Original Article

Enhancing Obstructive Apnea Disease Detection Using Dual-Tree Complex Wavelet Transform-Based Features and the Hybrid "K-Means, Recursive Least-Squares" Learning for the Radial Basis Function Network

Javad Ostadieh, Mehdi Chehel Amirani, Morteza Valizadeh[1]

Departments of Electrical Engineering and [1]Electrical and Computer Engineering, Urmia University, Urmia, Iran

Abstract

Background: The obstructive sleep apnea (OSA) detection has become a hot research topic because of the high risk of this disease. In this paper, we tested some powerful and low computational signal processing techniques for this task and compared their results with the recent achievements in OSA detection. **Methods:** The Dual-tree complex wavelet transform (DT-CWT) is used in this paper to extract feature coefficients. From these coefficients, eight non-linear features are extracted and then reduced by the Multi-cluster feature selection (MCFS) algorithm. The remaining features are applied to the hybrid "K-means, RLS" RBF network which is a low computational rival for the Support vector machine (SVM) networks family. **Results and Conclusion:** The results showed suitable OSA detection percentage near 96% with a reduced complexity of nearly one third of the previously presented SVM based methods.

Keywords: *Classification, feature reduction, hybrid K-means recursive least-squares, multi-cluster feature selection, obstructive sleep apnea, single-lead electrocardiogram*

Submitted: 29-Dec-2019 **Revised:** 30-Apr-2020 **Accepted:** 03-Jun-2020 **Published:** 11-Nov-2020

Introduction

There is a close relationship between the heartbeat and the breathing process. Therefore, we can use the electrocardiogram (ECG) signals to detect the breathing problems. Obstructive sleep apnea (OSA) disease is one of the most dangerous breathing deficiencies that happen during sleep and can be detected directly from the ECG signals using the signal processing techniques.[1] The usage of ECG signals in OSA detection is only one way to do this task and the apnea can be detected using respiration and other kinds of signals.[2] However, in this article, we only considered the OSA detection using the single-lead ECG. Several separated OSA detection methods have been proposed up until now.[1,3-18] Most of these methods have consisted of feature extraction, feature selection, and classifier parts. In Figure 1, we can see the collective flowchart of an OSA detection approach based on the ECG signal processing:

Here, we describe some of the proposed methods in the apnea detection. Khandoker *et al*.[3] have proposed the usage of the wavelet transform for ECG feature extraction. Furthermore, Rachim *et al*.,[9] Zarei *et al*.,[5] Avcı and Akbaş *et al*.,[8] and many other researchers have proposed the discrete wavelet transform (DWT)-based ECG decomposition for the OSA detection. Furthermore, the Tunable Q-factor wavelet transform is proposed in[7] by Nishad *et al*. Some researchers, including Hassan *et al*.[6,19] and Thomas *et al*.,[20] have proposed the usage of the dual-tree complex wavelet transform (DT-CWT) to extract the transform coefficients from the ECG signal. After, collecting the transform features, the usual path is to extract the statistical features from these coefficients. Based on these works, in this article, we used the DT-CWT for feature extraction of the apnea ECG signal. However, the nonlinear feature extraction from the DT-CWT coefficients is something that have not been experimented in Hilmisson *et al*.[1] and Hassan *et al*.[6]

How to cite this article: Ostadieh J, Amirani MC, Valizadeh M. Enhancing obstructive apnea disease detection using dual-tree complex wavelet transform-based features and the hybrid "K-means, recursive least-squares" learning for the radial basis function network. J Med Sign Sens 2020;10:219-27.

Address for correspondence:
Dr. Javad Ostadieh, Department of Electrical Engineering, Urmia University, Urmia, Iran.
E-mail: j.ostadieh@urmia.ac.ir

Access this article online
Website: www.jmssjournal.net
DOI: 10.4103/jmss.JMSS_69_19
Quick Response Code:

After feature extraction, feature reduction is necessary to reduce the computational complexity of the proposed method. Zarei *et al*.[5] have used the sequential forward feature selection method, whereas the principal component analysis have been suggested in Avcı and Akbaş *et al*.[8] and Rachim *et al*.[9] In this article, we proposed the multi-cluster feature selection (MCFS)[10] method for the feature reduction as much as possible for the best results.

The final part of the OSA detection process is the classification. Many researchers including Zarei *et al*.[5] have proposed the usage of the support vector machines (SVMs) for classifying between the apnea and normal ECG signal. Other classifiers like the Neural networks (NNs) have been proposed by Khandoker, *et al*.,[3] the Random forest by[7] and[8] the Adaboost classifier by,[11] Rusboost by,[12] Boot-strap by,[13] Convolutional NNs (CNNs) by[14-17] and Deep NNs (DNNs) by[18] for OSA detection. However, the SVMs are still the most prevalently used classifiers in this topic and we compared our results to these classifiers. It is important to mention that the articles that have proposed the usage of DNNs and CNNs, usually report near 100% OSA detection results that are not very credible due to the high computational complexity of theses network, also the deep networks are usually biased to the train data and do not perform well for the new data.[5] The proposed classifier in this article is the Hybrid Radial basis function (RBF) network with the "K-means-recursive least-squares (RLS)" learning algorithm.[21,22] The selection of this network is to compare its results with that of the SVM networks in the OSA detection. These networks have been compared before in other tasks,[21] and it has been shown that the hybrid RBF network is superior to the SVM network.

