Introducing the hybrid "K-means, RLS" learning for the RBF network in obstructive apnea disease detection using Dual-tree complex wavelet transform based features

Javad Ostadieh [1,2], Mehdi Chehel Amirani [2]

1. Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran.
2. E-mail any correspondence to: j.ostadieh@urmia.ac.ir

Abstract

Apnea is one of the deadliest diseases that can be prevented and cured if it is detected in time. In this paper, we propose a precise method for early detection of the obstructive sleep apnea (OSA) disease using the latest feature selection and extraction methods. The feature selection in this paper is based on the Dual tree complex wavelet (DT-CWT) coefficients of the ECG signals of several patients. The feature extraction from these coefficients is done using frequency and time techniques. The Feature selection is done using the spectral regression discriminant analysis (SRDA) algorithm and the classification is performed using the hybrid RBF network. A hybrid RBF neural network is introduced in this paper for detecting apnea that is much less computationally demanding than the previously presented SVM networks. Our findings showed a 3 percent improvement in the detection and at least a 30 percent reduction in the computational complexity in comparison with methods that have been presented recently.

Keywords: Electrocardiogram; Sleep Apnea; RBF network; Wavelet; DT-CWT; Feature extraction.

Introduction

Sleeping takes at least a third of the lifetime of every healthy human being and its quality affects the daily actions directly. Many people suffer from the respiration disorders during their sleep. The most alarming case of these disorders is the obstructive sleep apnea (OSA) [1]. The OSA can be described by the respiration cessation during sleep.

Table 1. List of the used abbreviations.

Abbreviations	Descriptions
OSA	Obstructive sleep apnea
ECG	Electrocardiogram
EDR	ECG-Derived Respiration
AHI	Apnea-Hypopnea Index
HMM	Hidden Markov model
RUSBoost	Random under-sampling Boost
Adaboost	Adaptive boost
DWT	Discrete wavelet transform
TQWT	Tunable Q-factor wavelet transform
LDA/QDA	Linear/Quadratic Discriminant Analysis
SFFS	Sequential forward feature selection
SRDA	Spectral regression discriminant analysis
DNN/CNN	Deep/Convolutional neural network
DT classifier	Decision tree classifier
RBF	Radial basis function
SVM	Support vector machine
RLS	Recursive least squares
GS	Gram-Schmidt
STLF	Short-time load forecasting
DT-CWT	Dual-tree complex wavelet transform

This problem can cause heart and brain strokes during sleep and many daytime problems. The improvement in detecting the OSA can improve the daily life of many patients and save the lives of many others. The OSA detection has been the investigation topic of many researchers [1-19]. As the ECG

signal acquisition is much faster and less computationally demanding, the OSA can be detected from analysis of short durations of the ECG signals and this helps the designing of the home setting of handled devices that can easily be used by the patients. The detection of the OSA with the single lead ECG signals can be performed using various strategies. The main measurement to distinguish between the Apnea and normal ECG signals is the Apnea-Hypopnea Index (AHI). AHI is defined by dividing the total number of apnea events by the total number of minutes of sleep time, multiplied by 60:

$$AHI = \frac{Total\ number\ of\ (Apnea+Hypoapnea)}{Total\ sleep\ time\ (in\ minutes)} \times 60 \quad (1)$$

Using this index we can mark the apnea ECG signals when $AHI > 5$ and normal ECG when $AHI \leq 5$. Here, we address some of the methods that have been suggested in the literature for the OSA detection process:

Reference [1] studies the usage of ECG signals in the detection of OSA. In this reference, the method is based on extracting 8 levels of wavelet features from ECG signals and then calculating 12 statistical and entropy-based features from these coefficients. The resulting features are then fed to a family of SVM classifiers (namely the least-square SVM or the Least-squares support vector machine (LS-SVM) and the library for support vector machine (LIB-SVM) which uses the linear kernels) and the results have been very satisfying.

In reference [2] the authors have proposed the Hidden Markov model (HMM) for extracting the correlation feature from the ECG signals and the SVM is used to classify these features.

In [3] the Tunable Q-factor wavelet transform (TQWT) is used instead of the ordinary wavelet transform and the statistical features are extracted from the coefficients of this transform. The RUSBoost classifier has been applied to these features.

The authors of [4] have proposed the usage of several frequency-based features like the Cepstrum, filter bank and Detrended Fluctuation Analysis (DFA) for detecting the OSA. The Logistic Regression (LR), Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) have been used for classification.

The contribution of [5] in OSA detection is the usage of Bootstrap classifier.

The authors in [6] have used the TQWT features and the Adaboost classifier in their method.

In [7] the proposed method is based on the time features of the ECG like the R-R interval and QRS properties and 5 different classifiers have been used for classifying these features.

In reference [8] have used the frequency features like the Spectogram and the statistical analysis to classify these features.

In [9] the ECG-Derived Respiration (EDR) signals along with the RR intervals are extracted from the ECG and the multiple classifiers such as the KNN, SVM, NN and LD-QD have been used to evaluate the results.

