Journal of Computer Science Research
https://journals.bilpubgroup.com/index.php/jcsr

ARTICLE

Development of New Machine Learning Based Algorithm for the Diagnosis of Obstructive Sleep Apnea from ECG Data

Erdem Tuncer

Faculty of Technology, Biomedical Eng. Department of Kocaeli University, Kocaeli, 41001, Turkey

ABSTRACT

In this study, a machine learning algorithm is proposed to be used in the detection of Obstructive Sleep Apnea (OSA) from the analysis of single-channel ECG recordings. Eighteen ECG recordings from the PhysioNet Apnea-ECG dataset were used in the study. In the feature extraction stage, dynamic time warping and median frequency features were obtained from the coefficients obtained from different frequency bands of the ECG data by using the wavelet transform-based algorithm. In the classification phase, OSA patients and normal ECG recordings were classified using Random Forest (RF) and Long Short-Term Memory (LSTM) classifier algorithms. The performance of the classifiers was evaluated as 90% training and 10% testing. According to this evaluation, the accuracy of the RF classifier was 82.43% and the accuracy of the LSTM classifier was 77.60%. Considering the results obtained, it is thought that it may be possible to use the proposed features and classifier algorithms in OSA classification and maybe a different alternative to existing machine learning methods. The proposed method and the feature set used are promising because they can be implemented effectively thanks to low computing overhead.

Keywords: ECG; Sleep apnea; Classification; Dynamic time warping; Median frequency

1. Introduction

Obstructive sleep apnea (OSA) is a sleep-related breathing disorder. It becomes evident with the obstructions in the upper respiratory tract during sleep and the waking periods following these obstructions. OSA can seriously reduce a person's quality of daily life and cause the development of many cardiovascular diseases. Therefore, early diagnosis and treatment of obstructive sleep apnea is important. Electrocardi-

*CORRESPONDING AUTHOR:
Erdem Tuncer, Faculty of Technology, Biomedical Eng. Department of Kocaeli University, Kocaeli, 41001, Turkey; Email: erdemtuncerr@gmail.com

ARTICLE INFO
Received: 6 June 2023 | Revised: 4 July 2023 | Accepted: 5 July 2023 | Published Online: 14 July 2023
DOI: https://doi.org/10.30564/jcsr.v5i3.5762

CITATION
Tuncer, E., 2023. Development of New Machine Learning Based Algorithm for the Diagnosis of Obstructive Sleep Apnea from ECG Data. Journal of Computer Science Research. 5(3): 15-21. DOI: https://doi.org/10.30564/jcsr.v5i3.5762

ogram (ECG) is the process of recording the electrical activity of the heart. In today's conditions, ECG signals are used in the diagnosis of OSA. Apnea diagnosis from the ECG signal is measured by heart rate variability. It will be economical and practical to determine whether a person has OSA syndrome with the proposed machine learning technique using single-channel ECG recordings. Because with such a system, there will be no need for environments such as sleep laboratories [1,2]. There are many studies in the literature on the detection of OSA from ECG using methods. In the study conducted by Yildiz [3], obstructive sleep apnea data from ECG recordings were classified. Twelve features were obtained using wavelet transform and they achieved the highest success rate of 98.3% with the support vector machine/artificial neural network classifier algorithms. In the study by Faal et al. [4], they presented a new feature generation method using autoregressive integrated moving average and exponential generalized autoregressive conditional heteroscedasticity model in the time domain from ECG signals. ECG signals were analyzed in one-minute segments. The results were evaluated using five different classifiers (support vector machine, neural network, quadratic separation analysis, linear separation analysis and k-nearest neighbor). As a result of the classification, a success rate of 81.43% was achieved. Tyagi et al. [5] proposed a new approach to cascade two different types of restricted boltzmann machines in the deep belief networks method for sleep apnea classification using electrocardiogram signals. They achieved a success rate of 89.11% from the ECG data examined in one-minute epochs. Yang et al. [6] proposed a one-dimensional compression and excitation residual group network for sleep apnea detection. With the proposed method, an accuracy rate of 90.3% was achieved. Thus, they argued that cheap and useful sleep apnea detectors can be integrated with wearable devices.

