pal S (2016) Differentiation of COPD from normal population using ECG derived respiration: a pilot observation. Pulmo-Face XV I(1):12–18
29. Rodgers JL, Nicewander WA (1988) Thirteen ways to look at the correlation coefficient. Am Stat 42(1):59–66
30. Hastie T, Tibshirani R, Friedman J (2009) Elements of statistical learning, 2nd edn. Springer, New York
31. Christianini N, Taylor JS (2000) An introduction to support vector machines and other kernel-based learning methods, 1st edn. Cambridge University Press, Cambridge
32. Sahin D, Ubeyli ED, Ilbay G, Sahin M, Yasar AB (2010) Diagnosis of airway obstruction or restrictive spirometric patterns by multiclass support vector machines. J Med Syst 34:967–973
33. Vapnik VN (1999) The nature of statistical learning theory, 2nd edn. Springer, New York
34. Parvin H, Alizadeh H, Minati B (2010) A modification on K-nearest neighbor classifier. Glob J Comput Sci Technol 10(14):37–41
35. Piotrowski Z, Szypulska M (2017) Classification of falling asleep states using HRV analysis. Biocybern Biomed Eng 37:290–301
36. Kim KS, Choi HH, Moon CS, Mun CW (2011) Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. Curr Appl Phys 11:740–745
37. Reardon M, Malik M (1996) Changes in heart rate variability with age. Pace 19:1863–1866
38. Cysarz D, Zerm R, Bettermann H, Fruhwirth M, Moser M, Kroz M (2008) Comparison of respiratory rates derived from heart rate variability, ECG amplitude, and nasal/oral airflow. Ann Biomed Eng 36(12):2085–2094
39. van Ravenswaaij CMA, Kollee LAA, Hopman JCW, Stoelinga GBA, van Geijn HP (1993) Heart rate variability. Ann Intern Med 118:436–447
40. Phillips C, Parthaláin NM, Syed Y, Deganello D, Claypole T, Lewis K (2014) Short-term intra-subject variation in exhaled volatile organic compounds (VOCs) in COPD patients and healthy controls and its effect on disease classification. Metabolites 4:300–318
41. Patel S, Mancinelli C, Bonato P, Healey J, Moy M (2009) Using wearable sensors to monitor physical activities of patients with COPD: a comparison of classifier performance. Body Sens Netw 234–239.
42. Caliskan SG, Bilgin MD, Polatli M (2018) Nonlinear analysis of electrodermal activity signals for healthy subjects and patients with chronic obstructive pulmonary disease. Australas Phys Eng Sci Med 41:487–494
43. Malmberg LP, Pesu L, Sovijarvi ARA (1995) Significant differences in flow standardised breath sound spectra in patients with chronic obstructive pulmonary disease, stable asthma, and healthy lungs. Thorax 50:1285–1291
44. Fard PJM, Moradi MH, Saber S (2015) Chaos to randomness: distinguishing between healthy and non-healthy lung sound behaviour. Australas Phys Eng Sci Med 38:47–54
45. de Chazel P, Heneghan C, Sheridan E, Reilly R, Nolan P, O'Malley M (2003) Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. IEEE Trans Biomed Eng 50(6):686–696
46. Pichot V, Chouchou F, Pepin JL, Tamisier R, Levy P, Fortune IC, Sfroza E, Barthelemy JC, Roche F (2015) ECG-derived respiration: a promising tool for sleep-disordered breathing diagnosis in chronic heart failure patients. Int J Cardiol 186:7–9
47. Volterrani M, Scalvini S, Mazzuero G, Lanfranchi P, Colombo R, Clark AL, Levi G (1994) Decreased heart rate variability in patients with chronic obstructive pulmonary disease. Chest 106(5):1432–1437
48. Skyba P, Joppa P, Orolín M, Tkáčová R (2007) Blood pressure and heart rate variability response to noninvasive ventilation in patients with exacerbations of chronic obstructive pulmonary disease. Physiol Res 56:527–533
49. Bartels MN, Jelic S, Ngai P, Banser RC, DeMeersman RE (2003) High-frequency modulation of heart rate variability during exercise in patients with COPD. Chest 124:863–869
50. Sarkar S, Pal S, Bhattacharyya P (2019) A comparative analysis between EDR and respiration signal: a pilot study with normal subjects. Modelling and simulation in science, technology and engineering mathematics, advances in intelligent systems and computing. Springer, Cham.
51. Parreira VF, Bueno CJ, França DC, Vieira DS, Pereira DR, Britto RR (2010) Breathing pattern and thoracoabdominal motion in healthy individuals: influence of age and sex. Braz J Phys Ther 14(5):411–416
52. Gabriel L, Hoffman M, Mendes L, Samora G, Rattes C, Dornelas A, Britto R, Parreira V (2015) Comparison of breathing pattern

and thoracoabdominal motion of healthy elderly. Eur Respir Soc 46:PA4217
53. Mendes LPDS, Vieira DSR, Gabriel LS, Ribeiro-Samora GA, De Andrade AD, Brandão DC, Goes MC, Fregonezi GAF, Britto RR, Parreira VF (2019) Influence of posture, sex, and age on breathing pattern and chest wall motion in healthy subjects. Braz J Phys Ther. https://doi.org/10.1016/j.bjpt.2019.02.007