International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.10, (2016), pp.335-344
http://dx.doi.org/10.14257/ijsip.2016.9.10.32

An Automatic Detection of Sleep using Different Statistical Parameters of Single Channel EEG Signals

Anand Prakash and Vandana Roy

Department of ECE
GGITS, Jabalpur, India

Abstract

The present work deals with the automatic detection of the sleep stages from the single-channel EEG data. Various stages of sleep are Awake, sleep stage 1, 2, 3 and 4 and rapid eye movement. Statistical attributes are extracted with the help of Ensemble Empirical Mode Decomposition, Hjorth parameter and zero-crossing rate. Ten-cross fold classification process is followed after best ranked attribute selection. After attributes are selected, the data is classified using bagging classifier. Accuracies of 98.46%, 95.62%, 93.87%, 93.17% and 91.93% for two-stages, three-stages, four-stages, five-stages and six-stages classification respectively. This classifier can be used for the real life application due to higher accuracies.

Keyword: *EEG, sleep stages, Bagging, EEMD, Hjorth parameter*

1. Introduction

Electroencephalograms (EEG) are the signals which are used for the study of the various types of processes and related problem of the brain. EEG signals are essential part of mental process such as learning and memory association, problem solving and dreaming. Sleep disorder or lack of sleep can cause severe health problem [[1]]. EEG along with electromyogram (EMG), electrooculogram (EOG) and eletrocardigram (ECG) are used for the determination of sleep stages. The sleep stage's study helps in the knowledge of the related disorder and its medication process.

Rech-tschaffen and Kale's (R&K) [[2]] and American Academy of Sleep Medicine (ASAM) [[3]] guidelines are two methods to determine the sleep stages. Expert visually score polysomnographic (PSG) recording for different stages. This study deals with the 6-stage classification according to the R&K criteria. 6-stages classification consists of Awake (Awa), stage 1 (S1), stage 2 (S2), stage 3 (S3), stage 4 (S4) and rapid eye movement (REM). In 5-stages classification, S3 and S4 of 6-stages combines to form slow wave sleep (SWS), 4-stages classification combines S1 and S2 of 5-stages and S1, S2, S3, S4 combines to form non- REM (NREM) in 3-stages classifications. Manual scoring can be a time consuming factor for the human. Also these score may consist of human error and variation with different expert scorer [[4]]. So, based on the statistical, spectral and non-linear features the classification process of different stages are done with the machine learning. Basic advantages of the machine learning are less error and time saving.

There are different techniques for the sleep stages analysis and classification. In major studies, the sleep stage classification is done with the help of Electroencephalogram (EEG), Electromyogram (EMG) and Electrooculogram (EOG) signals. [[5]] used two channel EEG, two channel EOG and one channel EMG and their spectral features were extracted. With these features and K-mean clustering of six-stage classification, 80.6% accuracy was obtained. Anderer *et. al.*, [[6]] implemented different spectral features of two EEG, two EOG and 1-chin EMG with linear discriminant analyser (LDA) for the experiment. Two stages, three stages, four stages and six stages classification experiment

ISSN: 2005-4254 IJSIP

were performed. Charbonnier *et. al.*, [[7]] also used the EEG, EMG and EOG data with various spectral and statistical features. Multi-layer perceptron gave 85.5% accuracy for the five-stage classification. Chaptot *et. al.*, [[8]] used four EEG data, one EMG data and one EOG data. Different statistical, spectral and non-linear features used with the multiple layer perceptron classification with 82% accuracy. Multichannel data is limited with the subject's movement and not favourable with the ambulatory environment.

