Information Sciences with Applications, Vol. 4, 2024
https://doi.org/10.61356/j.iswa.2024.4435

Sleep Apnea Detection

Mohamed Elkholy [1], Maryam Abdelfattah [2], Maryam Amer [2] and Hussam Elbehiery [3,*]

[1] Faculty of Computer Science & Engineering, Alamein International University, Egypt; melkholy@aiu.edu.eg.
[2] Faculty of Information Systems and Computer Science, October 6th University, Giza, 12585, Egypt.
Emails: maryamali047@icloud.com; maryaammohamedd@gmail.com.
[3] Vanridge University, USA; drhussam@vru-edu.net.

* Correspondence: drhussam@vru-edu.net.

Abstract: Sleep apnea is one of the most common sleep disorders, posing a significant health risk. The proposed paper provides a prospective technique by applying machine learning to enable the early prediction of sleep apnea. Two methodologies were used for this study: Firstly, the hybrid model was used in analyzing the electrocardiogram records by combining VGG16 and Long Short-Term Memory (LSTM) networks for the extraction of meaningful features from the ECG signals for modeling the process of aiding detection. Second, lifestyle patterns were assessed in their relationship with the disorder. Many lifestyle factors were analyzed in search of critical indicators that may, at an early age, indicate the onset of the case of sleep apnea. The fusion of deep learning methods and analysis of lifestyle patterns offers a comprehensive framework for the solution to the complexity of sleep apnea detection. More exactly, this kind of interdisciplinary approach should permit improvement in effectiveness and accuracy for early diagnosis, therefore enabling timely intervention and treatment of that effect. Proposed here is an innovative methodology that could lead to better management and recovery of the affected.

Keywords: Artificial Intelligence; Electrocardiogram (ECG); Sleep Diseases; Deep learning.

1. Introduction

One-third of human life is spent in the sleeping process. Hence, sleep has a huge impact on maintaining health and overall well-being [1, 2]. The sleep cycle includes two primary stages. Firstly, there is rapid eye movement, which is characterized by high speed in eye movement and includes dreams. Secondly, there is the non-rapid eye movement stage in sleep [3]. REM sleep includes an increased peak activity in the sympathetic nervous system. Increased hemodynamic changes are also a part of REM. On the other hand, NREM sleep reduces oxygen consumption, significantly decreases heart rate, and causes a decrement in blood pressure [5, 6].

Sleep disorders are increasingly becoming common, with approximately seventy million adults within the United States affected. Such disorder in the sleeping process leads to higher rates of mortality. Among the various sleep disorders are insomnia, central disorders of hypersomnolence, parasomnias, and movement disorders to mention a few. Sleeping Breath Disorder (SBD) is also considered one of the popular sleeping disorders. There are several types of SBDs: obstructive sleep apnea, sleep hypoventilation, and central sleep apnea. OSA affects one billion individuals all over the world.

Advanced technologies have made it feasible to develop a few techniques for sleep apnea detection. Most of the techniques use ECG signals in the detection of certain abnormal features while one is sleeping. Initially, conventional machine-learning techniques were used, but there has been a shift to advanced Deep-Learning models since they are good at feature extraction. However, this complication makes it very hard to compare the performance of such algorithms due to differences in data sets, physiological signals used, training, and evaluation criteria.

2. Background

Accurate and efficient detection of sleep apnea is critical to treat and reduce its impact. Several approaches are used in monitoring sleep apnea, which include questionnaires, medical imaging, and signal-processing techniques [16]. As much as questionnaires are cost-effective in determining patients with sleep disorders, medical imaging could be efficient in serious conditions by monitoring several activities of anatomies while sleeping. Currently, the gold standard for the diagnosis of sleep apnea is PSG, since it involves multiple biological signals. The method of PSG, however, tends to be restricted to special sleep laboratories and thus is not appropriate for home application.

There is a rising interest in wearable, non-invasive technologies that autonomously can monitor and manage sleep [18, 19]. Several signals connected with physiological activities have been discussed for the detection of sleep apnea, and among these, ECG has emerged as one of the potential candidates in wearable devices. In this respect, machine learning and deep learning techniques are used for the analysis of different human signals. Machine learning depends on manually defined features; on the other hand, deep learning methods do an auto feature extraction.

