Electronics and Electrical Engineering

AIMS Electronics and Electrical Engineering, 3(4): 347–358.
DOI: 10.3934/ElectrEng.2019.4.347
Received: 01 July 2019
Accepted: 21 October 2019
Published: 08 November 2019

http://www.aimspress.com/journal/ElectrEng

Research article

A machine learning approach using EEG signals to measure sleep quality

Maryam Ravan*

Department of Electrical & Computer Engineering, New York Institute of Technology (NYIT), 1855 Broadway, New York, NY 10023, USA

* **Correspondence:** Email: mravan@nyit.edu.

Abstract: Sleep quality has a vital effect on good health and well-being throughout a life. Getting enough sleep at the right times can help protect mental health, physical health, quality of life, and safety. In this study, an electroencephalography (EEG)-based machine-learning approach is proposed to measure sleep quality. The advantages of this approach over standard Polysomnography (PSG) method are: 1) it measures sleep quality by recognizing three sleep categories rather than five sleep stages, thus higher accuracy can be expected; 2) three sleep categories are recognized by analyzing EEG signals only, so the user experience is improved because fewer sensors are attached to the body during sleep. Using quantitative features obtained from EEG signals, we developed a new automatic sleep-staging framework that consists of a multi-class support vector machine (SVM) classification based on a decision tree approach. We used polysomnographic data from PhysioBank database to train and evaluate and test the performance of the framework, where the sleep stages have been visually annotated. The results demonstrated that the proposed approach achieves high classification performance, which helps to measure sleep quality accurately. This framework can provide a robust and accurate sleep quality assessment that helps clinicians to determine the presence and severity of sleep disorders, and also evaluate the efficacy of treatments.

Keywords: automatic sleep quality measurement; electroencephalography (EEG); machine learning; time-frequency features; dendrogram-support vector machine classifier

Abbreviations: EEG: electroencephalography; PSG: polysomnographic; SVM: support vector machine; EOG: electrooculogram; EMG: electromyogram; ECG: electrocardiogram; AASM: American academy of sleep medicine; REM: rapid eye movement; ANN: artificial neural networks;

KNN: K-nearest neighbor; HMM: hidden Markov model; LDA: linear discriminant analysis; DT: decision trees; NB: naive Bayes; RF: random forests; AW: awake; DS: deep sleep; LS: light sleep

1. Introduction

Sleep quality plays an important role in a person's learning ability, physical movement, and performance [1]. With the fast pace of modern life, millions of people suffer from poor sleep quality. Therefore, automated sleep quality measurement is of utmost interest and can help in evaluating the treatment progress in patients with common sleep disorders such as restless legs syndrome, insomnia, narcolepsy, and obstructive sleep apnea.

Sleep is characterized by continuous changes in respiration and heart beat rate, eye movement, muscles movements, and brain activity. Traditional polysomnographic (PSG) records different types of physiological data including the electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG) and electrocardiogram (ECG) to measure the sleep quality. The PSG recording then divides into 30 sec sequence of non-overlapping time windows (segments) based on the American academy of sleep medicine (AASM) manual [2] recommendation. Each segment is subsequently classified to one of the five sleep stages: 1) wakefulness (W), 2) rapid eye movement (REM), 3) stage 1 (S1); 4) stage 2 (S2); and 5) deep sleep, or slow wave sleep (SWS = S3 + S4) [2]. This segment duration is appropriate for hand scoring and can accurately reflect the macrostructure and time course of normal sleep, without the risk of having many stage shifts. In the case where two or more sleep stages occur on one segment, the stage that comprising the majority of the segment will be considered as the stage of the segment. Sleep stage scoring is the gold standard for analyzing human sleep [1], [3–7] that helps to identify the sleep stages that are vital in diagnosing and treating sleep disorders.

Sleep staging is usually conducted by specialized experts. This process, however, is hard, time consuming and error prone [8]. Many methods have been proposed for automatic sleep staging in order to reduce the required time and effort and reduce the number of errors [8]. In automatic sleep staging, classifiers are first trained using features associated with each segment of sleep data and its corresponding stage that is manually annotated by sleep specialists or neurologists. Then, the trained classifiers are used to automatically determine the sleep stage corresponding to each segment.