Authors also worked on the single channel EEG. Fraiwen *et. al.*, [[9]] used Choi William distribution, Hilbert-Huang transform and continuous wavelet transform methods for the feature extraction in EEG datasets. Random forest was applied for the five-stage classification purpose and 83% accuracy was obtained with CWT method. Berthomier *et. al.*, [[10]] applied the iterative fuzzy logic method for the two-stage, three-stage, four-stage and five-stage classification using single channel EEG data. Spectral and temporal features are used for the feature extraction. Ronzhina *et. al.*, [[11]] used adaptive neural network with the spectral and statistical features of the single channel EEG data. Two-stage, three-stage, four-stage and six stage classifications and the accuracy of 96.7%, 88.97%, 81.42 and 76.7% was obtained respectively. Zhu *et. al.*, applied the visibility graph method for the feature extraction and LIBSVM classification method with RBF kernel. Two-stage, three-stage, four-stage, five-stage and six stage classification experiments was performed and maximum 97.9% accuracy was obtained. Hassan *et. al.*, [[13]] applied TQWT for the spectral feature extraction. These features were used with random forest algorithm. In another study Hassan *et. al.*, [[14]] used EMD process in extraction of statistical moment features and AdaBoost classifier. In both study, author performed the two-stage, three-stage, four-stage, five-stage and six stage classification. Author got the better result than previous authors in all cases except two-stage classification.

This paper presents the enhanced method of EMD. Ensemble Empirical Mode Decomposition is used in this paper for the statistical feature extraction along with the zero crossing rate and Hjorth parameter. Structure of this paper is presented as: Section 2 gives details of experimental datasets. Section 3 explain about the detailed methodology *i.e.*, feature extraction with EEMD, Hjorth parameters and Zero crossing rates, classifier, Chi square method, evaluation parameters and platform used for the experiment. Section 4 provides the experimental result and discussion based on these results. At last section 5 gives the conclusion about all the experiment.

2. Experimental Dataset

The dataset used in this experiment is taken from Physionet's Sleep-EDF (expanded) database [[15]-[17]]. Four subjects for this experiment were chosen randomly. The subject datasets consist of the two EEG channels (Pz-Oz and Fpz-Cz) and one EOG channel with 100 Hz sampling rate. These signals were digitised in the form of EDF data [[18]]. Expert scoring has been provided in the hyphnogram file. The expert scoring are according to the R&K criteria. So, 3000 samples are used for 30 s data epochs. The data needed to be converted into the ASCII format, so that it can be used in the further analysis. It is done with the help of polyman software. The scoring is provided in eight states: AWA, S1, S2, S3, S4, REM, Movement state and unknown state. Only Sub 2 shows the movement state. In the experiment, AWA, S1, S2, S3, S4 and REM are considered. The experiment is performed in the Pz-Oz channel for the better accuracy [[10]-[14]]. Details of the dataset are provided in the Table 1.

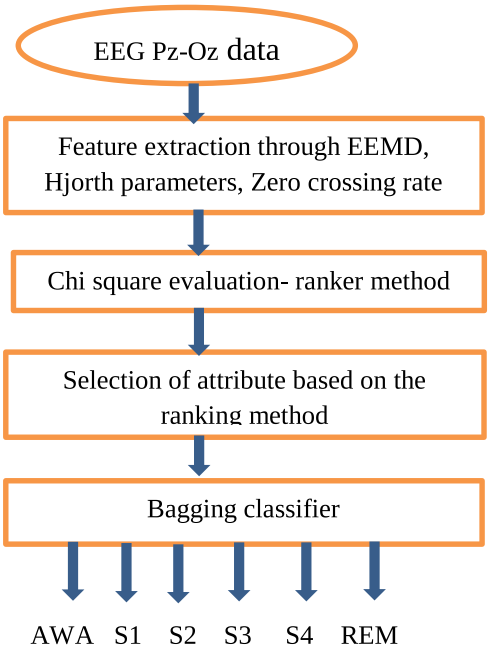
Table 1. Details of the Dataset

Subject	number of epochs
Sub 1 (ST4001e0)	2650
Sub 2 (ST4002e0)	2828
Sub 3 (ST4022e0)	2797
Sub 4 (ST4112e0)	2780
Total	11055

3. Methodology

Basic steps of proposed method for classification are shown in flowchart 1. All epochs are passed through the EEMD, Hjorth parameter and zero crossing rate. Features from these are extracted and evaluated using Chi-square distribution with ranker method. Best attribute is chosen from these. At last all the features are used with the Bagging algorithm for the two stages, three stages, four stages, five stages and six stages classification.
Details of each steps is explained in the upcoming section.