Deep learning, due to recent developments, has transformed the detection technique of sleep apnea. Many authors have reported high performance in this area using Convolutional Neural Networks (CNNs) [20, 21]. Other studies have assessed the performance of various deep learning architectures and feature extraction methods to increase accuracy in detection.

Some innovative technologies in motion detection methods and single-lead ECG wearable devices have been proposed for real-time sleep apnea monitoring and detection. The healthcare system of the future is likely to include these kinds of wearable devices, which provide the capability of automatic sleep condition monitoring and management.

3. Experimental Setup

3.1 Dataset

Two different datasets were used in the proposed experiments:

3.1.1 ECG Records

For ECG records, a dataset containing seventy records is used. Such data is divided into a learning set, containing thirty-five records, while the rest are used for the testing set. The records were chosen in different lengths for up to 10 hours. Each record has only a single continuous ECG signal. In addition, there is an annotation about sleep apnea that is based on specialized human decisions. Every record is supplemented with a collection of automated annotations to distinguish between normal and up-normal signals. A collection of eight records is supplemented with additional signals associated with abdominal respiratory signal airflow rate and oxygen saturation.

3.1.2 Tabular Data

In the work, the Sleep Health and Lifestyle Dataset has been used, with four hundred entries that denote sleep health. The thirteen columns represent sleeping behavior and daily habits; it also contains personal data, such as age, gender, amount of activity and stress levels one undergoes, the total number of sleeping hours, heart rate, and blood pressure—features deemed important in predicting sleeping apnea.

Dataset Features:

Sleeping metrics tell about the quality of sleep and sleeping duration, among other factors classifying sleep patterns. Personal data analyzes physical activities and the stress levels of different people. Cardiovascular health is in charge of measuring changes in blood pressure and heart rate. Sleep disorder measurements are used to detect the occurrence of sleep disorders.

3.2 Preprocessing

Two different datasets were used in the proposed experiments:

3.2.1 ECG Records

The ECG signals were divided into 1-minute intervals after which R-R Intervals were extracted using the Hamilton R-peak algorithm to detect normality. Chen et al used a median filter in reducing physiologically meaningful points. The subsequent R-R Intervals underwent essential processing before inputting into the machine learning algorithms.

The level of R-peaks was recorded near the R-R Interims and passed towards a profound learning algorithm. In arrange to test both R-peak value as well as R-R interims at a rise to evaluate a cubic introduction at 3 Hz was utilized. Hence, the added R-R Interims and R-peak value are bolstered into profound learning algorithms. To acquire the exact transient highlights of DRNN data while using it, the data was divided into n segments of 60/n seconds.

3.2.2 Tabular Data

This one had an amazingly simple preprocessing step where we cleaned up inconsistent data, replaced missing data, and encoded categorical data to make life easier for the model.

4. Conventional Machine Learning

Traditional machine learning methods include the following three steps. As depicted in Figure 1, the proposed steps in this work are classifying a set of selected features. Hence, the proposed work uses feature engineering methodologies to enhance algorithm performance. In this work, we presented a unified feature engineering framework, specially designed for conventional machine learning algorithms. Preprocessing was first applied to ECG signals. Then time, frequency, and nonlinear features were extracted. Dimensionality was reduced by PCA, followed by feature extraction. Afterward, the classification of ECG signals into apnea and normal episodes was realized.

4.1 Steps of Extracting Features

In detecting apnea by ECG signals, it has been proven to be HRV parameter dependent. The proposed work presents various machine learning models that efficiently detect sleep apnea by making use of different preprocessed data sets. Extracted features include time domain, frequency domain, and HRV features.

4.1.1 Time Domain

The statistical, time-domain features are divided into two main categories. The first is the long-term time domain while the second is the short-term time domain. Both features are extracted from R-R intervals. The extracted features include the minimum value, range value, and standard deviation value.

4.1.2 Frequency-Domain Features

Frequency-domain features provide information related to power frequency, for instance, total power, low-frequency power, and high frequency, and their ratio.