The rest of this paper prepared as follows:

In part II, the signal preprocessing steps are described. Part III, explains the feature extraction and selection methods from the apnea ECG signals. Part IV 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 investigations.

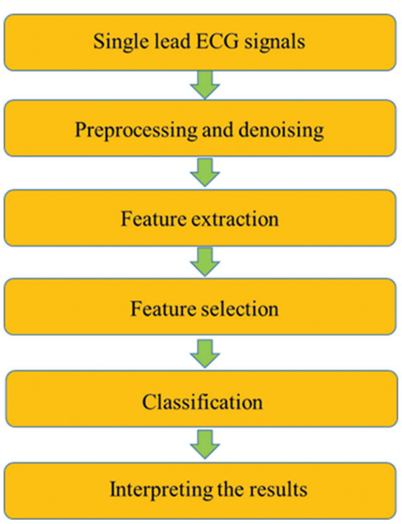

Figure 1: The overall steps of the obstructive sleep apnea detection using the single-lead electrocardiogram signals

Dealing with the Electrocardiogram Signal

In this section, we focus on the preliminaries of the ECG signal processing for the OSA detection. First, we introduce the data base that is used in this article because there are other ECG data bases that contain the heart signals of the apnea diseased patients. Then, we proceed to the preprocessing and signal preparation techniques.

Data base

The Physionet data base that is used for ECG apnea signals consists of the sleep duration parts of the 70 patients. In this data base, we have 35 records for the train set with 13 healthy (normal with Apnea Hypopnea Index ≤5) and 22 apnea participants. For the test set, we have 35 records with 12 healthy and 23 apnea cases. Furthermore, in this text file, the apnea or healthy condition of each segment is pointed out. Based on these we have presented our results in Table 1, with the segment by segment (minute by minute) assumption. As we mentioned, there are other apnea data bases[23] that we did not consider in this paper because of the comparison of the results with a long list of references.

Signal preparation

Before processing an ECG signal, we must first remove the power line interference and baseline wandering from it using a Chebyshev band-pass filter with the frequency range of 0.5–48 Hz. In Figure 2, we can see a typical ECG signal before and after interference cancelation and normalization.

Segmentation and segment selection

Not all the parts of the recorded ECG signal is useful for the OSA detection. Some parts become useless due to the movements of the patients and other parts may be contaminated with high noise levels. After we performed filtering, the weight calculation approach is applied for deleting the noisy segments. In[5] 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 pairwise similarity as the metric. The similarity values are then given as:

$$d_{st} = \frac{\left(X - \bar{X}_s\right)}{\sqrt{\left(X_s - \bar{X}_s\right)\left(X_s - \bar{X}_s\right)'}\sqrt{\left(X_t - \bar{X}_t\right)\left(X_t - \bar{X}_t\right)'}} \quad (1)$$

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

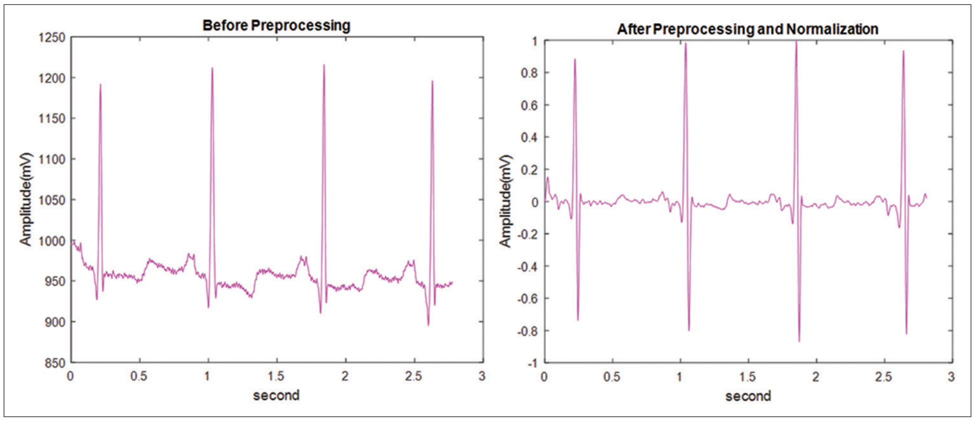

Figure 2: The first 3 s of a typical normalized electrocardiogram signal before (left) and after (right) preprocessing

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 values. For our method, all the d_{st} segments with a calculated weight lower than 0.8 were pointed out as the noisy segments and deleted from the tested segments.

Feature Extraction and Selection

The feature extraction of the ECG signal segments in this article consists of three parts: In Part A, we explain the DT-CWT feature extraction and in Part B, we introduce the methods of nonlinear feature extraction from the DT-CWT coefficients. In part C, the feature reduction or selection method which is the MCFS is explained.