Flowchart 1: Methodology for the experiment



1.1. Feature Extraction

A. Ensemble Empirical Mode Distribution

The Empirical Mode Decomposition [[19]] decomposes a signal into a number of intrinsic mode functions through an iterative method termed as sifting. At first level the IMF1 is mean of upper and lower envelop of original EEG signal X(t) . Then residual signal is obtained by subtracting IMF1 from X(t). This process is iterated till stopping criterion is fulfilled (Residual signal energy content is close to zero). The remaining residual signal is

$$P_n(t) = P_{n-1}(t) - IMF_n(t) \quad (1)$$

where, $P_n(t) = X(t)$.
Finally, the signal is reconstructed by adding all IMFs and residual signal as

$$X(t) = P_n(t) + \sum_{i=1}^{N} IMF_i(t) \tag{2}$$

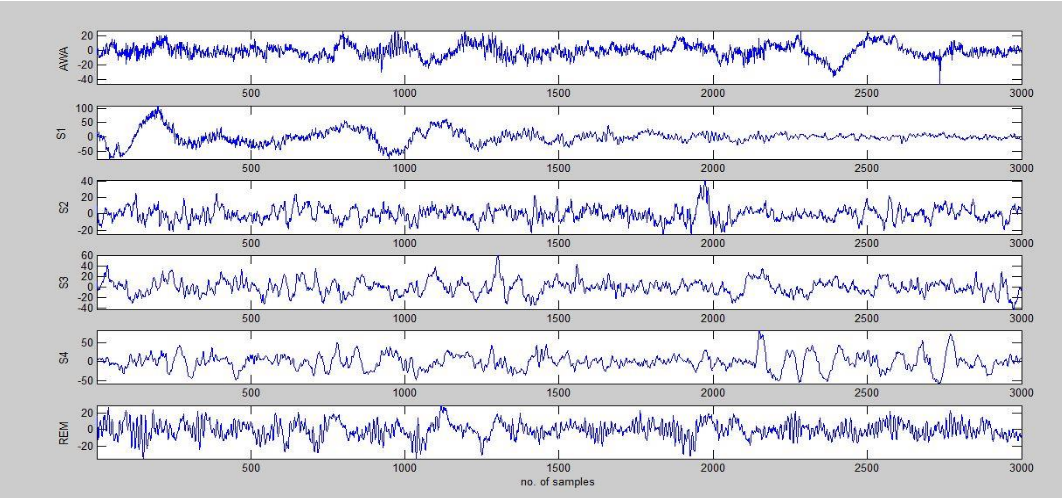


Figure 1. EEG Epochs of Different Stages

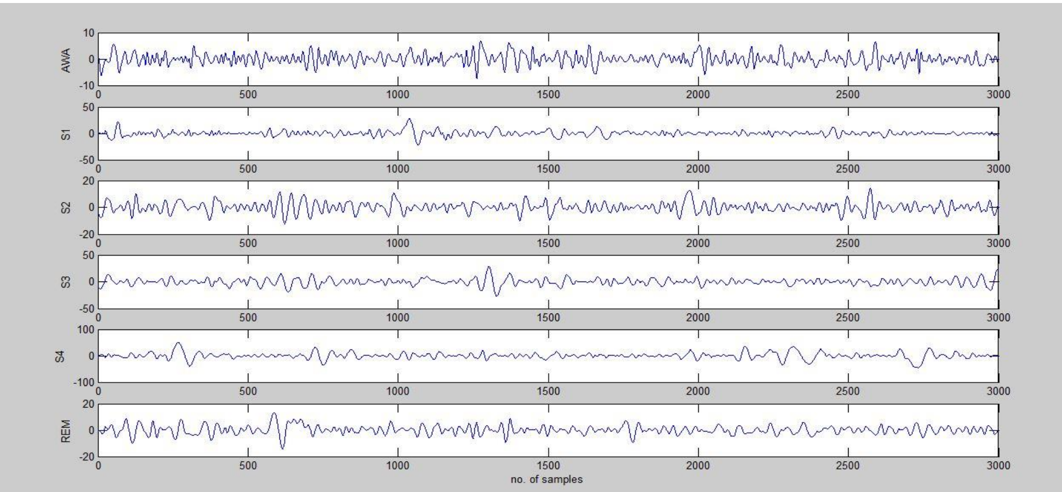


Figure 2. imf 3 from Each Stage Epochs

B. Statistical Feature Extraction with EEMD

We get k imf form the EEMD decomposition of epoch. Figure 1, shows the signal of different stage EEG signal and fig 2 shows the imf 3 component of respective data. These imf are shown in fig. These imf can be used to determine the statistical moment. Let us consider $p_i = [p_1, p_2, \ldots p_n]$, imf of n sample points. For each N imf, we can calculate the following:

- Mean (m) shows the central tendency of any data. So the central tendency of each k imf gives the detail of the signal. m can be calculated as given in eqn. (3)

$$m = \frac{1}{n}\sum_{i=1}^{n} p_i \tag{3}$$