The aim of this study is to present an automatic machine-learning method that can detect OSA from ECG recordings. In the proposed method, a wavelet transform-based algorithm is proposed. Unlike the studies in the literature, it is the examination of the effect of different features on apnea data instead of the features frequently used in the literature.

2. Materials and methods

2.1 Data set

The ECG recordings used in the study were taken from the PhysioNet Apnea-ECG dataset. There are 70 ECG recordings in total. Recordings can take up to 10 hours in length. All of the sleep recordings were taken from 32 subjects. The age range of the subjects was between 27 and 63 years. The standard V2 lead was used for the placement of the electrodes on the body surface during recording. ECGs were digitized by sampling at 16 bits per sample and 100 Hz. ECG signals with 16-bit resolution. Evaluation of whether the ECG recordings belong to people with obstructive sleep apnea was made according to the sleep study technique [7]. In this study, 18 ECG recordings of 10 randomly selected patients (a01, a02, a03, a04, a05, a06, a07, a08, a09, a10) were used. The randomly selected apnea and normal ECG data signal form is given in **Figure 1**. In **Figure 1(a)**, heart rate variability is visually striking after the 4000th sample. In **Figure 1(b)**, the normal one-minute ECG signal form is given.

2.2 Feature selection

Discrete wavelet transform

The discrete wavelet transform aims to solve the fixed width window source problem of the fourier transform by using a scalable wavelet function. Thus, optimum time-frequency resolution is provided in different frequency ranges for the biomedical signals to be analyzed. With the discrete wavelet transform, it is aimed to eliminate the excessive computational load. Since an efficient algorithm based on filters has been developed in the discrete wavelet transform, the calculation of the wavelet coefficients is made for discrete values at certain points. This algorithm, called multiple resolution, consists of sequential high-pass and low-pass filter pairs [8,9]. The lower fre-

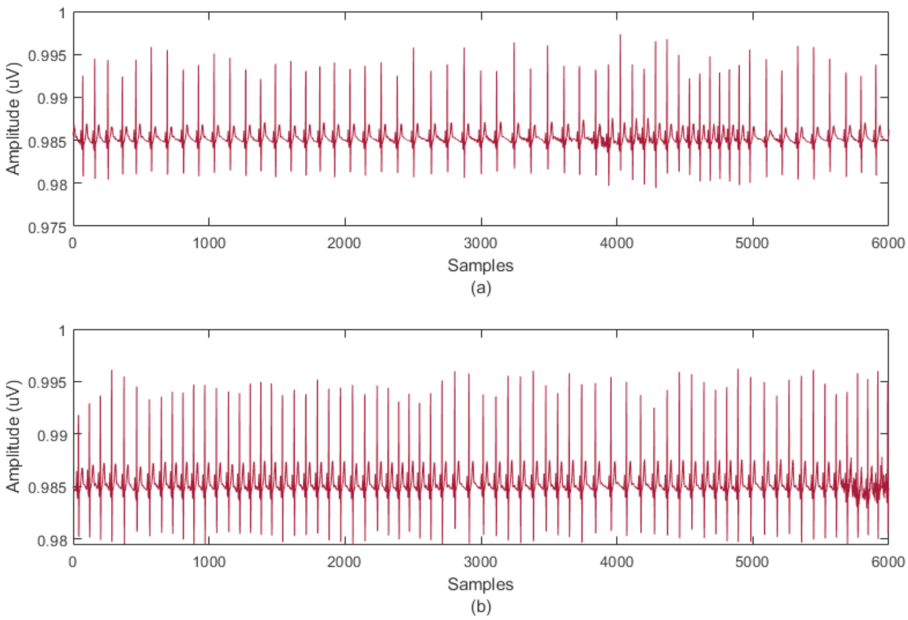

Figure 1. (a) ECG sign with apnea, (b) Normal ECG sign.

quency bands of the ECG data used in the study are given in **Table 1**. As shown in **Table 1**, a six-level wavelet transform is used.