4.1.3 Nonlinear Features

Seven nonlinear features, that were SD1, SD2, the ratio of SD2/SD1, CVI, CSI, modified CVI, and permutation entropy, were extracted.

4.2 Dimension Reduction

We applied PCA and performed dimensionality reduction to get rid of the curse of dimensionality, classifying only important features.

4.3 Classification

Following this, several machine-learning techniques were applied to the detection and classification of sleep apnea. One major challenge that remains is the lack of a common and fair comparison framework between these many techniques. The proposed study will, therefore, address this gap by performing a comparative evaluation of some prominent machine learning methods.

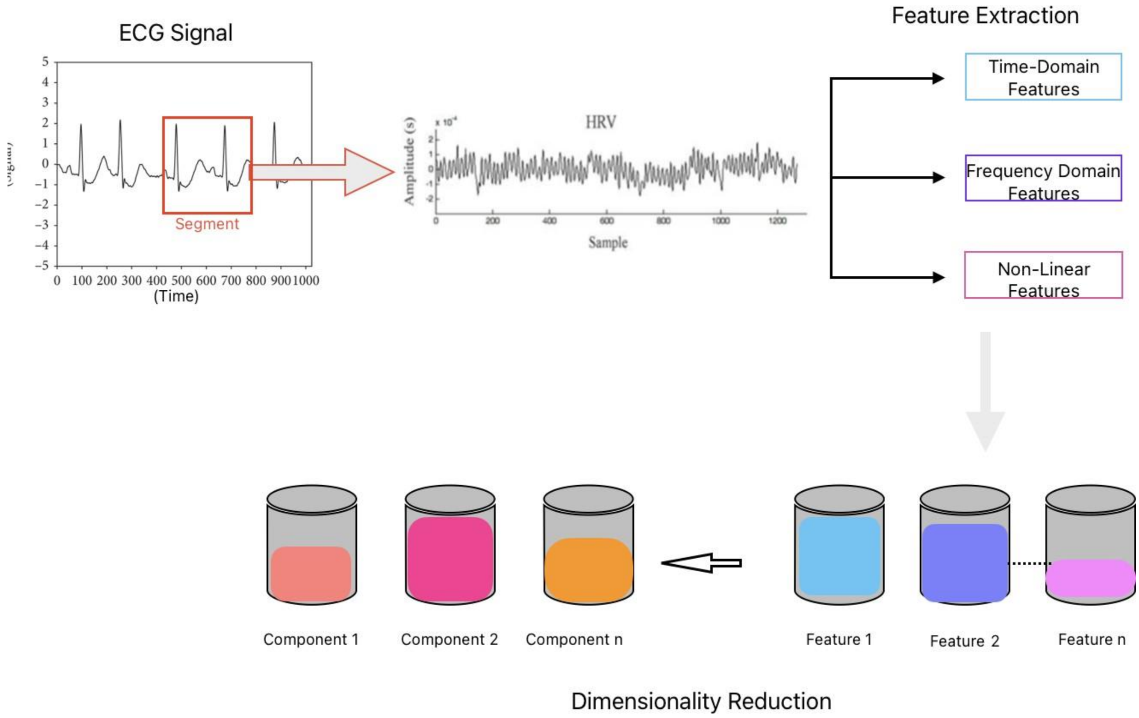

Figure 1. A Flowchart representing the process of using ML in detecting sleep disorder from ECG.

1) Linear Discriminant Analysis: LDA offers the most basic method of classification but very essential in creating linear decision boundaries by fitting Gaussian density per class, ruled by Bayes' rule. The major hyperparameters for LDA entail a solver model and parameters associated with tolerance; it sets the threshold of significance of singular Values.
2) Quadratic Discriminant Analysis: QDA through quadratic decision bound- emanates from the generalization of LDA. Ordinarily, using the SVD solver, the main parameter for QDA is the statement of tolerance, which is essential for efficient training.
3) Logistic Regression: LR includes the solver algorithm and choices for convergence tolerance.
4) Gaussian Naïve Bayes (GNB): GNB is an application of Bayes' theorem assuming independence between features. The key hyperparameter for GNB would be that of variance tuning for better stability, which is sensitive to the ratio of the biggest variance across different features.
5) Gaussian Process (GP): It is a non-parametric approach. GP uses kernel functions to identify class labels. Important hyperparameters for GP include the choice of kernel and algorithm for optimization.
6) Support Vector Machine: SVM uses kernel functions to project the data into higher-dimensional spaces. The key hyperparameters are about the type of kernel to be used and their associated parameters.

