Australasian Physical & Engineering Sciences in Medicine
https://doi.org/10.1007/s13246-019-00796-9

SCIENTIFIC PAPER

In obstructive sleep apnea patients, automatic determination of respiratory arrests by photoplethysmography signal and heart rate variability

Mehmet Recep Bozkurt[1] · Muhammed Kürşad Uçar[1] · Ferda Bozkurt[2] · Cahit Bilgin[3]

Received: 9 July 2019 / Accepted: 30 August 2019

Abstract
Obstructive sleep apnea is a disease that occurs in connection to pauses in respiration during sleep. Detection of the disease is achieved using a polysomnography device, which is the gold standard in diagnosis. Diagnosis is made by the steps of sleep staging and respiration scoring. Respiration scoring is performed with at least four signals. Technical knowledge is required for attaching the electrodes. Additionally, the electrodes are disturbing to an extent that will delay the patient's sleep. It is needed to have systems as alternatives for polysomnography devices that will bring a solution to these issues. This study proposes a new approach for the process of respiration scoring which is one of the diagnostic steps for the disease. A machine-learning-based apnea detection algorithm was developed for the process of respiration scoring. The study used Photoplethysmography (PPG) signal and Heart Rate Variability (HRV) that is derived from this signal. The PPG records obtained from the patient and control groups were cleaned out using a digital filter. Then, the HRV parameter was derived from the PPG signal. Later, 46 features were derived from the PPG signal and 40 features were derived from the HRV. The derived features were classified with reduced machine-learning techniques using the F-score feature-selection algorithm. The evaluation was made in a multifaceted manner. Besides, Principal Component Analysis was performed to reduce system input (features). According to the results, if a real-time embedded system is designed, the system can operate with 16 PPG feature 95%, four PPG feature 93.81% accuracy rate. These success rates are highly sufficient for the system to work. Considering all these values, it is possible to realize a practical respiration scoring system. With this study, it was agreed upon that PPG signal may be used in the diagnosis of this disease by processing it with machine learning and signal processing techniques.

Keywords Biomedical signal processing · Respiratory arrests · Photoplethysmography · Obstructive sleep apnea · Automatic respiratory staging · Apnea detection · Heart rate variability · Ensemble classification

Introduction

Obstructive sleep apnea (OSA) is a respiratory disorder that occurs in connection to pauses of respiration during sleep. Diagnosis of the disease is made with the help of a polysomnography (PSG) device which collects bioelectric signals, with the steps of sleep staging and respiration scoring based on the guide by the American Academia of Sleep Medicine (AASM) [1]. Sleep staging is for detection of the time the patient spends asleep, and respiration scoring is for detection of the number and duration of abnormal respiratory events that occur in sleep. Sleep staging can only be carried out by a specialist doctor using Electroencephalography (EEG), Electrooculogram (EOG) and Electromyogram (EMG) signals with at least 16 channels [1, 2]. Respiration scoring is achieved with at least 4 channels of signals using blood oxygen saturation, oral-nasal air flow signals, thorax and abdominal respiration movement signals [1]. As a result of these steps, diagnosis of the disease is made by calculation of the Apnea Hypopnea Index (AHI) by dividing the number of abnormal respiratory events by the time spent in sleep. If $AHI < 5$, the individual is considered to be normal. If $AHI > 5$, the individual has OSA and the severity of the

✉ Muhammed Kürşad Uçar
mucar@sakarya.edu.tr

1 Electrical-Electronics Engineering, Faculty of Engineering, Sakarya University, Sakarya, Turkey

2 Computer Programming, Vocational School of Adapazarı, Sakarya University of Applied Sciences, Sakarya, Turkey

3 Faculty of Medicine, Sakarya University, Sakarya, Turkey

Published online: 12 September 2019

disease is determined based on the value of AHI. In general, sleep staging and respiration scoring has crucial value for diagnosis. Several disadvantages of the system such as the inconvenient diagnosis process, unsuitable nature of the device for home usage, prevention of the patients from their natural sleep environment due to the high number of electrodes, the need for a technician to connect the patient to the system and the necessity of a specialist doctor to examine the data, trigger development of new systems [3, 4].

Respiration scoring is still carried out in practice by interpretation of PSG records by a specialist doctor based on the guide by AASM. However, in order to improve the current systems, studies have been conducted in the literature using new signals such as respiration signal, peripheral oxygen saturation (SpO_2), electrocardiogram (ECG), photoplethysmography (PPG) and Heart Rate Variability (HRV) [3, 5–13]. Signals may be ranked based on ease of obtaining them as PPG, SpO_2, HRV, ECG and respiration signal. If the diagnosis of OSA is achieved by disturbing the patient less, the discomfort created by the PSG device may be minimize.

There are many approaches in the literature on OSA diagnosis with ECG [6, 14–16]. A recent study on diagnosing OSA using ECG records for the entire period the patient spends in bed [6]. However, based on the definition of OSA, it is required to detect the abnormal respiratory events that occur during sleep. While even ECG signals differ between the states of being awake and asleep, usage of all records affects the reliability of the system negatively. A different study analyzed OSA records minute by minute [16]. However, more than one 10 s respiratory pauses may occur in a minute. As the number of apneas would change in this case, accurate OSA diagnosis would not be possible. The most noticeable shortcoming in other OSA diagnosis studies in the literature is that they are far from the criteria of AASM [3, 8, 11]. The most important reason for this issue is that engineers are not familiar with the OSA diagnosis process, they do not work in a multidisciplinary nature, and studies are far from a doctor's supervision.

Different studies emphasized that PPG signal may be used in OSA diagnosis [8, 11, 17]. However, a clear system using PPG was developed only recently [5]. This system also has shortcomings. The study not only distributed the data of the apnea and control groups unevenly, but the number of patients was low. Thus, accuracy rates could be distorted. Additionally, the study used to many PPG features. It is very difficult to realize a system in practice with so many features. Usability of the system may be increased by increasing the number of features even more and making selections later. Despite all this research, there is still a need for more practical systems [5, 18]. In addition to these studies, the OSA respiration rate determination method based on Characteristic Moment Waveform has been developed [19].

Another method developed in the literature for the separation of respiratory signals is the Voice Activity Detection (VAD) algorithm [20]. According to this method, breath-holding and regular breath exchange are determined by the energy of the acoustic breathing signal.

This study developed a machine-learning-based apnea detection algorithm for the process of respiration scoring, which is an important step of OSA diagnosis. Apnea is generally defined as the stoppage of air flow from the nose and the mouth for at least 10 h [21]. PPG signal and HRV that is derived from this signal was used for diagnosis. The PPG records taken from the patient and control groups were cleaned with a digital filter. Then, features were derived from PPG and HRV in the domain of time and frequency. The derived features were reduced using feature selection algorithms and classified using machine learning techniques.

The flow of the article is as follows: the 2nd section defines the database used in the study, explains the signal processing and feature derivation steps for PPG signals and introduces the machine learning methods used in the study. The 3rd section provides simulation results, and the 4th section interprets the results.

Materials and methods

The signal processing steps in the study were carried out based on the stages shown in Fig. 1. Based on these steps, firstly, a database was established with the PPG records of the individuals. Then digital filtering was done for cleaning the noise on the PPG signals, and HRV values were created from the PPG signals. Later, feature derivation was carried

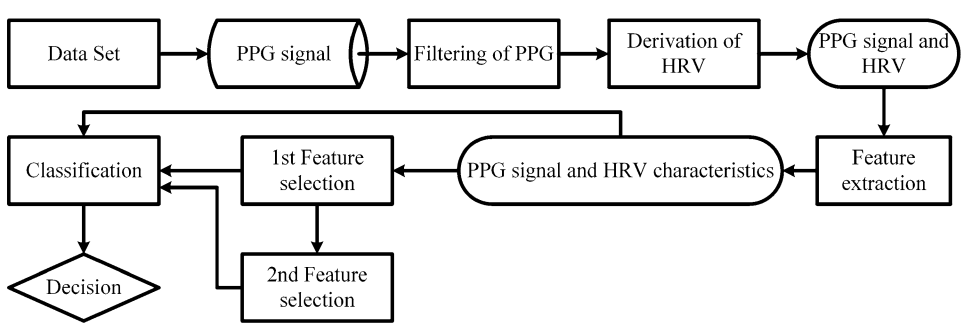

Fig. 1 General signal processing flow diagram

Table 1 Demographic information of apnea and control groups, and data distributions

Information	Female	Male	All individuals
	$n_1 = 5$	$n_2 = 5$	$n = n_1 + n_2 = 10$
Age (year)	59 ± 5	53 ± 11.31	56 ± 8.79
Weihgt (kg)	103.6 ± 10	102.02 ± 6.8	102.81 ± 8.28
Height (cm)	162 ± 3	173 ± 2.83	167.5 ± 6.43
BMI (kg/m^2)	39.46 ± 3	34.14 ± 3.06	36.8 ± 4.05
AHI	9.52 ± 6	24.38 ± 13.21	16.95 ± 12.52

IN	1	2	3	4	5	6	7	12	19	27
Sex	Female	Male	Male	Male	Male	Male	Female	Female	Female	Female
Apnea	122	234	191	268	19	57	100	78	48	8
Control	31	48	59	21	49	256	72	234	187	276

Group	Epoch numbers and durations			
Apnea	Epoch	356	769	1125
	Duration	32.0219 ± 14.3459	23.8934 ± 10.2011	26.4656 ± 12.2641
Normal	Epoch	800	433	1233
	Duration	24.8703 ± 11.858	25.987 ± 12.6832	25.2625 ± 12.1607