Reference [10] is based on cardiopulmonary coupling (CPC) and Respiratory event index (REI) features and statistical analysis for classification.

Reference [11] is based on the feature extraction with TQWT and Random forest classifier. Recently several references have proposed the usage of deep neural networks (DNNs) and convolutional neural networks (CNNs) for the OSA classification. The presented results with these classifiers have been very high, however, the computational complexity of these classifiers makes their usage unpopular.

Reference [12] has used the CNN to classify the R-R interval base features of the ECG signal.

Also, reference [13] has proposed CNN classifier based on the Scalogram features to detect OSA.

In [14] and [15] each has used a modified CNN to classify ECG time features.

Reference [16] has proposed the filter bank based features and their entropies to classify ECG signals with KNN, SVM, Decision tree (DT) classifiers.

Reference [17] is based on the Ensemble, Adaboost and Random forest classifiers.

Reference [18] has used the PCA feature reduction to reduce the DWT extracted features from the ECG and give them to the SVM classifier to detect OSA.

Reference [19] is based on wavelet packet decomposition (WPD), Entropy and R-R interval features and the random forest classifier.

Finally, reference [20] has proposed the feature extraction from ECG signals and in this paper, based on this novelty, we proposed the Dual-tree complex wavelet transform (DT-CWT) extracted features from the ECG to detect the OSA.

The DT-CWT is designed to compute the complex transform of a signal by two individual DWT decompositions. These transforms form two branches as tree **a** and tree **b**. If the filters in these branches are designed in a special manner, the output of one of the branches becomes the real coefficients of the DWT and the other branch produces the imaginary coefficients. By this transform, we get extra information in comparison with the conventional DWT. However, this extra information is achieved by extra computational demand that we can control by choosing the proper number of DT-CWT levels.

To the best of our knowledge, this is the first time to use DT-CWT for apnea detection purposes. While the DT-CWT extracted features are more promising than those of the DWT, the main hindrance in the usage of the DT-CWT extracted features instead of DWT features, is the computational complexity of the DT-CWT. However, because we wanted to compare our results with [1] and in that reference, the features have been extracted from 8 levels of DWT, we show in this paper that the suitable features with

the DT-CWT can be achieved with only 3 levels of this transform and with this we compensate the double computations complexity in comparison with 8 levels of DWT. The overall flowchart of the proposed method in this paper is given in Fig. 1:
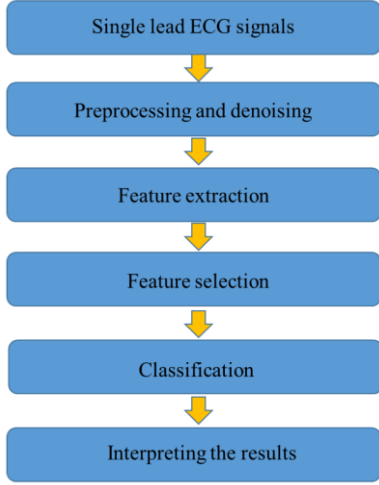

Fig. 1. The overall steps of the OSA detection with the help of ECG signals.

We tried to use novel methods in all the parts of the proposed method. In the feature selection part, we used the powerful SRDA algorithm [23] and for the classification, we used the hybrid RBF network using the 'K-means, RLS' learning [21,22] that is more powerful than the SVM network. The rest of this paper designed as follows:

In part II, we explain the feature extraction and selection techniques from our databases. Part III is dedicated to the explanation of the hybrid RBF classifier and its differences with the SVM network. Part IV presents our OSA detection results and part V consists of our concluding remarks and the suggestions for the future research.

ECG signal processing

The database that we implemented our proposed method on is the Physionet apnea ECG data set [24]. The age of subjects in this database is in the range of 27–63 years, and the weights of the subjects are between 35–135 kg. The AHI range in the extracted ECG signals is between 0 and 93.5. There are totally 70 records in this dataset that are clustered into two groups: train (called released-set) and test (called withheld-set). The released-set has 35 records with the indexes of a01-a20, b01-b05 and c01-c10, the withheld-set has 35 records with the indexes of x01-x35.
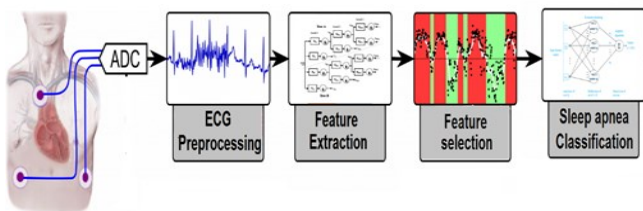

Fig. 2. The proposed OSA detection method.