The traditional PSG approach use several sensors to measure EEG, EOG, EMG and ECG signals [9]. This can make users feel uncomfortable during sleep since a lot of sensors are attached on their body and scalp. On the other hand, the EEG signals are able to provide information about brain activities based on electrical recordings taken on the scalp of a subject. The EEG signals at different frequency sub-bands of alpha, beta, delta and theta show different characteristics during different sleep stages. This makes EEG signals as the most important signals in sleep stage classification regardless of manual or automatic classification [1]. Therefore to improve the user experience, automatic sleep staging based on measuring only EEG signals is of utmost interest among the sleep research community during the last decade [1,3–6].

Many different machine learning-based methods for automatic sleep stage classification (ASSC) have been proposed in the past. Approximately 31% of the ASSC methods use classification schemes that are based on support vector machine (SVM) classifiers, 22% based on artificial neural networks (ANN) classifiers, 11% based on linear discriminant analysis (LDA), 10% based on K-nearest neighbor (KNN), 5% based on decision trees (DT) and the remaining 21% based on other types such as naive Bayes (NB), hidden Markov model (HMM), fuzzy classification, and combined classification [1].

"Sleep-EDF Database [Expanded]" dataset is one of the most widely used sleep datasets in the literature for evaluating automatic sleep staging algorithms that is publicly available on the Physionet website (https://physionet.org/physiobank/database/sleep-edfx/) [10]. Here we briefly compare the performance of some of the most popular procedures based on "Sleep-EDF Database [Expanded]" dataset with highest classification performance among the available literature. We must note that comparison with the studies using other sleep databases and PSG signals is very difficult and is not considered in this study.

In study by [3] Zhu et al. used multiclass SVM classifier to classify the six sleep stages of W, S1, S2, S3, S4, and REM, where the algorithm achieved 87.5% classification accuracy. Liu et al. [4] performed sleep stage classification based on ANN classifier. They achieved an optimal classification accuracy of 89.5% to classify W, S1 + REM, S2, and SWS. Sanders et al. [5] used LDA classifier for sleep stage classification. Their proposed method correctly classified the five stages of W, S1, S2, SWS, and REM with an average accuracy of 75%. Phan et al. [6] used KNN to develop an ASSC system to classify the four sleep stages of W, S1 + REM, S2, and SWS. The classifier provided 94.49% accuracy. Aboalayon et al. [1] compared the performance of DT, SVM, ANN, and KNN to classify six sleep stages of W, S1, S2, S3, S4, and REM. DT classifier obtained the best overall classification accuracy with an average accuracy of 93.13%. DT classifier was followed by the SVM (92.37%), ANN (91.70%), and KNN (89.38%) in terms of classification accuracy.

In order to reliably estimate sleep disorders, it is essential to precisely estimate sleep quality parameters. Given the sleep cycles overnight, we can measure the sleep quality using the following three main parameters: 1) sleep latency, 2) sleep efficiency, and 3) percentage of deep sleep [11,12]. Specifically, sleep latency is the time that it takes to fall asleep after going to bed. Sleep efficiency is the ratio of the amount of time spent asleep to the total amount of time in bed. Percentage of deep sleep is the ratio of deep sleep to the all sleep stages. In order to calculate these parameters, in contrast to the previous studies on the classification of sleep stages [1,3–7], only three sleep categories need to be distinguished: wakefulness, light sleep + REM (S1, S2, REM), and deep sleep (S3, S4) [11]. Therefore, assessment of sleep quality can be made with significant reductions in cost and complexity. On the other hand, the sleep literature show that EEG signals of S1 and REM sleep are very similar [13]. In addition, there is a high variability in the EEG signals between and within subjects, especially in stages S1 and REM sleep [14,15]. Therefore, we attempted to classify the three sleep categories based on the EEG signals alone. We used an automatic sleep staging framework that consists of a multi-class support vector machine (SVM) classification based on a decision tree approach. We first trained and evaluate the performance of the SVM classifier by using polysomnographic datasets of first night of 8 healthy subjects from PhysioBank database with annotated sleep stages [10]. Then we test the performance of the classifier using the remaining 110 sleep datasets (nights) from the remaining 67 subjects. The performance of the classifier is also compared with widely used random forests classification approach.

The rest of the paper is organized as follow. Section 2 describes the dataset and the method. The classification results are discussed in Section 3. Conclusions are given in Section 4.