Table 1. Ranges of frequency bands in wavelet transform decomposition of ECG signal.

Sub-bands	Frequency ranges (Hz)
D1	25-50
D2	12.5-50
D3	6.25-12.5
D4	3.125-6.25
D5	1.5625-3.125
D6	0.78125-1.5625
A6	0-0.78125

Dynamic time warping algorithm

Dynamic time warping algorithm is a classification algorithm that uses similarity measurement of time series. Biomedical signals sampled over a period of time form a time series. The similarity between the series can be calculated by finding the sum of the Euclidean distances between the elements of each element of two discrete time series. The closer the Euclidean distance sum is to zero, the more similar the time series are. Today, the dynamic time-warping algorithm is used in many areas from image processing to audio processing [10,11].

$$Q = q1, q2, …., qn\text{-}1, qn \quad (1)$$

$$C = c1, c2, …., cm\text{-}1, cm \quad (2)$$

Q and C in Equation (1) and Equation (2) represent two different signals or data; n and m indicate the lengths of these signals. The similarity ratio between the Q and C signals is calculated using the Euclidean length as in Equation (3).

$$d(q_i, c_j) = (q_i, c_j)^2 \quad (3)$$

After obtaining the (i, j) matrix for Q and C, the accumulated distance matrix is calculated using this matrix. d represents the accumulated cost matrix and is calculated recursively [12].

Median frequency

Power spectral density is the frequency domain equivalent of the power content of the signal. It is used to characterize broadband random signals. The median frequency represents the midpoint of the power spectral density distribution and is the name given to the frequencies above and below that make-up 50% of the total power in the ECG [13,14].

2.3 Classification

Random forest

Random Forest (RF) is a very popular learning algorithm for classification and regression problems.

The RF algorithm is to generate a large number of unbiased decision trees where each tree votes for a class. The Gini index is used to construct the decision trees and determine the last class in each tree. Therefore, the Gini index Gini (v) at node v measures the purity of v. It is expressed by the formula in Equation (4) [15,16].

$$\text{Gini (v)} = \sum_{i=1}^{k} f_i(1 - f_i) \tag{4}$$

Here fi is the fraction of class i recorded at node v.

Long short-term memory

Introduced by Hochreiter and Schmidhuber, Long Short-Term Memory (LSTM) is an advanced variant of the Recurrent Neural Network (RNN) architecture. The basic structure of LSTM is that it uses a memory cell to remember and explicitly span unit outputs at different time steps. The memory cell of LSTM uses cell states to remember the information of temporal contexts. It has a forget gate, an entry gate and an exit gate to control the flow of information between different time steps. The three gates of LSTM make it easy to organize long-term memory. LSTM models can learn the temporal dependence between data. Due to its ability to learn long-term correlations in a sequence, LSTM networks are capable of accurately modeling complex multivariate sequences such as the ECG signal [17,18].

2.4 Evaluation of classification models

One of the performance metrics for the machine learning classification problem is the confusion matrix. **Table 2** contains four different combinations of the value to be estimated and the actual values are called the confusion matrix [19].

Table 2. Confusion matrix.

	Predicted: No	Predicted: Yes
Actual: No	True Negative	False Positive
Actual: Yes	False Negative	True Positive

Here, TP: True positive, TN: True negative, FP: False positive, FN: False negative. Some of the metrics we can calculate with the terms in **Table 2** are accuracy, precision and recall. Their mathematical equations are given in Equations (5), (6) and (7).

$$\text{Accuracy} = TP + TN/TP + FP + TN + FN \tag{5}$$

$$\text{Precision} = TP/TP + FP \tag{6}$$

$$\text{Recall} = TP/TP + FN \tag{7}$$

In addition, the test performance of each classifier was evaluated by calculating statistical parameters.

3. Results

In this study, the machine learning method that can predict the automatic detection of OSA disease, which is time-consuming and costly to diagnose, from single-channel ECG recordings is presented. The flow chart of the proposed method is shown in **Figure 2**.