Distributions are shown as the mean ± standard deviation

BMI Body Mass Index, *AHI* Apnea Hypopnea Index, *IN* individual number

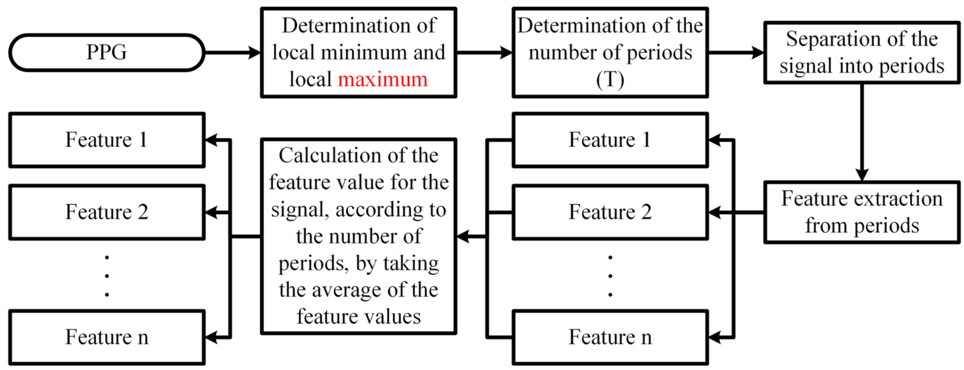

Fig. 2 Feature extraction flow diagram

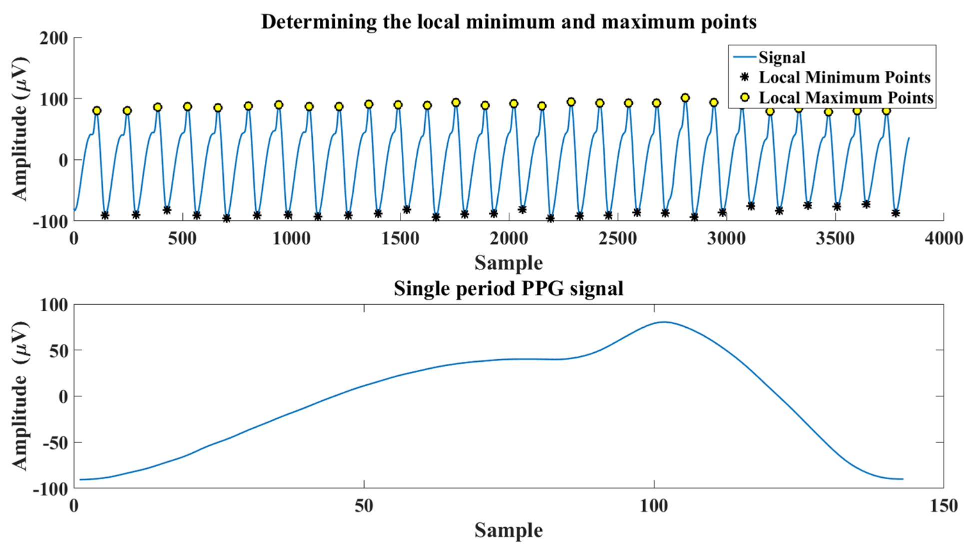

Fig. 3 Determining the local minimum and maximum points and a single period PPG signal

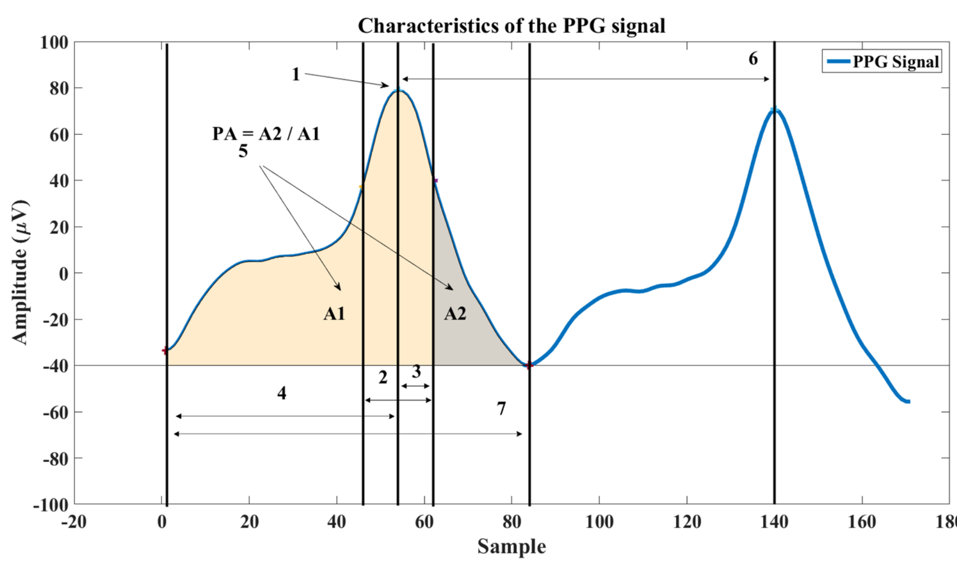

Fig. 4 Characteristic features of the PPG signal

Table 2 Equation for time domain features

No	Feature	Equation	No
8	Mean	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} = \frac{1}{n}(x_1 + \cdots + x_n)$	1
9	Standard deviation	$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})}$	2
10	Average curve length	$CL = \frac{1}{n}\sum_{i=2}^{n}\lvert x_i - x_{i-1}\rvert$	3
11	Average energy	$E = \frac{1}{n}\sum_{i=1}^{n} x_i^2$	4
12	Average teager energy	$TE = \frac{1}{n}\sum_{i=3}^{n}(x_{i-1}^2 - x_i x_{i-2})$	5
13	Activity—Hjort parameters	$A = S^2$	6
14	Mobility—Hjort parameters	$M = S_1^2/S^2$	7
15	Complexity—Hjort parameters	$C = \sqrt{(S_2^2/S_1^2)^2 - (S_1^2/S^2)^2}$	8
–	* Maximum	$x_{max} = max(x_i)$	9
16	Skewness	$x_{ske} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)S^3}$	10
17	Kurtosis	$x_{kur} = \frac{\sum_{i=1}^{n}(x(i) - \bar{x})^4}{(n-1)S^4}$	11
18	Shape factor	$SF = X_{rms} \Big/ \left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{\lvert x_i \rvert}\right)$	12
19	* Minimum	$x_{min} = min(x_i)$	13
20	Root mean squared	$X_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert^2}$	14
21	* Singular value decomposition	$SVD = svd(x)$	15
22	Median	$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & : x \text{ odd} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & : x \text{ even} \end{cases}$	16
23	Geometric mean	$G = \sqrt[n]{x_1 + \cdots + x_n}$	17
24	Harmonic mean	$H = n \Big/ \left(\frac{1}{x_1} + \cdots + \frac{1}{x_n}\right)$	18
25	* 25% trimmed mean	$T25 = trimmean\,(x, 25)$	19
26	* 50% trimmed mean	$T50 = trimmean\,(x, 50)$	20
27	Range	$R = x_n - x_1$	21
28	* IQR	$IQR = iqr(x)$	22
29	* Mean absolute deviation	$MAD = mad(x)$	23
30	* Central moments	$CM = moment(x, 10)$	24
31	DK	$DK = (S/\bar{x})100$	25
–	* Normality test p	$[p, h] = kstest(x)$	26
–	* Normality test $h_{1,0}$		27
32	* Sign test p	$[p, h] = signtest(x)$	28
33	* Sign test $h_{1,0}$		29
34	Standard error	$S_{\bar{x}} = S \Big/ \sqrt{n}$	30
35	Count of maximum	YMaks	–
36	Count of minimum	YMin	–

IQR interquartile range, *DK* coefficient of variation

*The property was computed using MATLAB

Table 3 Formulas for features in frequency domain

No	Equation	Feature	Equation	No
37	E_{PPG}	Energy rate	E_{HRV}	31
–	–		$E_{HRV_{VLF}}$	32
38	$E_{PPG_{LF}}$		$E_{HRV_{LF}}$	33
39	$E_{PPG_{MF}}$		–	–
40	$E_{PPG_{HF}}$		$E_{HRV_{HF}}$	34
41	$E_{PPG_{LF}}/E_{PPG}$		$E_{HRV_{VLF}}/E_{HRV}$	35
42	$E_{PPG_{MF}}/E_{PPG}$		$E_{HRV_{LF}}/E_{HRV}$	36
43	$E_{PPG_{HF}}/E_{PPG}$		$E_{HRV_{HF}}/E_{HRV}$	37
44	$E_{PPG_{LF}}/E_{PPG_{MF}}$		$E_{HRV_{VLF}}/E_{HRV_{LF}}$	38
45	$E_{PPG_{LF}}/E_{PPG_{HF}}$		$E_{HRV_{VLF}}/E_{HRV_{HF}}$	39
46	$E_{PPG_{MF}}/E_{PPG_{HF}}$		$E_{HRV_{LF}}/E_{HRV_{HF}}$	40

NF number of features, *S* sensitivity, *Sp* specificity, *A* accuracy

out on the PPG and HRV values. Based on the relationship between the derived features and apnea, features that were in harmony with the feature selection algorithm were selected. They were then classified with machine learning techniques and the system performance was tested.

Receiving records

The data in the scope of the study were recorded using a SOMNOscreen Plus PSG device at the Sleep Laboratory of the Department of Pulmonary Diseases at Sakarya Hendek State Hospital. The study was conducted using PPG signals with a sampling frequency of 128 Hz. The demographic information for the individuals is given in Table 1. The records were examined by a specialist doctor, respiration scoring and sleep staging steps were carried out. At the stage after this, PPG records during respiration pauses were recorded. The study used the records of respiration pauses that were labeled as obstructive apnea and hypopnea. As a study found that there was no significant difference between apnea and hypopnea, this study used apnea as an umbrella term for both apnea and hypopnea [22]. As correspondence to apnea records, control group records were created. The statistical information regarding the records collected from the individuals for apnea and control groups is given in Table 1. Individual number is a unique number given to the patients by the system. The apnea and control labels provide the numbers of records obtained from the patient. For example, 122 apnea epochs determined by the doctor were obtained from the patient with the number 1. Additionally, 31 epochs were taken from this patient where they were asleep, and their respiration was regular. Each epoch contains a minimum 10 s of PPG record as necessitated by the definition of apnea.