2. Materials and method

2.1. Subjects

In this study we used sleep dataset from SC Sleep-EDF Database [Expanded] that is freely

available through Physionet at "https://physionet.org/physiobank/ database /sleep-edfx/" for training, evaluation, and testing purposes [10]. We selected EEG signals recorded from 67 healthy subjects (female (n = 34, 50.74%), male (n = 33, 49.25%), mean age of 57.13 years (age range: 25–101 years with the standard deviation of 23.03 years)) without any medication for 24 hours sampled at 100 Hz. For each subject the EEG dataset were recorded for two nights. However, we considered one-night dataset for 13 of the subjects, since the dataset from the other night were noisy and could not be considered in the study. Also 3 of the subjects had just one-night dataset. Thus the total number of datasets is 118. Sleep stages have been scored manually according to Rechtschaffen & Kales (R & K) criteria [16] based on 30 sec segments of recordings. We selected Fpz-Cz and Pz-Oz EEG electrodes in our evaluations.

2.2. Feature extraction

As EEG signals are dynamic and mostly nonstationary for analysis, their frequency components are needed to know with the times at which they occur. Time-frequency analysis is especially suitable for addressing such issues [17]. We usually need more time accuracy in high frequency locations with transient waves, and more frequency resolution for slow waves. Such an analysis can be performed using wavelet transform (WT). We designed a wavelet packet tree (WPT) with 7 levels for this purpose. Daubechies wavelet of order 2 (db2) was applied to each 30 sec segments of EEG signal [18]. The frequency ranges of the EEG signal were divided into Delta (below 3.5 Hz), Theta (4–7 Hz), Alpha (8–13 Hz), and Beta (14–30 Hz) bands [19]. We also considered another frequency band named spindle frequency band in the sleep EEG dataset, because of presence of sleep spindles. We then manually selected the wavelet coefficients vectors (sub-bands) C_n (n = 1:38) that containing frequency information of the following 6 bands (Figure 1):

1. Delta: {0.39–3.13 Hz}, Wavelet coefficients = [C_{38}, C_{30}, C_{31}, C_{32}] = B1
2. Theta: {3.13–8.46 Hz}, Wavelet coefficients = [C_{33}, C_{34}, C_{22}, C_{23}, C_{35}] =B2
3. Alpha: {8.46–10.93 Hz}, Wavelet coefficients = [C_{36}, C_{25}] = B3
4. Spindle: {10.9–315.63 Hz}, Wavelet coefficients = [C_{26}, C_{27}, C_{28}] = B4
5. Beta1: {15.63–21.88 Hz}, Wavelet coefficients = [C_{16}, C_{17}] = B5
6. Beta2: {21.88–37.50 Hz}, Wavelet coefficients = [C_{18}, C_5] = B6

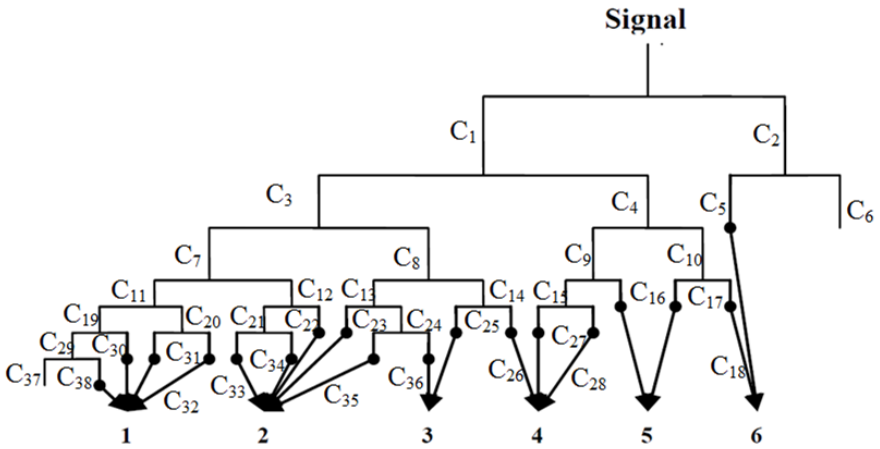


Figure 1. WPT and selected sub-bands.

For example, the frequency range of Delta frequency band is {0.39–3.13 Hz}. Based on the frequency range for each vector C_n, we need to concatenate several vectors to cover the frequency range for this specific band. Considering that C_{38} covers the frequency range of {0.39–0.78 Hz}, C_{30} covers the frequency range of {0.78–1.56 Hz}, C_{31} covers the frequency range of {1.56–2.34 Hz}, and C_{32} covers the frequency range of {2.34–3.13 Hz}, concatenating these vectors will cover the whole frequency range for Delta frequency band.

The following 32 statistical features were then extracted from each EEG segment to represent the time-frequency distribution of the segment for each electrode. Considering two EEG electrodes (Fpz-Cz and Pz-Oz) in this study, the total number of features are $2 \times 32 = 64$.