Preprocessing

Signal segmentation with the time duration of 60 s is the first stage of preprocessing. It has been shown in [25] that the 60 seconds duration is the most suitable period for apnea classification. After this, we have the noise cancelation procedure where the baseline wandering and power line interference of the signals are canceled using a Chebyshev type II band-pass filter with the low and high cut-off frequencies of 0.5 Hz and 48 Hz, respectively. Fig. 3, shows only 3 seconds of a model record of the Physionet apnea database [24].
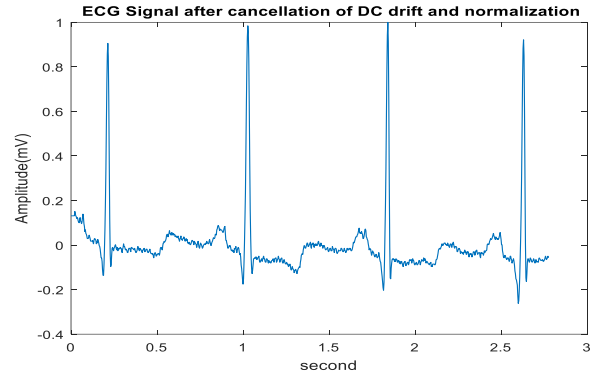

Fig. 3. The first 3 seconds of the apnea ECG from an example record.

Segmentation

After we performed filtering, the weight calculation approach is applied for deleting the noisy segments. In [1] a simple method is proposed for the automatic cancelation of the noisy parts. In this method, a weight (W) is calculated for each segment based on the similarity of its Autocorrelation Function (ACF) with other segments ACF, by taking into account the cosine pair-wise similarity as the metric. The similarity values are then given as:

$$d_{st} = \frac{(X_s - \bar{X}_s)(X_t - \bar{X}_t)'}{\sqrt{(X_s - \bar{X}_s)(X_s - \bar{X}_s)'}\sqrt{(X_t - \bar{X}_t)(X_t - \bar{X}_t)'}} \qquad (2)$$

where, d_{st} is the correlation distance, and X_s and X_t are the ACF of two different segments. $\bar{X}_s$ and $\bar{X}_t$ are the mean value of the X_s and X_t, respectively. The output of (1) is a vector showing the similarity of each segment. W of each segment is calculated with the normalized summation of all the d_{st} values. For our method, all the segments with a calculated weight lower than 0.8 were pointed out as the noisy segments. As the results of this paper are presented for the segment-by-segment (minute-by-minute) case rather than the subject-by-subject case, the selection of these segments is very important.

Feature extraction and selection

Feature extraction is a vital part of the computerized disease detection process. This section consists of three parts, in part A, we explain the DT-CWT feature extraction, in part B, we introduce the methods of statistical feature extraction from

the DT-CWT coefficients. In part C, the feature reduction or selection method which is the SRDA is explained.

The Dual-Tree Complex Wavelet Transform (DT-CWT)

In [20], the usage of DT-CWT in ECG feature extraction is proposed. The main deficiency of DWT based feature extraction in analyzing 1D ECG signals is the lack of shift invariance. It means that the amplitude of the wavelet coefficients varies substantially as the input signal is shifted a little. This happens because of the down sampling operation at each level. A better way of achieving shift invariance is to implement the undecimated form of the dyadic filter tree. However, this method has heavy computation demands and high redundancy in the output.

The DT-CWT tackles this problem with a redundancy factor for 1D signal, which is significantly lower than the undecimated DWT. In [20], the authors have explained the shift invariance property of DT-CWT in detail. The DT-CWT implements two trees of real filters (Tree A and Tree B) as shown in Fig. 1. The two trees correspond to the real and complex part of the complex wavelet transform. The DTCWT of a signal $x(n)$ (ECG) is implemented using two critically sampled DWTs in parallel to the same data. The filters are designed so that the sub band signals of the upper DWT can be interpreted as the real part of a complex wavelet transform and sub band signals of the lower DWT can be interpreted as the imaginary part. When the transform is designed in this manner, the DT-DWT is approximately shift invariant, unlike the critically sampled DWT. The filters implemented in each stage are of length 10. The sets of filter coefficients (H) used in this transform are given in [20]. The selected transform coefficients are x_{1a}, x_{01a}, x_{001a}, x_{000a}, x_{1b}, x_{01b}, x_{001b} and x_{000b}.
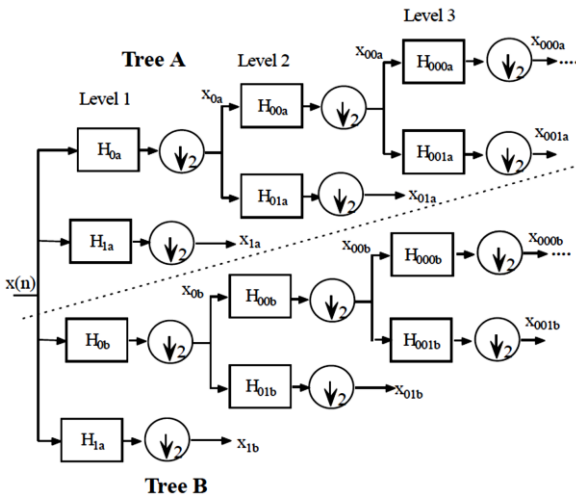

Fig. 4. The three level dual-tree complex wavelet transform.