Signal pre-processing

In order to clean the noise on the PPG signals, the signals were treated with a 0.1–20 Hz IIR-Chebyshev Type II band

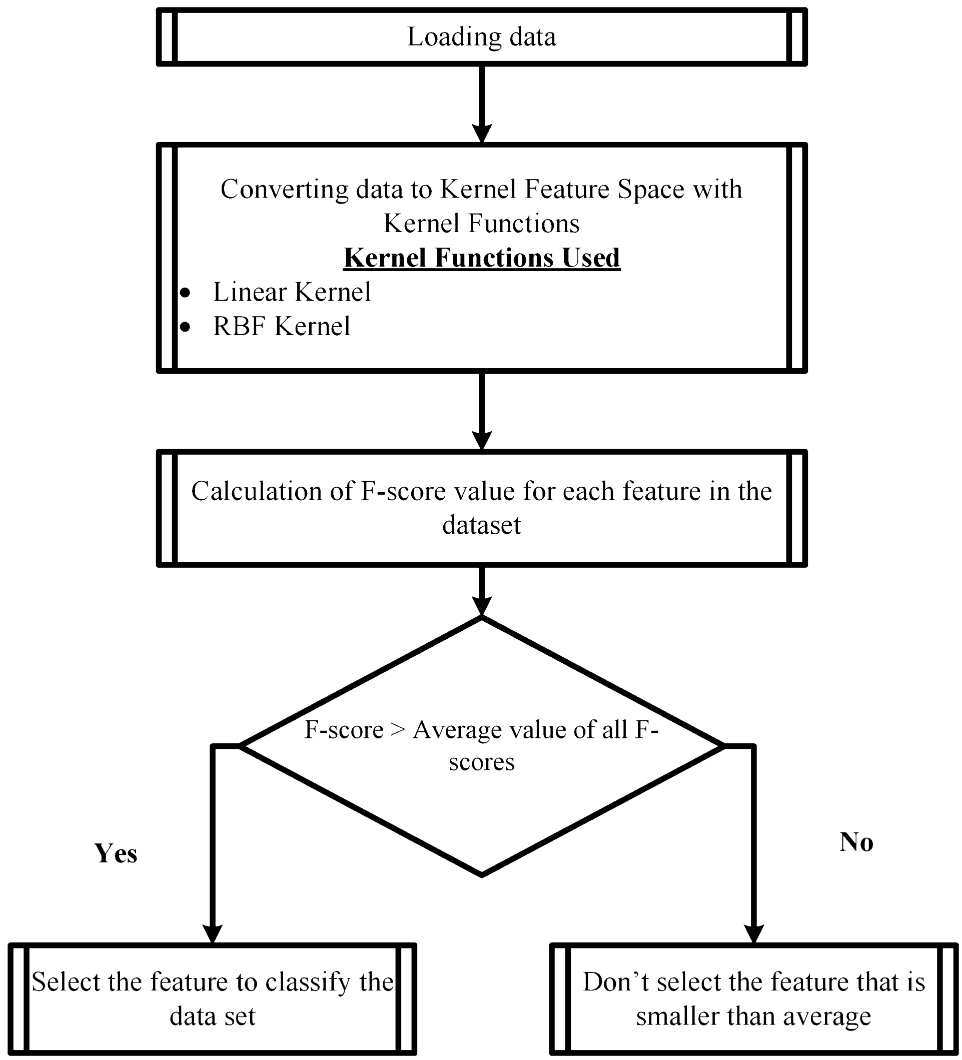

Fig. 5 F-score feature selection processing steps

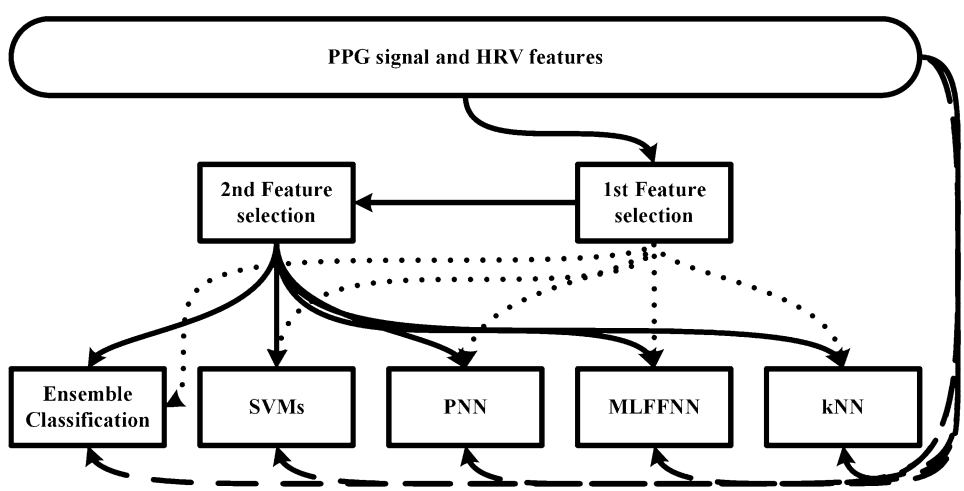

Fig. 6 Classification flow diagram

pass moving average filter. Filters are designed and applied in Matlab environment. Later, HRV parameters were created by using the cleaned PPG signals. While obtaining each HRV parameters from PPG signal, peak points of the PPG signal are determined. Corresponding to N peak points detected, $N-1$ HRV parameters are calculated. A HRV parameter corresponds to the time interval between PPG's two subsequent peak points.

Feature extraction for signal processing

46 (Time domain 36, frequency domain 10) features from the PPG and 40 (Time domain 30, frequency domain 10) from the HRV were excluded.

To get a feature (Fig. 2), first, we divided the signal into periods. The number of periods is $T = LOCMIN - 1$. The number of the periods in the signal is calculated by $T = LOCMIN - 1$, where T is the number of periods, and $LOCMIN$ is the number of minima. Figure 3 shows determination of the minimum and maximum points for a 30-s PPG signal.

The desired feature was derived from each period of the PPG signal, its average value was taken, and recorded as the feature of the relevant epoch. This way, it is ensured that the features are obtained with minimum error. For example, while computing the feature of standard deviation, the standard deviation values are computed separately in each period obtained, their average is taken, and a single standard deviation value is obtained for the epoch. The steps were repeated for all the features derived.

In the study the first 7 features among the 46 features derived from the PPG signal were characteristic features, and they are shown on Fig. 4 as numbered.

In the time domain, 29 and 30 statistical properties were derived from the PPG signal and HRV respectively (Table 2).

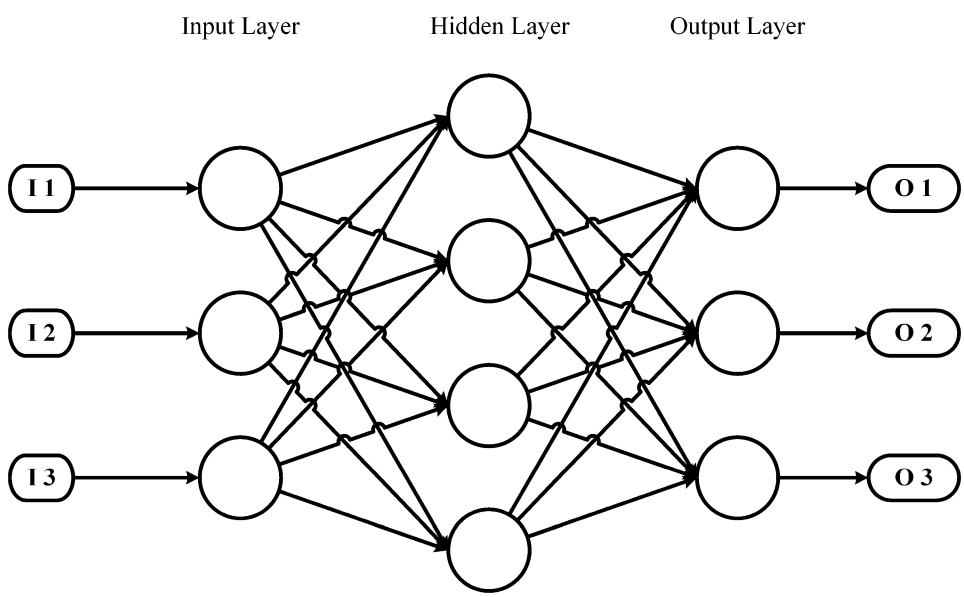

Fig. 7 General network structure for MLFFNN and PNN

Table 4 Network operation parameters

Training algorithm		Number of neurons	Iteration
Levenberg–Marquardt	trainlm	1	10
BFGS Quasi-Newton	trainbfg	2	
Resilient backpropagation	trainrp	3	
Scaled conjugate gradient	trainscg	4	
Conjugate gradient with Powell/Beale restarts	traincgb	5	
Fletcher–Powell conjugate gradient	traincgf	6	
Polak–Ribiére conjugate gradient	traincgp	7	
One step secant	trainoss	–	
Variable learning rate gradient descent	traingdx	100	

The features marked "*" were calculated using MATLAB library [23]. The x shown in formulae represents the signal.