- Mean quadratic value or Energy (E1, E2, …, E6) of wavelet packet (WP) coefficients for each of the 6 bands (features 1–6)
- Total Energy (E7) (feature 7)
- Mean of the absolute values of the coefficients in each sub-band (features 8–13)
- Standard deviation of the coefficients in each sub-band (features 14–19)
- Ratio of different mean absolute values in different sub-bands (features 20–24)
- Shanon entropy of the vector B = [B1, B2, B3, B4, B5, B6] (feature 25)
- Permutation entropy [20]. (feature 26)
- Mean of each segment (feature 27)
- Maximum of each segment (feature 28)
- Minimum of each segment (feature 29)
- Median of each segment (feature 30)
- Standard deviation of each segment (feature 31)
- Mean of absolute differences (MAD) of each segment (feature 32)

$$MAD = \frac{1}{N}\sum_k |x(k) - x(k-1)| \quad (1)$$

where $x(k)$ is the kth time sample of segment x and N is the total number of samples (here $N = 30 \times 100 = 3000$).

Features 1–13 display the frequency distribution of the signal, features 14–24 show the amount of variation in the distribution of the frequency. Feature 25, Shannon entropy, describes the energy distribution of the wavelet coefficients. Shannon entropy can be used in sleep EEG signal processing since it has high values in wakefulness and REM sleep stages, and low values in SWS stages [21]. Feature 26, permutation entropy, is used to measure quantitative complexity for a dynamical time series. Features 27–32 mainly contain statistical measures that directly applied to the time series.

2.3. Feature selection

The second step in the machine learning process is feature selection, which is critical to the performance of the classifier. The goal in this step is to find a set of N_r features which are most relevant to differentiating between the three sleep categories. We tested widely used feature selection algorithms such as minimum redundancy maximum relevance (MRmR) [22]; fast correlation based feature selection (FCBF) [23]; t-test; and Fisher score algorithms to select most discriminating features between the three categories. The best performance was obtained using MRmR approach.

In order to avoid choosing features that are dominant in just a few patterns, a 10-fold cross validation procedure was used to select the best N_r features, where each fold contains approximately

the same number of segments for each sleep category. The 10-fold cross validation is an iterative process, where in each iteration, 9 folds are used for feature selection. In the proposed scheme, for each iteration, MRmR method is used to determine a list of the best kN_r, $k > 1$ features. The value of k is considered to be 2. After completing all iterations, the N_r features with the highest number of repetitions (probability of appearance) were selected as the final set of selected features. To avoid over-fitting, it is desirable to select N_r as small as possible.

2.4. Dendrogram multi-class SVM

In this study, we used a decision-tree-based support vector machine classifier named Dendrogram-SVM (DSVM) for classifying the three sleep categories [7]. The kernel function is Gaussian Radial basis function and the optimization technique is sequential minimal optimization [24] using the Statistics and Machine Learning Toolbox in MATLAB R2016. The rationale behind it is that combining decision tree architecture with binary SVMs benefits the advantages of the efficient computation of decision trees and the high classification accuracy of SVMs. The performance of DSVM approach is then compared with widely used random forests classification approach.

3. Results

3.1. Dendrogram generation

Figure 2 shows the hierarchical cluster analysis step yielded the dendrogram. At the top of the tree (i.e. the root node), the first binary classifier (SVM1) is trained to classify the awake class (a terminal node) as a negative class and the remaining merged two classes as positive class. Similarly, the second binary classifier in the tree (SVM2) is trained to classify the elements of deep sleep as negative class and the elements of light sleep + REM as positive class. In this approach, the hierarchical cluster tree is created using the smallest distance between elements in the two clusters, where pairwise distance between pairs of observations is correlation, which is one minus the sample correlation between them.

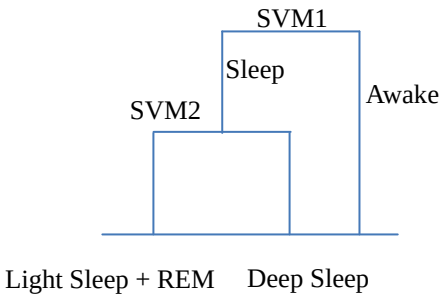


Figure 2. Dendrogram shows the multiple SVM classification generated for the three classes (awake, light sleep + REM, and deep sleep).

