IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 2, 2018, accepted August 29, 2018, date of publication September 20, 2018, date of current version October 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2871091

Automatic Quantitative Analysis of Human Respired Carbon Dioxide Waveform for Asthma and Non-Asthma Classification Using Support Vector Machine

OM PRAKASH SINGH[1], RAMASWAMY PALANIAPPAN[2], (Senior Member, IEEE), AND MB MALARVILI[1]

[1]Bio-Signal Processing Research Group, Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia
[2]Data Science (E-Health) Research Group, School of Computing, University of Kent, Medway ME4 4AG, U.K.

Corresponding author: Ramaswamy Palaniappan (R.Palani@kent.ac.uk)

This work was supported by the Ministry of Higher Education under the Prototype Research Grant Scheme (PRGS) under Grant R.J130000.7845.4L669.

ABSTRACT Currently, carbon dioxide (CO_2) waveforms measured by capnography are used to estimate respiratory rate and end-tidal CO_2 ($EtCO_2$) in the clinic. However, the shape of the CO_2 signal carries significant diagnostic information about the asthmatic condition. Previous studies have shown a strong correlation between various features that quantitatively characterize the shape of CO_2 signal and are used to discriminate asthma from non-asthma using pulmonary function tests, but no reliable progress was made, and no translation into clinical practice has been achieved. Therefore, this paper reports a relatively simple signal processing algorithm for automatic differentiation of asthma and non-asthma. CO_2 signals were recorded from 30 non-asthmatic and 43 asthmatic patients. Each breath cycle was decomposed into subcycles, and features were computationally extracted. Thereafter, feature selection was performed using the area (A_z) under the receiver operating characteristics curve analysis. A classification was performed via a leave-one-out cross-validation procedure by employing a support vector machine. Our results show maximum screening capabilities for upward expiration (AR_1), downward inspiration (AR_2), and the sum of AR_1 and AR_2, with an A_z of 0.892, 0.803, and 0.793, respectively. The proposed method obtained an average accuracy of 94.52%, sensitivity of 97.67%, and specificity of 90% for discrimination of asthma and non-asthma. The proposed method allows for automatic classification of asthma and non-asthma condition by analyzing the shape of the CO_2 waveform. The developed method may possibly be incorporated in real-time for assessment and management of the asthmatic conditions.

INDEX TERMS Area, quantitative, feature, classifier, asthma.

I. INTRODUCTION

Quantitative analysis of the respired carbon dioxide (CO_2) waveform shape carries significant diagnostic information for the classification of asthma exacerbation and its severity level [1]–[7]. During an asthma attack, the inflamed or sensitive trachea causes the tracheal muscle to repeatedly contract, and an excess amount of mucus to be secreted, which leads to airway obstruction, and symptoms such as difficulty breathing, chest pain, and coughing [8]. In addition, asthma is presumed to be a high-risk and non-curable disease; hence, early screening or prediction of asthma may possibly control and minimize the morbidity and mortality rate of asthma attacks [9]. The existing method for early identification of asthma is based on the healthcare provider's physical assessment utilizing a spirometer or peak flow meter, which are manual and unreliable if patients are noncooperative [10]. In addition, they require a set of instructions to be followed and deep, quick, and complete in-and-out breaths to be used during the maneuver [11], [12]. Thus, young, elderly, injured, anesthetized, sore, and ill patients are often unable to perform these tests [13]. Therefore, capnography has been proposed as a patient-independent asthma assessment device.

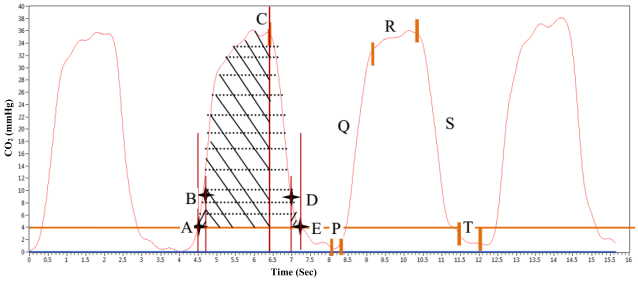

FIGURE 1. Four breath cycles extracted from a CO_2 signal recorded for 2.5 minutes. A-B, A-C, A-B-C-D-E, and D-E, represent the upward expiratory (4mmHg <= CO_2 <= 10mmHg), absolute expiratory (4mmHg <= CO_2 = Max (A-B-C-D-E)), complete breath cycle (4mmHg <= CO_2 <= 4 mmHg), and downward inspiratory phase (10 mmHg <= CO_2 <= 4 mmHg) respectively, refer to cycle 2. P, Q, R, and T indicate the gas of anatomic dead space (airway), vacating of alveoli gas progressing, alveolar gas, and inspiratory down-stroke of gas. α- angle lies between Q and R, while β- angle lies between R and T, refer to cycle 1.

The capnography device measures and displays the CO_2 signal, end-tidal CO_2 ($EtCO_2$), respiratory rate (RR) and inspired CO_2 (ICO_2) as a waveform and as numeric values continuously from the carbon dioxide partial pressure in a sample of expiratory air ($PeCO_2$) [1]. The normal CO_2 signal looks like a square waveform as presented in Fig. 1 (refer to breath cycle three), followed by an alternating inspiration (ICO_2 equal to zero) and expiration phase [14]–[16].

The expired portion of each breath cycle consists of three phases: 1) the latency phase, ''P'', which represents the exhalation of the anatomical dead space ($PeCO_2 = 0$) and is difficult to differentiate from prior inhalation; 2) the upward expiration phase, ''Q'', which has a rapid increase in $PeCO_2$ indicating the exhalation of mixed air; 3) the alveolar plateau phase that consists of ''R'', which reveals the purging of alveolar air (a slight increase in $PeCO_2$) that consists of a peak at the end of exhalation, ''S'', known as end-tidal carbon dioxide ($EtCO_2$), which is immediately followed by inspiration, ''T'' and ''U'' [1] that are both angles (α and β) and indicate the changes between Q, and R and R and T as shown in Fig. 1 (refer to breath cycle 3). It is believed that features provided by these phases, specifically the alveolar and upward expiration phases, can reveal asthma exacerbation [1]–[3], [12], [17], [18].

Several studies have been conducted with the subjects suffering from asthma that have demonstrated significant correlations between CO_2 signal indices computed using capnography and the standard spirometry and peak flow meter index [1]–[3], [17], [18]. These studies are of great interest since capnographic measurements are patient independent and can be applied with young, elderly, injured or even unconscious patients and could, therefore, be used as an alternative to the standard pulmonary function test to monitor airways obstruction in a number of clinical conditions.

The study conducted by You *et al.* [17] (1992) revealed that computation of the slope of the alveolar phase of a CO_2 signal has a good correlation between the capnographic index and spirometric parameters. Furthermore, You *et al.* [1] (1994) proposed multiple indices to provide a better representation of bronchial obstruction. Specifically, they measured eight indices (S1, S2, S3, SR, AR, SD1, SD2 and SD3) from each valid breath cycle and compared the mean values of these indices with spirometric indices and found that the angle (α) between the ascending phase (Q, refer to Fig. 1, breath cycle 3) and the alveolar plateau (R, refer to Fig. 1, breath cycle 3), was the most significant. However, real-time implementation is still challenging due to the random time-based setting criteria. Later, the study conducted by Yaron *et al.* [2] (1996) found that the derivative of the alveolar phase differed significantly between asthmatic and non-asthmatic patients. However, this study used manual methods to extract the features. Thereafter, an apparent trend has developed for the extraction of CO_2 signal features in the belief that the incorporation of these features into CO_2 monitoring devices may provide a patient independent method to understand asthma exacerbation.

In 2009, Hisamuddin *et al.* [18] reported that the slope of phase 3 (R, refer to Fig. 1) and α-angle (refer to Fig. 1) differed significantly between asthma and non-asthma patients, which coincides with the finding of Howe *et al.* [3] (2011). However, the above-mentioned studies were conducted at a low sampling rate (10 Hz) and used setting criteria and linear trend methods in order to calculate the slope, which required the sample distribution to be linear. In fact, obtaining simple linear data sets in real time is quite difficult due to the uneven removal of CO_2 samples from the alveoli. In addition, the studies conducted by Langhan *et al.* [19] and Howe *et al.* [3] divulged that time-based setting criteria are challenging to implement in the real-time environment. These authors also stated that the quantification and analysis of the CO_2 waveform cannot be easily employed in the emergency department and are not clinically useful. Furthermore, Kazemi *et al.* (2013) computed the power spectral density (PSD) from each breath cycle and proposed a new feature for differentiating the asthmatic condition, which may possibly overcome the limitations of previous studies [6]. However, computation of the PSD is complex and requires higher level calculations. In addition, the computation time is relatively slow and requires more memory compared to other time domain features.