In addition to the statistical features, the study also determined the energy amounts in the frequency bands of signal and used these as features of the frequency domain. While deriving the frequency domain features, firstly the sub-frequency bands of the signals were derived. The PPG signal has three different sub-frequency bands as the low frequency band at 0.04–0.15 Hz (LF), the medium frequency band at 0.09–0.15 Hz (MF) and the high frequency band at 0.15–0.6 Hz (HF) [24, 25]. PPG_{LF}, PPG_{MF} and PPG_{HF} are the PPG signals in LF, MF and HF bands respectively. HRV was divided into three bands as the very low frequency band at 0.0033–0.04 Hz (VLF), the low frequency band at 0.04–0.15 Hz (LF) and the high frequency band at 0.15–0.4 Hz [24, 25]. HRV represents HRV, HRV_{VLF} represents the VLF band, HRV_{LF} represents the LF band, and HRV_{HF} represents the HF band.

In order to obtain the PPG and HRV sub-frequency bands, a IIR Chebyshev Type II filter was used. As a result of the application, three sub-frequency bands were acquired. With the six frequency bands, a total of eight vectors were obtained. In order to calculate frequency features, the energies of these signals were calculated. Energy calculation was made with Eq. 1. Here, x represents the signal whose energy is calculated.

$$E = \sum_{i=-\infty}^{+\infty} |x[i]|^2 \tag{1}$$

The energy of the signal was calculated (Table 3). E_{PPG}, $E_{PPG_{LF}}$, $E_{PPG_{MF}}$, $E_{PPG_{HF}}$, E_{HRV}, $E_{HRV_{VLF}}$, $E_{HRV_{LF}}$ and $E_{HRV_{HF}}$ are energy of the PPG signal, the LF band, the MF band, the HF band, the HRV, the VLF band, LF band and the HF band respectively.

Property selection algorithm

F-score is a feature selection algorithm that can be used to reduce the number of features and select important features (Fig. 5) [26]. The F-score value is calculated for each property according to the method 2 [26]. The threshold value for the feature selection is the average of the F-score of all properties. Properties on the threshold value are selected.

In Eq. 2, $x_{k,i}$ is the feature vector, $k = 1, 2 \ldots, m$ and $m = n_+ + n_-$ are the total numbers elements in the positive (+) and negative (−) classes, i is the number of features. n_- and n_+ are the number of samples in the negative (−) class and the positive (+) class respectively.

$\bar{x}_i$, $\bar{x}_i^{(-)}$ and $\bar{x}_i^{(+)}$ are the average value, the average value in the negative class and positive class respectively. $x_{k,i}^{(+)}$ and $x_{k,i}^{(-)}$ are kth positive and negative sample for ith feature respectively.

$$F(i) = \frac{(\bar{x}_i^{(+)} - \bar{x}_i)^2 + (\bar{x}_i^{(-)} - \bar{x}_i)^2}{\frac{1}{n_+-1}\sum_{k=1}^{n_+}(x_{k,i}^{(+)} - \bar{x}_i^{(+)})^2 + \frac{1}{n-1}\sum_{k=1}^{n_-}(x_{k,i}^{(-)} - \bar{x}_i^{(-)})^2} \tag{2}$$

Classification stage

The purpose in the process of classification is to achieve the respiration scoring process based on machine learning with the help of PPG signal and HRV features. Classification was made in two steps as seen in Fig. 6. In the first stage, the features were classified without subjecting them to any process. Later, they were treated with the F-score feature selection algorithm two times, and classified in each step.

Table 5 Three different kernel functions in SVM classifier

Kernel function		BoxConstraint	Standardized	
Gaussian or radial basis function (RBF) kernel	rbf	1	1	True
		2		
Linear kernel	Linear	3		
		4	0	False
Polynomial kernel	Polynomial	–		
		100		

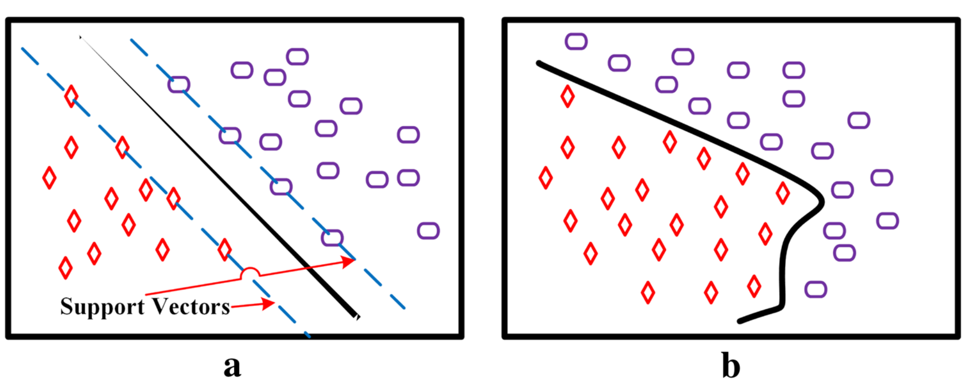

Fig. 8 Separation of classes by **a** linear and **b** nonlinear lines

Four different machine learning techniques were used for the classification operation. These are the k-nearest neighbor algorithm (kNN), the multilayer feed forward artificial neural networks (MLFFNN), probabilistic artificial neural networks (PNN) and support vector machines (SVMs). Additionally, the study used an ensemble classifier which worked based on the common decision of all these classifiers.

In order to interfere with machine learning parameters, algorithms are prepared without using Matlab toolbox.

k-Nearest neighbors classification

kNN is a controller learning method which solves the problem of classification [27]. The important thing in the method is that the features of each class have been clearly determined beforehand. The performance of the method is influenced by k nearest neighbors, similarity measurement and sufficient sample in the datasets. The k value is selected at the start. Selecting a high k value may result in gathering groups of data that are not similar to each other. Studies usually prefer k values of 3, 5 or 7 [28].

The kNN classifier works as the following. A k value is determined at the start. $k = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ were used here. Then, the distance between the data with known class label and those without known class label is calculated. Distance calculation may be made using various distance calculation formulae. This study used 11 different distance calculation formulae. These were: Chebychev, City-block, Correlation, Cosine, Euclidean, Hamming, Jaccard, Mahalanobis, Minkowski, Seuclidean and Spearman. Using the distance formulae, the k nearest neighbors are found. The majority is determined using the k nearest neighbors. The majority label is determined as the label for the data without previously known label. The study used 10 different k values and 11 different distance calculation formulae. This way, 110 different structures of kNN networks were established for classification of a single data set, and the most efficient result was determined.

Table 6 Data distribution in training and test phases in the classifiers

Class (%)	Respiratory scoring		
	Apnea (%)	Control (%)	Total (%)
Training	562	617	1179
(50)	(47.67)	(52.33)	(100)
Test	563	616	1179
(50)	(47.75)	(52.25)	(100)
Total	1125	1233	
(100)	(100)	(100)	

Multi-layer feed forward artificial neural networks

Artificial neural networks are architectures that are formed by connecting artificial neurons [29]. This architecture consists of one-way channels and interconnected elements for processing data [30]. There are three types of networks as feed forward networks, cascade networks and back-propagation networks. A feed forward network consists of three fundamental layers for a certain task as the input layer, the hidden layer and the output layer as seen in Fig. 7. The data start at the input layer and follow the order of the hidden layer and finally the output layer in one direction.

There are various initial parameters for the MLFFNN classifier. These are summarized in Table 4. The study used nine different training algorithms and neuron number (1–100) parameters for MLFFNN. A MLFFNN with any parameter does not produce the same result two times in a row. In order to produce better results in the study, the MLFFNN was run ten times with the same parameters. For example, a total of $9 \times 100 \times 10 = 9000$ different networks were run for classification of the PPG data, and the best network results were selected. Considering there were nine different data sets, the study formed 81000 different MLFFNNs.

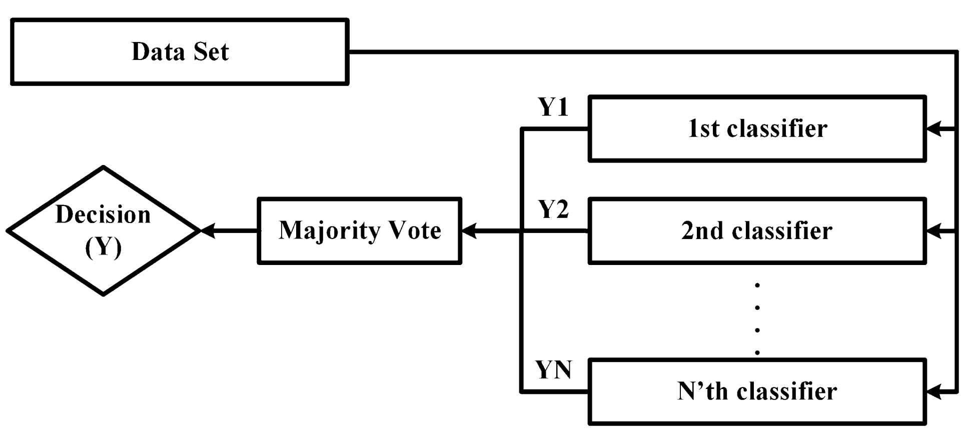

Fig. 9 Ensemble classifier working algorithm

Probabilistic artificial neural networks

PNN is the application of a statistical algorithm for ideal classification problems [31]. In classification processes, a PNN is a network that is based on consideration of all points [32]. For a point to be used for classification, distances to all other points are calculated. As the basis of the function is the radius distance, it was named a radial-based function.

PNN is similar to the structure of multi-layer artificial neural networks. Figure 7 shows the general structure of the network. In this structure, the number of neurons in the input layer is equal to the number of features assigned to the network. The study used numbers of features that could be used as network input as the numbers of features selected (4, 5, 11, 16 and 28) as summarized in Table 7. The network structure had two hidden layers. The number of neurons in the first one was same as the number of samples fed to the network. The second one had two neurons. As the study had two different output values (Apnea/Control), the output layer had two neurons.