3.2. Evaluate the classification performance

To train DSVM and random forests classifiers, we try to find the least number of training samples for each sleep category with the best classification performance. To do that, we first

randomly selected 8 datasets from first night of 8 different subjects in PhysioBank database. The number 8 is the minimum number of subjects needed to have adequate number of training samples (i.e. non-overlapping 30 sec segments of subjects' EEG recordings), for each sleep category. We then used the EEG data of these 8 subjects and selected different set of training samples, where in each set the number of training samples N for the three sleep categories are the same. We started from N = 20 training samples for each sleep category and increased the numbers with the step of 20. In the third step, for each set of training samples, we evaluated the classification performance with the remaining samples from the same selected 8 subjects using the bootstrap approach when different number of best features are used starting from the best selected feature to all 64 features. The best features are selected using MRmR method with 10-fold cross validation. Using different set of training samples and number of features, we found that the best classification performance is achieved when 800 samples of each sleep category and $N_r = 15$ best features are used for training.

For the illustrative purpose, Figure 3 shows the average total classification accuracy for different numbers of training samples for both DSVM and random forests approaches using $N_r = 15$ best features. From Figure 3, the minimum number of training samples for each sleep category to achieve best evaluation performance is 800 for DSVM approach with the average classification accuracy of 94.2% and 750 for random forests with the average classification accuracy of 93.5%. Table 1 shows the corresponding sensitivity and specificity for each sleep category along with the total classification accuracy over all three sleep categories using $3 \times 800 = 2400$ training samples (800 sample for each category) for DSVM and random forests approaches. The total number of evaluation segments is 11676 including 4173 light sleep + REM, 526 deep sleep, and 6977 awake segments. The results show that the DSVM method could classify the three categories with higher sensitivity, specificity, and total accuracy. Furthermore, Figure 4 shows the total classification performance versus number of selected features for DSVM and random forests approaches using 2400 samples for training. From the figure, the least number of features with the best classification performance is $N_r = 15$. The performance is approximately the same with increasing the number of features up to 45 for DSVM approach and then the performance is decreasing by increasing the number of features from 45 to 64 that may be due to over-fitting for this classifier. For random forests approach, the performance remains approximately at the same level by increasing the number of features from 15 to 64. The classification performance versus number of training samples for other numbers of features and versus number of features for other numbers of training samples are not shown for the sake of brevity. The set of 15 most relevant features selected by the MRmR procedure is shown in Table 2, sorted in terms of the optimized MRmR value. This number of features is much lower than 2400 training samples to prevent over-fitting (the feature to training sample ratio is $15/2400 \times 100 = 0.625\%$).

Table 1. DSVM and random forests classification performance using 800 samples in each sleep category for training and the remaining 4173 light sleep + REM, 526 deep sleep, and 6977 awake samples for evaluation.

Class	Sensitivity DSVM	Sensitivity RF	Specificity DSVM	Specificity RF	Total accuracy DSVM	Total accuracy RF
LS + REM	89.0	87.3	97.3	97.2	94.2	93.5
DS	94.5	90.8	96.7	96.3		
AW	97.3	97.2	97.9	97.1		

Table 2. The N_r = 15 discriminating features.

Feature #	Feature	MRmR
1	Standard deviation of each segment (Pz-Oz)	0.9245
2	Standard deviation of sub band S6 (Pz-Oz)	0.9238
3	Mean quadratic value or Energy in sub band S6 (Pz-Oz)	0.9220
4	Mean quadratic value or Energy in sub band S1 (Pz-Oz)	0.8989
5	Max of the segment (Pz-Oz)	0.8906
6	Mean absolute value of sub band S5 (Pz-Oz)	0.8895
7	Mean of absolute differences of each segment (Pz-Oz)	0.8867
8	Mean absolute value of sub band S4 (Pz-Oz)	0.8841
9	Mean absolute value of sub band S1 (Pz-Oz)	0.8816
10	Mean quadratic value or Energy in sub band S2 (Pz-Oz)	0.8199
11	Permutation entropy (Fpz-Cz)	0.7602
12	Permutation entropy (Pz-Oz)	0.7446
13	The ratio of mean absolute value of sub band S3 to sub band S4 (Pz-Oz)	0.7013
14	The ratio of mean absolute value of sub band S5 to sub band S6 (Fpz-Cz)	0.6752
15	Mean absolute value of sub band S6 (Fpz-Cz)	0.5428