To the best of our knowledge, currently, no significant progress has been made regarding the incorporation of the above-discussed features into a real-time CO_2 measurement device, nor has a translation to clinical practice been achieved. Therefore, here we propose the area (AR_i) and derivative $\left(\frac{dco_2}{dt}\right)$ as new prognostic indices for the screening of asthma and non-asthma, which may possibly be easily incorporated into real-time CO_2 measurement devices due to the simplicity of the algorithm. We employed Simpson's rule to compute the AR_i of the upward expiration (AR_1, A-B, refer to breath cycle 2), downward inspiration (AR_2, D-E, refer to breath cycle 2), absolute expiration (AR_3, A-B-C, refer to breath cycle 2), complete breath cycle (AR_4, A-B-C-D-E,

refer to breath cycle 2), sum of upward expiration and inspiration ($AR_1 + AR_2$), and the derivative $\left(\frac{dco_2}{dt}\right)$ of the complete expiratory phase (refer to Fig. 1, breath cycle 2, A-C) for the discrimination of asthma and non-asthma. Thereafter, the optimum features were selected based on the non-parametric test, receiver operating characteristics (ROC) curve. The selected features were then fed to support vector machine (SVM), k-Nearest Neighbor (k-NN) and Naive Bayes (NB) classifiers to obtain the maximum accuracy with a minimum set of features.

The work of the paper is organized as follows. Section II outlines (A) data recording, preprocessing, outlier removal and rectification, (B) automatic CO_2 waveform segmentation that includes valid breath detection and segmentation of each valid breath cycle into subcycles, (C) feature extraction, (D) feature selection, and (E) the statistical supervised classifier method, followed by a performance measurement of the proposed method presented in (F). The findings and discussion of the proposed system are presented in Section III. A conclusion is made in Section IV.

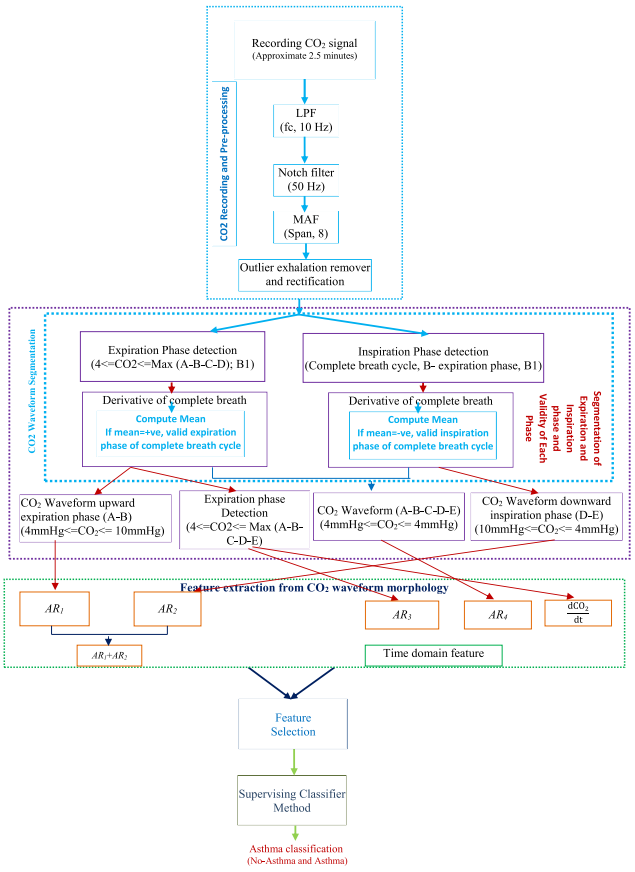

FIGURE 2. Overview of the proposed system for asthma screening based on CO_2 signal features.

II. METHODS

The proposed method for the classification of asthma and non-asthma based on the CO_2 signal's features is summarized in Fig. 2. The proposed method consists of a series of processing steps, namely, recording and preprocessing to reduce the bandwidth and smooth the CO_2signal, CO_2 waveform segmentation that includes automatic segmentation of the expiration and inspiration phase and valid breath detection steps of each breath cycle of the CO_2 signal, followed by subdivision of expiration and inspiration phases, feature extraction and selection from the CO_2 signal for discriminating the asthmatic condition, and classification of asthma using the supervising statistical classifier. Each of these steps is explained in detail in the following sections.

A. DATA RECORDING AND PREPROCESSING

The CO_2 data used in this study were collected from the Emergency Department of Penang Hospital, Penang and Pusat Kesihatan, UTM, Malaysia via a simple sampling method [3]. The data was recorded using a newly developed real-time human respiration CO_2 measurement device [20] that was designed based on sidestream technology, with a sampling rate of 100 Hz for 2.5 minutes. For this, a nasal cannula was placed into the nose of the patients and they were instructed to breathe in and out in a relaxed manner at their own comfort.

Seventy-three subjects, aged 13-84 years, participated in this study, 43 of these had an asthmatic disease that was identified and annotated by emergency physicians from the Penang Hospital and Pusat Kesihatan, UTM, Malaysia, while 30 were non-asthmatic. The inclusion criteria for the study were patients who had a complaint wheezing, increased work of breathing, cough, or shortness of breath and those suffering from asthma. Patients were excluded from the study in the case of an unclear diagnosis of asthma or severe life-threatening conditions that required immediate attention and treatment. The study protocol was approved by the Medical Research and Ethics Committee (MREC), Ministry of Health Malaysia (Ref: (13) KKM/NIHSEC/P17-1027). Additionally, in order to record their demographic information, an informed consent form was collected from all the participants before being involved in the study.

1) PREPROCESSING

Sixteen consecutive breaths with regular shapes (approximately 72 s, may vary with respiratory rate) were extracted from the two and half minutes of recorded CO_2 data from each subject, as presented in Fig. 3. A mean value was derived from the sixteen breaths for further processing. The complete data set was divided into two parts, the CO_2 signal of the non-asthmatic patient (CO2NAP) and the CO_2 signal of the asthmatic patient (CO2AP), with numbers added to the end of these codes to represent specific patients (e.g., CO2AP4 represents the data of the fourth asthmatic patient).

The CO_2 signals were filtered via a low pass filter (fc, 10 Hz) to limit the bandwidth of the signal. As reported by Yang *et al.* [21] (2010), human respiration CO_2 signals lie within 10 Hz. Thereafter, a 50 Hz notch filter was applied to remove power line interferences, followed by moving average filter (span, 8) to smooth the shape of the signal [22].

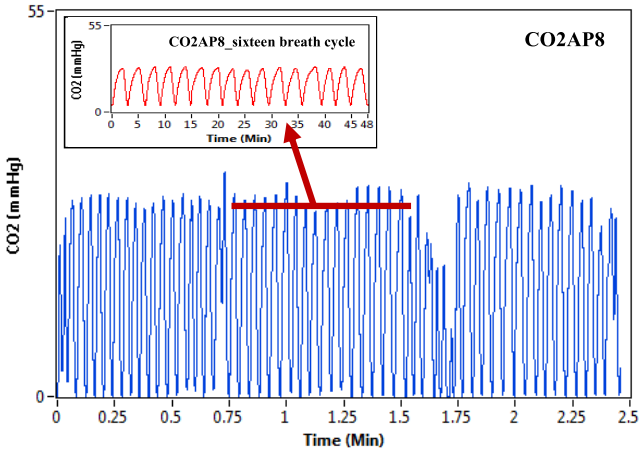

FIGURE 3. The pre-processed CO_2 signal recorded for two and half minutes from each subject, CO2AP8 indicates the 8th asthmatic patient CO_2 signal. The upper left CO_2 signals, restricted between the red lines, indicate sixteen breath cycle extracted from the whole signal (CO2AP8).

The process of the moving average filter was applied using Equation 1:

$$y(p) = \frac{1}{2p+1}\left(y(p+q)+y(p+q-1)\ldots\ldots+y(p-\ q)\right) \tag{1}$$

where y (p) indicates the result of q-smooth points, p is the adjacent data values on either side of y (p) and 2p + 1 represents span width.

2) OUTLIER REMOVAL AND RECTIFICATION

Outlier removal involves the pathologic CO_2 signal that seems to be more irregular and chaotic than the healthy CO_2 signal, which tends to be more consistent and is found more often in asthmatic patients. The outliers were removed by applying exhalation inclusion criteria that included exhibiting a positive or negative deviation from the mean expiration above a specified multiple of the standard deviation. Additionally, a feature was validated only if it did not diverge more than a specified value of the standard deviation from the record mean, as advocated by Asher *et al.* [23] (2014). In addition, a sliding window lasting for 4.5 s was created to extract the maximum and minimum CO_2 values from each breath cycle, because typically each breathing cycle ranges from 1.5 to 4.5 s for pediatric, young adults and elderly subjects [24], [25].