For the PNN classifier, it is possible to manipulate the spread starting parameter only. As the spread parameter approaches zero, it starts to behave like the nearest neighbor classifier [33]. As this value get further away from zero, the classifier makes a classification by considering a few vectors that separate classifying data from each other [33]. This value was designed with a total of 5000 different values in the range of 0.001–5 with increments of 0.001. At the end of the study, the network parameters and performance criteria that provided the best performance were calculated.

Support vector machines

SVMs were proposed by Cortes [34]. In addition to classification problems, SVMs are also effectively used in regression analyses. SVMs, in general, aim to divide two classes with linear or non-linear lines. As an example, Fig. 8 shows how the data are divided as (a) linear and (b) non-linear.

SVMs are a training algorithm for classification operations. The purpose of the algorithm is to be able to separate the data sets on the hyperplane in the most suitable way and classify the data with maximum accuracy rate [35]. The learning data that are the nearest to the hyperplane are named support vectors. Figure 8a shows support vectors. Curves is solution that divides the datasets into two parts (Fig. 8).

The network parameters used in the SVMs designed in the study are summarized in Table 5. As the network parameters, there were three different kernel functions, BoxConstraint parameter in the range of 1–100, and the parameters of whether the data were standardized (normalized) or not (two different states). Considering these parameters, at total of $3 \times 100 \times 2 = 600$ different networks were defined in the study for classification, the best-performing network was determined, and performance evaluation criteria for the network were calculated.

Ensemble classifier

The community classifier consists of many classifiers. The aim of the community classifier is to have better performance than single classifiers [36]. The working structure and flow chart of the ensemble classifier are presented in Fig. 9. The system is formed with N classifiers. N may be odd or even. While making classification based on the features vector, each classifier produces an output value for the 1st feature vector. The output values are counted. In the following stage, with majority of votes, the decision of the ensemble classifier is reached. If the number of classifiers is even, the

Table 7 Feature selection results for respiratory scoring with F-score

FSN		1st FS		2nd FS	
Signal	NF	NSF	SFN	NSF	SFN
PPG	46	16	1 9 10 15 18 19 20 21 23 27 28 29 32 33 34 43	4	18 32 33 43
HRV	40	11	2 5 7 8 9 13 21 22 23 25 30	5	2 9 21 23 25
PPG HRV	86	13 + 15 = 28	PPG 1 9 15 18 19 20 21 27 29 32 33 34 43	3 + 8 = 11	18 32 33
			HRV 2 5 6 7 8 9 13 14 16 20 21 22 23 25 28		2 7 8 9 21 23 25 28

FSN feature selection number, *FS* feature selection, *NSF* number of selected features, *SFN* selected feature numbers

Table 8 Classifier results of PPG features for respiration scoring

k-Nearest neighbor algorithm (kNN)									
NP	k = 3, distance function = 'cosine'			k = 5, distance function= 'cityblock'			k = 9, distance function = 'seuclidean'		
Class	WoFSI			WFSI			WFSI		
	NF = 46			NF = 16			NF = 4		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.80	0.84	82.27	0.74	0.83	79.13	0.69	0.77	73.11
Control	0.84	0.80		0.83	0.74		0.77	0.69	
AUC	0.82			0.79			0.73		
Kappa	0.64			0.58			0.46		
F-measure	0.82			0.79			0.73		
k(10)-fold (%)	79.30			75.83			68.96		

Multi-layer artificial neural networks (MLFFNN)									
NP	NN = 92, TA = 'trainlm'			NN = 92, TA = 'trainlm'			NN = 79, TA = 'trainlm'		
Class	WoFSI			WFSI			WFSI		
	NF = 46			NF = 16			NF = 4		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.80	0.88	83.80	0.80	0.83	82.02	0.68	0.79	73.79
Control	0.88	0.80		0.83	0.80		0.79	0.68	
AUC	0.84			0.82			0.74		
Kappa	0.67			0.64			0.47		
F-measure	0.83			0.82			0.73		
k(10)-fold (%)	–			–			–		

Probabilistic artificial neural networks (PNN)									
NP	Spread = 0.3410			Spread = 0.0910			Spread = 0.0980		
Class	WoFSI			WFSI			WFSI		
	NF = 46			NF = 16			NF = 4		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.71	0.88	80.24	0.73	0.81	77.18	0.67	0.76	71.93
Control	0.88	0.71		0.81	0.73		0.76	0.67	
AUC	0.80			0.77			0.72		
Kappa	0.60			0.54			0.44		
F-measure	0.79			0.77			0.71		
k(10)-fold (%)	–			–			–		

Support vector machines (SMVs)									
NP	Kernel = 'rbf', BC = 70			Kernel = 'rbf', BC = 1			Kernel = 'rbf', BC = 1		
Class	WoFSI			WFSI			WFSI		
	NF = 46			NF = 16			NF = 4		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.82	0.86	83.80	0.79	0.80	79.64	0.73	0.70	71.67
Control	0.86	0.82		0.80	0.79		0.70	0.73	
AUC	0.84			0.80			0.72		
Kappa	0.67			0.59			0.43		
F-measure	0.84			0.80			0.72		
k(10)-fold (%)	76.68			76.84			69.80		

Table 8 (continued)

Ensemble classifier									
Class	WoFSI			WFSI			WFSI		
	NF = 46			NF = 16			NF = 4		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.84	0.97	90.67	0.93	0.96	95.00	0.89	0.98	93.81
Control	0.97	0.84		0.96	0.93		0.98	0.89	
AUC	0.90			0.95			0.93		
Kappa	0.81			0.90			0.87		
F-measure	0.90			0.95			0.93		
k(10)-fold (%)	–			–			–		

NP network parameters, *WoFSI* without feature selection implementation, *WFSI* with feature selection implementation, *NF* number of features, *S* sensitivity, *Sp* specificity, *A* accuracy, *NN* number of neurons, *TA* training algorithm, *BC* BoxConstraint

average of the decision values of the classifiers is calculated, rounded and the decision of the ensemble classifier is determined. This operation is applied on all feature vectors.

Let us assume that the output values of an ensemble classifier with four classifiers are 1: Normal and 2: Apnea. If the four classifiers have produced the outputs of 1 1 2 1 in order, the decision of the ensemble classifier would be 1 by majority vote. If the outputs are 1 1 2 2 in order, the average is calculated as 1.5. When this number is rounded up, the output value of the ensemble classifier becomes 2.

The ensemble classifier was prepared in MATLAB environment using four different classifiers as kNN, MLFFNN, PNN, SVMs [23].

In each classifier, network training is done with training and test sets. Sets are created according to systematic sampling theory (Table 6) [37].

The used performance criteria

The performance of the developed systems was evaluated with the following parameters. There were k-fold cross-validation accuracy rate, sensitivity, specificity, accuracy rates, ROC—Receiver Operating Characteristic, kappa coefficient and AUC—Area Under ROC [38, 39].

Principal component analysis

Principal Component Analysis (PCA) was used to validate the data in this study. This process was performed differently from the normal process. PCA analyzed the entire property matrix, and the key components were identified. 86 features are described in 100% with only one essential component.

An essential component obtained as a result of PCA has been classified. The data set for the classification process has three parts: Training (70%), Test (20%) and Verification (10%). In order to evaluate the results of PCA, all data (86 features) were divided into three parts and re-classified as Training (70%), Test (20%) and Verification (10%) as in PCA.

Results

This section presents the results obtained in the study. The study developed a new approach alternative to the respiration scoring process which is one of the diagnosis steps of the disease OSA.

For the respiration scoring process, a machine-learning-based system was developed using features derived from PPG and HRV. Moreover, in order to increase the performances of the classifier, the ones that were effective on the result among the 86 features derived from PPG and HRV were selected by the F-score method and utilized. The features were treated with the F-score method two times, classification was made in each step, and the effect of the F-score method on different levels was investigated. Table 7 shows the PPG and HRV features that were selected after the F-score method. The numbers of total PPG and HRV features are given in the "Number of Features" column in Table 7. The derived features were used by combination in the "PPG HRV" column. The 46 features derived from PPG were reduced to 16 features after the first F-score treatment. The feature numbers assigned to these 16 features are shown in the "Selected Feature Numbers" column. After the second F-score treatment, the 16 features were reduced up to four features. The case was the same for HRV. The 40 features that were derived were reduced to 11 after the first F-score treatment and to five after the second. The total of 86 features that were found when PPG and HRV were combined,

Table 9 Classifier results of HRV features for respiration scoring

k-nearest neighbors classification algorithm (kNN)									
NP	k = 3, distance function = 'cosine'			k = 8, distance function= 'mahalanobis'			k=10. distance function = 'mahalanobis'		
Class	WoFSI			WFSI			WFSI		
	NF = 40			NF = 11			NF = 5		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.75	0.85	80.15	0.77	0.77	76.76	0.74	0.76	75.32
Control	0.85	0.75		0.77	0.77		0.76	0.74	
AUC	0.80			0.77			0.75		
Kappa	0.60			0.53			0.51		
F-measure	0.80			0.77			0.75		
k(10)-fold (%)	79.05			76.34			72.35		
Multi-layer artificial neural networks (MLFFNN)									
NP	NN = 3, TA = 'trainlm'			NN = 29, TA = 'trainlm'			NN = 87, TA = 'trainlm'		
Class	WoFSI			WFSI			WFSI		
	NF = 40			NF = 11			NF = 5		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apneaa	0.81	0.84	82.70	0.69	0.84	76.93	0.67	0.79	73.45
Control	0.84	0.81		0.84	0.69		0.79	0.67	
AUC	0.83			0.77			0.73		
Kappa	0.65			0.53			0.47		
F-measure	0.83			0.76			0.73		
k(10)-fold (%)	–			–			–		
Probabilistic artificial neural networks (PNN)									
NP	Spread = 0.1580			Spread = 0.3440			Spread = 0.2450		
Class	WoFSI			WFSI			WFSI		
	NF = 40			NF = 11			NF = 5		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.55	0.86	71.50	0.48	0.84	67.18	0.51	0.85	69.04
Control	0.86	0.55		0.84	0.48		0.85	0.51	
AUC	0.71			0.66			0.68		
Kappa	0.42			0.33			0.37		
F-measure	0.67			0.61			0.64		
k(10)-fold (%)	–			–			–		
Support vector machines (SMVs)									
NP	Kernel = 'rbf', BoxConstraint = 2			Kernel = 'rbf', BoxConstraint = 1			Kernel = 'rbf', BoxConstraint = 1		
Class	WoFSI			WFSI			WFSI		
	NF = 40			NF = 11			NF = 5		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.77	0.85	81.00	0.76	0.81	78.29	0.78	0.73	75.66
Control	0.85	0.77		0.81	0.76		0.73	0.78	
AUC	0.81			0.78			0.76		
Kappa	0.62			0.56			0.51		
F-measure	0.81			0.78			0.76		
k(10)-fold (%)	79.90			79.05			73.71		