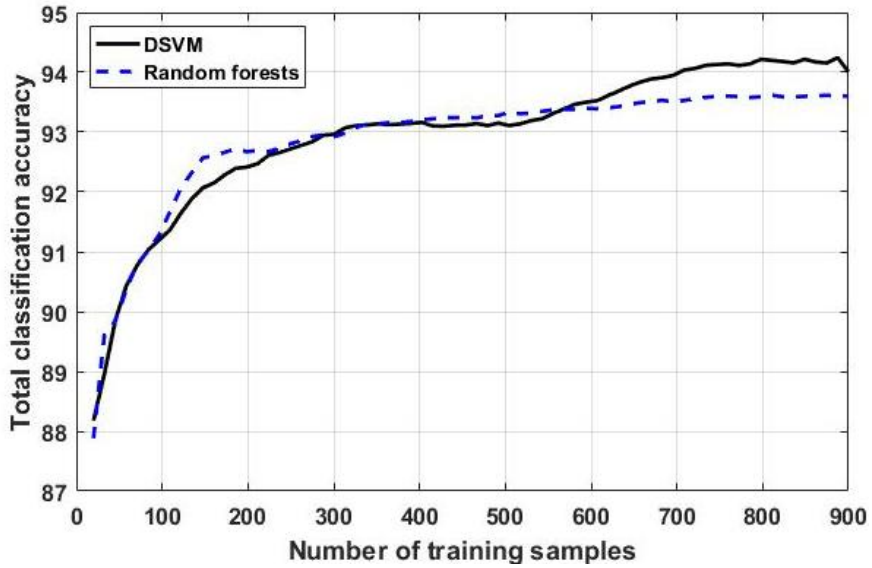


Figure 3. Total classification accuracy versus number of training samples in each sleep category using DSVM and random forests algorithms.

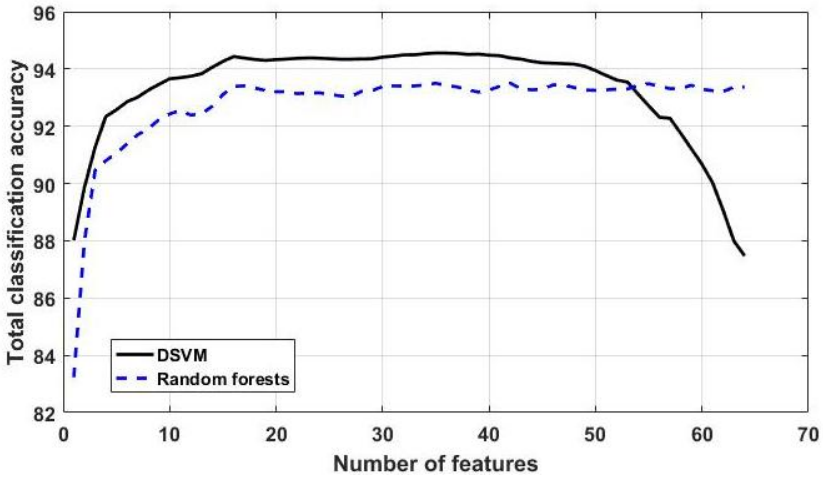


Figure 4. The classification performance versus number of features using 800 training samples for each sleep category.

3.3. Test the classification performance

In this section we test the trained classifier for the remaining 110 datasets (nights) from 67 subjects that are not used for training and evaluation of the classifiers. The classification performance of the DSVM and random forests approaches for three sleep categories of awake, light sleep + REM, and deep sleep using 91285 light sleep + REM, 10022 deep sleep, and 133246 awake segments for 110 test datasets are shown in Tables 3 and 4, respectively. From Tables 3 and 4, DSVM is capable of discriminating the three categories with the total accuracy of 91.4%, which is about 2.4% more than random forests approach (with the total accuracy of 89.0%) using only test dataset segments. Furthermore, the test accuracy for both approaches are very close to the training accuracy (Table 1). This confirms that the over-fitting has not occurred. Comparing with the preceding works using the same database and EEG signals [1,3–6], the performance of the proposed procedure obtained high accuracy rate even for the case where the test dataset are from 110 nights that were not used for training.

Table 3. DSVM classification performance using 91285 light sleep + REM, 10022 deep sleep, and 133246 awake segments for 110 test datasets.