In Fig. 5 and 6 we depicted the sub band signals for three levels of the tree A and B respectively. It is important to mention that all of these signals are depicted for the same model record of the Fig. 3 of the Physionet database. It is important to mention that for the following figures from Fig.5 to 8, the horizontal axis depicts the number of signal samples and the vertical axis shows the amplitude of the signal.
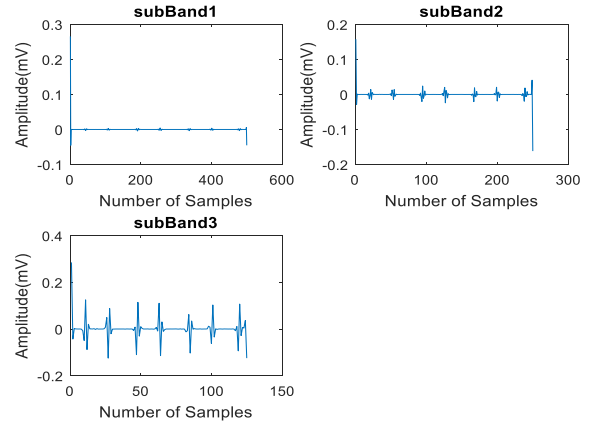

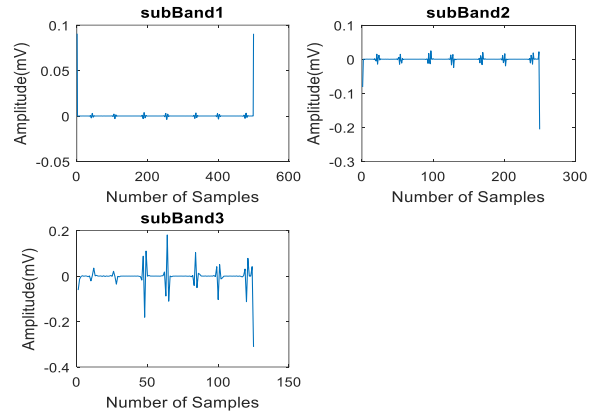Fig. 5. The sub bands of the ECG signal for Tree A.

Fig. 6. The sub bands of the ECG signal for Tree B.

The absolute energy of the signal x_{000a} is depicted in Fig. 7, and the absolute energy of x_{000b} is depicted in Fig. 8.
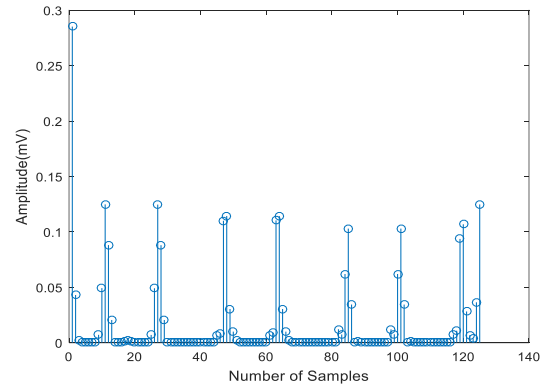

Fig. 7. The absolute energy of the sub band signal x_{000a}.
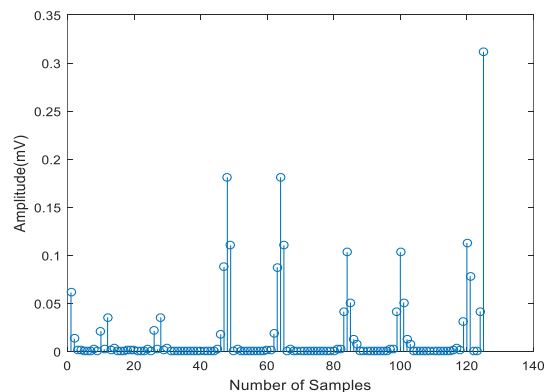

Fig. 8. The absolute energy of the sub band signal x_{000b}.

The non-linear feature extraction from the DT-CWT coefficients

After extracting the subbands of the DT-CWT from the selected ECG segments, we calculate some non-linear features based on the extracted transform coefficients. In [1] it has been shown that the ApEn, FE, IQR, RP and Poincare plot features make large differences among the two classes (Apnea and Normal). These features are collected in Table 2 and as they are explained in [1], we do not present their theoretical calculations here.

Table 2. List of non-linear features that are extracted from the DT-CWT coefficients in this paper.

Features	Description
FE	Fuzzy Entropy
ApEn	Approximate Entropy
IQR	Interquartile Range
RP	Recurrence Plot
SD1, SD2, SD1/SD2	Poincare Plot

Using these 7 feature extraction methods and with the 8 DT-CWT coefficients that are explained in part III, we have 56 features for each ECG to be fed to the classifier. However, we used the feature reduction to reduce these features as much as possible.