- Variance (v^2) gives the dispersion of signal data against mean value (m). This is helpful in classification of REM sleep from S1 and S2. v^2 of each N imf can be calculated as given in eqn. no. (4):

$$v^2 = \frac{1}{n}\sum_{i=1}^{n}(p_i - m)^2 \tag{4}$$

- Skewness (s) provides asymmetry of the different signal. S of each N imf can be calculated as given in eqn. no. (5):

$$s = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_i - m}{v}\right)^3 \tag{5}$$

- Kurtosis (k) provides the peakedness value of data. N of each imf can be calculated as given in eqn no (6):

$$k = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_i - m}{v}\right)^4 \tag{6}$$

C. Hjorth Parameters

Hjorth parameters provide us activity, mobility and complexity of the signal [[7]]. Activity shows the signal power, the variance of the amplitude. Mobility provides a measure of standard deviation of slope with respect to standard deviation of amplitude. Complexity provides the number of standard slope actually generated during the average time required for generation of single standard amplitude as given by the mobility. Equation shows the formula for activity (A), mobility (M) and complexity (C). Hjorth parameter is applied in the EEG signal $p(n)$.
Activity (A) of $p(n)$ is given in eqn. (8).

$$A = var(p(n)) \tag{8}$$

Mobility (M) of $p(n)$ is given in eqn. (9).

$$M = \sqrt{\frac{var(p(n)\frac{dp}{dn})}{var(p(n))}} \tag{9}$$

Complexity (C) of $p(n)$ is given in eqn. (10).

$$C = \frac{M(p(n)\frac{dp}{dn})}{M(p(n))} \tag{10}$$

D. Zero Crossing Rate

Zero crossing rate is the number of time-domain zero-crossings within a defined region of signal, divided by the number of samples of that region.