The dual-tree complex wavelet transform

In[6,20] the usage of DT-CWT in ECG feature extraction is proposed. The main deficiency of DWT-based feature extraction in analyzing 1D ECG signal 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[21] 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 Figure 3. The two trees correspond to the real and complex part of the CWT. 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 CWT and subband 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.[21] The selected transform coefficients are x_{1a}, x_{01a}, x_{001a}, x_{000a}, x_{1b}, x_{01b}, x_{001b}, and x_{000b}.

In Figures 4 and 5, we depicted the subband 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 a01 record of the Physionet database.

The absolute energy of the signal x_{000b}, is depicted in Figure 6.

The absolute energy of the signal x_{000b}, is depicted in Figure 7.

The nonlinear feature extraction from the dual-tree complex wavelet transform coefficients

After extracting the subbands of the DT-CWT from the selected ECG segments, we calculate some nonlinear features based on the extracted transform coefficients. In[5] it has been shown that the ApEn, FE, interquartile range, RP and Poincare plot features make large differences among the two classes (Apnea and Normal). These features are collected in Table 1 and as they are explained in,[5] we do not present their theoretical calculations here.

Using these seven feature extraction methods 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.

The multicluster feature selection algorithm

The MCFS[10] is an algorithm that can reduce feature dimensionality without their class labels. It is a highly powerful algorithm for reducing the correlation between the features. Here, we describe this algorithm:

We assume there are N training samples each with N_p features, i.e., $x_i \in R^{N_p}$ that construct the matrix $X = (x_1, x_2, \ldots, x_N)$. The function of MCFS is to determine

