

Sleep Stage Classification using Laplacian Score Feature Selection Method by Single Channel EEG

Mahtab Vaezi¹, Mehdi Nasri^{2*}

1- Department of Biomedical Engineering, Khomeinishahr Branch, Islamic Azad University, Isfahan, Iran.

Email: mahtab.vaezi@iaukhsh.ac.ir

2- Department of Electrical Engineering, Khomeinishahr Branch, Islamic Azad University, Isfahan, Iran.

Email: nasri_me@iaukhsh.ac.ir (Corresponding author)

Received: February 2020

Revised: May 2020

Accepted: July 2020

ABSTRACT:

Sleep is a normal state in humans and the subconscious level of brain activity increases during sleep. The brain plays a prominent role during sleep, so a variety of mental and brain-related diseases can be identified through sleep analysis. A complete sleep period according to the two world standards R&K and AASM consists of seven and five steps, respectively. To diagnose diseases through sleep, it is necessary to identify different stages of sleep because the disorder at each stage indicates a certain disease. On the other hand, efficient and useful features should be selected to increase the accuracy of sleep stage classification. In this paper, at first, different statistical, entropy, and chaotic features are extracted from sleep data. Afterwards, by introducing and using the Laplacian score selector, the best feature set is selected. At the end, some conventional classification algorithms such as SVM, ANN and KNN are used to classify different sleep stages. Simulation results confirm the superiority of the proposed method based on the classification results. With the proposed algorithm, 2, 3, 4, 5 and 6 stages of sleep were classified by SVM and decision tree with 98.0%, 98.0%, 97.3%, 96.6%, and 95.0% accuracy, which are more superior to previous method's results.

KEYWORDS: Sleep Stage Classification, EEG, Laplacian Score, Chaotic Features.

1. INTRODUCTION

Sleep is a phenomenon of physiology in humans [1], and the unconscious level of the person increases in this case [2]. For this reason, it is possible to diagnose mental illness by carefully study sleep [1]. By studying sleep, you can diagnose various illnesses such as respiratory apnea [3], depressive [4], epilepsy [5]. To diagnose different diseases through sleep, it is necessary to identify the sleep stages at the beginning. For the first time in 1968, Rechtschaffen defined seven different stages for a complete sleep period. According to international standard Rechtschaffen and Kale (R&K) [6], which is the oldest and most complete sleep classification, a hole night sleep includes seven stages: Wake, movement time, stage 1, stage 2, stage 3, stage 4, Rapid Eye Movement (REM). In 2007, the American Academy of Sleep medicine presented a new classification for the sleep stages. According to American Academy of Sleep Manual [7], a whole night sleep contains five stages: wake, REM, stage 1, stage 2 and stage 3 + stage 4 (SWS). Disorder in each stage of sleep represents a particular condition. Brain frequencies appear in sleep during several stages. The

four brain waves that occur during sleep are Alpha, Beta, Theta and Delta. Alpha is a wave with a frequency of 8-13 (Hz) and it can appear in wake, S1 and REM stages. Beta is a wave with a frequency of more than 13 (Hz) and it can appear in wake stage. Theta is a wave with a frequency of 4-8 (Hz) and it can appear in S1, S2, S3 and S4 stages. Delta is a wave with a frequency of 0-4 (Hz) and it can appear in S3 and S4 stages. Because if this, S3 and S4 combine into one stage in AASM standard that its name is SWS [8]. Since a frequency appears at different stages of sleep, diagnosing and separating different stages of sleep require efficient algorithms to be able to correctly identify the stages of sleep.

It is obvious that, checking sleep process manually may reduce the accuracy of sleep stages classification. Therefore, in recent years many studies have been done on automatic classification of sleep stages [9-14]. In this process, inefficient features can greatly reduce the classification accuracy. Because of this, many studies have been done to select a suitable feature set for automatic classification of sleep stages. In sleep studies with attention to the use of labeled and unlabeled data, feature selector are separated in two groups of

supervised and unsupervised ones. In the supervised method, three feature selection categories are used: wrapper methods, embedded method and filter based method [8].

Wrapper based feature selection, operates on the basis of features and classifier and interacts with variables [9]. Seifpour et al. [8] by using statistical behavior of local Extrema signal and multi-cluster/class feature selection algorithm, was able to classify six stages of sleep. Pejman Memar et al. [10] to study the sleep stages, used the minimal-redundancy-maximal-relevance algorithm to choose the best features. By using wrapper method feature selectors, the type of classifier used has while targeting the useful features before the data classification. The second category of feature selection is embedded method. The embedded method has great influence on the correct selection of features and the selection of efficient features but it depends on the choice of the type of classification [11]. This greatly eliminates the dependency of wrapper method on the class type, but the type of classification also has a great influence on the accuracy of the detection of useful features and it raises the problem of automatic classification of sleep stages. Rahman et al. [12] by using the statistical features and neighborhood component analysis as feature selection, were able to automatically classify the six sleep stages by used EOG signal. In research, we need a property selector to provide useful features independently before classifying operations, so we go to the third method.