7) K-Nearest Neighbors: KNN will enable the classification of the data by neighbors; several distance metrics are at one's disposal.
8) Decision Tree: DT is another nonparametric classifier that predicts classes by learning simple decisions based on input features.
   Extract rules from features. The most important hyperparameters include the maximum depth of the tree, the strategy to perform the splits, and the criterion to evaluate the split quality.
9) Random Forest RF — this method generalizes DT by averaging several trees on sub-sampling with replacement. The main hyperparameters are maximum depth, the lowest count of samples needed to split an internal node, the criterion by which the quality of the split is measured, and the count of trees, or estimators.
10) Extra Trees: ET is a variant of RF that uses the whole original sample (rather than bootstrapping it), and it also randomly selects the cut points. This makes ET a great deal more computationally efficient than RF.
11) AdaBoost (AB): AB is another ensemble technique that combines multiple weak learners to come up with a strong classifier. Contrary to RF's bagging technique, AB works on boosting, where the weights of the misclassified samples are increased in every iteration. The most important hyperparameters are the count of estimators and the rate of learning.
12) Gradient Boosting, GB: Generalizes boosting as an optimization problem. In this process of sequentially adding weak learners—mostly DTs, GB minimizes model loss by using a gradient descent-like technique. Some major hyperparameters are the loss function, highest value of depth, count of estimators, sample fraction for base learner fitting, and split criterion.
13) Majority Voting (MV): Being itself an ensemble method, this process enforces the combination of classifiers by a vote. In this way, MV pools predictions from different conventional machine-learning algorithms to bolster classification performance.

5. Deep Learning

We have used the VGG16 architecture in its hybrid version in our ECG model. In the proposed model, multiple convolutional layers are combined with pooling operations as follows:

- In the first stage, two convolutional layers are using 8*8 filters with size of 3. After that a max-pooling layer with two size windows.
- At the second stage, 128 extractor filter is used followed by the same max-pooling layer as the first stage.
- The third stage involves the application of two convolutional layers, by increasing the number of filters to 265.
- The fourth stage consists of two convolutional layers and the number of filters is also duplicated to be 512.

After that, the output from these convolutional and pooling layers was fed into a one-layer stacked Deep Recurrent Neural Network with Long Short-Term Memory cells. Each LSTM cell processed a 2-D input of size 512 × 5 512×5. The DRNN produced an output with dimensions 128 × 2 128×2.
The algorithms used in the proposed work are listed below.

Algorithm 1 Main function

```
1: MAIN ()
2:   specific path /to file location
3:   LOAD data from path (data folder path)
       Process data from the rr interval folder and rp folder.
4:   if ¬rp data ∨ ¬rr_intervals data then
5:     return
6: end if
7:   A, B ←...   % properties
8:   B←tf. Keras. utils. categorical (Y, num_classes = 2)
9:   kfold ← StratifiedKFold (n_splits = 5, shuffle = True, random state = 7)
10: for data set (training set) i n kfold. split (A, B. argmax (1)) do
11:     m ← create new_model ()
12:     TRAIN_MODEL (m, A [train], B [train])
13:     EVALUATE_MODEL (model, A [test], B [test])
14: end for
```

Algorithm 2 Load and preprocess data

```
1:  Function LOAD data (data folder_path)
2:   rpeaks data ← {}
3:   rr_intervals data ← {}
4: for filename found in os. listdir do
5:     file_path ← os. join path.
6:     ...
7: end for
8:   return r_peaks_list, rr_intervals_list
```

Algorithm 3 Create a deep learning model.

1: **Function** CREATE_
new model ()
2: *model* ←
Sequential ()
3: ...