Spectral regression discriminant analysis (SRDA) algorithm

After the extraction of the final features from the DT-CWT coefficients, it is time to select the most suitable ones to reduce the volume of the features and the amount of computation that is needed for processing them. In [1], the sequential forward feature selection (SFFS) method has been proposed for this task that is based on the detection results. Here we propose a data based feature reduction method. SRDA is one of the most efficient methods for feature reduction. We implemented this technique in our proposed method. To start the SRDA algorithm, suppose we have a set of data points $x_1, \ldots x_m \in \mathbb{R}^N$ that belong to Nc different classes and m_k denotes the number of training samples of kth class $(\sum_{k=1}^{Nc} m_k = m)$. The steps of SRDA can be summarized as [23]:

① Let

$$y_k = \left[\underbrace{0, \ldots, 0}_{\sum_{i=1}^{k-1} m_i}, \underbrace{1, \ldots 1}_{m_k}, \underbrace{0, \ldots, 0}_{\sum_{i=k+1}^{Nc} m_i} \right]^T \quad k = 1, \ldots, Nc \quad (3)$$

and $y_0 = [1, 1, \ldots, 1]^T$ represents the vector of ones. The Gram-Schmidt process [23] is utilized for orthogonalization of $\{y_k\}$. Since y_0 is in the subspace described by $\{y_k\}$, the $Nc - 1$ vectors are obtained as:

$$\{\overline{y}_k\}_{k=1}^{Nc} , \left(\overline{y}_i^T y_0 = 0 \; where \; \overline{y}_i^T \overline{y}_j = 0, i \neq j \right) \quad (4)$$

② In this step, a new entry "l" is appended to each x_i which will also be shown as x_i. Then, $Nc - 1$ verctors $\{a_k\}_{k=1}^{Nc-1} \in \mathbb{R}^{N+1}$ are constructed, where a_k is the solution of the regularized least squares problem given in (5):

$$a_k = \arg\min_a \left(\sum_{i=l}^{m} \left(a^T x_i - \bar{y}_i^k \right)^2 + \alpha \|a\|^2 \right) \quad (5)$$

where $\bar{y}_i^k$ is the ith element of $\overline{y}_k$ and $\alpha \geq 0$ is a parameter to control the amount of reduction.

③ The $Nc - 1$ verctors $\{a_k\}$ are the basis vectors of SRDA. Let $A = [a_1, \ldots, a_{Nc-1}]$, which is a $(N + 1) \times (Nc - 1)$ transformation matrix. The x can be embedded into z in the $(Nc - 1)$ dimension subspace by:

$$z = A^T \begin{bmatrix} x \\ 1 \end{bmatrix} \quad (6)$$

The SRDA method reduces the 56 features that are mentioned in part *B*, based on the value of α which is between [0, 1].

The classifier and detector

We compared the results of our proposed method with that of several classifiers in part IV. Explaining the operation of all these classifiers would increase the volume of the paper inordinately. Therefore, we addressed them accordingly in Table 3 for the interested researchers to find their explanation in the references. Here we only explain the performance of our proposed classifying network:

RBF Classifier with hybrid 'K-means, RLS' learning

The SVMs are the most prevalently used classifiers in the field of the disease detection. The RBF networks, on the other hand, are not used as much as SVMs. The hybrid RBF network is the solution for this, because, they can rival the SVMs. The hybrid RBF [21,22] consists of three layers and the middle and the output layers work with the K-means and the RLS algorithms, respectively and for this reason, the hybrid adjective is attributed to them.

In this part, we describe the RBF classifier with hybrid learning that is our proposed classifying tool in this paper. We call the proposed classifier as hybrid RBF because it has a hybrid learning procedure with two stages as follows:

Stage 1: implements the K-means clustering algorithm to train the hidden layer in an unsupervised scheme. Usually, the number of clusters and the computational units in the hidden layer is notably smaller than the size of the train sample.

Stage 2: implements the recursive least-squares (RLS) algorithm (or another adaptive algorithm) to determine the weight vector of the linear output layer.
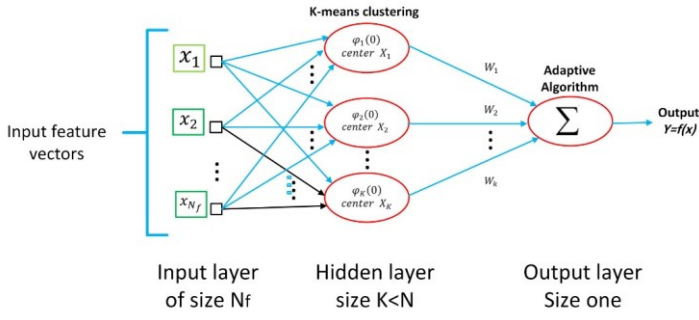
Fig. 9. The proposed hybrid RBF classifier.

The two-stage design procedure has some desirable features such as low computational complexity and fast convergence.

The RBF network consists of 3 layers as in Fig. 9. Here we describe them briefly [21]:

1. Input layer, which contains the source nodes that connected the network to its inputs. The inputs of the network for classification are the features vectors.