Table 9 (continued)

Ensemble classifier									
Class	WoFSI			WFSI			WFSI		
	NF = 40			NF = 11			NF = 5		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.86	0.97	92.28	0.78	0.98	88.04	0.82	0.98	90.16
Control	0.97	0.86		0.98	0.78		0.98	0.82	
AUC	0.92			0.88			0.90		
Kappa	0.84			0.76			0.80		
F-measure	0.91			0.87			0.89		
k(10)-fold (%)	–			–			–		

NP network parameters, *WoFSI* without feature selection implementation, *WFSI* with feature selection implementation, *NF* number of features, *S* sensitivity, *Sp* specificity, *A* accuracy, *NN* number of neurons, *TA* training algorithm, *BC* BoxConstraint

were reduced to 28 after the first F-score treatment and to 11 after the second.

In this study, the results of comparison of the newly proposed system to the reference system for respiration scoring are shown in different tables for PPG and HRV. A total of 86 features were derived in the study. However, the difficulty of deriving so many features in systems that work in real-time was considered, and it was aimed to reduce the number of features and improve the system. Classification results are given for PPG features in Table 8, for HRV features in Table 9, and PPG and HRV features in Table 10. Considering the results in Table 8 for PPG, when all features are used, the respiration scoring success rate was 80% or higher for all classifiers. Likewise, the classifiers' specificity and sensitivity rates, that is, their capacities of distinguishing between asleep and awake states were also around 80%. Considering the results in Table 9 for HRV, the success rates of the kNN, MLFFNN, SVMs and ensemble classifiers were 80% or higher. When the number of features were reduced, the accuracy rates of the classifiers dropped, while on the other hand, the success rate of the ensemble classifier was 90.16%. This rate was achieved with only five features. It may be argued that this was an impressive performance for a system that may be used in practice.

In addition to the accuracy rates of the classifiers, their F-Measure and AUC values were also around 80 for all features. The reliability of the system was reinforced with the parameters. Another performance criterion is the ROC curve. The ROC curves of the systems developed for respiration scoring are shown in Fig. 10. Nine different ROC curves are shown for the nine different data sets that were classified. While evaluating the curves, the ideal ROC curve in the plot is used as a reference. It may be stated that the classifier that is the closest to this curve is the best classifier. The ROC curves of the developed classifiers were highly close to the ideal. In addition to all these good aspects of the system, the Kappa values for the classifiers were very low. In this sense, the system may be improved a bit more. To improve the system, different features that may represent abnormal respiratory events may be derived from PPG and HRV signals. The database may also be broadened.

After the feature selection process was completed, classification was made. Classification operations were made separately for PPG and HRV, followed by a combined operation. This way, all the features were utilized.

The results were provided in the Table in detail for each classifier (Tables 8, 9, 10). The classification operations were carried out in order and the results were as the following. Firstly, the 46 PPG features were classified without being subjected to any feature selection algorithm, performance parameters for measuring the performance of the classifier were calculated, and recorded in the relevant column. Later, 46 features were reduced to 16 with the first selection algorithm and the same process was repeated. The second feature selection algorithm reduced 21 features to four and classification was made again. The performance criteria for the classifiers were calculated and are shown in the Table. The numbers of features obtained after each F-score treatment are also included. In the tab named "Network Parameters" for each classifier, their network parameters were calculated. Additionally, the ROC curves for the classifiers were prepared and are shown in Fig. 10. An ROC curve may be interpreted as the following: if the curve is closer to the left axis, it is able to diagnose apnea better; if the curve is closer to the upper axis, it is able to identify the control group better.

PCA analysis was performed to reduce the number of features (Table 11). The table has two columns. The first shows the classification results for all properties, and the second shows the classification results for PCA analysis. Considering all the features, system performance (Test or Validation) varies between 79.83 and 86.92%. As a result of PCA analysis, performance (Test or Validation) ranged from 46.35 to

Table 10 Classifier results of PPG and HRV features for respiration scoring

k-nearest neighbors classification algorithm (kNN)									
NP	k = 1, distance function = 'cosine'			k = 3, distance function = 'cosine'			k = 8, distance function = 'mahalanobis'		
Class	WoFSI			WFSI			WFSI		
	NF = 86			NF = 28			NF = 11		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.80	0.89	84.56	0.80	0.84	82.02	0.78	0.77	77.27
Control	0.89	0.80		0.84	0.80		0.77	0.78	
AUC	0.84			0.82			0.77		
Kappa	0.69			0.64			0.55		
F-measure	0.84			0.82			0.77		
k(10)-fold (%)	81.59			78.88			76.51		
Multi-layer artificial neural networks (MLFFNN)									
NP	NN = 14, TA = 'trainlm'			NN = 63, TA = 'trainlm'			NN = 4, TA = 'trainlm'		
Class	WoFSI			WFSI			WFSI		
	NF = 86			NF = 28			NF = 11		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.81	0.90	85.41	0.81	0.84	82.44	0.72	0.83	77.69
Control	0.90	0.81		0.84	0.81		0.83	0.72	
AUC	0.85			0.82			0.77		
Kappa	0.71			0.65			0.55		
F-measure	0.85			0.82			0.77		
k(10)-fold (%)	–			–			–		
Probabilistic artificial neural networks (PNN)									
NP	Spread = 0.4500			Spread = 0.2700			Spread = 0.1030		
Class	WoFSI			WFSI			WFSI		
	NF = 86			NF = 28			NF = 11		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.74	0.89	81.51	0.69	0.86	77.69	0.73	0.80	76.51
Control	0.89	0.74		0.86	0.69		0.80	0.73	
AUC	0.81			0.77			0.76		
Kappa	0.63			0.55			0.53		
F-measure	0.80			0.76			0.76		
k(10)-fold (%)	–			–			–		
Support vector machines (SMVs)									
NP	Kernel = 'rbf', BoxConstraint = 2			Kernel = 'rbf', BoxConstraint = 3			Kernel = 'rbf', BoxConstraint = 2		
Class	WoFSI			WFSI			WFSI		
	NF = 86			NF = 28			NF = 11		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.86	0.88	87.36	0.83	0.84	83.97	0.75	0.80	77.35
Control	0.88	0.86		0.84	0.83		0.80	0.75	
AUC	0.87			0.84			0.77		
Kappa	0.75			0.68			0.55		
F-measure	0.87			0.84			0.77		
k(10)-fold (%)	81.17			81.26			76.42		

Table 10 (continued)

Ensemble classifier									
Class	WoFSI			WFSI			WFSI		
	NF = 86			NF = 28			NF = 11		
	S	Sp	A (%)	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.89	0.95	92.54	0.87	0.96	91.69	0.84	0.97	90.59
Control	0.95	0.89		0.96	0.87		0.97	0.84	
AUC	0.92			0.91			0.91		
Kappa	0.85			0.83			0.81		
F-measure	0.92			0.91			0.90		
k(10)-fold (%)	–			–			–		

NP network parameters, *WoFSI* without feature selection implementation, *WFSI* with feature selection implementation, *NF* number of features, *S* sensitivity, *Sp* specificity, *A* accuracy, *NN* number of neurons, *TA* training algorithm, *BC* BoxConstraint

64.77%. As a result of both analyzes, the Ensemble classifier was able to tolerate the weaknesses of weak classifiers.

The general assessment of classification results is given in Table 12 and their plot is given in Fig. 11.

In the graph, the x axis represents the performance criteria, and the y axis represents the performance values. The best performance value is 1. The graph with red marking shows the performance values of the system with four PPG features, and the graph with blue marking shows the performance values of the system with 16 PPG features.

According to Table 11, the best performance for respiration scoring was obtained with the ensemble classifier. While HRV may be used for this process by itself, PPG and HRV may also be used in combination. Combined usage of PPG and HRV not only increased the performance of classification, but it also provided a decrease in the number of features, which is a benefit. The best performance for respiration scoring was achieved by the ensemble classifier using PPG signal features. This operation may be carried out using either 16 or four PPG features. While using 4 features in the system reduces the workload, using 16 features improves the performance despite increasing the workload. It is possible to develop a physical system based on the respiration scoring results obtained here.

Discussion

Studies in the literature which used PPG for respiration scoring are a minute amount [17]. Nevertheless, even this study aimed to detect the arousals during apnea, rather than apnea itself. In addition to this, the literature frequently reported PPG number of respirations and effort tests [17, 24, 40, 41]. This study is expected to fill the gap in the literature regarding the process of respiration scoring using PPG.

According to the results obtained in this study, it was determined that PPG features and features of HRV, which may be obtained through PPG signals, may be used in processes of respiration scoring and these will provide significant results. Easy acquisition of PPG signals and derivation of HRV from PPG signals paves the way for possibility of respiration scoring with a single signal. Easy measurement and processing of the signal in real-time systems will increase the systems' practicality. In OSA diagnosis, there is a need for at least three channels of signal. Usage of PPG instead of these will reduce the workload.