The filter method is well capable of detecting superior features even for very large data sets and it chooses the superior feature regardless of the type of classifier, which is a great advantage over the two previous methods [13]. Cho et al. [14] used the normalized mutual information feature selection method to select the best feature for automatically classify four sleep stages. Shahin Akhter et al. [15] used sequential forward feature selection to identify the best features to diagnose sleep apnea by snore sound. Depending on the type of class used, it is difficult to select the optimal property. In sleep study, we need a property selector to provide useful features independently before classifying operations. In addition, feature selections that can work on labeled and unlabeled data can be very useful and comprehensive.

In this paper, we use “Laplacian score” (LS) for feature selection in sleep study. To do this, a complete set of different statistical features, entropy features and chaotic features are extracted from the sleep data, at first. In sleep studies due to the high volume of sleep data, we need optimized features in automatic sleep monitoring algorithms to provide the best information in the shortest time possible. Therefore, in this study, Laplacian feature

selector is used to automatically classify sleep stages to improve classification accuracy and computational speed by identifying superior features. LS is a filter based method and is used for labeled and unlabeled data. This selector acts as a robust guide to class selection for labeled and unlabeled data and solves the problem of difficult clustering. LS on labeled data, selecting the most appropriate features and through the selection of optimal features improves classification accuracy [16].

The structure of the paper is as follows. At first, in Section 2, statistical, entropy and chaotic features are introduced on labeled data, and LS method will be examined. In the next step, in Section 3, different classification algorithms are used to classify the selected features and review the results. Finally, the paper is concluded in Section 4.

2. MATERIAL AND METHODS

In this section, the used database is explained in the first step. In the second step, we examine the proposed method. In proposed method, a complete set of different statistical, entropy and chaotic features are used in feature extraction stage. Fig. 1 shows the flowchart of the proposed method.

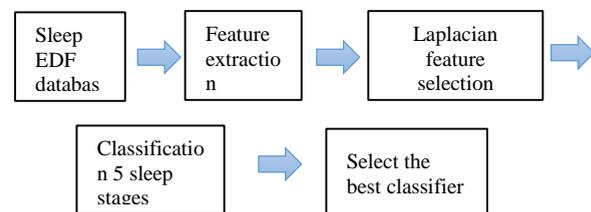


Fig. 1. Flowchart of the proposed method.

2.1. Database

To study the sleep stages and apply the suggested method, the Sleep EDF database, which is available on the Physionet¹ site, is used [17]. The database includes the bio signals such as EEG, EOG, EMG and body temperature of 20 hours during two subsequent day-night periods at the subjects' homes (Sleep Cassette (SC)), 153 healthy Caucasian men and women (aged 25-101 years) and 44 Caucasian males and females with taking temazepam (Sleep Telemetry (ST)). EEG signal is effective for sleep study. In EDF database, EEG signal recorded from two channel Pz-Oz and Fpz-Cz. Pz-Oz channel have high accuracy than Fpz-Cz channel according to [18, 19]. In the proposed method, we used single channel EEG (Pz-Oz) from 100 healthy persons with 100 Hz sample rate.

Due to the high duration of recording of sleep signals consider every 30 seconds of the signal in this database as an epoch. So, one hour is considered as 120 epoch.

¹ www.physionet.org

2.2. Feature Extraction

To identify the different stages of sleep, the second step is the extraction of the feature. The selection of effective features plays an important role in identifying the correct stages of sleep. In this section, different types of features are introduced, and used in sleep stage classification.

2.2.1. Statistical features

The first category of statistical features are mean, variance, skewness and kurtosis. Mean is a central tendency of a collection and the variance determines the extent of data dispersal from this mean point. Skewness asymmetry in a set and kurtosis describes the general form of the probability distribution. Mean, variance, skewness and kurtosis are calculated according to Equations 1-4 respectively.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (2)$$

$$x_s = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \mu}{\sigma} \right)^3 \quad (3)$$

$$x_K = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \mu}{\sigma} \right)^4 \quad (4)$$

In Equation 1-4, N displays the number of signal samples and x_i time series.

The second category of statistical features is Hjorth parameters [20]. It was first introduced by Bo Hjorth in 1970, and based on the definition of variance and use for sleep study in recent researches [21, 22]. The Hjorth parameters are used for statistical analysis of the EEG signal and include 3 parameters of activity, complexity and mobility which are calculated according to Equations 5-7 for $x(n)$ signal.

$$Hjorth_activity = \sigma_x^2 \quad (5)$$

$$Hjorth_mobility = \frac{\sigma_{x'}}{\sigma_x} \quad (6)$$

$$Hjorth_complexity = \left(\frac{\sigma_{x''}}{\sigma_{x'}} \right) \div \left(\frac{\sigma_{x'}}{\sigma_x} \right) \quad (7)$$

In Equations 5-7, σ_x is standard division of $x(n)$ signal derivative and $\sigma_{x'}$ is standard deviation of the second derivative of the $x(n)$ signal.