$$ZCR = \frac{1}{N-1}\sum_{n=1}^{N} ¶\{p_n p_{n-1} < 0\} \tag{11}$$

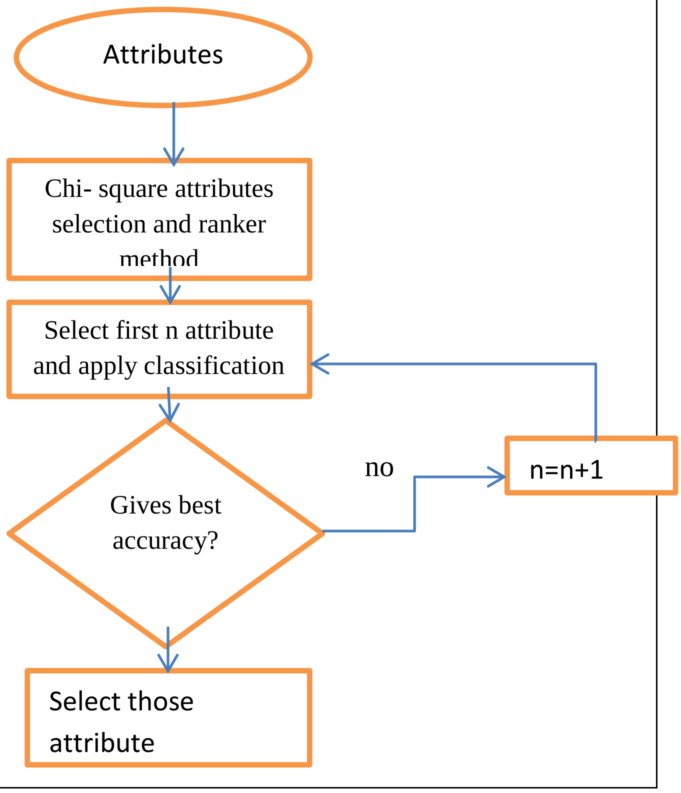
Here p is the signal and N is the length of the signal.

3.2. Chi Square Evaluation- Rankers Method

This method evaluates the worth of an attribute by computing the value of the chi-squared statistic with respect to the class. Ranker give ranks attributes by their individual evaluations.

Best attribute are selected on the basis of the ranking but we do not how much attribute is needed. So, this process is done with the help of flowchart 2.

Flowchart 2: Selection of best attribute



After the above flowchart evaluation, the best ranked attribute with the best result we got. These attribute will provide the best result out of total result.

3.3. Bagging

Bagging or bootstrap algorithm is a machine learning algorithm which is meant to provide the stability and accuracy of the classification. Each training data is used to create a unique class determiner. The basic of this flowchart is give below.

- Determine the number of cross fold, K.
- For i=1,2,…K
- Select 1/K as test data, and (K-1)/K as training data.
- Choose number of iterations, N
  - ❖ For j=1,2,….N
  - ❖ Select REP learner algorithm with 3 fold
  - ❖ j=j+1;
  - ❖ end
- K=K+1
- End

3.4. Interface

All the experiment is performed on the computer with Windows 10 operating system, Intel Pentium CPU N3540 2.16 GHz, MATLAB 2013a and WEKA 3.6.

Feature extraction process done with the help of the MALAB. MATLAB provides reading a various range format data. EEMD, Hjorth parameter and zero crossing rate on the epochs were applied with the help of MATLAB.

WEKA 3.6 is a machine learning software provided by The University of Waikato, New Zealand. This software provides the pre-processing, classification and clustering and attributes selection. In this study, the attribute selection, pre-processing and classification is used. Chi-square evaluation ranker method and classification with bagging is performed.

4. Result and Discussion

In this study the 10 cross fold validation is followed. In this process the data is divided into 10 subsets, 9 subsets are taken as training and 1 is for the test. This process is repeated 10 times. The data for this process is done from table 2 data. After applying Chi-square evaluation on the attribute, the data is passed through the different classifier present in the WEKA. This is done to find the best classifier out of them. According to no free lunch theorem some attribute may give best result with one class some gives with the other [[22]].

Table 2. Data Epochs of Different Stage

	AWA	S1	S2	S3	S4	REM
Number of epochs	7886	227	1559	360	369	654

SVM is used with the RBF polynomial, Naïve Bayes and logistic regression uses ridge (l2) with c=1 and random forest uses 10 tree method for the classification purpose. The comparison of different classifier is shown in Figure 2. This comparison chart gives the clear figure of performance of different classifier. Linear regression and random forest is approximately equal for each classification. Worst performance in these attributes is given by the SVM followed by the naïve bayes. Bagging algorithm gave the best result in all category when compared to all other algorithm.