Respiration scoring may be achieved only by PPG records at a 93–95% accuracy rate. The system may be realized with 16 PPG features and 95% accuracy rate. If one wants to be economical in code-writing, the number of features may be reduced to 4. Which will provide a success rate of 93.81%. This success rate is highly sufficient for the system to work. Considering all these value, it is possible to create a practical respiration scoring system. Figure 11 and Table 12 present the graphical summary of the numerical value mentioned above.

In order to reduce the number of features with PCA, we found that the performance was not good enough. By PCA analysis, 86 properties are explained 100% with only one essential component. However, with an essential component, a performance value of approximately 40–50% was taken. The purpose of PCA is to reduce system input. However, this is not available according to performance values.

The ensemble classifier provided a superior performance in classification of all groups of data. The reason for this is that, while making classifications one by one, the mistake made by a classifier is compensated for by another classifier. This way, the individual performances were combined, and the power of the system was increased. Furthermore, the distribution of the data was arranged to prevent differences among the groups. The normal and regular distribution of the data provided a positive influence of a better operation of the system [42].

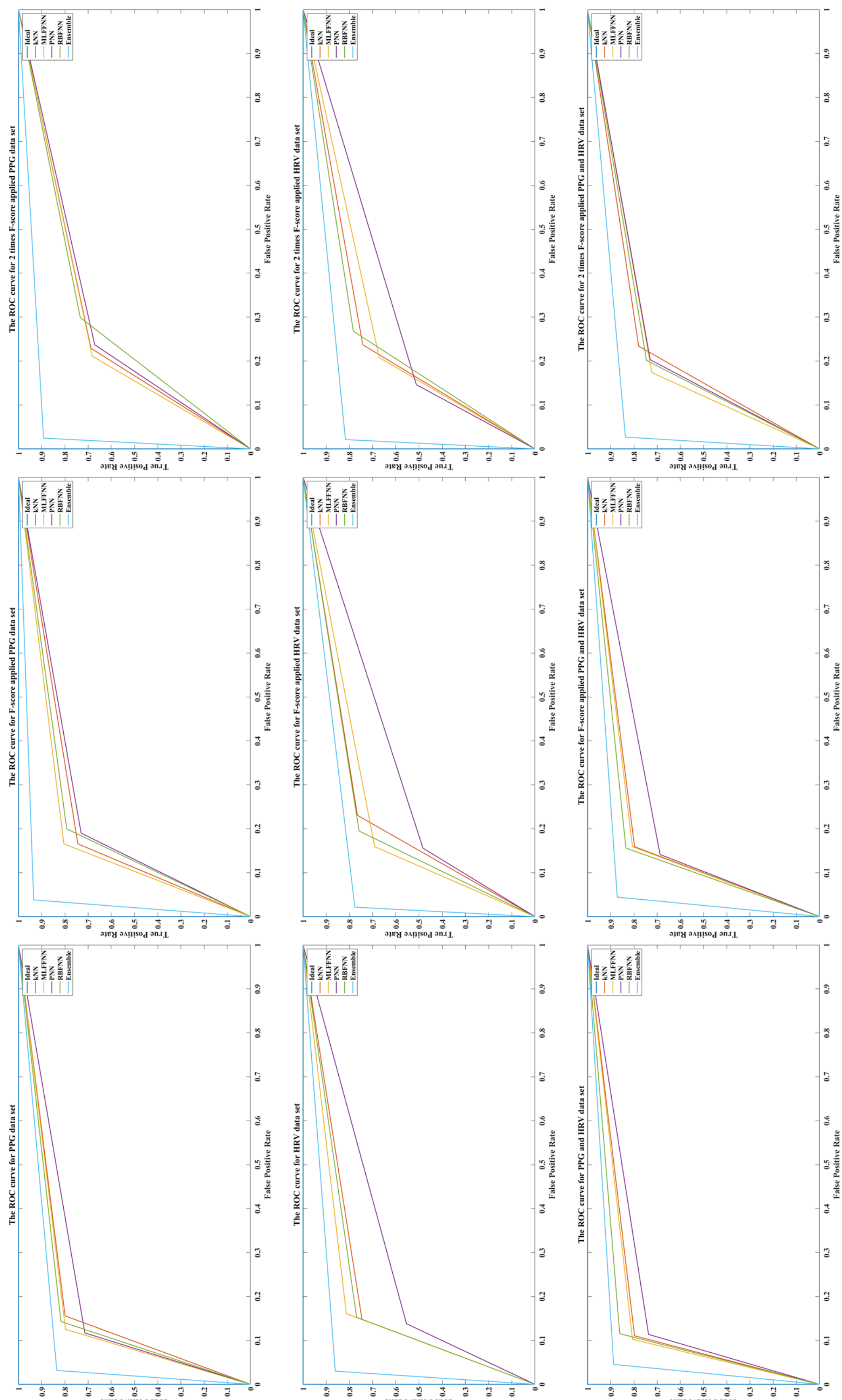

Fig. 10 The ROC curve for all datasets

Table 11 Classification results for principal component analysis

Method	All features					Using PCA				
	Train (70%)					Train (70%)				
Classifier	S	Spe	A (%)	K	FM	S	Spe	A (%)	K	FM
kNN	0.816	0.883	85.28	0.702	0.848	1	0.576	76.86	0.552	0.731
MLFFNN	0.793	0.922	86.37	0.723	0.853	0.024	0.991	55.18	0.016	0.047
PNN	1	1	100	1	1	0.828	0.693	75.41	0.512	0.754
SVMs	0.897	0.927	91.34	0.825	0.912	0.861	0.155	47.61	0.016	0.263
Ensemble	0.855	0.964	91.46	0.826	0.906	0.7	0.815	76.26	0.518	0.753
Method	All features					Using PCA				
	Test (20%)					Test (20%)				
Classifier	S	Spe	A (%)	K	FM	S	Spe	A (%)	K	FM
kNN	0.836	0.871	85.23	0.705	0.853	0.744	0.388	57.59	0.135	0.51
MLFFNN	0.764	0.875	81.65	0.634	0.816	0.036	0.973	47.89	0.009	0.069
PNN	0.848	0.853	85.02	0.7	0.85	0.524	0.786	64.77	0.305	0.629
SVMs	0.848	0.893	86.92	0.738	0.87	0.844	0.188	53.38	0.033	0.307
Ensemble	0.8	0.946	86.92	0.74	0.867	0.464	0.808	62.66	0.266	0.589
Method	All features					Using PCA				
	Validation (10%)					Validation (10%)				
Classifier	S	Spe	A (%)	K	FM	S	Spe	A (%)	K	FM
kNN	0.8	0.898	84.55	0.692	0.846	0.768	0.463	62.66	0.235	0.578
MLFFNN	0.752	0.852	79.83	0.598	0.799	0.048	0.944	46.35	− 0.01	0.091
PNN	0.88	0.843	86.27	0.724	0.861	0.496	0.787	63.09	0.276	0.609
SVMs	0.808	0.88	84.12	0.683	0.842	0.848	0.176	53.65	0.025	0.291
Ensemble	0.784	0.926	84.98	0.702	0.849	0.48	0.796	62.66	0.269	0.599

kNN k-nearest neighbor algorithm, *MLFFNN* multilayer feed forward artificial neural networks, *PNN* probabilistic artificial neural networks, *SVMs* support vector machines, *S* sensitivitiy, *Spe* specificity, *A* (%) accuracy, *K* kappa, *FM* F-measure

Table 12 Summary classifier results of PPG and HRV properties

Ensemble classifier						
Signal	PPG					
	NF = 16			NF = 4		
	S	Sp	A (%)	S	Sp	A (%)
Apnea	0.93	0.96	95	0.89	0.98	93.81
Control	0.96	0.93		0.98	0.89	
AUC	0.95			0.93		
Kappa	0.9			0.87		
F-measure	0.95			0.93		
k(10)-fold (%)	–			–		

NF number of features, *S* sensitivity, *Sp* specificity, *A* accuracy

This study reached the conclusion that PPG signals may be used in OSA diagnosis by processing them with machine learning and signal processing methods. Many different signals have been used in the literature to diagnose OSA [1]. PPG measurement is easier than other signals. Thus, the patient will not be disturbed during the measurement.

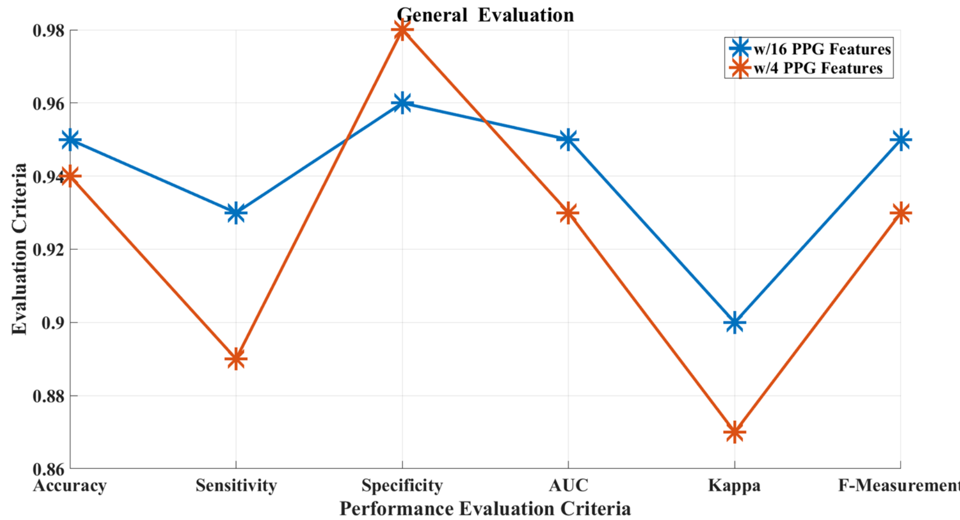

Fig. 11 General evaluation chart for classification results

Conclusions

This study suggests that the PPG signal and HRV can be used to detection respiratory arrestsby machine learning and signal processing techniques. In the literature, quite different signals and combinations are used for this processing. However, measuring the signal to be used with easy and noninvasive methods will reduce the discomfort experienced by the patient.