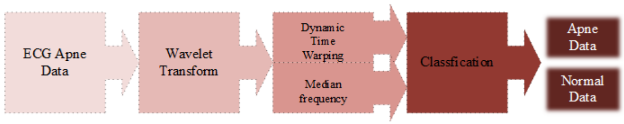

Figure 2. Flow chart of the proposed model.

In the presented method, ECG data were analyzed in one-minute windows. The coefficients of the lower frequency bands were obtained from each window data by using the wavelet transform (6-level Symlet2 wavelet). After applying the dynamic time-warping algorithm to the wavelet coefficients in different frequency bands, the results obtained are recorded in the feature matrix. The relationship of the A6 coefficients with the other coefficients was evaluated with the dynamic time-warping algorithm. Another parameter calculated as a feature is the median frequency. The median frequency values of the wavelet coefficients obtained from all lower frequency bands were calculated. As shown in **Table 3**, a total of 13 features were extracted and given as input to the classifier algorithms.

In this study, two different classifier algorithms were evaluated. One is the deep learning architecture LSTM and the other is the traditional learning algorithm RF. The architecture of the model created in the LSTM classifier is shown in **Figure 3**. LSTM architecture layers are composed of input layer, LSTM layer, dropout layer, LSTM layer, dropout layer and output layer, respectively. The LSTM layer contains 50 units per layer. These units use the Corrected

Table 3. Feature list.

No	Feature name	Difference between wavelet coefficients	No	Feature name	Wavelet coefficient
1	Dynamic Time Warping	A6, D1	7	Median frequency	D1
2	Dynamic Time Warping	A6, D5	8	Median frequency	D2
3	Dynamic Time Warping	A6, D4	9	Median frequency	D3
4	Dynamic Time Warping	A6, D3	10	Median frequency	D4
5	Dynamic Time Warping	A6, D2	11	Median frequency	D5
6	Dynamic Time Warping	A6, D1	12	Median frequency	D6
			13	Median frequency	A6

Linear Unit (ReLU) activation function and give a different output for each time step. The reason for using ReLU is that it is generally less costly to train the model in terms of computational load and can achieve better performance than other models. In addition, ReLU can avoid the vanishing gradient problem, which is an advantage over the tanh function.

After the first LSTM layer, the dropout layer (with a value of 0.2) is applied to reduce overfitting. The next layer is a new LSTM layer containing 50 units and ReLU activation functions, followed by the dropout layer. Finally, the value containing the classification result is estimated after a sigmoid activation function is used to estimate the result with the output layer.

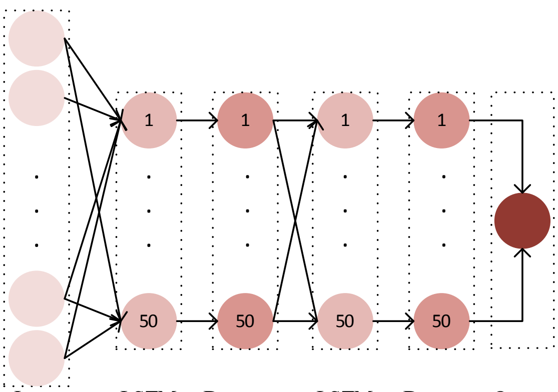

Figure 3. Diagram of the LSTM architecture.

The classification accuracy of the RF method depends on user-defined parameters such as the number of trees and the number of parameters. Therefore, the selection of the most appropriate parameter for the data increases the classification accuracy. In the study, multiple combinations were tested to find the optimum parameters of tree and parameter numbers. The success rates obtained for combinations of different tree and parameter numbers are shown in **Table 4**. As can be seen in **Table 4**, the number of trees with the highest success rate was selected as 250 and the number of parameters as two for the classification of apnea data. Since increasing the number of trees does not increase the performance of the model, the model with the highest performance with the least number of trees was selected.

Table 4. RF algorithm success results by parameters.