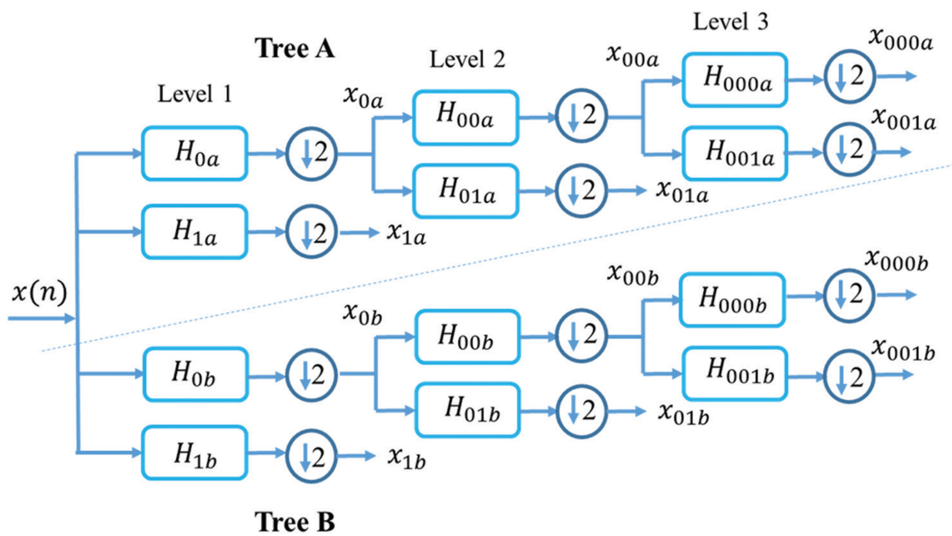

Figure 3: The three level dual-tree complex wavelet transform

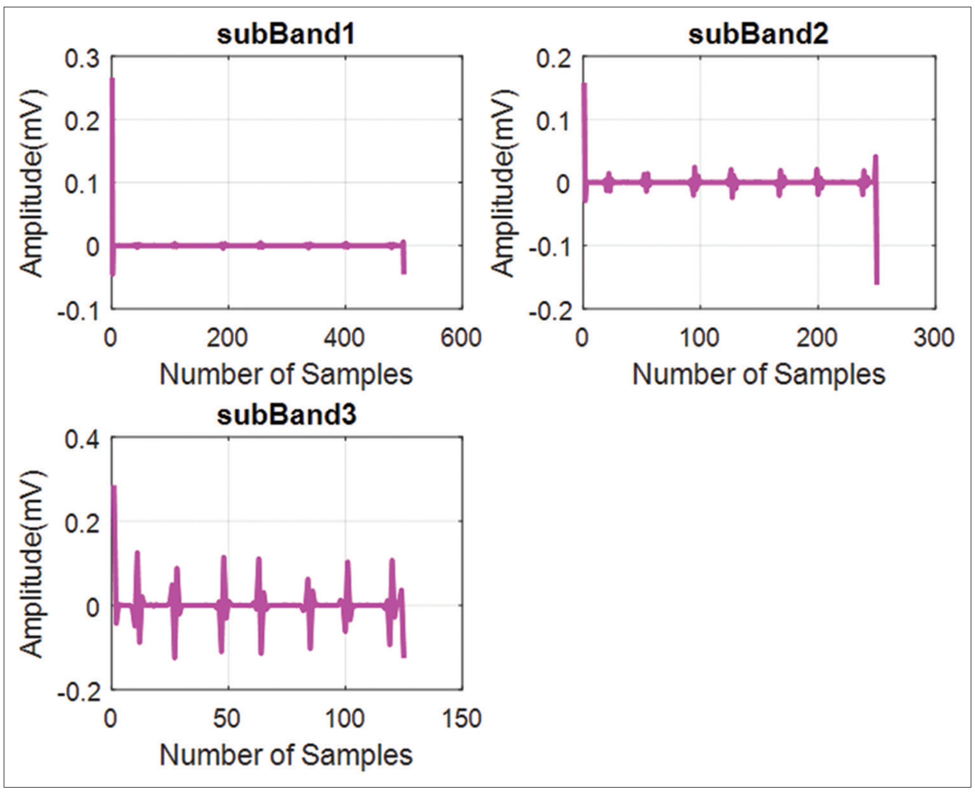

Figure 4: The sub bands of the electrocardiogram signal for Tree A

the feature subset with N_F information features from X. In other words, MCFS is able to generate the points $\left(x_1^{'}, x_2^{'}, \ldots, x_N^{'}\right)$ in R^{N_F} which preserves the geometric structure as the data in the original N_p-dimension space. Training samples build a graph with N vertices in which, each of the vertexes correspond to a training sample. MCFS at first finds p nearest for each training sample x_i and put edge between x_i and its neighbors. Then, we could define the weight matrix W on the graph. Next, diagonal matrix D with its column being the sum of W which was mentioned before is defined and the graph Laplacian is calculated in the manner:

$$L = W - D \tag{2}$$

To find the flat embedding for the data points, the generalized Eigen problem mentioned below must be solved:

$$Ly = \lambda Dy \tag{3}$$

Each row of = (y_1, , y_c, flat embedding for each data point, where y_i's are the eigen vectors of (2) with respect