Statistical features examine the signal for dispersion, symmetry or asymmetry and other statistical tests. Now

let's examine the entropy set of features that examine the signal in terms of disorganization.

2.2.2. Entropy features

The entropy features examine the complexity and integrity of the EEG signal in various ways. Normalized Signal Entropy (NSE) is a kind of entropy that can be calculated from the normal power spectrum of the signal [23]. To classify the sleep stages, the first entropy feature used is NSE. In other words, this entropy is calculated by the short-time Fourier transform of the signal and indicates the irregularity of the power spectrum. NSE is calculated according to Equation 8.

$$NSE = \frac{1}{\log(N_f)} \sum_{f=F_1}^{F_2} S(f) \left(\log \frac{1}{S(f)} \right) \quad (8)$$

In Equation 8, F_1 and F_2 are high and low frequency power spectrum, S is spectrum power and N is the average frequencies.

The second entropy feature which is used for sleep stages classification is Shannon entropy [24]. Shannon entropy is defined as the index of variety and calculated according to Equation 9.

$$sh_entropy = \frac{1}{1-M} \log \left(\sum_{i=1}^N P_i^m \right) \quad (9)$$

In Equation 9, N is the number of samples, P is the sequence of signal probability and M is equal to 2 [10]. The third entropy feature which is used for sleep stages classification is Boltzmann's entropy. This entropy was first introduced by Ludwig Boltzmann in 1872 [25, 26]. The correlation of the parameters can be described in terms of order and disorder respectively [27]. Boltzmann's entropy calculated according to Equation 10.

$$Boltzmann_entropy = k \ln(W) \quad (10)$$

In Equation 10, k refers to Boltzmann constant equal to 1.38065×10^{-23} J/K and W refers to thermodynamic probability. After examining statistical and entropy features, we turn to chaotic features. In the signal, there are apparently irregularities in some places. The chaos examines the hidden rules of these irregularities that are rooted in the initial conditions.

2.2.3. Chaotic features

These features on EEG signals describe the basic concepts of nonlinear dynamics and the measurement of complexity and stability of signal.

Using the fractal dimension, can measure the complexity of a signal in nonlinear behavior, Petrosian Fractal Dimension (PFD) method is a quick and easy

calculation of fractal dimension through the Petrosian algorithm [10]. PDF is calculated according to Equation 11.

$$PFD = \frac{\log N}{\log N + \log N + 0.4M} \quad (11)$$

In Equation 11, N is number of sample and M is the number of signal changes in the derivative of the signal. The Normalized Line Length (NLL) method is another method for calculating the fractal dimension, which was first proposed by Esteller [10]. NLL can be calculated according to Equation 12.

$$NLL(n) = \frac{1}{M} \sum_{m=n-N}^n |x(m) - x(m-1)| \quad (12)$$

In Equation 12, N is the window length, n is number of sample and M denotes the number of movable and variable windows.

2.3. Feature Selection

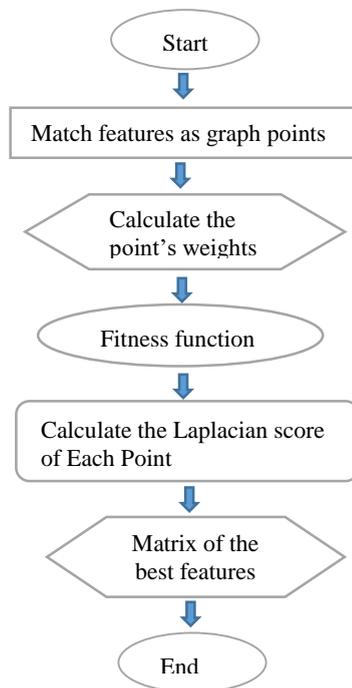


Fig. 2. Laplacian score algorithm for feature selection.

After extracting sleep signal features, inefficient features should be removed and effective features should be selected. In the proposed method, we use ‘Laplacian score’ (LS) for feature selection. LS is a kind of filter based feature selection method which is used for unsupervised and supervised data and taken from

Eigenmaps [28]. In the Laplacian method, each feature is considered as a point of the graph and the Laplacian graph is obtained. Fig. 2 shows the Laplacian algorithm [16].

In this algorithm, for two neighbor i, j , the weight matrix s can be calculated according to Equation 13.

$$S_{ij} = e^{-\frac{|x_i - x_j|^2}{t}} \quad (13)$$

In Equation 13, t is a constant, which is considered 1 in this paper. t is a global coefficient which does not affect the weight matrix in discrete Laplacian [28]. If $i=j$, s is calculated according to Equation 13 and otherwise $s=0$.