2. The second layer, consisting of hidden units, implements a nonlinear transformation from the input space to hidden (feature) space. For most applications, the dimensionality of the only hidden layer of the network is high; this layer is trained in an unsupervised manner using stage 1 of the hybrid learning scheme. Each unit in the hidden layer is described mathematically by a radial basis function:

$$\varphi_j(x) = \varphi\big(\|x - x_j\|\big) \qquad j = 1,2,\dots,N \qquad (7)$$

The jth input data point x_j identifies the center of the radial basis function and the vector x is the signal (pattern) applied to the input layer. Therefore, the links connecting the source nodes to the hidden units are direct connections with no weights. There are multiple radial basis functions for using in the hidden layer but we implement the Gaussian function for the sake of comparison between SVM and RBF in [1].

3. The output layer is linear and provides the response of the network to the activation pattern implemented to the input layer; this layer is trained in a supervised fashion using stage 2 of the hybrid scheme. There is no limitation on the size of the output layer, except that typically, the size of the output layer is much smaller than that of the hidden layer.

Here we describe learning algorithms of RBF:

K-means clustering
K-means is a method that utilizes distances for clustering with two steps [21, 22]:

Step 1: The total cluster variance is minimized with respect to the assigned set of cluster means $\{\hat{\mu}\}_{j=1}^K$, the following minimization must be performed:

$$\min_{\{\hat{\mu}\}_{j=1}^K} \sum_{j=1}^K \sum_{C(i)=j} \|x_i - \hat{\mu}_j\|^2 \qquad for\ a\ given\ C \qquad (8)$$

Step 2: After computing the optimized cluster means $\{\hat{\mu}\}_{j=1}^K$, we optimize the encoder as follows:

$$C(i) = arg\ \min_{1\le j\le K} \|x_i - \hat{\mu}_j\|^2 \qquad (9)$$

The RLS algorithm in hybrid learning
Adaptive algorithms have been designed to converge to certain weights. These weights in the RBF network are adjusted in the learning phase. The RLS algorithm is one of the most powerful adaptive algorithms. In this section, we explain the role of RLS in the output layer of the RBF network [21]. Let the $K \times 1$ vector:

$$\boldsymbol{\Phi}(x_i) = \begin{bmatrix} \varphi(x_i, \mu_1) \\ \varphi(x_i, \mu_2) \\ \vdots \\ \varphi(x_i, \mu_K) \end{bmatrix} \qquad (10)$$

represent the outputs of the K units in the hidden layer. This vector is constructed to respond to the stimulus $x_i, i = 1,2,\dots,N$. Thus, insofar as the supervised training of the output layer is concerned, the training sample is defined by $\{\boldsymbol{\Phi}(i), d(i)\}_{i=1}^N$, where d_i is the desired response at the overall output of the RBF network for input x_i. This training is implemented by the RLS algorithm described as below [21]:

Given the training sample $\{\boldsymbol{\Phi}(i), d(i)\}_{i=1}^N$, do the following calculations for iterations $n = 1,2,\dots,N$:

$$\boldsymbol{P}(n) = \boldsymbol{P}(n-1) - \frac{\boldsymbol{P}(n-1)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\boldsymbol{P}(n-1)}{1+\boldsymbol{\Phi}^T(n)\boldsymbol{P}(n-1)\boldsymbol{\Phi}(n)} \qquad (11)$$

$$\boldsymbol{g}(n) = \boldsymbol{P}(n)\boldsymbol{\Phi}(n) \qquad (12)$$

$$\alpha(n) = d(n) - \hat{\boldsymbol{w}}^T(n-1)\boldsymbol{\Phi}(n) \qquad (13)$$

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) + \boldsymbol{g}(n)\alpha(n) \qquad (14)$$

To initialize the algorithm, we have $\hat{\boldsymbol{w}}(0) = \boldsymbol{0}$ and $\boldsymbol{P}(0) = \lambda^{-1}\boldsymbol{I}$ where λ is a small positive constant.

In [22] a complete analysis was made to show the superiority of hybrid RBF to the SVM classifier both computationally and with respect to accuracy. Also, at least a 30% percent time saving is guaranteed using RBF in comparison with SVM [21]. This is important because we compared the results of our proposed method with that of the reference [1].

Ethical approval
The conducted research is not related to either human or animal use.

OSA detection results
We mentioned that after feature extraction we get 56 features from each ECG signal. One of the contributions of this paper is the usage of the SRDA feature reduction algorithm that gives the best results for the $0.1 \le \frac{\alpha}{\alpha+1} \le 0.4$

[23]. Using these values the number of features for the ECG signals reduces to 5 to 10 features which are substantially lesser than the 18 features of the proposed method in [1]. Our OSA detection results are presented based on the accuracy, sensitivity and specificity of the proposed methods that are given as follows [26]:

$$Accuracy \ (ACC) = \frac{Tp+TN}{TP+TN+FP+FN} \quad (15)$$

$$Sensititivity \ (Sen) = \frac{TP}{TP+FN} \quad (16)$$

$$Specificity \ (Spec) = \frac{TN}{TN+FP} \quad (17)$$

where, TP, TN, FP and FN denotes true positive, true negative, false positive and false negative, respectively. The comparison of the segment-by-segment (minute-by-minute) results of the proposed method and the results in several recent references are given in Table 3.