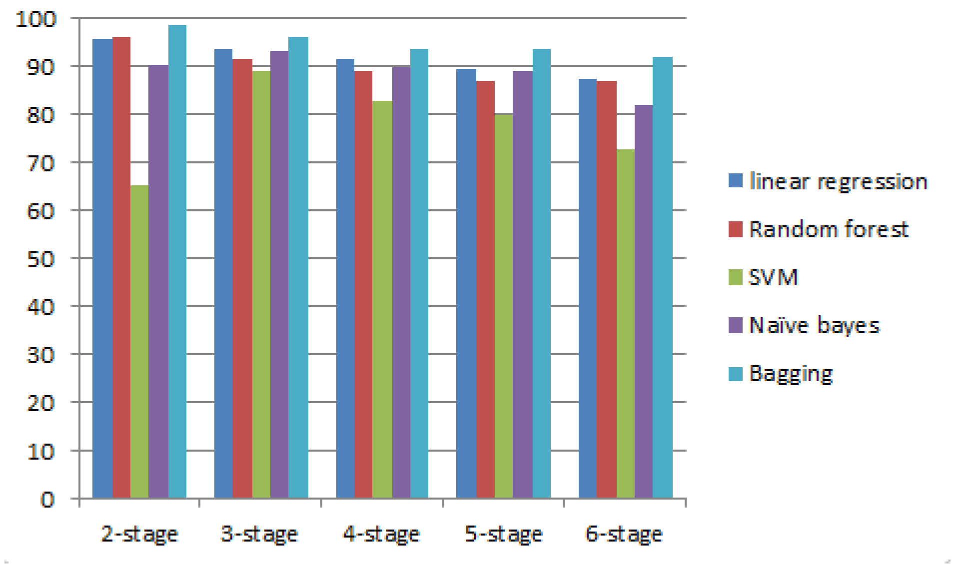


Figure 3. Comparison of Accuracies Different Classifier

After classification our method is compared with the existing method proposed by other author with different features and classifiers. The proporsed method is compared with Ronzhina *et. al.*, Berthomier *et. al.*, Zhu *et. al.* and Hassan *et. al.* classification methods. From Table 3, it can be observed that the proposed method is 0.48% more accurate in 2-stages classification process as compared to the [[12]] which has the best classification as compared to other authors. In 3-stages, 4-stages, 5-stages and 6-stages [[14]] had the best accuracy, but the proposed study experiment shows better accuracy as compared to the author in each stages. The 2-stage, 3-stage, 4- stage, 5-stage and 6-stage accuracy is 98.46%, 95.62%, 93.87%, 93.17% and 91.93% respectively.

Table 3. Performance Evaluation Compared with Other Method

Stages	Ronzhina et al [[11]]	Berthomier et al [[10]]	Zhu et.al [[12]]	Hassan et al [[14]]	Proposed method
2-stage	96.90%	95.40%	97.90%	97.73%	**98.46%**
3-stage	88.97%	88.30%	92.60%	93.35%	**96%**
4-stage	81.42%	74.50%	89.30%	91.20%	**93.75%**
5-stage	-	71.20%	88.90%	90.11%	**93.47%**
6-stage	76.70%	-	87.50%	88.62%	**92.06%**

In six stages classification, confusion matrix shows the classification of each stage category. This gives the actual picture of the classification. Sensitivity of any data shows the percentage of the data classified. In table 4, the sensitivity of the stage AWA is best; it means that the classification of the AWA is best. Only 0.72% of AWA data is classified in other category. For the S2 and S4, the data categories have good sensitivity of 88.71% and 79.95%. For REM stage it shows good sensitivity of 75.53%. It shows that, in S2 and S4 the misclassification is very less with 11.29% and 21.05% respectively. S1 and S4 have worst sensitivity. From table, it can be observed that the 90% data of S1 and 58.06% data have been misclassified in the other category. The attribute for these stages are not sufficient for the classification of these stages in the six-stage classification.