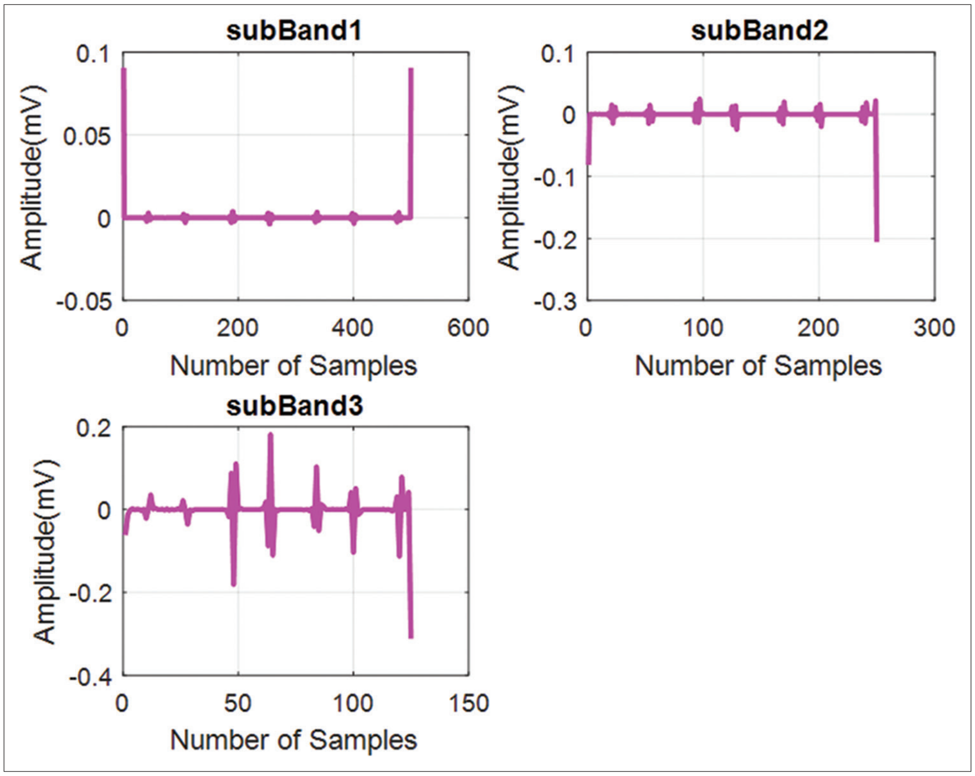

Figure 5: The sub bands of the electrocardiogram signal for Tree B

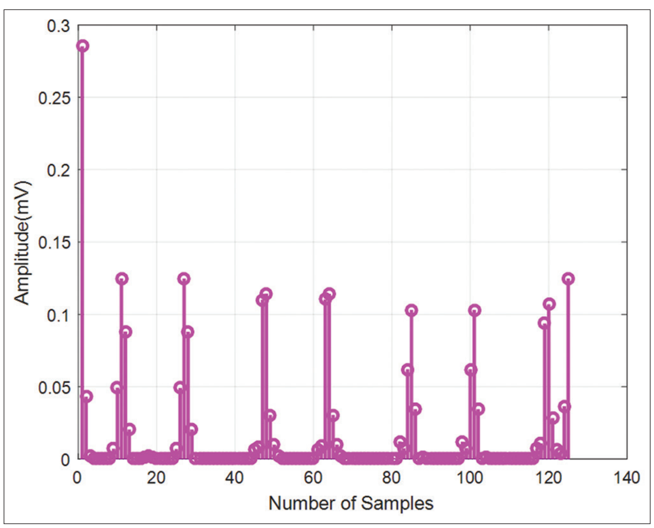

Figure 6: The absolute energy of the sub band signal x_{000a}

to smallest eigenvalue, C is the intrinsic dimensionality of data and data distribution reflected as y_i along the corresponding dimension. Minimizing the fitting error as follows could result in finding the relevant subset of the features:

$$\left\| y_k - X^T a_k \right\| \; Subject\ to \left| a_k \right| \leq \gamma \tag{4}$$

Where a_k is a N_p-dimension vector, $\left| a_k \right| = \sum_{j=1}^{N_p} \left| a_{k,j} \right|$ denotes the L1-norm of $a_k, a_{k,j}$ is the *j*th element of the vector a_k and γ is a parameter to be quantified. The least angel regression algorithm can be used to solve the optimization problem in (4). For every feature *j*, the MCFS score is defined as:

$$MCFS\ Score(j) = \left| a_{k,j} \right| \tag{5}$$

Then, all properties according to their MCFS scores are then sorted in the descending order with top N_F features being selected. The MCFS method reduces the 56 features that are mentioned in Part *B*.

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 2 for the interested researchers to find their explanation in the references. Here, we only explain the performance of our proposed classifying network:

Radial basis function classifier with hybrid "K-means, recursive least-squares" learning

The SVMs are the most prevalently used classifiers in the field of disease detection and classification. The RBF networks, on the other hand, are not used as much as SVMs. The hybrid RBF network[22] is the solution for this, because, they can rival the SVMs. The hybrid RBF 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 RBF classifier with the hybrid learning scheme that is our proposed classifying tool. We call the proposed classifier as hybrid RBF because it has a hybrid-learning procedure with two stages as follows:[22]

- 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 are notably smaller than the size of the train sample
- Stage 2: Implements the RLS algorithm (or another adaptive algorithm) to determine the weight vector of the linear output layer.

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

As we mentioned, the RBF network consists of three layers as in Figure 8. Here, we describe them briefly:

Table 1: List of nonlinear features that are extracted from the dual-tree complex wavelet transform coefficients in this article

Features	Description
FE[5]	Fuzzy entropy
ApEn[5]	Approximate entropy
IQR[5]	Interquartile range
RP[5]	Recurrence plot
SD1, SD2, SD1/SD2[5]	Poincare plot

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 hidden layer is described mathematically by a RBF:

$$\varphi_j(\boldsymbol{x}) = \varphi(\boldsymbol{x} - \boldsymbol{x}_j)\ j = 1, 2, \ldots, N \tag{6}$$

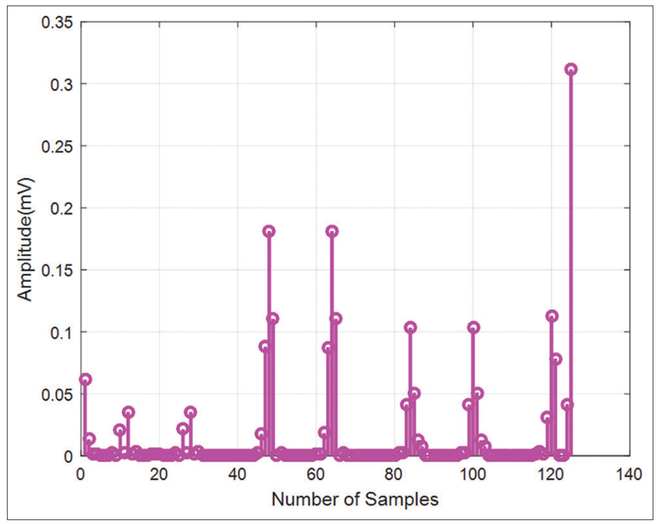

Figure 7: The absolute energy of the sub band signal x_{000b}

Table 2: The comparison of the obstructive sleep apnea detection results based on various methods