For r^{th} feature, the Laplacian score (Laplacian graph) is calculated according to Equations 14-18 [16].

$$f_r = [f_{r1}, f_{r2}, \dots, f_{rm}]^T \quad (14)$$

$$D = \text{diag}(S1) \quad (15)$$

$$L = D - S \quad (16)$$

$$f_r = f_r - \frac{f_r^T D 1}{1^T D 1} 1 \quad (17)$$

$$L_r = \frac{f_r^T L f_r}{f_r^T D f_r} \quad (18)$$

In Equation 18, L is the Laplacian graph and $\mathbf{1}$ is $\mathbf{1} = [1, 1, \dots, 1]^T$.

Laplacian feature selection gives us a matrix of best feature, so we can reduce the dimensions of the feature matrix. Reducing feature space will remove inefficient features. In the proposed method, at first features are extracted, and then the top features are selected by the supervised Laplacian score algorithm. LS is used for both unsupervised clustering and supervised classification problems, but it is usually used in a supervised manner in recent researches [29, 30]. The selected features by LS algorithm are Normalized Spectral Entropy (NSE), Boltzmann Entropy, Petrosian Algorithm (PFD), complexity, mobility, activity and Shannon entropy.

In general, the performance of other algorithms such as Fisher [31] is similar to LS algorithm, but the Fisher algorithm has a lower computational speed and does not clearly specify the final answer and because of its low

computational speed, it cannot be a good option for processing high-volume sleep data. Moreover, in the recent studies in the field of sleep, it has been shown that the Fisher feature selector cannot be a viable option for signal processing operations [32-34].

3. RESULTS

Sleep stages can be divided into 5 different classes, based on the number of stages that are identified and classified in each class. Classes 4th and 5th are defined by two global standards AASM and R&K which are fully described in the Introduction section. AASM Standard is newer and more logical than R&K, and R&K Standard is a more comprehensive classification of sleep. In the proposed method, a complete sleep period can be classified into different forms depending on the importance of each class for clinical use. The different classification for sleep stages in this paper is shown in Table 1.

Table 1. The challenge of different classification for sleep stages in this paper.

Number of class	Stages
6 class (R&K)	Wake, REM, S1, S2, S3, S4
5 class (AASM)	Wake, REM, S1, S2, SWS (S3+S4)
4 class	Wake, REM, S1 (S1+S2), SWS (S3+S4)
3 class	Wake, REM, NREM (S1-S4)
2 class	Wake, sleep (REM-NREM)

3.1. Evaluation Criteria

To evaluate the simulation results, three criteria of accuracy, sensitivity and specificity are used.

Sensitivity, specificity and accuracy are calculated according to Equations 19-21.

$$\text{Sensitivity} = \frac{T_{pos}}{T_{pos} + F_{neg}} \quad (19)$$

$$\text{Specificity} = \frac{T_{neg}}{T_{neg} + F_{pos}} \quad (20)$$

$$\text{Accuracy} = \frac{T_{pos} + T_{neg}}{T_{neg} + F_{neg} + T_{pos} + F_{pos}} \quad (21)$$

In Equations 19-21, T_{pos} is the number of target class that are correctly identified, T_{neg} is the number of non-target classes that are correctly identified, F_{pos} is

the number of target class that are incorrectly identified, and finally F_{neg} is the number of non-target classes that are incorrectly identified.

Table 2, shows the value of sensitivity, specificity and accuracy for five sleep stages by the proposed method and SVM classifier. In the simulation results section, we will investigate different classifiers, and their results on the proposed method. If all the results are correctly detected, the three values of accuracy, sensitivity, and accuracy are given 1, so the best result of sensitivity is for stage REM, the best result of specificity and accuracy are for stage 2 according to Table 2.

Table 2. Value of sensitivity, specificity and accuracy for five sleep stages by and proposed method with SVM classifier.

Stages	Sensitivity	Specificity	Accuracy
Wake	0.880	0.989	0.96
REM	0.887	0.987	0.95
S1	0.873	0.992	0.97
S2	0.860	0.997	0.99
SWS	0.880	0.989	0.96

For the best comparison between sensitivity, specificity, and accuracy of different five sleep stages, Fig. 3 shows the bar graph of the value of sensitivity, specificity and accuracy for five sleep stages by the proposed method. In the proposed method, for five stages of sleep, the best result of sensitivity is for stage REM and the best result of specificity and accuracy is for stage 2.