The work is open to improvement. The features derived in this study may be used to establish OSA diagnosis system. This way, the required workload for diagnosis can be reduced. This method can be applied for real-time analysis. Additionally, the patient can use the device alone. Home-usability of the system is a different benefit. Fast diagnosis without waiting in line will help faster start of the treatment. This way, the harm created by OSA on the human body in time will be prevented.

The study has several advantages. The advantages can be summarized as follows. The system has an embedded system design that can work in real-time. Real-time operation is the most significant advantage of the system. It has advanced signal processing algorithm infrastructure. The workload for a real-time system is minimized with 16 PPG features. Compared to other signals in the literature, PPG can be easily measured. In this way, the system provides comfort to the patient. The system can be used at home. The patient can make connections without the need for technical knowledge. When all advantages are evaluated, the system can be used in practice.

Future work

This study can be improved by combining different methods. The most important methods are advanced signal processing methods, feature extraction methods, feature selection algorithms and machine learning algorithms.

The studies planned for the future are as follows.

- Advanced signal processing and feature extraction technique design
- Detection of respiratory arrest with deep learning
- Converting signal into frequency domain and using it in image format

Funding This research was supported by The Scientific and Technical Research Council of Turkey (TUBITAK) through The Research Support Programs Directorate (ARDEB) with project number of 115E657, and project name of "A New System for Diagnosing Obstructive Sleep Apnea Syndrome by Automatic Sleep Staging Using Photoplethysmography (PPG) Signals and Breathing Scoring" and by The Coordination Unit of Scientific Research Projects of Sakarya University.

Compliance with ethical standards

Conflict of interest There is no conflict of interest between the authors.

Ethical approval The ethics committee report numbered 16214662/050.01.04/70 from Sakarya University Deanship of Faculty of Medicine, and the data use permission numbered 94556916/904/151.5815 from T.C. Ministry of Health Turkey Public Hospitals Authority Sakarya Province General Secretariat of Association of Public Hospitals were received to perform the study. 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. Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF, Redline S, Strohl KP, Davidson Ward SL, Tangredi MM (2012) Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med 8(5):597–619
2. Al-Khammasi S, Aboalayon KAI, Daneshzand M, Faezipour M, Faezipour M (2017) Hardware-based FIR filter implementations for ECG signal denoising: A monitoring framework from industrial electronics perspective. In: 2016 annual connecticut conference on industrial electronics, technology and automation, CT-IETA 2016, IEEE, pp 1–6
3. Changyue S, Kaibo L, Xi Z, Lili C, Xiaochen X (2015) An obstructive sleep apnea detection approach using a discriminative hidden markov model from ECG signals. IEEE Trans Biomed Eng 63:1532
4. Bruyneel M, Ninane V (2014) Unattended home-based polysomnography for sleep disordered breathing: current concepts and perspectives. Sleep Med Rev 18(4):341–7
5. Uçar MK, Bozkurt MR, Bilgin C, Polat K (2016) Automatic detection of respiratory arrests in OSA patients using PPG and machine learning techniques. Neural Comput Appl 28:2931–2945
6. Jezzini A, Ayache M, Elkhansa L, Ibrahim ZAA (2015) ECG classification for sleep apnea detection. In: 2015 international conference on advances in biomedical engineering (ICABME), IEEE, pp 301–304
7. Xie B, Minn H (2012) Real-time sleep apnea detection by classifier combination. IEEE Trans Inf Technol Biomed 16(3):469–477
8. Lazaro J, Gil E, Vergara JM, Laguna P (2012) OSAS detection in children by using PPG amplitude fluctuation decreases and pulse rate variability. Comput Cardiol 2012:185–188
9. Kim MS, Cho YC, Seo S-T, Son C-S, Kim Y-N (2012) Comparison of heart rate variability (HRV) and nasal pressure in obstructive sleep apnea (OSA) patients during sleep apnea. Measurement 45(5):993–1000
10. Penzel T, Kantelhardt JW, Becker HF, Peter JH, Bunde A (2003) Detrended fluctuation analysis and spectral analysis of heart rate variability for sleep stage and sleep apnea identification. In: Computers in cardiology, 2003, IEEE, pp 307–310
11. Gaurav G, Mohanasankar S, Jagadeesh Kumar V (2013) Apnea sensing using photoplethysmography. In: 2013 seventh international conference on sensing technology (ICST), IEEE, pp 285–288
12. Yilmaz B, Asyali MH, Arikan E, Yetkin S, Ozgen F (2010) Sleep stage and obstructive apneaic epoch classification using single-lead ECG. Biomed Eng 9(1):39
13. Rolón RE, Larrateguy LD, Di Persia LE, Spies RD, Rufiner HL (2017) Discriminative methods based on sparse representations of pulse oximetry signals for sleep apnea-hypopnea detection. Biomed Signal Process Control 33:358–367
14. Hassan AR (2016) Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting
15. Hassan AR, Aynal Haque M (2017) An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting. Neurocomputing 235:122–130
16. Sharma H, Sharma KK (2016) An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions. Comput Biol Med 77:116–124
17. Karmakar C, Khandoker A, Penzel T, Schobel C, Palaniswami M (2014) Detection of respiratory arousals using photoplethysmography (PPG) signal in sleep apnea patients. IEEE J Biomed Health Inform 18(3):1065–1073
18. Annakkaya AN, Balbay Ö, Bilgin C, Erbafi M, Arbak P (2004) Outcomes of the eight-channel ambulatory recording unit in a sleep laboratory. Eurasian J Pulmonol 6(1):12–23
19. Fang Y, Jiang Z, Wang H (2018) A novel sleep respiratory rate detection method for obstructive sleep apnea based on characteristic moment waveform. J Healthc Eng 2018:1–10
20. Almazaydeh L, Elleithy K, Faezipour M, Abushakra A (2013) Apnea detection based on respiratory signal classification. Procedia Comput Sci 21:310–316
21. Köktürk O (1998) Uykuda solunum bozuklukları. Tarihçe, tanımlar, hastalık spektrumu ve boyutu. Tüberküloz ve Toraks Dergisi 46(2):187–192
22. Ucar MK, Bozkurt MR, Polat K, Bilgin C (2015) Investigation of effects of time domain features of the photoplethysmography (PPG) signal on sleep respiratory arrests. In: 2015 23nd signal processing and communications applications conference (SIU), IEEE, pp 124–127
23. Wallisch P, Lusignan ME, Benayoun MD, Baker TI, Dickey AS, Hatsopoulos NG (2014) MATLAB for neuroscientists. Elsevier, Amsterdam
24. Shi P, Zhu Y, Allen J, Sijung H (2009) Analysis of pulse rate variability derived from photoplethysmography with the combination of lagged Poincaré plots and spectral characteristics. Med Eng Phys 31(7):866–71
25. Dehkordi P, Garde A, Karlen W, Wensley D, Ansermino JM, Dumont GA (2014) Sleep stage classification in children using photoplethysmogram pulse rate variability. In: Computing in cardiology conference (CinC), IEEE, pp 297–300
26. Polat K, Güneş S (2009) A new feature selection method on classification of medical datasets: kernel F-score feature selection. Expert Syst Appl 36(7):10367–10373
27. Şahan S, Polat K, Kodaz H, Güneş S (2007) A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. Comput Biol Med 37(3):415–23
28. Khan M, Ding Q, Perrizo W (2002) k-Nearest neighbor classification on spatial data streams using P-trees. Springer, Berlin, pp 517–528
29. Farajzadeh J, Fard AF, Lotfi S (2014) Modeling of monthly rainfall and runoff of Urmia lake basin using "feed-forward neural network" and "time series analysis" model. Water Resour Ind 7–8:38–48
30. Zounemat-kermani M, Kisi O, Rajaee T (2013) Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff. Appl Soft Comput 13(12):4633–4644
31. Khamis A, Shareef H, Mohamed A, Bizkevelci E (2015) Islanding detection in a distributed generation integrated power system using phase space technique and probabilistic neural network. Neurocomputing 148:587–599
32. Parzen E (1962) On estimation of a probability density function and mode. Ann Math Stat 33(3):1065–1076
33. Wasserman PD (1993) Advanced methods in neural computing, 1st edn. Wiley, New York
34. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297
35. Mandhala VN, Sujatha V, B Renuka Devi (2014) Scene classification using support vector machines. In: 2014 IEEE international conference on advanced communications, control and computing technologies, IEEE, pp 1807–1810

36. Rokach L, Schclar A, Itach E (2014) Ensemble methods for multi-label classification. Expert Syst Appl 41(16):7507–7523
37. Ramachandran KM, Tsokos CP (2015) Mathematical statistics with applications in R. Elsevier, Boston
38. Alpar R (2010) Applied statistic and validation—reliability. Detay Publishing, Ankara
39. Uçar MK, Bozkurt MR, Bilgin C, Polat K (2018) Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques. Neural Comput Appl 29:1–16
40. Addison PS (2017) Respiratory effort from the photoplethysmogram. Med Eng Phys 41:9–18
41. Orphanidou C (2017) Derivation of respiration rate from ambulatory ECG and PPG using ensemble empirical mode decomposition: comparison and fusion. Comput Biol Med 81:45–54
42. Liu Z, Sun J, Zhang Y, Rolfe P (2016) Sleep staging from the EEG signal using multi-domain feature extraction. Biomed Signal Process Control 30:86–97