As we can see the proposed method can detect the OSA with overcoming results in comparison with the conventional classifiers and can closely rival the results of the computationally complex CNN classifiers. The main purpose of this paper was to improve the results in [1]. By comparing the results, we can see that an average of 31% improvement is achieved in all the performance metrics.

Concluding remarks

This paper aimed to propose the hybrid RBF classifier along with some novel ECG signal processing techniques to improve the OSA detection. The proposed methods are all selected in order to reduce the computational complexity and the time consumption of the detection process.

For the feature selection, we used 3 levels of the Dual Tree Complex Wavelet Transform (DT-CWT) instead of 8 levels of the DWT features. In addition, the computational complexity of the proposed spectral regression discriminant analysis (SRDA) feature selection algorithm and the hybrid RBF classifier can guarantee at least 30% of complexity reduction plus the improvement of the OSA detection accuracy. The result presented an average of 3% improvement in detection and a 30% of time and complexity reduction when compared to the previously presented methods.

In future works, we will consider the usage of more advanced feature extraction and reduction methods. Also, the usage of DNNs and CNNs based on the features that are extracted in this paper may cause the perfect OSA detection results.

Conflict of interest

Authors state no conflict of interest.

TABLE 3. The comparison of the OSA detection results based on various methods.

References	Feature extraction/selection method	Classifier	Results		
			ACC%	Sens%	Spec%
[1] Zarei 2018	DWT+SFFS	SVM (RBF kernel)	92.98	91.74	93.75
[2] Song 2016	HMM	HMM+SVM	86.2	82.6	88.4
[3] Hassan 2017	TQWT	RUSBoost	88.88	87.58	91.49
[4] Gonzalez 2017	Cepstrum+ Filter bank	QDA	84.76	81.45	86.82
[5] Hassan 2016	Statistical and spectral	Bootstrap aggregating	85.97	84.14	86.83
[6] Hassan 2016	Normal invers Gaussian modeling	AdaBoost	87.33	81.99	90.72
[7] Sharma 2016	QRS features	LS-SVM (RBF kernel)	83.8	79.5	88.4
[8] Hilmisson 2018	Frequency features	Statistical analysis	93	100	81
[9] Janbakhshi 2018	Time domain feaures+PSD	SVM-KNN-NN-LD-QD	90.9	89.6	91.8
[10] Ma 2019	Statistical features	Statistical analysis	87	89	79
[11] Nishad 2018	Tunable-Q wavelet transform features	Random Forest	92.78	93.91	90.95
[12] Wang 2019	RR-intervals	CNN (LeNet-5)	92.3	90.9	100
[13] Singh 2019	Time-frequency Scalogram features	CNN (AlexNet)	86.22	90	100
[14] Urtnasan 2018	RR-intervals	CNN	96	96	96
[15] Wang 2018	RR-intervals	CNN	97.8	100	93
[16] Sharma 2019	Fuzzy-entropy (FUEN) and the Log of signal-energy (LOEN)	KNN-DT-SVM	90.87	92.43	88.33
[17] Avci 2015	DWT+PCA	Random forest	92-98	-	-
[18] Rachim 2014	DWT+PCA	SVM	94.3	92.65	92.2
Proposed method	DT-CWT+SRDA	Hybrid "k-means, RLS" RBF	**95.62**	**96.37**	**96**

References

1. A. Zarei, B. Mohammadzadeh Asl, "Automatic Detection of Obstructive Sleep Apnea Using Wavelet Transform and Entropy based Features from Single-Lead ECG Signal", IEEE Biomedical and Health Informatics Journal, vol. 23, no. 3, pp. 1011-1021, 2019. https://doi.org/10.1109/jbhi.2018.2842919

2. C. Song, K. Liu, X. Zhang, L. Chen, and X. Xian, "An obstructive sleep apnea detection approach using a discriminative hidden markov model from ECG signals," IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1532–1542, 2016. https://doi.org/10.1109/tbme.2015.2498199

3. A. R. Hassan and M. A. Haque, "An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting," Neurocomputing, vol. 235, pp. 122–130, 2017. https://doi.org/10.1016/j.neucom.2016.12.062

4. S. Martín-González, J. L. Navarro-Mesa, G. Juliá-Serdá, J. F. Kraemer, N. Wessel, and A. G. Ravelo-García, "Heart rate variability feature selection in the presence of sleep apnea: An expert system for the characterization and detection of the disorder," Computers in Biology and Medicine, vol. 91, pp. 47-58, 2017. https://doi.org/10.1016/j.compbiomed.2017.10.004

5. A. R. Hassan and M. A. Haque, "Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating," Biocybernetics and Biomedical Engineering, vol. 36, no. 1, pp. 256–266, 2016. https://doi.org/10.1016/j.bbe.2015.11.003