References	Feature extraction/selection method	Classifier	Results		
			Accuracy %	Sensitivity %	Specificity %
Hilmisson *et al*., 2018[1]	Frequency features	Statistical analysis	93	100	81
Janbakhshi *et al*., 2018[2]	EDR	SVM-KNN-NN-LD-QD	90.9	89.6	91.8
Ma *et al*., 2019[4]	Statistical features	Statistical analysis	87	89	79
Zarei and Asl 2018[5]	DWT + SFFS	SVM (RBF kernel)	92.98	91.74	93.75
Hassan *et al*., 2017[6]	DT-CWT	AdaBoost	84.4	90.38	74.38
Nishad *et al*., 2018[7]	Tunable-Q wavelet transform features	Random Forest	92.78	93.91	90.95
Avcı and Akbaş 2015[8]	DWT + PCA	Random forest	92-98	-	-
Rachim *et al*., 2014[9]	DWT + PCA	SVM	94.3	92.65	92.2
Hassan and Haque 2017[11]	Normal invers Gaussian modeling	AdaBoost	87.33	81.99	90.72
Hassan and Haque 2016[12]	TQWT	RUSBoost	88.88	87.58	91.49
Hassan 2016[13]	Statistical and spectral	Bootstrap aggregating	85.97	84.14	86.83
Wang *et al*., 2019[14]	RR-intervals	CNN (LeNet-5)	92.3	90.9	100
Singh *et al*., 2019[15]	Time-frequency Scalogram features	CNN (AlexNet)	86.22	90	100
Urtnasan *et al*., 2018[16]	RR-intervals	CNN	96	96	96
Wang *et al*., 2018[17]	RR-intervals	CNN	97.8	100	93
Wang *et al*., 2019[18]	RR-intervals and frequency features	DNN	97.1	100	91.7
Proposed method	DT-CWT + MCFS	Hybrid “k-means, RLS” RBF	95.62	96.37	96

ECG – Electrocardiogram; EDR – ECG derived respiratory; DWT – Discrete wavelet transform; SFFS – Sequential forward feature selection; PCA – Principal component analysis; TQWT – Tunable Q-factor wavelet transform; MCFS – Multi-cluster feature selection; SVM – Support vector machine; RBF – Radial basis function; CNN – Convolutional neural network; RLS – Recursive least-square; KNN – K-nearest neighbor; NN – Neural network; LD – Linear discriminant; QD – Quadratic discriminant

The j^{th} input data point x_j identifies the center of the RBF 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 RBFs for using in the hidden layer, but we implement the Gaussian function for the sake of comparison between SVM and RBF in[5]

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 the learning algorithms of RBF:

K-means clustering

K-means is a method that utilizes distances for clustering with two steps:

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

$$\min_{(\hat{\mu})_{j=1}^{K}} \sum_{j=1}^{K} \sum_{C(i)=j} \boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_j^{\,2} \text{ for a given } C \tag{7}$$

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

$$C(i) = arg\ \min_{1\le j\le K} \left\| \boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_j^{\,2} \right\| \tag{8}$$

The recursive least-squares algorithm in hybrid learning

Adaptive algorithms have been designed to converge to certain weight vectors. These weights in 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 RBF network.[22] Let the K × 1 vector:

$$\boldsymbol{\Phi}(\boldsymbol{x}_i) = \begin{bmatrix} \varphi(\boldsymbol{x}_i, \boldsymbol{\mu}_1) \\ \varphi(\boldsymbol{x}_i, \boldsymbol{\mu}_2) \\ \vdots \\ \varphi(\boldsymbol{x}_i, \boldsymbol{\mu}_K) \end{bmatrix} \tag{9}$$

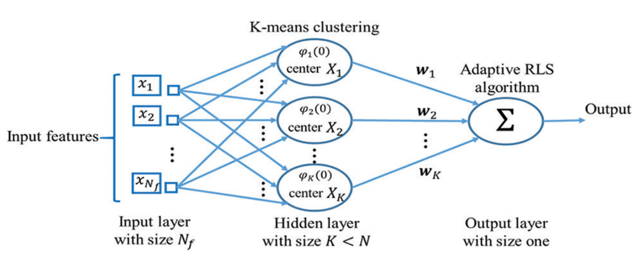

Figure 8: The implemented hybrid radial basis function network for the obstructive sleep apnea detection

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,…., N. Thus, insofar as the supervised training of the output layer is concerned, the training sample is defined by $\left(\boldsymbol{\Phi}[i], d[i]\right)_{i=1}^{N}$, 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:[22]

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

$$\boldsymbol{P}(n) = \boldsymbol{P}(n-1) - \frac{\boldsymbol{P}(n-1)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^{T}(n)\boldsymbol{P}(n-1)}{1+\boldsymbol{\Phi}^{T}(n)\boldsymbol{P}(n-1)\boldsymbol{\Phi}(n)} \tag{10}$$

$$\boldsymbol{g}(n) = \boldsymbol{P}(n)\boldsymbol{\Phi}(n) \tag{11}$$

$$\alpha(n) = d(n) - \hat{\boldsymbol{w}}^{T}(n-1)\boldsymbol{\Phi}(n) \tag{12}$$

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) + \boldsymbol{g}(n)\alpha(n) \tag{13}$$

To initialize the algorithm, we have $\hat{\boldsymbol{w}}(0) = 0$ and $P(0) = \lambda^{-1}\boldsymbol{I}$ where λ is a small positive constant. In[21] a complete analysis was made to show the superiority of hybrid RBF to the SVM classifier both computationally and with respect to accuracy. Furthermore, at least a 30% percent time-saving is guaranteed using RBF in comparison with SVM.[22] This is important because we compared the results of our proposed method with that of the reference.[5]