Furthermore, a sub-sliding window lasting between 1.5 to 4s was chosen to extract exactly the shape of the CO_2 signal for feature extraction and to avoid unnecessary CO_2waveform patterns that may be caused due to an inept expiratory valve, where the exhaled breath is reinspired. As reported by Landis and Romano [26], this occurs when subjects take breaths superseding mechanical ventilation, or when cardiogenic oscillations caused by the rhythmic increase and decrease in intrathoracic volume with each cardiac cycle. Figs. 4 A and B present the CO_2 signal recorded

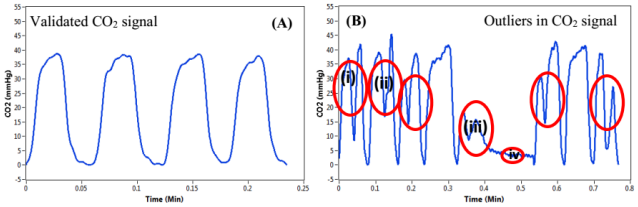

FIGURE 4. (A) validated CO_2 signal (B) Outlier CO_2 signal (red circled CO_2 signals represent i) rebreathing condition; ii) overriding mechanical ventilation during breathing; iii) cardiogenic oscillations; iv) breath-hold or short halt inspiration and then resumed breath).

from a healthy subject with red circle differentiating the validated CO_2 signal and outlier CO_2 signal, respectively. The results revealed that validated CO_2 signal appears consistent across breaths, while the outlier CO_2 signal deviates appreciably from the remaining CO_2 signal.

B. AUTOMATIC CO_2 WAVEFORM SEGMENTATION

The procedure for the segmentation of the expiration and inspiration phases, valid breath detection, and subdivision of the expiration and inspiration into subparts, from the complete breath cycle is presented in Fig. 1 (refer to breath cycle 2, A-B, A-C, D-E, and A-E). For this, the maximum CO_2 value was calculated from each breath cycle (recording time 4.5 s) by applying the simple max-min algorithm. The algorithm works based on the filtering or rule to detect the maximum CO_2 values. According to this, the maximum CO_2 value was considered as the greatest value in either of the following cases: (a) the maximum CO_2 value present during the restricted recording time of each breath cycle, or (b) the maximum CO_2 value for the previous breath if reduced by the maximum allowable breath to breath CO_2 values (1-4 mmHg) as advocated by Jaffe and Orr [27]. Further, the inspiratory phase of the breath was computed by subtracting the template of the maximum CO_2 waveform (refer breath cycle 2 of Fig. 1, A-B-C) from the complete waveform (refer breathe cycle 2 of Fig. 1, A-B-C-D-E).

Furthermore, breath validation was performed for both the exhalation and inspiration phases by computing the derivative$\left(\frac{dCO_2}{dt}\right)$ and a mean of each phase. The exhalation phase was assumed valid when the mean was found to be positive; on the contrary, a negative mean value denotes the inspiration phase as presented in Fig. 5, which shows the CO_2 signal (blue solid line) and its $\frac{dCO_2}{dt}$ (green solid line) for a healthy subject (CO2NAP4). The $\frac{dCO_2}{dt}$ of the expiration phase seems to be positive, followed by negative slope values for the inspiratory phase. Thus, valid breath detection was performed for the expiration and inspiration phases prior to the sub-division of these phases.

Thereafter, a threshold algorithm was adopted to automatically segment each phase into sub-phases in order to extract the features as suggested in an earlier study [1], [3], [17], [18], [28]. The existing methods that are either manual or criteria based include time, which may

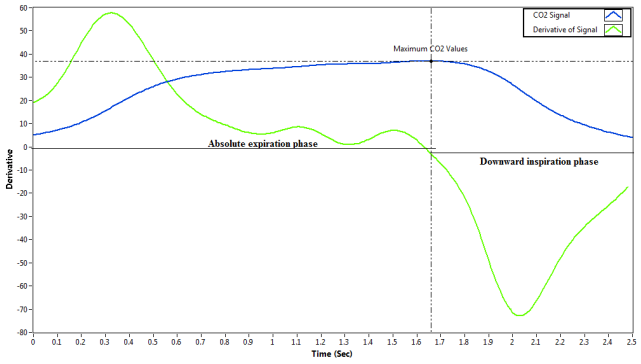

FIGURE 5. The CO_2 signal (blue solid line) and its $\left(\frac{dco_2}{dt}\right)$ (green solid line), intersection point of black dot lines represent the maximum CO_2 values, the derivative (green dash line) seems to be positive until the maximum CO_2 that represents the absolute expiration phase, followed by negative derivative (blue dot line) that reveal the downward inspiration phase.

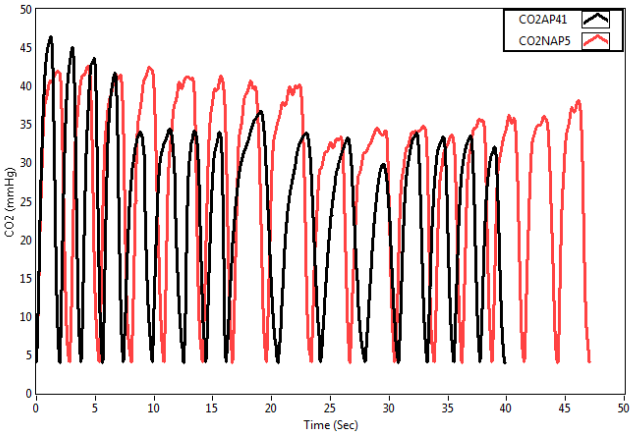

FIGURE 6. CO_2 signal of sixteen consecutive breath of non-asthmatic and asthmatic subjects, the red solid line indicates the CO2NAP5 – a CO_2 signal of the 5th non-asthmatic patient; black solid line represents the CO_2 signal of the 41th asthmatic subject.

exclude significant waveform information as reported by Howe *et al.* [3]. Hence, we consider dividing these phases using the simple threshold method by considering the y-axis (CO_2, mmHg) rather than the x-axis (time). Here, we first limited each breathing cycle to be between 4 to 4 mmHg by applying threshold method as suggested by You *et al.* [1] and Howe *et al.* [3] who state that expired CO_2 must have come from the lungs to reach a level of 4 mm Hg. Thereafter, a threshold was created lasting between 4 to 10 mmHg, which resembles part of the upward expiratory phase (A-B, refer to Fig. 2). Further, the absolute expiratory phase (A-B-C, refer to Fig. 2) was separated from the complete breath cycle by restricting the waveform until the maximum value of CO_2 as shown in Equation 2. Thereafter, the downward phase (D-E, refer to Fig. 1) was detected from the complete breath cycle by confining between 10 to 4 mmHg using the threshold method. Thereafter, features were extracted based on the segmentation of each breath cycle.

Absolute expiration phase $\gg 4$

$$<= CO_2 = \text{Max}(CO_2\text{value, A-B-C-D-E}) \quad (2)$$

C. FEATURE EXTRACTION OF CO_2 SIGNAL

The time domain features were extracted from the segmented part of the CO_2 signals in order to differentiate asthma and non-asthma conditions. The five features (AR_1, AR_2, AR_3, AR_4, and $\frac{dCO_2}{dt}$) were estimated from the CO_2 signal as illustrated in Fig. 2. AR_1, AR_2, AR_3, and AR_4 represent the AR_i of upward expiration, downward inspiration, absolute expiration, and complete breath cycle, respectively, while $\frac{dCO_2}{dt}$ indicates the derivative of the entire expiratory phase. In addition, the sum $(AR_1 + AR_2)$ was derived from AR_1 and AR_2 in order to classify the asthmatic condition. We chose to compute the AR and $\frac{dCO_2}{dt}$ because it is believed that these phases may increase or decrease during an asthma attack as presented in Fig. 6.

The AR_i of each quantified breath cycle was computed using Simpson's rule comparing with the trapezoidal because it is more accurate as it approximates the CO_2 signal with a sequence of quadratic parabolic segments instead of straight lines. Thus, the result is more accurate and may provide significant information about asthmatic changes. The procedure for the computation of the AR_i is presented in Fig. 7. Equation 3 was used to calculate the AR_i as follows:

$$A_i = \frac{dt}{6} \sum_{j=0}^{i} \left(C_{j-1}(t) + 4C_j(t) + C_{j+1}(t) \right) \quad (3)$$

Where, $C(t)$ and dt represent the CO_2 signal and sampling interval, respectively.