Table 4. Confusion Matrix of Six Stage Classification

Proposed method								
Expert scoring		**AWA**	**S1**	**S2**	**S3**	**S4**	**REM**	**Sensitivity**
	AWA	7830	1	19	3	3	30	99.28%
	S1	71	25	53	0	0	78	11.01%
	S2	15	2	1383	63	21	75	88.71%
	S3	5	0	120	151	82	2	41.94%
	S4	1	0	24	48	295	1	79.95%
	REM	36	4	121	0	0	493	75.53%

Table 5, shows the confusion matrix of 5-stages classification. In this stage the S3 and S4 is combined to form the SWS. For the AWA, the sensitivity is best. Only 58 out of 7886 data are misclassified in other category. For S1, the sensitivity is worst for 5-stage classification also. Most of them are misclassified in the AWA, S2 and REM categories. But when S4 is combined with S3 to form the SWS the sensitivity increased to 82.85%. S2 and REM have good sensitivity in 5-stage classification and shows improvement than 6-stage classification.

Table 5. Confusion Matrix of 5-Stage Classification

Proposed method							
Expert scoring		AWA	S1	S2	SWS	REM	Sensitivity
	AWA	7828	1	20	7	30	99.26%
	S1	62	27	57	2	79	11.9%
	S2	10	2	1372	104	71	88.01%
	SWS	4	0	18	604	3	82.85%
	REM	32	3	117	0	502	76.75%

4. Conclusion

In this study statistical feature extraction is done with the help of EEMD statistical moments, Hjorth parameters and Zero crossing rate. Data used for the feature extraction of sleep stages is single channel EEG. The proposed method gives better accuracy than other method. It also gives good sensitivity for AWA, S2 and S4 for six stage classification and AWA, S2, SWS and REM for the 5 stage classification. The accuracy was best in the all classification as compared to the other author. We get an overall improvement of 0.73% in 2-stage, 2.65% in 3-stage, 2.55% in 4-stage, 3.36% in 5-stage and 3.44% in 6- stage classification, we got the lower sensitivity for S1 in both 6-stage and 5-stage an S3 in 6 stage. The better sensitivity towards the SWS makes it the good classifier. This classification can be implemented in the real world application. In future the better attribute may be added to get the good sensitivity for S1 in 5-stages and 6-stages an S4 in 6-stages classification.