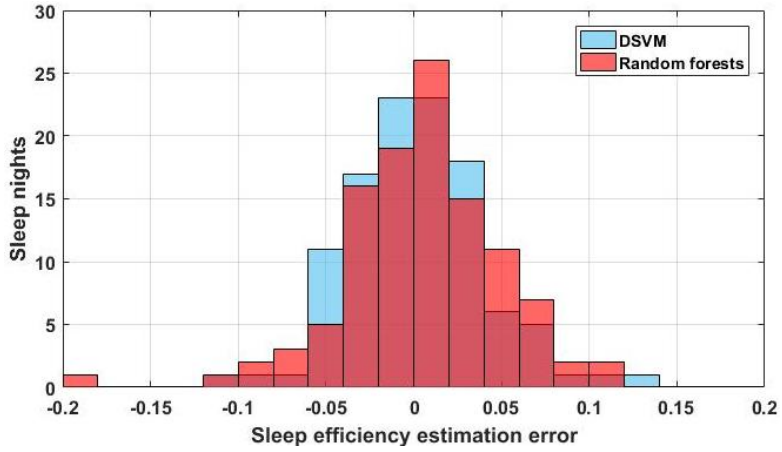
Class	LS + REM	DS	AW	Sensitivity	Specificity	Total accuracy
LS + REM	76914	6729	7642	84.3	96.1	
DS	991	8960	71	89.4	96.9	91.4
AW	4548	178	128520	96.5	92.4	

Table 4. Random forests classification performance using 91285 light sleep + REM, 10022 deep sleep, and 133246 awake segments for 110 test datasets.

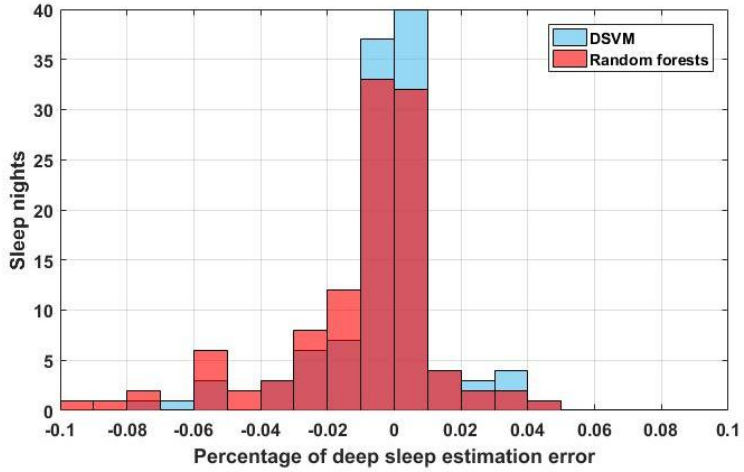
Class	LS + REM	DS	AW	Sensitivity	Specificity	Total accuracy
LS + REM	73119	8229	9937	80.1	95.5	
DS	1303	8633	86	86.1	95.8	89.0
AW	5185	1132	126929	95.3	90.1	

3.4. Sleep quality evaluation

To evaluate the accuracy for estimating the sleep quality, we calculate the sleep quality criteria using the true (annotated) and estimated (using machine learning) sleep categories. Of the three criteria of measuring sleep quality (i.e. sleep latency, sleep efficiency, and percentage of deep sleep), sleep latency is not recorded for the subjects in this dataset. Due to the lack of this information, sleep efficiency was calculated as the ratio of the time spent asleep to the total sleep time (from falling asleep time) detected by the EEG dataset. Figures 5(a) and (b) show the histograms of estimation errors of sleep efficiency and percentage of deep sleep, respectively for all 110 test datasets (sleep nights) using DSVM and random forests approaches. From the figures, the sleep efficiency error is less than 0.02 for 63 and 62 sleep nights for DSVM and random forests approaches, respectively. Furthermore, the percentage of deep sleep estimation error is less than 0.02 for 82 and 75 sleep nights for DSVM and random forests approaches, respectively. The mean ± mean absolute deviation (MAD) of the difference between true and estimated sleep efficiency and percentage of deep sleep for all 110 test datasets are shown in Table 5. From Table 5 and Figure 5, DSVM can estimate sleep quality with higher accuracy than random forests approach.



(a)



(b)

Figure 5. Histograms of (a) sleep efficiency estimation error and (b) percentage of deep sleep.

Table 5. Mean (±MAD) values of the difference between true and estimated sleep efficiency and percentage of deep sleep across all 110 test datasets.