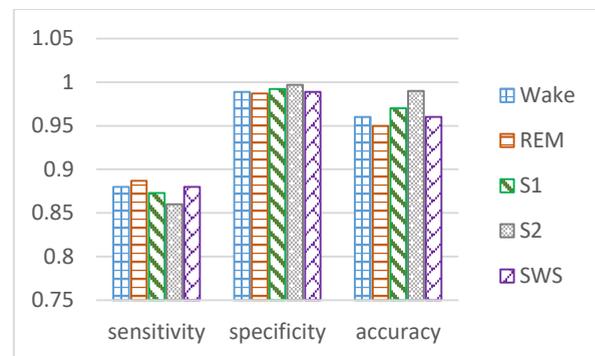


Fig. 3. Bar-graph of the value of sensitivity, specificity and accuracy for five sleep stages by the proposed method.

Table 3 shows the value of sensitivity, specificity and accuracy of the six stages of sleep classification by SVM classifier and Laplacian feature selector. According to Table 3, the best result of specificity and accuracy is for S2 and the best result of sensitivity is for stage S1.

Table 3. Value of sensitivity, specificity and accuracy for seven sleep stages.

Stages	Sensitivity	Specificity	Accuracy
Wake	0.776	0.987	0.94
REM	0.772	0.989	0.95
S1	0.786	0.985	0.96
S2	0.768	0.991	0.96
S3	0.772	0.989	0.95
S4	0.776	0.987	0.94

Fig. 4 shows the bar graph of the value of sensitivity, specificity and accuracy for six sleep stages by the proposed method. According to Fig. 4, the best result of sensitivity and accuracy is for stage 1 and the best result of accuracy and specificity is for stage 1. Despite the difficulty of detecting stage 1 of sleep, the proposed algorithm is capable of identifying stage 1 with high accuracy.

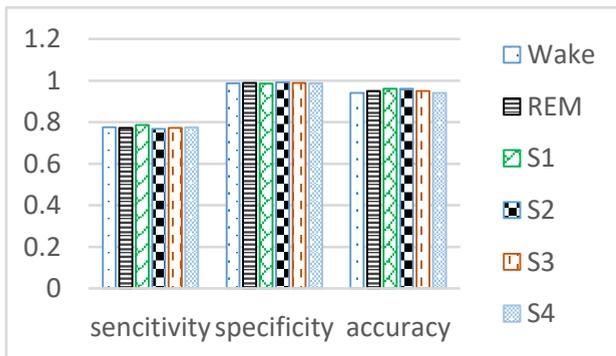


Fig. 4. Bar graph of the value of sensitivity, specificity and accuracy for six sleep stages by the proposed method.

Cohen's kappa is one of the most important criteria for agreement between two or more subject [35], [36] and is presented as an evaluation criterion in recent research on sleep stages [8, 37, 38]. Equation 22 shows the calculation of Kappa coefficient.

$$k = 1 - \frac{1 - p_o}{1 - p_e} \quad (22)$$

In Equation 11, P_o is the relative observed agreement among raters and P_e is hypothetical probability of chance agreement. The most ideal case for the kappa coefficient is 1. Comparison between kappa of recent studies and the proposed method is shown in Table 4. According to Table 4, the best kappa in 5 class (AASM) is for the proposed method.

Table 4. Comparison between kappa coefficient of recent studies and proposed method by SVM classifier.

Methods	2 class	3 class	4 class	5 class	6 class
Ref. [39]	0.98	0.947	0.93	0.8405	0.88
Ref. [40]	0.9435	0.8933	0.863	0.8543	0.836
Ref. [41]	-	-	-	0.865	-
Proposed method	0.961	0.959	0.956	0.957	0.9343

Table 5 shows performance comparison of recent studies on sleep study, and the results of the proposed method are compared with previous methods. As shown in Table 5, proposed method improved the accuracy of automatic classification of 2-6 sleep stages compared to previous research by selecting and using the effective features and using SVM classifier.

Table 5. Performance comparison of recent studies on sleep study and the results of the proposed method are compared with previous methods.

Methods	feature	3 class	4 class	5 class	6 class
Ref. [42]	Statistical feature	93.5 %	91.2 %	90.1 %	88.6 %
Ref. [43]	Wavelet	-	88.7 %	-	-
Ref. [40]	Statistical + wavelet	96.5 %	94.3 %	93.7 %	92.4 %
Ref. [8]	Local statistical behavior	94.5 %	92.8 %	91.8 %	90.6 %
Proposed method	Statistical + entropy + chaotic	98.0 %	97.3 %	96.6 %	95.0 %

3.2. Simulation Results

The simulation of the proposed method is done in the MATLAB 2018a. In the proposed method, SVM classifier is used to classify the superior features which are identified by LS algorithm. The confusion matrix of SVM with Laplacian score feature selection is shown in

Table 6, for the classification of the five stages of sleep according to AASM standard. According to Table 6, with 93.8% accuracy, sleep stages can be automatically classified by SVM. In addition, it can be seen that the proposed algorithm has a good ability to identify the SWS stage.

Table 6. The confusion matrix of SVM for the classification of the five stages of sleep according to AASM standard.