Furthermore, $\frac{dCO_2}{dt}$ of the absolute expiration phase was calculated using Equation 4 and the mean slope of each breath cycle was determined. Thereafter, the best-suited features were identified by applying the feature selection algorithm as reported in section D.

$$y_i = \frac{(x_i - x_{i-1})}{dt} \quad (4)$$

Where, i and dt indicate the number of CO_2 samples and sampling time between the samples, respectively.

D. FEATURE SELECTION

Feature selection is necessary in order to select the optimum features from a set of extracted features that are both pertinent and non-redundant [29]. It is believed that some of the features may consist of redundant and/or irrelevant information that may degrade the performance of the classifier if not removed [30].

The feature selection method is divided into three categories, namely, the embedded, wrapper and filter method [31], [32]. The wrapper and embedded interact with the classifier and include interaction between feature subsets and model selection. In addition, they are able to consider

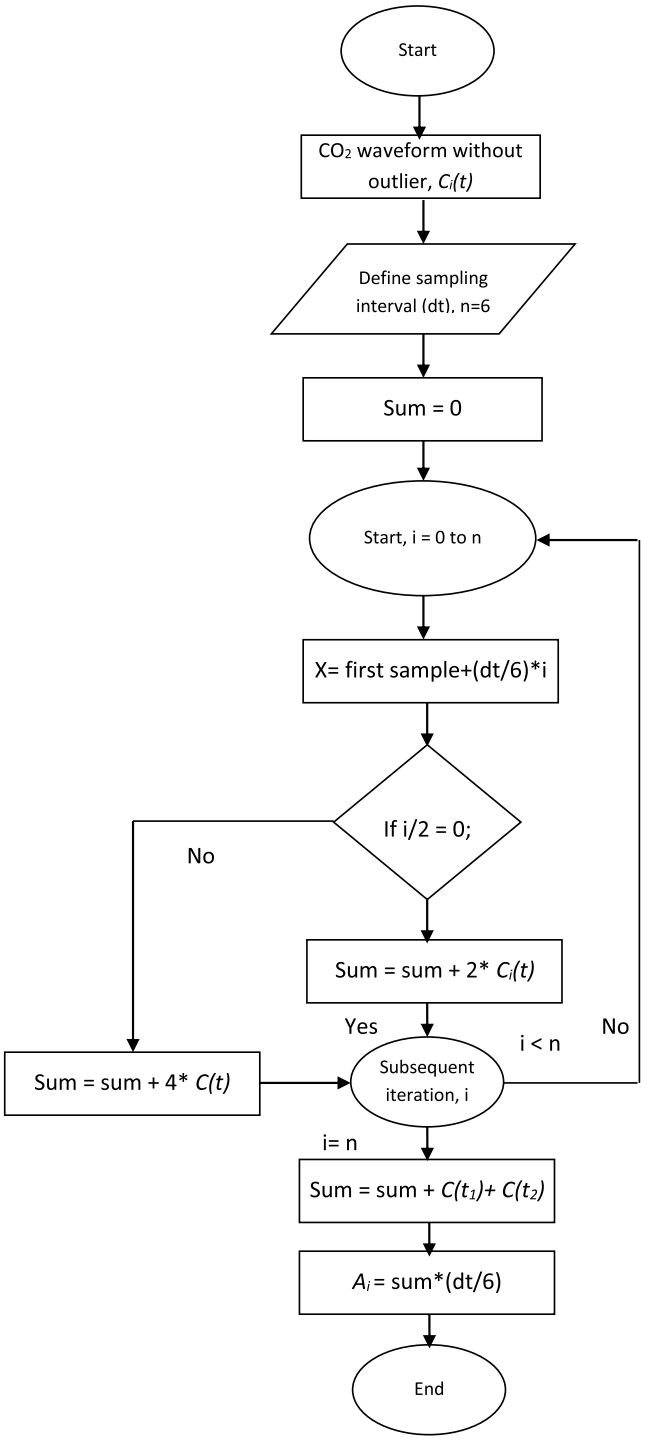

FIGURE 7. Procedure to compute the AR_i for each quantized breath cycle.

feature dependencies. However, these methods have some common limitations, such as a higher risk of overfitting more than filter techniques and being computationally expensive. In contrast to these methods, the filter technique easily scales to a large datasheet, is computationally simple, fast, and does not depend upon the classifier. Additionally, the feature selection algorithm can be performed only once, and more than one classifier method can be evaluated. Hence, we proposed to use a filter-based feature selection algorithm presented earlier in [33] to select the optimum set of CO_2 signal features for differentiating asthma and non-asthma conditions.

The filter technique evaluates the features by considering the intrinsic properties of the CO_2 signal. It is based on the score of relevant features and the features that have low scores are supposed to be removed from the machine learning. Hence, we employed a ROC test [33], which is a fraction of true positives and false positives, for a binary classification system. The ROC considered two specific groups, asthmatic and non-asthmatic, for each feature. Thus, the ROC curve coincides with the diagonal AUC (A_z, 0.5), but the features can differentiate two groups whenever the ROC curve reaches the upper left corner, i.e., is close to 1.

The feature is not considered to be statistically significant if the area under the curve (AUC, A_z) is found to be less than or equal to 0.5, which is in agreement with earlier studies [32], [34]. Usually, an A_z of 0.5, 0.7 to 0.8, 0.8 to 0.9, and more than 0.9 advocates no discrimination, acceptable, excellent, and outstanding, respectively, as suggested by Hosmer and Lemeshow [35], Ware *et al.* [36]. Thus, based on the maximum and minimum value of A_z, the features were chosen for the classification purpose.

E. STATISTICAL SUPERVISED CLASSIFIER METHOD

In this study, statistical supervised classifiers were used to categorize asthma and non-asthma conditions based on the AR_i and $\frac{dCO_2}{dt}$. The approach incorporates a set of d-dimensional feature vectors as an input to a statistical supervised classifier. Here, we chose to integrate the SVM classifier based on the radial basis function (RBF) for the classification of asthma and non-asthma [5], [37]. This approach is considered to be a good classifier that has been used in bioinformatics, engineering, computer vision, and other fields. It uses an optimum hyperplane to differentiate the asthmatic and non-asthmatic groups by maximizing the distance from the decision boundaries [38]–[40]. It can also tackle simple, linear and nonlinear classification tasks, as well as more complex tasks. It enables mapping of the input data point from the input space to a high dimensional feature space where the classification problem can be easily simplified. The SVM is executed using the C++ library known as LIBSVM [41], and the parameters of the RBF kernel function (kernel width, γ, and penalty constant, C) are optimized by a grid-search to achieve a maximum result [42]. The process of the grid-search algorithm can be found elsewhere [42]. Furthermore, for the sake of comparison, the k-NN with k = 1, 2, and 3 and NB classifiers were also applied [6], [40].

The performance of the classifiers was evaluated using leave-one-out (LOO) cross-validation [43] rather than k-fold because it allowed all possible partitions to be explicitly tested. This approach includes data from the original data as test data and the leftover data is considered as training data. The procedure is performed by dividing the P dataset into

k mutually exclusive subsets $P_1, P_2, \ldots, P_k$. The classifier is trained and tested k-times; each time $n = 1, \ldots, k$, it is trained on P/P_k, and tested on P_k. Although the training and testing sets are taken from the same datasets, cross-validation is not considered to produce biased results [44], [45]. In addition, LOO is assumed to be statistically efficient for small data sizes and hence it is applied here as it suits these circumstances [43].

F. PERFORMANCE MEASUREMENT

The performance of the classifier was measured in terms of sensitivity, specificity, precision, and accuracy using the confusion matrices [46]. The sensitivity evaluates if the asthmatic CO_2 signal data is correctly identified by the classifier, while specificity is the proportion of the non-asthmatic CO_2 signal correctly recognized by the classifier. The precision provides information about the consistency of classifier's sensitivity and specificity when repeated multiple times, whereas the accuracy compares the actual result to what it supposed to be [46]. The sensitivity, specificity, precision, and accuracy are expressed mathematically as follows:

$$\text{Sensitivy (SEN)} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \tag{5}$$

$$\text{Specificity (SPE)} = \frac{(\text{TN})}{(\text{TN} + \text{FP})} \tag{6}$$

$$\text{Precision (Pre)} = \frac{(\text{TP})}{(\text{TP} + \text{FP})} \tag{7}$$

$$\text{Accuracy (ACC)} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{8}$$

Where, TP, TN, FP, and FN indicate the true positive, true negative, false positive and false negative, respectively. In addition, the error rate was calculated using the following expression:

$$\text{E}_\text{r} = 1 - \frac{(\text{Sen} + \text{Spc})}{2} \tag{9}$$

III. RESULTS AND DISCUSSION

Here, we have proposed an approach that quantitatively analyzes respired CO_2 waveforms for classification of asthma and non-asthma using a new prognostic index (AR_i) and $\frac{dco_2}{dt}$, which can possibly be implemented in real time due to the ease and simplicity of the algorithm. A total of 73 sets of CO_2 data were collected, including data from 43 asthmatics with a mean+standard deviation (SD) of 31.64+14.60, and 30 non-asthmatics with mean+SD (24.86+5.25). Further, five shape features (AR_1, AR_2, AR_3, AR_4, $\frac{dco_2}{dt}$) were extracted from sixteen-consecutive regular-shaped CO_2 waveforms for each subject. Additionally, the feature ($AR_1 + AR_2$) was derived from AR_1 and AR_2. The mean and SD of the extracted features are provided in Table 1.