4: **return** *model*

Algorithm 4 Train and evaluate the model.

1: **Function** TRAIN_MODEL (*model*, *A*, *B*)
2: ...
3: **Function** EVALUATE_MODEL (*model*, *A*, *B)*
4: ...

A. Tabular Data

Algorithm 5 Random Forest Classifier

1: **Input**: Data *A*, target *B*
2: **Output**: Feature categories
3: Random Forest initialization
4: Data fitting
5: Get feature categories
6: Create a Data Frame to store feature categories
7: Sort the data frame by categories
8: Print top features

Algorithm 6 DBSCAN Clustering

1: **Input**: Data *A*
2: **Output**: Clusters
3: Extract features and target
4: Perform DBSCAN clustering
5: Visualize clusters

Algorithm 7 Regression Models

1: **Input**: Data *A*, target *B*
2: **Output**: Model MSEs
3: Split the data into training and testing sets
4: Initialize regression models (Linear, Ridge, Lasso, Polynomial)
5: Fit models to training data
6: Predict on testing data
7: Calculate MSE for each model

5.1 Results Comparison and Evaluation

Since difficulties were used in getting results from the ECG model, the expected results are compared rather than the real ones. The experimental results are listed in Table 1.

Table 1. Proposed experimental results.

	metrics	ECG model Tabular data	
1	Accuracy	88.02%	93%
2	Recall	83.34%	91%
3	F-score	84.24%	92%

6. Conclusion

Unfortunately, in the case of the ECG model, problems extracting R-peaks made it hard to feed the model with our data since it would not derive actual results, hence the failure of our attempt in trying to detect sleep apnea using ECG records. After several trials, it was concluded that the R-peaks files were corrupted and could not be processed. Off these tabular data, processing of this data can lead us to the conclusion that high-stress jobs and high blood pressure are two of the biggest risk factors for sleep apnea. Overall, we couldn't derive real comparison results since we couldn't even get results from the ECG model, while the tabular model results are not indicative of better performance compared to the expected ECG, since we had two different kinds of data, which makes it hard to give any objective conclusion.

Declarations

Ethics Approval and Consent to Participate

The results/data/figures in this manuscript have not been published elsewhere, nor are they under consideration by another publisher. All the material is owned by the authors, and/or no permissions are required.

Consent for Publication

This article does not contain any studies with human participants or animals performed by any of the authors.

Availability of Data and Materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing Interests

The authors declare no competing interests in the research.

Funding

This research was not supported by any funding agency or institute.

Author Contribution

All authors contributed equally to this research.

Acknowledgment

The author is grateful to the editorial and reviewers, as well as the correspondent author, who offered assistance in the form of advice, assessment, and checking during the study period.

References

[1] M. Bahrami and M. Forouzanfar. Detection of sleep apnea from single-lead ECG: Comparison of deep learning algorithms. IEEE Int. Symp. Med. Meas. Appl., pages 1–5, Jun 2021.

[2] Mahsa Bahrami and Mohamad Forouzanfar. Sleep Apnea Detection from Single-Lead ECG: A Comprehensive Analysis of Machine Learning and Deep Learning Algorithms. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 71, 2022.

[3] Amaral L. Glass L. Hausdorff-J. Ivanov P. C. Mark R. ... Stanley H. E. Goldberger, A. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online], 101, 2000.
[4] Herbert F Jelinek. A typical ECG signal showing the RR interval. ResearchGate.
[5] RG Mark AL Goldberger JH Peter T Penzel, GB moody. The Apnea ECG Database. Computers In Cardiology, 27:255–258, 2000.
[6] LAKSIKA THARMALINGAM. Sleep Health and Lifestyle Dataset. Kaggle.
[7] Sung-Hyun Yang Zubaer Md. Abdullah Al, Keshav Thapa. Improving R Peak Detection in ECG Signal Using Dynamic Mode Selected Energy and Adaptive Window Sizing Algorithm with Decision Tree Algorithm. MDPI.

Received: 21 Aug 2024, **Revised:** 30 Oct 2024,

Accepted: 24 Nov 2024, **Available online:** 26 Nov 2024.