True Class	Wake	96%	4%	0%	0%	0%
	REM	2%	95%	3%	0%	0%
	S1	0%	2%	97%	1%	0%
	S2	0%	0%	1%	99%	0%
	SWS	1%	1%	0%	2%	96%
		Wake	REM	S1	S2	SWS
Predicted class						

Table 7 shows the confusion matrix of six stages of sleep classification according to R&K standard by SVM with Laplacian score feature selection. According to Table 7, six stages of sleep are classified by decision tree classifier and Laplacian feature selection, with 90.5% accuracy.

Table 7. Confusion matrix of six stages of sleep classification according to R&K standard by DA with Laplacian score feature selection.

True Class	Wa ke	94%	5%	0%	0%	0%	1%
	RE M	2%	95%	2%	0%	0%	1%
	S1	0%	1%	96 %	3%	0%	3%
	S2	0%	0%	2%	96 %	0%	2%
	S3	1%	0%	0%	2%	95 %	2%
	S4	0%	1%	2%	3%	0%	94 %
	Wak e	RE M	S1	S2	S3	S4	

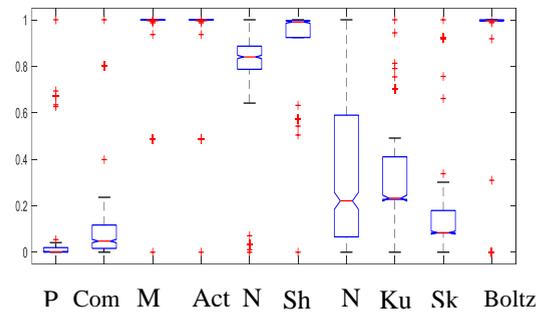


Fig. 5. The box plot of the features.

The box plot of the features is shown in Fig. 5. A box plot shows the dispersion of a feature and the line inside each box represents the median of its values. The larger the size of a box, the greater the dispersion in its values. To draw this box plot, we first normalized the values of all features to fit the values between 0 and 1. As shown in this figure, the best features are Normalized Spectral Entropy (NSE), Boltzmann entropy, Petrosian Algorithm (PFD), complexity, mobility, activity and Shannon entropy. Normalized Line Length (NLL), skewness and kurtosis are not suitable for algorithm.

3.3. Result of Different Classifiers

In this Section, four different classifiers named Support Vector Machine (SVM) with quadratic kernel, K Nearest Neighbor (KNN) with 9 number of neighbors, Multi-Layer Perceptron neural network, and Decision Tree (DT) with 100 maximum number of splits are used to classify sleep stages. Table 8 shows different classifier accuracies on the extracted features, without using the feature selector for classifying different sleep stages. As it can be seen in this table, the automatic sleep stages classification without using the feature selector and with a complete set of features has a low accuracy value.

Table 8. Accuracy of different classifiers used for automatic classification of sleep stages without using feature selection.

classifie r	2 class	3 class	4 class	5 class	6 class
SVM	96.0 %	91.9 %	96.0 %	93.0 %	91.7 %
KNN	96.0 %	80.5 %	80.0 %	82.8 %	77.6 %
ANN	94.7 %	91.0 %	88.9 %	91.3 %	84.7 %
Decisio n Tree	90.0 %	91.9 %	89.2 %	91.9 %	78.0 %

Table 9 shows four different classifiers on the superior features of using Laplacian score for feature selection. Laplacian score gives a matrix of best feature

and increases the accuracy of classification. According to Tables 8 and 9, the best classifier for automatic classification of sleep stages is SVM with quadratic kernel.

Table 9. Accuracy of different classifiers used for automatic classification of sleep stages with Laplacian score feature selection.

classifier	2 class	3 class	4 class	5 class	6 class
SVM	98.0 %	98.0 %	97.3 %	96.6 %	95.0 %
KNN	95.5 %	85.3 %	86.0 %	87.6 %	88.3 %
ANN	98.9 %	98.0 %	97.4 %	96.0 %	93.2 %
Decision tree	91.5 %	96.0 %	93.0 %	95.0 %	90.2 %

4. CONCLUSIONS AND FUTURE WORKS

In this paper, statistical, entropy and chaotic features were used to investigate and automatically classify sleep stages, and then the optimal properties were selected by the Laplacian algorithm, which is a filter-based feature selector. The use of superior features and the removal of ineffective features have been used to increase the accuracy of sleep classification. In the proposed method, different classifiers were used, and finally the SVM classifier stimulation results were evaluated as the best classifier. In the proposed algorithm, using Laplacian feature selector and SVM classifier, 5 and 6 stages of sleep were classified according to AASM and R&K standards with accuracy of 96.6% and 95.0% respectively. In general, the proposed algorithm was able to improve the accuracy of classification of different stages of sleep compared to previous research (classic research and deep learning research) and demonstrate the ability of filter-based feature selectors in sleep studies. It is very difficult to detect stage 1 in automatic algorithms in sleep study because of its high resemblance to wake and REM stages, but the proposed algorithm in classification of stage 1 sleep identified the first stage more accurately than other stages. Increasing the accuracy of sleep classification by increasing the number of statistical, entropy, chaotic and distance-based features and the use of efficient feature selectors are the topics of future researches in this field.