Number of trees	Number of parameters	Accuracy rate (%)
10	2	78.15
20	2	80.13
30	2	80.46
40	2	80.57
50	2	80.79
70	2	81.22
100	2	81.66
150	2	81.99
200	2	81.88
250	2	82.43
500	2	82.43

The success rates obtained as a result of LSTM architecture and RF architecture are given in **Table 5**. As can be seen from **Table 5**, the optimized RF algorithm performed better than the LSTM architecture. Therefore, the LSTM architecture has the best performance.

Table 5. Classifier performances.

Dataset	Accuracy (%)		Precision (%)		Recall (%)	
	RF	**LSTM**	**RF**	**LSTM**	**RF**	**LSTM**
ECG Apnea	82.43	77.60	82.10	76.70	82.40	77.50

4. Discussion

The analyzed results show that it gives the highest accuracy with 82.43% accuracy with the RF algorithm. High classification performance was achieved with thirteen features obtained by using two features from ECG data. When the studies in the literature were examined, the norm entropy values of each wavelet level were calculated by using the twelve-level wavelet transform of the obstructive sleep apnea data from the ECG recordings in the study conducted by Yildiz [3]. The obtained features were applied to the support vector machine/artificial neural network classifier algorithms and the highest success rate of 98.3% was obtained. In the study by Faal et al. [4], they presented a new feature generation method. As a result of five different classifier algorithms, a success rate of 81.43% was achieved. Tyagi et al. [5] proposed a new approach and achieved a success rate of 89.11%. Yang et al. [6] proposed a one-dimensional compression and excitation residual group network and 90.3% accuracy was achieved with the proposed method. In the study by Razi et al. [20], ten-time domain features were extracted and reduced to five features. Principal component analysis and discriminant linear analysis were used for size reduction. RF algorithm is proposed for classification and the results are compared with other classifier algorithms. The highest success rate detected is 95.01%.

When the studies in the literature are examined, it is observed that the success rates are generally higher than the study of this article. Most of the studies aimed to reach a higher success rate by using similar methods and techniques. However, in the field of machine learning, the goal is not only to increase classification success but also to develop different features and method techniques. From this point of view, our article differs from the studies in the literature. A previously unused feature set is suggested on ECG apnea data. At the same time, the change in success rates with the optimization of the classifier algorithms was examined. It is possible to reach higher success rates by diversifying and optimizing the parameters of the machine learning model.

5. Conclusions

This article discusses the estimation of apnea diagnosis from ECG data. We propose a binary classification machine learning method to support physicians' decisions in clinical practice. For decision support applications, modeling using the RF algorithm as a classifier and classification of patients' apnea data are recommended. It has been seen that the feature method selected with the RF algorithm is successful. In the classification made with the used feature set and RF algorithm optimization, a successful prediction was made with 13 features with an accuracy rate of 82.43%. The feature set and method we used in our study give hope for higher future success rates. In further studies, it is aimed to evaluate the efficiency of the feature set by expanding the dataset.

Conflict of Interest

The author has no conflicts of interest to declare.

Funding

This research received no external funding.

References

[1] Wiegand, L., Zwillich, C.W., 1994. Obstructive sleep apnea. Disease-a-Month. 40(4), 202-252. DOI: https://doi.org/10.1016/0011-5029(94)90013-2

[2] Paiva, T., Attarian, H., 2014. Obstructive sleep apnea and other sleep-related syndromes. Handbook of clinical neurology. Elsevier: Amsterdam. pp. 251-271.

[3] Yildiz, A., 2017. Tek kanallı EKG kayıtları analizinden uyku apne tespiti (Turkish) [Detection of sleep apnea from analysis of single-channel ECG recordings]. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi. 8(1), 111-122.