Table 1 shows that the areas (AR_1, AR_2, AR_3, AR_4, and $AR_1 + AR_2$) for the asthmatic CO_2 signal possess a higher mean value than the non-asthmatic CO_2 signal, while the $\frac{dco_2}{dt}$ of the expired phase is deceased is asthmatic patients

TABLE 1. Mean and standard deviation (SD) of the estimated features for the CO2NAP (30) and CO2AP (43) CO_2 signals.

Feature	Mean±SD (CO2NAP)	Mean±SD (CO2AP)
AR_1	0.84±0.16	1.56± 0.96
AR_2	1.73 ± 0.35	2.22±0.67
AR_3	77.19±20.16	88.55±52.47
AR_4	52.36±15.65	61.65±41.28
AR_1+AR_2	2.57±0.45	3.79±1.37
$\frac{dCO_2}{dt}$	19.81±3.43	15.66±8.70

AR_1, AR_2, AR_3, and AR_4: Area (AR_i) of the upward expiratory, downward inspiratory, absolute expiratory and complete breath cycle, respectively. AR_1+AR_2 represent the sum of the area of upward expiratory and downward inspiratory phase, respectively and $\frac{dCO_2}{dt}$ indicate the derivative of whole expiratory phase. CO2NAP and CO2AP represent the CO_2 signal for non-asthmatic and asthmatic subjects.

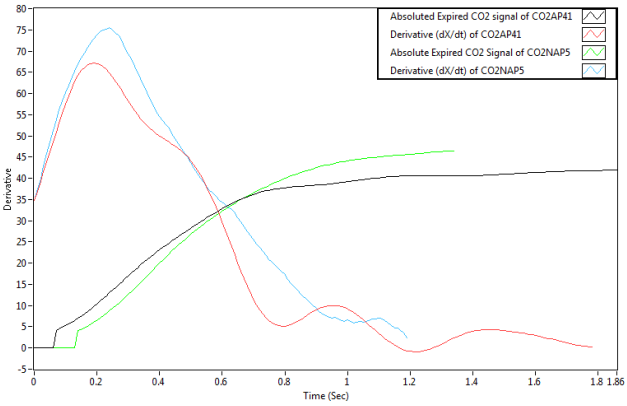

FIGURE 8. Absolute expiration phase of the asthmatic patient (black solid line) and its $\frac{dCO_2}{dt}$ (red solid line); complete expiration phase of the non-asthmatic patient (green solid line) and its derivative (light blue solid line); the slope of non-asthmatic possess higher peak than asthmatic.

compared to non-asthmatic patients. Further, the $\frac{dCO_2}{dt}$ of the absolute expiratory phase revealed that the mean value of the slope for the asthmatic group was lower than that of the non-asthmatic group. Fig. 8 depicts the derivative of the absolute expiration phase for the asthmatic and non-asthmatic patients.

Table 2 lists the p-values, area (A_z) under the ROC curve and standard error (SE) values that aid in ranking the features, which can be used as feature vectors for the classifier. The p-values reveal that all the features were statistically significant ($p < 0.01$) for the comparison of asthma and non-asthma groups. However, the A_z and corresponding SE values indicate that AR_1, AR_2, and $AR_1 + AR_2$ and $\frac{dCO_2}{dt}$ were the strongest indices among all those whose mean values were significantly different between asthmatic and non-asthmatic. The A_z and corresponding SEvalues indicate that the AR_i of the upward expiratory phase ($AR_1, A_z = 0.89$ with SE $= 0.04$), downward inspiratory phase ($AR_2, A_z = 0.80$ with SE $= 0.06$), sum of upward expiratory and downward inspiratory phase ($AR_1 + AR_2, A_z = 0.79$ with SE $= 0.06$), and derivative ($\frac{dCO_2}{dt}, A_z = 0.71$ with SE $= 0.06$) are more

TABLE 2. Statistical significance of distinguishable A_z, p-value, and standard error (SE) of the features used for the classification of the asthmatic and non-asthmatic condition.

Feature	A_z -values	p-value	SE
AR_1	0.89	1.40E-8	0.04
AR_2	0.80	1.2E-5	0.06
AR_3	0.68	8E-3	0.06
AR_4	0.64	3.8E-3	0.07
AR_1+AR_2	0.79	2.3E-5	0.06
$\frac{dCO_2}{dt}$	0.71	2E-3	0.06

A_z – Area under the receiver operating characteristic curve. AR_1, AR_2, AR_3, and AR_4: Area (AR_i) of the upward expiratory, downward inspiratory, absolute expiratory and complete breath cycle, respectively and $\frac{dCO_2}{dt}$ indicates the derivative of whole expiratory phase. AR_1+AR_2 represent the sum of the area of upward expiratory and downward inspiratory phase, respectively.

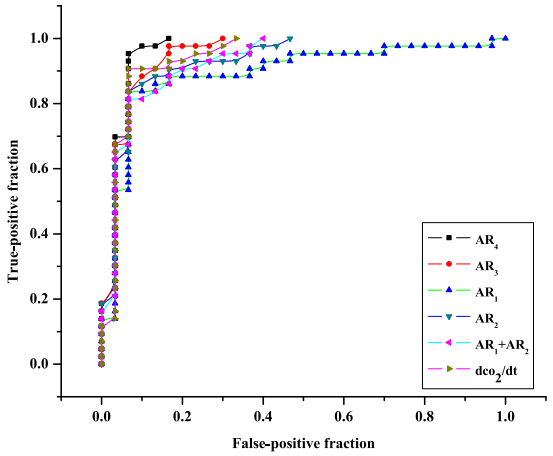

FIGURE 9. Illustration of ROC curves to identify the best-suited features for differentiating asthma and non-asthma conditions. The up triangle (AR_1) possesses the higher sensitivity and specificity, followed by down triangle (AR_2), left triangle ($AR_1 + AR_2$), right triangle $\frac{dCO_2}{dt}$, circle (AR_3), and square (AR_4).

highly statistically considerable than the rest of the features for the discrimination asthma and non-asthma conditions.

However, it should be noted that the rest of the features also exhibit the good A_z values, which can be used for the asthma classification. Thus, we have combined the features based on the A_z and p-value as feature vectors for the SVM, KNN, and NB classifier to maximize the accuracy of the discrimination of asthma and non-asthma. Fig. 9 depicts the feature capabilities based on their corresponding A_z values.

Further, all six were ranked and sets of three and four features were provided to the SVM, KNN and NB classifier via the LOO procedure for the automatic classification of asthma and non-asthma. The set of features AR_1, AR_2, $AR_1 + AR_2$ and AR_1, AR_2, $AR_1 + AR_2$, $\frac{dCO_2}{dt}$ exhibited the maximum accuracy $= 94.15\%$ and minimum error rate $= 5.47$, with an SVM RBF kernel function- width (γ, 0.1) and penalty constant (C, 1000) as presented in Fig. 10. Table 3 illustrates

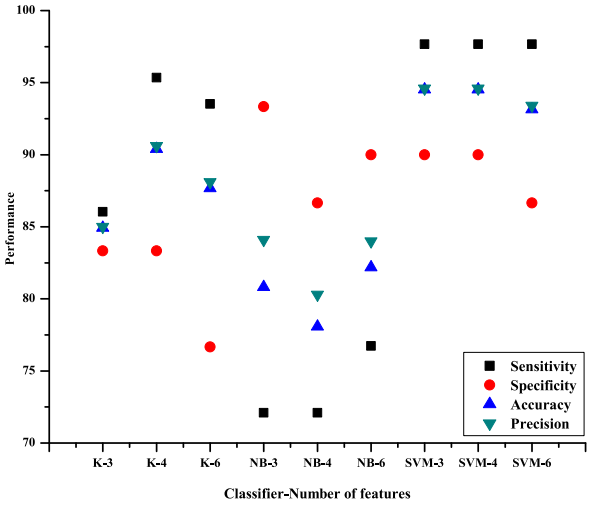

FIGURE 10. A plot of average accuracy (%), precision (%), sensitivity (%) and specificity (%) versus different classifiers with sets of the feature. SVM, K, and NB represent the support vector machine, k-NN, and Naive Bayes classifier, whereas 3, 4 and 6 indicate the set of features AR_1, AR_2, $AR_1 + AR_2$; AR_1, AR_2, $AR_1 + AR_2$, $\frac{dCO_2}{dt}$; and AR_1, AR_2, AR_3, AR_4, $AR_1 + AR_2$, $\frac{dCO_2}{dt}$, respectively. Up-ward triangle, downward triangle, rectangle, and circle elucidate accuracy, precision, sensitivity, and specificity, respectively.