Criterion	Estimated-True Mean (±MAD) DSVM	Estimated-True Mean (±MAD) RF
Sleep efficiency	2.93e-5 ±0.031	4.1e-3 ±0.033
Percentage of deep sleep	-3.3e-3 ±0.011	-1.2e-2 ±0.02

4. Conclusions

In this study, we developed a machine-learning algorithm based on Dendrogram Multi-Class SVM to detect the three sleep categories of light sleep+REM, deep sleep, and awake. Considering that standard PSG system may make users feel uncomfortable, our approach is specifically designed to recognize three sleep categories from two EEG electrodes only. We trained and evaluated the machine-learning algorithm using only 8 datasets from first night of 8 subjects available in Physiobank sleep database. We then tested the algorithm using the remaining 110 datasets from 67 subjects in comparison to widely used random forests algorithm. The results demonstrated that our approach can achieve high sensitivity, specificity and accuracy though only two EEG electrodes were used. The detected sleep categories were then being used to estimate the sleep quality. Comparing the estimated sleep quality criteria with the true ones (using visual annotations) revealed that the proposed approach could estimate sleep quality with high accuracy. This framework can help clinicians by reducing the analysis time of polysomnographic signals while enhancing the quantitative nature and robustness of the scoring procedure.

Conflict of interest

The author declares that there is no conflict of interest in this paper.

References

1. Aboalayon KAI, Faezipour M, Almuhammadi WS, et al. (2016) Sleep stage classification using EEG signal analysis: A Comprehensive Survey and New Investigation. *Entropy* 18: 272.
2. Iber C (2006) *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. 1st ed. Westchester, IL: American Academy of Sleep Medicine.
3. Zhu G, Li Y, Wen PP (2014) Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal. *IEEE J Biomed Health* 18: 1813–1821.
4. Liu Y, Yan L, Zeng B, et al. (2010) Automatic sleep stage scoring using Hilbert–Huang transform with BP neural network. *Proc IEEE Int Conf Bioinf Biomed Eng*, 1–4.
5. Sanders TH, McCurry M, Clements MA (2014) Sleep stage classification with cross frequency coupling. *Proc IEEE Ann Int Conf Eng Med Biol Soc*, 4579–4582.
6. Phan H, Do Q, Do TL, et al. (2013) Metric learning for automatic sleep stage classification. *Proc IEEE Ann Int Conf Eng Med Biol Soc*, 5025–5028.
7. Lajnef T, Chaibi S, Ruby P, et al. (2015) Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. *J Neurosci Meth* 250: 94–105.

8. Younes M (2107) The case for using digital EEG analysis in clinical sleep medicine. *Sleep Science and Practice* 1: 2.
9. Carden KA (2009) Recording sleep: The electrodes, 10/20 recording system, and sleep system specifications. *Sleep Medicine Clinics* 4: 333–341.
10. Goldberger AL, Amaral LAN, Glass L, et al. (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101: 215–220.
11. Zhao W, Wang X, Wang Y (2010) Automated sleep quality measurement using EEG signal: first step towards a domain specific music recommendation system. *Proc ACM Int Conf Multimedia*, 1079–1082.
12. Krystal AD, Edinger JD (2008) Measuring sleep quality. *Sleep Med* 9: S10–S17.
13. da Silva FHL, Niedermeyer E (1999) *Electroencephalography: basic principles, clinical applications, and related fields*. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins.
14. Buckelmuller J, Landolt HP, Stassen HH, et al. (2006) Trait-like individual differences in the human sleep electroencephalogram. *Neuroscience* 138: 351–356.
15. van Dongen HPA, Vitellaro KM, Dinges DF (2005) Individual differences in adult human sleep and wakefulness: Leitmotif for a research agenda. *Sleep* 28: 479–498.
16. Rechtschaffen A, Kales A (1969) A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. *Electroencephalography and Clinical Neurophysiology* 26: 644.
17. Thakor NV, Tong S (2004) Advances quantitative electroencephalogram analysis methods. *Annu Rev Biomed Eng* 6: 453–495.
18. Primer A, Burrus CS, Gopinath RA (1998) *Introduction to Wavelets and Wavelet Transforms*. Upper Saddle River, NJ: Prentice Hall.
19. Webster JG, Clark JW (1998) *Medical Instrumentation, Application and Design*. 3rd ed. Hoboken, NJ: Wiley.
20. Bandt C, Pompe B (2002) Permutation entropy: a natural complexity measure for time series. *Phys Rev Lett* 88: 174102.
21. Fell J, Roschke J, Mann K, et al. (1996) Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. *Electroencephalography and Clinical Neurophysiology* 98: 401–410.
22. Peng H, Long F, Ding C (2005) Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE T Pattern Anal* 27: 1226–1238.
23. Yu L, Liu H (2003) Feature selection for high-dimensional data: A fast correlation-based filter solution. *Proc Int Conf Machine Leaning*, Washington, 856–863.
24. Platt JC (1998) *Sequential minimal optimization: a fast algorithm for training support vector machines*. Technical Report.

AIMS Press