REFERENCES

- [1] S. Sanei, and J. A. Chambers, "EEG signal processing," 2007.
- [2] A. Sors, S. Bonnet, S. Mirek, L. Vercueil, J.-F. J. B. S. P. Payen, and Control, "A convolutional neural network for sleep stage scoring from raw single-

channel eeg," *Biomedical Signal Processing and Control*, Vol. 42, pp. 107-114, 2018.

- [3] E. Kakar, L. J. Corel, R. C. Tasker, R. de Goederen, E. B. Wolvius, I. M. Mathijssen, and K. F. J. S. m. Joosten, "Electrocardiographic variables in children with syndromic craniosynostosis and primary snoring to mild obstructive sleep apnea: significance of identifying respiratory arrhythmia during sleep," *Sleep medicine*, Vol. 45, pp. 1-6, 2018.
- [4] E. Toffol, M. Lahti-Pulkkinen, J. Lahti, J. Lipsanen, K. Heinonen, A.-K. Pesonen, E. Hämäläinen, E. Kajantie, H. Laivuori, and P. M. J. S. m. Villa, "Maternal depressive symptoms during and after pregnancy are associated with poorer sleep quantity and quality and sleep disorders in 3.5-year-old offspring," 2018.
- [5] S.-Y. Tsai, W.-T. Lee, S.-F. Jeng, C.-C. Lee, and W.-C. J. J. o. P. H. C. Weng, "Sleep and Behavior Problems in Children With Epilepsy," *Journal of Pediatric Health*, vol. 33, no. 2, pp. 138-145, 2019.
- [6] A. J. B. i. s. Rechtschaffen, "A manual for standardized terminology, techniques and scoring system for sleep stages in human subjects," *Brain information service*, 1968.
- [7] C. Iber, and C. Iber, "The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications", *American Academy of Sleep Medicine Westchester, IL*, 2007.
- [8] S. Seifpour, H. Niknazar, M. Mikaeili, and A. M. Nasrabadi, "A new automatic sleep staging system based on statistical behavior of local extrema using single channel EEG signal," *Expert Systems with Applications*, vol. 104, pp. 277-293, 2018.
- [9] Q. Al-Tashi, S. J. A. Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection," *IEEE Access*, Vol. 7, pp. 39496-39508, 2019.
- [10] P. Memar, and F. Faradji, "A novel multi-class EEG-based sleep stage classification system," *IEEE TRANSACTIONS ON NEURAL SYSTEMS*, Vol. 26, No. 1, pp. 84-95, 2018.
- [11] M. M. Rahman, M. I. H. Bhuiyan, and A. R. Hassan, "Sleep stage classification using single-channel EOG," *Computers in biology medicine*, Vol. 102, pp. 211-220, 2018.
- [12] N. El Aboudi, and L. Benhlima, "Review on wrapper feature selection approaches." pp. 1-5.
- [13] H. Rao, X. Shi, A. K. Rodrigue, J. Feng, Y. Xia, M. Elhoseny, X. Yuan, and L. Gu, "Feature selection based on artificial bee colony and gradient boosting decision tree," *Applied Soft Computing*, Vol. 74, pp. 634-642, 2019.
- [14] D. Cho, and B. Lee, "Optimized automatic sleep stage classification using the normalized mutual information feature selection (NMIFS) method." pp. 3094-3097.
- [15] S. Akhter, U. R. Abeyratne, V. Swarnkar, and C. J. J. o. C. S. M. Hukins, "Snore sound analysis can detect the presence of obstructive sleep apnea