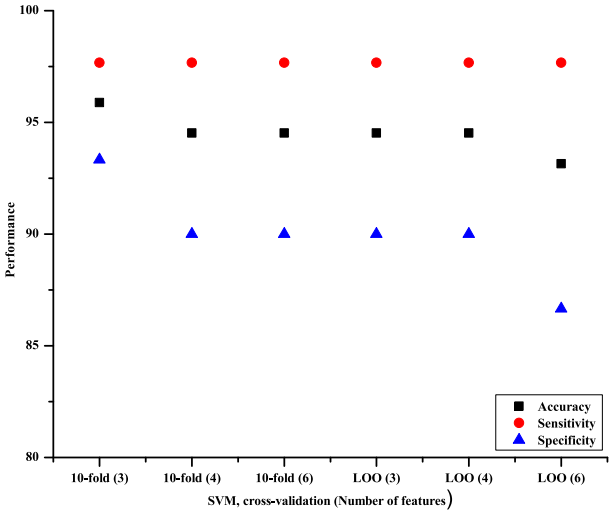

FIGURE 11. A plot of average accuracy (%), sensitivity (%) and specificity (%) versus different cross-validation (10-fold and LOO) for SVM classifier. 10-fold and LOO stand for k-fold and leave-one-out cross-validation while 3, 4 and 6 indicate the set of features AR_1, AR_2, $AR_1 + AR_2$; AR_1, AR_2, $AR_1 + AR_2$, $\frac{dCO_2}{dt}$; and AR_1, AR_2, AR_3, AR_4, $AR_1 + AR_2$, $\frac{dCO_2}{dt}$, respectively. Black rectangle red circle and upward triangle represent the accuracy and error of the classifier with 10-fold and LOO.

that the SVM classifier possesses higher sensitivity (97.67%) and specificity (90.0%) with an A_z of 0.938 than KNN and NB. Fig. 11 shows the plot of the average accuracy (%), sensitivity (%) and specificity (%) for the different sets of features using 10-fold and LOO cross-validation for the SVM classifier. The plot reveals that there is the slight change in the performance of the SVM using the 10-fold and LOO. The 10-folds cross-validation possesses slightly higher accuracy than the LOO. Hence, we chose to use LOO over the

TABLE 3. Average results obtained from the SVM, KNN and NB classifiers using AR_1, AR_2, AR_3, AR_4, $AR_1 + AR_2$ and $\frac{dCO_2}{dt}$.

Classifier	Features	Accuracy (%)	Error rate	Sensitivity (%)	Specificity (%)	Precision (%)	ROC
SVM/LOO	3	94.52	5.47	97.67	90.00	94.60	0.938
	4	94.52	5.47	97.67	90.00	94.60	0.938
	6	93.15	6.84	97.67	86.66	93.40	0.922
KNN/LOO	3	84.93	15.06	86.04	83.33	85.00	0.849
	4	90.41	9.589	95.34	83.33	90.60	0.895
	6	87.67	12.32	93.53	76.67	88.11	0.861
NB/LOO	3	80.82	19.17	72.09	93.33	84.10	0.82
	4	78.08	21.97	72.09	86.66	80.30	0.806
	6	82.19	17.80	76.74	90.00	84.00	0.834

SVM- Support vector machine, LOO – Leave-one-out, KNN - K-Nearest Neighbor, NB - Naive Bayes whereas 3, 4 and 6 indicate the set of features AR_1, AR_2, AR_1+AR_2; AR_1, AR_2, AR_1+AR_2, $\left(\frac{dco_2}{dt}\right)$; and AR_1, AR_2, AR_3, AR_4, AR_1+AR_2, $\left(\frac{dco_2}{dt}\right)$ respectively. AR_1- upward expiratory phase, AR_2 – downward inspiratory phase, AR_3 – Absolute expiratory phase, AR_4 – complete breath cycle, and $AR_1 + AR_2$ – Sum of expiratory and inspiratory phase, $\left(\frac{dco_2}{dt}\right)$- a derivative of the absolute expiratory breath.

10-fold cross-validation as it offers less bias and is efficient with smaller data sets.

In this work, a new indices-based method is proposed for the classification of asthma, which may possibly be implemented in real time due to ease and simplicity of the features computation procedure. Table 4 illustrates a comparison of the existing and proposed work aimed at discriminating asthma by extracting features from the shape of the CO_2 waveform. All the features extracted from our study exhibit good discriminatory capabilities for asthma and non-asthma, as presented in Table 2. The features AR_1 (upward expiratory phase), AR_2 (downward inspiratory phase), $AR_1 + AR_2$ (sum of the area of upward expiratory and downward inspiratory phases) and $\frac{dCO_2}{dt}$ possess higher A_z values (ranges: 0.714-0.892) and low standard error (ranges: 0.063-0.042) and hence were chosen as feature vectors for the classifier, in order to maximize the accuracy. These features were extracted using a simple threshold and uncomplicated mathematical Equation (refer to Equations 2, 3 and 4), contrary to setting criteria [1]–[3], [17], [18], which involve higher and more complicated mathematics, [5], [6] keeping in mind that this concept may possibly be easily integrated into real-time CO_2 measurement devices in the future.

The recording duration of the CO_2 signal for each breath was 1.5 to 4.5 s in order to extract the maximum and minimum CO_2 values, as the respiratory rate was found to be slower with elderly asthmatic subjects [23]. The details of the implemented algorithm for the computation of maximum and minimum CO_2 values can be found in section B and are in agreement with an earlier study [26].

We also incorporated an algorithm for the removal of outliers using a statistical method (i.e., mean and SD), which in agreement with earlier studies [11] reporting asthmatic data. Furthermore, the breath size was limited between 1.5 and 4 s for the extraction of the features as we analyzed both the expiration and inspiration phase. The reason to select this duration is that the studies conducted by You *et al.* (1994) considered absolute expired phase to be significant, which lasts between 0.8 s and 3 s, and Yaron et al. (1996) assumed that a valid expired breath ranged from 1.5 to 2.5 s. Hence, we considered each breath cycle valid with the duration of 1.5 s to 4 s.

Further, we found that AR_1 and AR_2 were increased independently for the asthmatic patient as per our hypothesis (Table 1). The mean values for the sixteen consecutive breaths of AR_1 and AR_2were 1.56 and 2.22, respectively, for the asthmatic patients compared with healthy individuals (AR_1, 0.84 and AR_2, 1.73). This contradicts the similar studies conducted by You *et al.* [1] and Kean *et al.* [12] However, they computed the area from the half of the end-tidal peak value above a threshold by limiting the x-axis (time, 0.25 s) compared to our proposed method (y-axis, limiting between 4 and 10 mmHg), which revealed that the area decreases with asthmatic subjects. The previously used method may possibly miss significant CO_2 values, whereas the proposed Simpson's rule is considered to be fairly accurate as it applies a sequence of quadratic parabolic segments to determine the AR_i. Furthermore, the derivative of absolute expiratory phase was found to be lower for the asthmatic group, which agrees with You *et al.* [1], where they measured the slope from 0 to 0.2 s using linear regression.

Further, the extracted features were sent to the different classifiers for the classification of asthma and non-asthma. Table 3 illustrates that the SVM classifier obtained higher accuracy, sensitivity, and specificity compared to k-NN and NB. In addition, the asthma discrimination capability of

TABLE 4. Studies carried out for the differentiation of asthmatic and non-asthmatic conditions using CO_2 signal features.