The Computational Complexity

Up until now, several researchers have tried to present different analyses about the computational comparison between the hybrid RBF and the SVM networks.[21,22,24,25] In all of these references, it has been declared that the RLS-based classifiers need less time to be trained and converge to the SVMs while having similar or better classifying results. Rifkin[21] has shown that the excess computational complexity of the SVM networks arises from the solving of the quadratic programming problem. The computational complexity of this optimization problem can be in the order of $O(N_s^3 + N_s^2 m + N_s nm)$ where N_s is the number of support vectors, n is the dimensions and m is the number of the input data. In the worst case, $Ns \approx m$ we have $O(nm^2)$. On the other hand, solving the linear system of equations using the RLS-based learning algorithm in the worst case can be bounded by $O(m^{2.376})$ Overall, a RLS solution can be obtained much faster than that computed by SVMs.[24] However, the "K-means, RLS" algorithm that has been presented in[22] and is used by us, could not be fully analyzed for the computational complexity. Therefore, we trust on the experimental results both in[22] and in our work that shows the time for training and processing of the hybrid RBF network is one third of the SVM network for the same problem and conditions.

It is important to mention that our comparisons were conveyed with the SVM networks with the RBF kernels that are from the strongest presented SVM schemes.

Results and Discussion

The comparison between the proposed methods for the OSA detection is usually based the accuracy and the complexity of the used signal processing techniques in each part of the task. As we emphasized earlier, the feature extraction, reduction, and detection are the main parts of the OSA detection. Therefore, in this part, we present them for each proposed method in our references along with the proposed detection results to facilitate the comparison. Some of the proposed methods (as for example in[4]) have obviously weak detection results while having less computational complexity. Other methods (like the ones associated with the DNNs) have satisfactory results but also have extracomputational complexity. The aim of this article is the reconciliation between the computational complexity and accuracy. We claim that, apart from the results from DNNs and CNNs, our proposed method is both accurate and less complex.

It is obvious that after feature extraction we get 56 features from each ECG signal (7 sub bands and 8 no-linear features). One of the contributions of this article is the usage of the MCFS feature reduction algorithm in apnea detection.[10] Using this algorithm, the number of features for the ECG signals reduces to ten features which is much lesser than the 18 features of the proposed method in.[5] Our OSA detection results are presented based on the accuracy, sensitivity, and specificity of the proposed methods that are given as follows:

$$Accuracy\left(ACC\right)=\frac{Tp+TN}{TP+TN+FP+FN} \tag{14}$$

$$Sensititivity\left(Sen\right)=\frac{TP}{TP+FN} \tag{15}$$

$$Specificity\left(Spec\right)=\frac{TN}{TN+FP} \tag{16}$$

Where, *TP, TN, FP* and *FN* denotes true positive, true negative, false positive, and false negative, respectively. The comparison of the results of the proposed method and the results in several recent references are given in Table 2.

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 article was to improve the results in.[5] By comparing the results, we can see that an average of 3% improvement is achieved in all the performance metrics. Furthermore, the computational complexity of our classifier is at least 30% lesser than the proposed SVM classifier in.[5]

Conclusion

In this article, we considered the OSA detection using various signal processing techniques to compare the results with the previously proposed methods. The feature extraction in this article is based on the nonlinear properties of the DT-CWT coefficients. After feature extraction, in order to reduce the computational complexity, we used the MCFS algorithm that shrinks the feature vector size to 10. Using these features and the “Hybrid RBF” network, we presented the results that were better than those of the previously presented SVM networks. Furthermore, the proposed method has less computational complexity that makes it a powerful rivalry to the computationally costly but very accurate DNNs and CNNs. In future works, we will consider other low computational signal processing methods for the OSA detection.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

References

1. Hilmisson H, Lange N, Duntley SP. Sleep apnea detection: Accuracy of using automated ECG analysis compared to manually scored polysomnography (apnea hypopnea index). Sleep Breath 2019;23:125-33.
2. Janbakhshi P, Shamsollahi MB. Sleep apnea detection from single-lead ECG using features based on ECG-derived respiration (EDR) signals. IRBM 2018;39:206-18.
3. Khandoker AH, Gubbi J, Palaniswami M. Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings. IEEE Trans Inf Technol Biomed 2009;13:1057-67.
4. Ma Y, Sun S, Zhang M, Guo D, Liu AR, Wei Y, *et al*. Electrocardiogram-based sleep analysis for sleep apnea screening and diagnosis. Sleep Breath 2019;24:231-40.
5. Zarei A, Asl BM. Automatic detection of obstructive sleep apnea using wavelet transform and entropy-based features from single-lead ECG signal. IEEE J Biomed Health Inform 2019;23:1011-21.
6. Hassan AR, Bashar SK, Bhuiyan MI. Computerized Obstructive Sleep Apnea Diagnosis from Single-Lead ECG Signals Using Dual-Tree Complex Wavelet Transform. IEEE Region 10 Humanitarian Technology Conference (R10-HTC, Dhaka, Bangladesh); 21-23 December, 2017.
7. Nishad A, Pachori RB, Acharya UR. Application of TQWT based filterbank for sleep apnea screening using ECG signals. J Ambient Intell Humaniz Comput 2018. doi.org/10.1007/s12652-018-0867-3.
8. Avcı C, Akbaş A. Sleep apnea classification based on respiration signals by using ensemble methods. Biomed Mater Eng 2015;26 Suppl 1:S1703-10.
9. Rachim VP, Li G, Chung WY. Sleep apnea classification using ECG-signal wavelet-PCA features. Biomed Mater Eng