- specific to NREM or REM sleep,” *Journal of Clinical Sleep*, Vol. 14, No. 06, pp. 991-1003, 2018.
- [16] X. He, D. Cai, and P. Niyogi, “**Laplacian score for feature selection.**” pp. 507-514.
- [17] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. J. C. Stanley, “**PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals,**” *AHA journals*, Vol. 101, No. 23, pp. e215-e220, 2000.
- [18] S.-F. Liang, C.-E. Kuo, Y.-H. Hu, Y.-H. Pan, and Y.-H. Wang, “**Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models,**” *IEEE Transactions on Instrumentation and Measurement*, Vol. 61, No. 6, pp. 1649-1657, 2012.
- [19] M. Ronzhina, O. Janoušek, J. Kolářová, M. Nováková, P. Honzík, and I. J. S. m. r. Provazník, “**Sleep scoring using artificial neural networks,**” Vol. 16, No. 3, pp. 251-263, 2012.
- [20] B. Hjorth, “**EEG analysis based on time domain properties,**” *Electroencephalography and clinical neurophysiology*, Vol. 29, No. 3, pp. 306-310, 1970.
- [21] S. Hartmann, and M. Baumert, “**Automatic a-phase detection of cyclic alternating patterns in sleep using dynamic temporal information,**” *IEEE Transactions on Neural Systems Rehabilitation Engineering*, Vol. 27, No. 9, pp. 1695-1703, 2019.
- [22] S. Bayatfar, S. Seifpour, M. A. Oskoei, and A. Khadem, “**An Automated System for Diagnosis of Sleep Apnea Syndrome Using Single-Channel EEG Signal.**” pp. 1829-1833.
- [23] T. Inouye, K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. Iyama, Y. Katsuda, and M. Hirano, “**Quantification of EEG irregularity by use of the entropy of the power spectrum,**” *Electroencephalography clinical neurophysiology*, Vol. 79, No. 3, pp. 204-210, 1991.
- [24] J. Lin, “**Divergence measures based on the Shannon entropy,**” *IEEE Transactions on Information theory*, Vol. 37, No. 1, pp. 145-151, 1991.
- [25] P. Perrot, *A to Z of Thermodynamics*: Oxford University Press on Demand, 1998.
- [26] J. L. Lebowitz, “**Boltzmann's entropy and time's arrow,**” *Physics today*, Vol. 46, pp. 32-32, 1993.
- [27] T. W. Deacon, “**Shannon–Boltzmann–Darwin: Redefining information (Part I),**” *Cognitive Semiotics*, Vol. 1, No. fall2007, pp. 123-148, 2007.
- [28] M. Belkin, and P. Niyogi, “**Laplacian eigenmaps and spectral techniques for embedding and clustering.**” pp. 585-591.
- [29] K. Benabdeslem, and M. Hindawi, “**Constrained laplacian score for semi-supervised feature selection.**” pp. 204-218.
- [30] J. Zhao, K. Lu, and X. He, “**Locality sensitive semi-supervised feature selection,**” *Neurocomputing*, Vol. 71, No. 10-12, pp. 1842-1849, 2008.
- [31] W. Malina, “**On an extended Fisher criterion for feature selection,**” *IEEE Transactions on pattern analysis machine intelligence*, no. 5, pp. 611-614, 1981.
- [32] Q. Gu, Z. Li, and J. Han, “**Generalized fisher score for feature selection,**” *arXiv preprint arXiv*, 2012.
- [33] Z. Zhao, L. Wang, and H. Liu, “**Efficient spectral feature selection with minimum redundancy.**”
- [34] R. J. Little, and D. B. Rubin, “**Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches,**” *Annual review of public health*, Vol. 21, No. 1, pp. 121-145, 2000.
- [35] K. J. Berry, and P. W. Mielke Jr, “**A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters,**” *Educational Psychological Measurement*, Vol. 48, No. 4, pp. 921-933, 1988.
- [36] J. Cohen, “**A coefficient of agreement for nominal scales,**” *Educational psychological measurement*, Vol. 20, No. 1, pp. 37-46, 1960.
- [37] M. Gaiduk, T. Penzel, J. A. Ortega, and R. Seepold, “**Automatic sleep stages classification using respiratory, heart rate and movement signals,**” *Physiological measurement*, Vol. 39, No. 12, pp. 124008, 2018.
- [38] S. Taran, P. C. Sharma, and V. Bajaj, “**Automatic sleep stages classification using optimize flexible analytic wavelet transform,**” *Knowledge-Based Systems*, Vol. 192, pp. 105367, 2020.
- [39] A. R. Hassan, and M. I. H. Bhuiyan, “**Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting,**” *Computer methods programs in biomedicine*, Vol. 140, pp. 201-210, 2017.
- [40] A. R. Hassan, and A. Subasi, “**A decision support system for automated identification of sleep stages from single-channel EEG signals,**” *Knowledge-Based Systems*, Vol. 128, pp. 115-124, 2017.
- [41] D. Jiang, Y.-n. Lu, M. Yu, and W. Yuanyuan, “**Robust sleep stage classification with single-channel EEG signals using multimodal decomposition and HMM-based refinement,**” *Expert Systems with Applications*, Vol. 121, pp. 188-203, 2019.
- [42] A. R. Hassan, M. I. H. J. B. Bhuiyan, and B. Engineering, “**Automatic sleep scoring using statistical features in the EMD domain and ensemble methods,**” *Biocybernetics and Biomedical Engineering*, Vol. 36, No. 1, pp. 248-255, 2016.
- [43] B. K. Kanoje, A. S. J. I. J. o. A. R. i. C. E. Shingare, and Technology, “**Automatic sleep stage detection of an EEG signal using an ensemble method,**” *International Journal of Advanced Research in Computer Engineering & Technology*, Vol. 3, No. 8, 2014.