6. A. R. Hassan, "Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting," Biomedical Signal Processing and Control, vol. 29, pp. 22–30, 2016. https://doi.org/10.1016/j.bspc.2016.05.009

7. H. Sharma and K. Sharma, "An algorithm for sleep apnea detection from single-lead ECG using hermite basis functions," Computers in Biology and Medicine, vol. 77, pp. 116–124, 2016. https://doi.org/10.1016/j.compbiomed.2016.08.012

8. H. Hilmisson, N. Lange, S.P. Duntley, "Sleep apnea detection: accuracy of using automated ECG analysis compared to manually scored polysomnography (apnea hypopnea index)" Sleep and Breathing, 23(1), pp. 125-133, 2019. https://doi.org/10.1007/s11325-018-1672-0

9. P. Janbakhshi, M.B. Shamsollahi. "Sleep Apnea Detection from Single-Lead ECG Using Features Based on ECG-Derived Respiration (EDR) Signals" IRBM 39, pp. 206–218, 2018. https://doi.org/10.1016/j.irbm.2018.03.002

10. Y. Ma, et al. "Electrocardiogram-based sleep analysis for sleep apnea screening and diagnosis" Sleep and Breathing, First Online: 21 June 2019. https://doi.org/10.1007/s11325-019-01874-8

11. A. Nishad, R.B. Pachori, U.R. Acharya "Application of TQWT based filter-bank for sleep apnea screening using ECG signals" Journal of Ambient Intelligence and Humanized Computing, pp 1–12, 2018. https://doi.org/10.1007/s12652-018-0867-3

12. T. Wang, et al. "Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network", PeerJ. 7: e7731, 2019. https://doi.org/10.7717/peerj.7731

13. S. A. Singh, et al. "A novel approach OSA detection using single lead ECG Scalogram based on deep neural network" Journal of Mechanics in Medicine and Biology, Vol. 19, No. 4, pp.1-18, 2019. https://doi.org/10.1142/s021951941950026x

14. E. Urtnasan, et al. "Automated Detection of Obstructive Sleep Apnea Events from a Single-Lead Electrocardiogram Using a Convolutional Neural Network", J Med Syst 42:104, 2018. https://doi.org/10.1007/s10916-018-0963-0

15. X. Wang, et al. "Obstructive sleep apnea detection using ecg-sensor with convolutional neural networks" Multimedia Tools Applications, First Online: 18 June 2018. https://doi.org/10.1007/s11042-018-6161-8

16. M. Sharma, M. Raval, U. R. Acharya "A new approach to identify obstructive sleep apnea using an optimal orthogonal wavelet filter bank with ECG signals" Informatics in Medicine Unlocked, 16, 2019. https://doi.org/10.1016/j.imu.2019.100170

17. C. Avcı and A. Akbaş, "Sleep apnea classification based on respiration signals by using ensemble methods", Bio-Medical Materials and Engineering, 26, pp. 1703–S1710, 2015. https://doi.org/10.3233/bme-151470

18. V. P. Rachim, G. Li and W. Chung, "Sleep apnea classification using ECG-signal wavelet-PCA features", Bio-Medical Materials and Engineering, 24, pp. 2875–2882, 2014. https://doi.org/10.3233/bme-141106

19. T. Li and M. Zhou, "ECG Classification Using Wavelet Packet Entropy and Random Forests", Entropy, 18, 285, 2016. https://doi.org/10.3390/e18080285

20. M. Thomas, M. Kr Das, S. Ari, "Automatic ECG arrhythmia classification using dual-tree complex wavelet based features", Int. J. Electron. Commun, Volume 69, Issue 4, Pages 715-721, April 2015. https://doi.org/10.1016/j.aeue.2014.12.013

21. S. Haykin, "Neural networks and learning machines". Prentice Hall, pp. 230-263, 2008.

22. R.M. Rifkin, "Everything old is new again: A fresh look at historical approaches in machine learning", Ph.D. thesis, MIT, 2002.

23. D. Cai, X. He, and Z. Han, "SRDA: An efficient algorithm for large-scale discriminant analysis", IEEE Trans. on Knowledge and Data Engineering, vol. 20, pp. 1-12, 2008. https://doi.org/10.1109/tkde.2007.190669

24. T. Penzel, G.B. Moody, R.G. Mark, AL Goldberger, JH Peter. "The Apnea-ECG Database." Computers in Cardiology, 27:255-258, 2000. https://doi.org/10.1109/cic.2000.898505

25. P. de Chazal, T. Penzel, and C. Heneghan, "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram," Physiological Measurement, vol. 25, no. 4, p. 967, 2004. https://doi.org/10.1088/0967-3334/25/4/015

26. A. Jafari Chashmi, M. C. Amirani, "An efficient and automatic ECG arrhythmia diagnosis system," J Electr Bioimp, 10, 47-54, 2019. https://doi.org/10.2478/joeb-2019-0007