2014;24:2875-82.
10. Cai D, Zhang C, He X. Unsupervised Feature Selection for Multi-Cluster Data. In Proceedings of 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2010. p. 333-42.
11. Hassan AR, Haque MA. An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting. Neurocomputing 2017;235:122-30.
12. Hassan AR, Haque MA. Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating. Biocybern Biomed Eng 2016;36:256-66.
13. Hassan AR. Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting. Biomed Signal Proc Control 2016;29:22-30.
14. Wang T, Lu C, Shen G, Hong F. Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network. PeerJ 2019;7:e7731.
15. Singh SA, Majumder S. A novel approach OSA detection using single lead ECG Scalogram based on deep neural network. J Mech Med Biol 2019;19:1-18.
16. Urtnasan E, Park JU, Joo EY, Lee KJ. Automated detection of obstructive sleep apnea events from a single-lead electrocardiogram using a convolutional neural network. J Med Syst 2018;42:104.
17. Wang X, Cheng M, Wang Y, Liu S, Tian Z, Jiang F, *et al*. Obstructive Sleep Apnea Detection using ECG-Sensor with Convolutional Neural Networks. Multimedia Tools and Application, First Online; 18 June, 2018.
18. Wang T, Lu C, Shen G. Detection of sleep apnea from single-lead ECG signal using a time window artificial neural network. Hindawi Bio Med Res Int 2019: Article ID 9768072. https://doi.org/10.1155/2019/9768072.
19. Hassan AR, Haque MA, *et al*. Computer-Aided Sleep Apnea Diagnosis from Single-Lead Electrocardiogram using Dual-Tree Complex Wavelet Transform and spectral features. 1st International Conference on Electrical & Electronic Engineering (ICEEE), RUET, Rajshahi, Bangladesh; 04-06 November, 2015.
20. Thomas M, Das MK, Ari S. Automatic ECG arrhythmia classification using dual-tree complex wavelet based features. Int J Electron Commun 2015;69:715-21.
21. Rifkin RM. Everything Old is New Again: A Fresh Look at Historical Approaches in Machine Learning. Ph.D. Thesis, MIT; 2002.
22. Haykin S. Neural Networks and Learning Machines. Pearson Education, Inc., Upper Saddle River, New Jersey: Prentice Hall; 2008. p. 230-63.
23. Papini GB, Fonseca P, Margarito J, van Gilst MM, Overeem S, Bergmans JWM, *et al*. On the generalizability of ECG-based obstructive sleep apnea monitoring: Merits and limitations of the Apnea-ECG database. Conf Proc IEEE Eng Med Biol Soc 2018;2018:6022-5.
24. Zhang P, Peng J. SVM vs. Regularized Least Squares Classification. Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK; August, 2004.
25. Debnath R, Takahashi H. Learning Capability: Classical RBF Network vs. SVM with Gaussian Kernel. Proceedings of International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems; 2002. p. 293-302.

BIOGRAPHIES

Javad Ostadieh was born in 1974, Urmia, Iran. He received his B.S. degree from Urmia Azad University and the M.S. degree from Najaf Abad-Isfahan Azad University, in 1999 and 2004, respectively, both in electrical engineering. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Urmia University. His research interests include medical Image and signal processing, pattern recognition, and classifying networks. He is a full-time academic staff of the Khoy Azad University.

Email: j.ostadieh@urmia.ac.ir

Mehdi Chehel Amirani received the B.Sc. degree in Electronic Engineering from Urmia University, Iran, in 1993 and the M. Sc. and Ph.D. degrees in Communications engineering from Iran University of Science and Technology (IUST) in 1998 and 2009, respectively. In 2009, he joined the department of electrical engineering at Urmia University, where he is currently an associate professor. His research interests include pattern recognition, digital and adaptive signal processing, and wireless communications. He has published more than 60 scientific papers in the signal processing field and advised numerous masters and Ph.D. students.

Email: m.amirani@urmia.ac.ir

Morteza Valizadeh received both his M.Sc. and Ph.D. degrees in electronic engineering from Tarbiat Modares University in 2008 and 2012, respectively. Currently he is an assistant professor of communication at Urmia University. His research interests include image processing, medical signal analysis, machine vision and deep learning.

Email: mo.valizadeh@urmia.ac.ir