[4] Faal, M., Almasganj, F., 2021. Obstructive sleep apnea screening from unprocessed ECG signals using statistical modelling. Biomedical Signal Processing and Control. 68, 102685. DOI: https://doi.org/10.1016/j.bspc.2021.102685

[5] Tyagi, P.K., Agrawal, D., 2023. Automatic detection of sleep apnea from single-lead ECG signal using enhanced-deep belief network model. Biomedical Signal Processing and Control. 80(2), 104401. DOI: https://doi.org/10.1016/j.bspc.2022.104401

[6] Yang, Q., Zou, L., Wei, K., et al., 2022. Obstructive sleep apnea detection from single-lead electrocardiogram signals using one-dimensional squeeze-and-excitation residual group network. Computers in Biology and Medicine. 140, 105124. DOI: https://doi.org/10.1016/j.compbiomed.2021.105124

[7] Penzel, T., Moody, G.B., Mark, R.G., et al. (editors), 2000. The Apnea-ECG database. Computers in Cardiology. 2000 September 24-27; Cambridge. USA: IEEE. p. 255-258.

[8] Tuncer, E., Bolat, E.D., 2022. Destek Vektör Makinaları ile EEG Sinyallerinden Epileptik Nöbet Sınıflandırması (Turkish) [Epileptic seizure classification from EEG signals with support vector machines]. Politeknik Dergisi. 25(1), 239-249. DOI: https://doi.org/10.2339/politeknik.672077

[9] Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 11(7), 674-693. DOI: http://dx.doi.org/10.1109/34.192463

[10] Zhang, Z., Tavenard, R., Bailly, A., et al., 2017. Dynamic time warping under limited warping path length. Information Sciences. 393, 91-107. DOI: https://doi.org/10.1016/j.ins.2017.02.018

[11] Jeong, Y.S., Jeong, M.K., Omitaomu, O.A., 2011. Weighted dynamic time warping for time series classification. Pattern Recognition. 44(9), 2231-2240. DOI: https://doi.org/10.1016/j.patcog.2010.09.022

[12] Bakir, C., 2016. Automatic speaker gender identification for the German language. Balkan Journal of Electrical and Computer Engineering. 4(2), 79-83. Available from: https://dergipark.org.tr/en/download/article-file/458645

[13] Brown, C.G., Griffith, R.F., Ligten, P.V., et al., 1991. Median frequency—a new parameter for predicting defibrillation success rate. Annals of Emergency Medicine. 20(7), 787-789. DOI: https://doi.org/10.1016/S0196-0644(05)80843-1

[14] Tonner, P.H., Bein, B., 2006. Classic electroencephalographic parameters: Median frequency, spectral edge frequency etc. Best Practice & Research Clinical Anaesthesiology. 20(1), 147-159. DOI: https://doi.org/10.1016/j.bpa.2005.08.008

[15] Masetic, Z., Subasi, A., 2016. Congestive heart failure detection using random forest classifier. Computer Methods and Programs in Biomedicine. 130, 54-64. DOI: https://doi.org/10.1016/j.cmpb.2016.03.020

[16] Coskun, G., Aytekin, I., 2021. Early detection of mastitis by using infrared thermography in holstein-friesian dairy cows via classification and regression tree (CART) Analysis. Selcuk Journal of Agriculture and Food Sciences. 35(2), 118-127.

[17] Tuncer, E., Bolat, E.D., 2022. Classification of epileptic seizures from electroencephalogram (EEG) data using bidirectional short-term memory (Bi-LSTM) network architecture. Biomedical Signal Processing and Control. 73, 103462. DOI: https://doi.org/10.1016/j.bspc.2021.103462

[18] Sowmya, S., Jose, D., 2022. Contemplate on ECG signals and classification of arrhythmia signals using CNN-LSTM deep learning model. Measurement: Sensors. 24, 100558. DOI: https://doi.org/10.1016/j.measen.2022.100558

[19] Dağli, E., Büber, M., Taspinar, Y.S., 2022. Detection of accident situation by machine learning methods using traffic announcements: The case of metropol Istanbul. International journal of applied mathematics electronics and computers. 10(3), 61-67.

[20] Razi, A.P., Einalou, Z., Manthouri, M., 2021. Sleep Apnea classification using random forest via ECG. Sleep and Vigilance. 5, 141-146. DOI: https://doi.org/10.1007/s41782-021-00138-4