Authors	Methods	Features	Number of features/ discrimination method/comparison/classifier	Performance
B. You *et al.* [17]	Time domain (manual)	ETS	One/mean value/correlation/nil	Demonstrated a good correlation with the proposed index for differentiating asthma and non-asthma
B. You *et al.* [1]	Time domain (setting criteria /manual)	S1, S2, S3, SR, AR, SD1, SD2 and SD3	Eight/mean value/correlation/nil	All features were statistically significant; however, angle α between the ascending phase and an alveolar plateau was strongly correlated.
M. Yaron *et al.* [2]	Time domain (manual)	$\left(\frac{dco_2}{dt}\right)$	One/mean/ correlation/nil	The derivative of the alveolar plateau phase was steeper with an asthmatic patient.
N.N. Hisamuddin *et al.* [18]	Time domain (computerized)	A slope of upward expiratory and alveolar phase, α-angle	Three/mean/ correlation/nil	The slope of the alveolar phase and α-angle were increased with the asthmatic patient and shown good correlation.
T.A. Howe *et al.* [3]	Time domain (computerized)	A slope of upward expiratory and alveolar phase, α-angle	Three/mean/ correlation/nil	Two features (slope of plateau and α-angle) increased with asthma
M. Kazemi *et al.* [6]	Frequency domain (computerized)	PSD, FFT	Five/nil/nil/RBF	Asthma prognostic index, Accuracy:95.65%; error rate:4.34%
J.P. Betancourt *et al.* [5]	Time-frequency domain (computerized)	Wavelet coefficient	8/nil/nil/SVM	Asthma prognostic index, Accuracy:96.52%; error rate:3.48%;sensitivity:100%; specificity:88.57%; execution time:0.3 feature estimation time- 0.35s
S.A. Malik *et al.* [47]	Time domain (computerized)	CO_2 concentration	1/mean/nil/Turns count	Early screening of asthma, Sensitivity:80%; specificity:90%
T.T. Kean *et al.* [12]	Time domain	You [1] and Hjorth Parameter (activity)	13/nil/nil/t-test /ROC	SR:p <0.0001, A_z-0.92, activity:p < 0.0001, A_z-0.90
Proposed work	Time domain (automatic computerized method)	AR_1, AR_2, $AR_1 + AR_2$,	3/mean/t-test/ROC/SVM	Accuracy:94.52%; error rate: 5.47%; sensitivity:97.67%; specificity:90.0%; execution time:0.03s average feature estimation time-0.24s

SVM is found to be outstanding [35] with an A_z of 0.938, compared to k-NN (0.895), and NB (0.834) for higher ranked features (refer to Table 2).

It can also be noticed that although the k-NN classifier exhibited an accuracy of 90% with a set of 4 features, its sensitivity (95.34%), specificity (83.33%), and A_z values were lower than the proposed machine learning method (SVM). In addition, the performance of the proposed method in terms of accuracy (94.52%), and sensitivity (97.67%) are close to the previous studies, reported by Betancourt *et al.* [5] and Kazemi *et al.* [6]. However, the specificity (90.0%) of the proposed method is slightly higher than that in the previous study [5]. Furthermore, the computation and execution time of each feature and the proposed classifier were 0.24 s and 0.03 s, respectively, which is lower in comparison with an earlier study [5], [48]. These qualities provide a platform for the implementation of the proposed methods into a low-cost processor and/or microcontroller.

This study presents a relatively simple signal processing algorithm-based method that may lead to implementation in the real-time environment to help the physicians screen for asthma. Here, for the first time, we developed a non-invasive and patient-independent method to automatically quantitate and analyze CO_2 waveform shapes, which has the potential to be implemented into real-time CO_2 measurement devices for automatic classification of asthmatic conditions. In the future, this approach should be extended to enable tracking of changes in asthma severity level and response to treatment over time. In addition, the proposed approach may be incorporated into a real-time human respiration CO_2 measurement device for asthma classification, which is in progress at the BSP research group.

IV. CONCLUSION

Asthma is a major noncommunicable and preventable disease that characterized by inflammation or swelling of smaller airways (bronchioles) of the lung in response to different stimuli. Early screening and continued treatment of asthma may possibly control the morbidity and mortality rate. This study shows how the shape of the CO_2 signal changes in the asthmatic condition; specifically, an area of upward expiration (AR_1), downward inspiration (AR_2), the sum of both phases

and the derivative ($\frac{dCO_2}{dt}$) significantly increase or decrease. Therefore, we can classify individuals with and without asthma by computing the area and derivative of these phases. This work reports an automatic asthma classification system that consists of estimating the AR_i from the respired CO_2 waveform. Using the proposed method, we have achieved an accuracy of 94.52%, sensitivity of 97.67%, and specificity of 90.0% by utilizing the SVM classifier via the LOO procedure. In addition, the average feature estimation and execution times were 0.24 s and 0.03 s, respectively. Thus, the proposed method can be implemented in a real-time environment or can be used to develop a CAD system, which can help physicians screen for asthmatic conditions. Further, the developed approach will be integrated into a real-time human respiration CO_2 measurement device for asthma classification, which is under development by the BSP research group in coordination with physicians.

DISCLOSURES

The author has no conflict of Interest.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Teo Aik Howe and Dr. Ismail Bin Ahmed for support for data collection. Special thanks to the Universiti Teknologi Malaysia for providing the facilities and laboratory equipment for the completion of the research.

REFERENCES

[1] B You, R. Peslin, C. Duvivier, V. D. Vu, and J. P. Grilliat, "Expiratory capnography in asthma: Evaluation of various shape indices," *Eur. Respiratory J.*, vol. 7, no. 2, pp. 318–323, Feb. 1994, doi: 10.1183/09031936.94.07020318.

[2] M. Yaron, P. Padyk, M. Hutsinpiller, and C. B. Cairns, "Utility of the expiratory capnogram in the assessment of bronchospasm," *Ann. Emergency Med.*, vol. 28, no. 4, pp. 403–407, Oct. 1996, doi: 10.1016/S0196-0644(96)70005-7.

[3] T. A. Howe, K. Jaalam, R. Ahmad, C. K. Sheng, and N. H. N. A. Rahman, "The use of end-tidal capnography to monitor non-intubated patients presenting with acute exacerbation of asthma in the emergency department," *J. Emergency Med.*, vol. 41, no. 6, pp. 581–589, Dec. 2011, doi: 10.1016/j.jemermed.2008.10.017.

[4] C. L. Herry, D. Townsend, G. C. Green, A. Bravi, and A. J. E. Seely, "Segmentation and classification of capnograms: Application in respiratory variability analysis," *Physiol. Meas.*, vol. 35, no. 12, p. 2343, Nov. 2014, doi: 10.1088/0967-3334/35/12/2343.

[5] J. P. Betancourt *et al.*, "Segmented wavelet decomposition for capnogram feature extraction in asthma classification," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 18, no. 4, pp. 480–488, Apr. 2014, doi: 10.20965/jaciii.2014.

[6] M. Kazemi, M. B. Krishnan, and T. A. Howe, "Frequency analysis of capnogram signals to differentiate asthmatic and non-asthmatic conditions using radial basis function neural networks," *Indian J. Allergy, Asthma Immunol.*, vol. 12, no. 3, pp. 236–246, 2013.

[7] M. B. Jaffe, "Using the features of the time and volumetric capnogram for classification and prediction," *J. Clin. Monit.*, vol. 31, no. 1, pp. 19–41, Feb. 2017, doi: 10.1007/s10877-016-9830-z.

[8] D.-B. Shieh and C.-H. Lin, "Portable asthma detection device and stand-alone portable asthma detection device," U.S. Patent 13 290 238, Jul. 11, 2011.

[9] S. E. Weinberger, B. A. Cockrill, and J. Mandel, *Principles of Pulmonary Medicine*. Amsterdam, The Netherlands: Elsevier, 2008.

[10] S. Lamba *et al.*, "Initial out-of-hospital end-tidal carbon dioxide measurements in adult asthmatic patients," *Ann. Emergency Med.*, vol. 54, no. 3, p. S51, Sep. 2009, doi: 10.1016/j.annemergmed.2009.06.193.

[11] J. Acker *et al.*, "End-tidal gas monitoring apparatus," U.S. Patent 14 129 760, Jun. 27, 2012.

[12] T. T. Kean and M. B. Malarvili, "Analysis of capnography for asthmatic patient," in *Proc. IEEE Int. Conf. Signal Image Process. Appl.*, Nov. 2009, pp. 464–467, doi: 10.1109/ICSIPA.2009.5478699.

[13] J. G. Goepp, "Effort-independent, portable, user-operated capnograph devices and related methods," U.S. Patent 12 265 574, Nov. 5, 2008.

[14] L. H. Greve, "Unequal ventilation," M.S. thesis, Univ. Medical School, Utrecht, The Netherlands.

[15] A. Berengo and A. Cutillo, "Single-breath analysis of carbon dioxide concentration records," *J. Appl. Physiol.*, vol. 16, no. 3, pp. 522–530, May 1961, doi: 10.1152/jappl.1961.16.3.522.

[16] B. Smalhout and Z. Kalenda, *An Atlas of Capnography*. Zeist, The Netherlands: Kasteel Kerckebosch, 1975.

[17] B. You, D. Mayeux, B. Rkiek, N. Autran, V. V. Dang, and J. P. Grilliat, "Expiratory capnography in asthma. Perspectives in the use and monitoring in children," *Revue Maladies Respiratoires*, vol. 9, no. 5, pp. 547–552, 1992.

[18] N. N. Hisamuddin *et al.*, "Correlations between capnographic waveforms and peak flow meter measurement in emergency department management of asthma," *Int. J. Emergency Med.*, vol. 2, no. 2, pp. 83–89, Jun. 2009, doi: 10.1007/s12245-009-0088-9.

[19] M. L. Langhan, M. R. Zonfrillo, and D. M. Spiro, "Quantitative end-tidal carbon dioxide in acute exacerbations of asthma," *J. Pediatrics*, vol. 152, no. 6, pp. 829–832, Jun. 2008, doi: 10.1016/j.jpeds.2007.11.032.

[20] O. P. Singh, T. A. Howe, and M. B. Balakrishnan, "Real-time human respiration carbon dioxide measurement device for cardiorespiratory assessment," *J. Breath Res.*, vol. 12, no. 2, pp. 1–13, Jan. 2018, doi: 10.1088/1752-7163/aa8dbd.

[21] J. Yang, K. An, B. Wang, and L. Wang, "New mainstream double-end carbon dioxide capnograph for human respiration," *J. Biomed. Opt.*, vol. 15, no. 6, p. 065007, Nov. 2010, doi: 10.1117/1.3523620.

[22] M. Kirkko-Jaakkola, J. Collin, and J. Takala, "Bias prediction for MEMS gyroscopes," *IEEE Sensors J.*, vol. 12, no. 6, pp. 2157–2163, Jun. 2012, doi: 10.1109/JSEN.2012.2185692.

[23] R. J. Asher, T. Heldt, B. S. Krauss, and G. C. Verghese, "Systems and methods for quantitative capnogram analysis," U.S. Patent 13 849 284, Apr. 22, 2014.

[24] S. L. DeBoer, *Emergency Newborn Care*. Victoria, BC, Canada: Trafford, 2004, p. 30.

[25] W. Q. Lindh, M. Pooler, C. D. Tamparo, B. M. Dahl, and J. Morris, *Delmar's Comprehensive Medical Assisting: Administrative and Clinical Competencies*. New York, NY, USA: Cengage, 2009, p. 573.

[26] B. Landis and P. M. Romano, "A scoring system for capnogram biofeedback: Preliminary findings," *Appl. Psychophysiol. Biofeedback*, vol. 23, no. 2, pp. 75–91, Jun. 1998, doi: 1090-0586/98/0600-0075.

[27] M. B. Jaffe and J. A. Orr, "End-tidal gas estimation system and method," U.S. Patent 8 166 971 B2, May 1, 2012.

[28] J. Corbo, P. Bijur, M. Lahn, and E. J. Gallagher, "Concordance between capnography and arterial blood gas measurements of carbon dioxide in acute asthma," *Ann. Emergency Med.*, vol. 46, no. 4, pp. 323–327, Oct. 2005, doi: 10.1016/j.annemergmed.2004.12.005.

[29] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, nos. 1–4, pp. 131–156, Jan. 1997.

[30] M. B. Malarvili, M. Mesbah, and B. Boashash, "HRV feature selection for neonatal seizure detection: A wrapper approach," in *Proc. IEEE Int. Conf. Signal Process. Commun.*, Nov. 2007, pp. 864–867, doi: 10.1109/ICSPC.2007.4728456.

[31] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, Oct. 2007, doi: 10.1093/bioinformatics/btm344.

[32] C. Lazar *et al.*, "A survey on filter techniques for feature selection in gene expression microarray analysis," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 9, no. 4, pp. 1106–1119, Jul. 2012, doi: 10.1109/TCBB.2012.33.

[33] R. M. Rangayyan and Y. Wu, "Analysis of vibroarthrographic signals with features related to signal variability and radial-basis functions," *Ann. Biomed. Eng.*, vol. 37, no. 1, pp. 156–163, Jan. 2009, doi: 10.1007/s10439-008-9601-1.

[34] J. N. Mandrekar, "Receiver operating characteristic curve in diagnostic test assessment," *J. Thoracic Oncol.*, vol. 5, no. 9, pp. 1315–1316, Sep. 2010, doi: 10.1097/JTO.0b013e3181ec173d.

[35] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, 2nd ed. New York, NY, USA: Wiley, 2000, ch. 5, pp. 160–164, doi: 10.1002/0471722146.

[36] C. B. John and C. H. David, *P values, Medical Uses of Statistics*, Arnold S. Relman, Ed. Hoboken, NJ, USA: Wiley, 2012, pp. 181–200.
[37] K. Soman, A. Sathiya, and N. Suganthi, "Classification of stress of automobile drivers using radial basis function kernel support vector machine," in *Proc. Int. Conf. Inf. Commun. Embedded Syst.*, Feb. 2014, pp. 1–5, doi: 10.1109/ICICES.2014.7034000.
[38] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York, NY, USA: Wiley, 1973.
[39] H. Xue, S. C. Chen, and Q. Yang, "Structural regularized support vector machine: A framework for structural large margin classifier," *IEEE Trans. Neural Netw.*, vol. 22, no. 4, pp. 573–587, Apr. 2011, doi: 10.1109/TNN.2011.2108315.
[40] U. R. Acharya *et al.*, "Computer-aided diagnosis of diabetic subjects by heart rate variability signals using discrete wavelet transform method," *Knowl.-Based Syst.*, vol. 81, pp. 56–64, Jun. 2015, doi: 10.1016/j.knosys.2015.02.005.
[41] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011, doi: 10.1145/1961189.1961199.
[42] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Dept. Comput. Sci., Nat. Taiwan Univ., Taipei, Taiwan, 2003, pp. 1–16.
[43] P. A. Devijver and I. Kittler, *Pattern Recognition: A Statistical Approach*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1982.
[44] B. Efron and R. Tibshirani, "Improvements on cross-validation: The 632+ bootstrap method," *J. Amer. Stat. Assoc.*, vol. 92, no. 438, pp. 548–560, Jun. 1997.
[45] R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2009.
[46] P. Armitage and T. Colton, *The Encyclopedia of Biostatistics*. New York, NY, USA: Wiley, 1998.
[47] S. A. Malik, O. P. Singh, and M. Balakrishnan, "Portable respiratory CO_2 monitoring device for early screening of asthma," in *Proc. ACEC*, 2016, pp. 90–94, doi: 10.15224/978-1-63248-113-9-61.
[48] J. P. Betancourt, "Intelligent methods for capnogram feature extraction applied to asthma classification," M.S. thesis, Dept. Comput. Intell. Syst. Sci., Interdiscipl. Graduate School Sci. Eng., Tokyo Inst. Technol., Tokyo, Japan.

OM PRAKASH SINGH received the bachelor's degree in science (physics, chemistry, and math), and the master's degree in physics and biomedical engineering. He is currently pursuing the Ph.D. degree in biomedical engineering with the School of Biomedical Engineering and Health Sciences, Universiti Teknologi Malaysia, Malaysia. In addition, he has obtained six years of teaching and three years of research experience. He has authored and co-authored for around 14 research and review manuscripts. His research interest includes designing of handy medical device using optical sensor, bio-signal processing, and machine learning.

RAMASWAMY (PALANI) PALANIAPPAN (SM'08) received the degree and M.Eng.Sc. degree in electrical engineering and the Ph.D. degree in microelectronics/ biomedical engineering in 1997, 1999, and 2002, respectively. He is currently a Reader with the School of Computing, University of Kent and heads the Data Science Research Group. He has supervised over half a dozen postgraduate students to completion and has 15 years of multi-disciplinary teaching experience in computer science and engineering (electrical and biomedical) disciplines. He has written two text books in engineering and published over 150 papers (citation indices: h-index of 23 and i-10 index of 45) in peer-reviewed journals, book chapters, and conference proceedings. His research interests include biological signal processing, brain-computer interfaces, biometrics, neural-networks, genetic-algorithms, and image processing. He is a member of the Institution of Engineering and Technology. He also serves as an editorial board member for several international journals and serves in EPSRC and MRC Peer Review Colleges and many other international grant funding bodies.

MB MALARVILI received the bachelor's and master's degrees in biomedical signal processing from Universiti Teknologi Malaysia (UTM), Malaysia and the Ph.D. degree in medical sciences engineering from the University of Queensland, Brisbane, Australia, in 2008. She is currently a Senior Lecturer with the School of Biomedical Engineering and Health Science and the Head of the Biosignal Processing Research Group, UTM. She has authored and co-authored over 100 research papers (citation indices: h-index of 9 and i-10 index of 8) in peer-reviewed journals, book chapters, and conference proceedings. Her research interest includes the areas of physiological signal processing, pattern recognition in biomedical applications, time-frequency signal analysis, multi-modal signal processing, computer aided medical diagnosis system, and biomedical data acquisition.