References

[1] K. N. Anderson, "An Update in Sleep Neurology: The Latest Bed Time Stories", J. Neurol., vol. 262, **(2015)**, pp. 487–491. doi:10.1007/s00415-014-7602-6.
[2] A. Rechtschaffen and A. Kales, "Manual of Standardized Terminology, Techniques and Scoring Systems for Sleep Stages of Human Subjects", U. G. P. Office, Washington DC Public Health Service, **(1968)**.
[3] A. L. C. C. Iber, S. A. Israel and S. F. Quan, "The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification", American Academy of Sleep Medicine, Westchester, USA, **(2005)**.
[4] N. A. Collop, "Scoring Variability Between Polysomnography Technologists in Different Sleep Laboratories", Sleep Medicine, vol. 3, **(2002)**, pp. 43–47.
[5] R. Agarwal and J. Gotman, "Computer-assisted Sleep Staging", IEEE Trans Biomed Eng, vol. 48, NO. 12, **(2012)**, pp. 1412–23.
[6] P. Anderer, G. Gruber, S. Parapatics, M. Woertz, T. Miazhynskaia, G. Kl¨osch, B. Saletu, J. Zeitlhofer, M. J. Barbanoj and H. Danker-Hopfe, "An E-health Solution for Automatic Sleep Classification According to Rechtschaffen and Kales: validation study of the Somnolyzer 24 × 7 utilizing the Siesta database", Neuropsychobiology, vol. 51, **(2005)**, pp. 115–133.
[7] S. Charbonnier, L. Zoubek, S. Lesecq and F. Chapotot, "Self-evaluated Automatic Classifier as a Decision-support Tool for Sleep/Wake staging", Comput. Biol. Med., vol. 41, **(2011)**, pp. 380–389.
[8] F. Chapotot and G. Becq, "Automated Sleep–Wake Staging Combining Robust Feature Extraction, Artificial Neural Network Classification, and Flexible Decision Rules", Int. J. Adaptive Control Signal Process, vol. 24, **(2010)**, pp. 409–423.
[9] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz and H. Dickhaus, "Automated Sleep Stage Identification System Based on Time–Frequency Analysis of a Single EEG Channel and Random Forest Classifier", Comput. Methods Programs Biomed., vol. 108, **(2012)**, pp. 10–19.
[10] C. Berthomier, X. Drouot, M. Herman-Stoca, P. Berthomier, J. Prado and D. Bokar-Thire, "Automatic Analysis of Single Channel Sleep EEG: Validation in Healthy Individuals", Sleep, vol. 30, no. 11, **(2007)**, pp. 1587–95.
[11] M. Ronzhina, O. Janouek, J. Kolrov, M. Novkov, P. Honzk and I. Provaznk, "Sleep Scoring using Artificial Neural Networks", Sleep Med Rev, vol. 16, no. 3, **(2012)**, pp. 251–63.

[12] G. Zhu, Y. Li and P. Wen, “Analysis and Classification of Sleep Stages Based on Difference Visibility Graphs from a Single Channel EEG Signal”, IEEE J Biomed Health Inform, vol. 99, no. 1, **(2014)**, pp. 1-8.

[13] A. R. Hassan, S. K. Bashar and M. I. H. Bhuiyan, “A Decision Support System for Automatic Sleep Staging from EEG Signals using Tunable Q-factor Wavelet Transform and Spectral Features”, Journal of Neuroscience Methods, vol. 271, **(2015)**, pp. 107–118.

[14] A. R. Hassan and M. I. H. Bhuiyan, “Automatic Sleep Scoring using Statistical Features in the EMD Domain and Ensemble Methods”, Biocybern. Biomed. Eng., vol. 36, no. 1, **(2016)**, pp. 248–255.

[15] B. Kemp. The Sleep-EDF Database [Onlne]. Available:http://www.physionet.org/physiobank/database/sleep-edf/**(2013)** June.

[16] B. Kemp, A. H. Zwinderman, B. Tuk, H. A. C. Kamphuisen and J. J. L. Oberye, “Analysis of a Sleep-Dependent Neuronal Feedback Loop: the Slow-Wave Microcontinuity of the EEG”, IEEE Trans. Biomed. Eng., vol. 47, no. 9, **(2000)** September, pp. 1185–1194.

[17] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng and H. E. Stanley, “PhysioBank, Physiotoolkit, and Physionet: Components of a New Research Resource for Complex Physiologic Signals,” Circulation, vol. 101, pp. e215–e220, **(2000)** June 13.

[18] B. Kemp, A. Varri, A. C. Rosa, K. D. Nielsen and J. Gade, “A Simple Format for Exchange of Digitized Polygraphic Recordings,” Electroencephalogr. Clinical Neurophysiol., vol. 82, **(1992)**, pp. 391–393.

[19] Z. Wu and N. E. Huang, “Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method,” Adv. Adapt. Data Anal., vol. 1, **(2009)**, pp. 1–41.

[20] A. Krakovska and K. Mezeiova, “Automatic Sleep Scoring: A Search for an Optimal Combination of Measures,” Artif. Int. Med., vol. 53, no. 1, **(2011)**, pp. 25–33.

[21] J. D. Jobson, “Applied Multivariate Data Analysis, Regression and Experimental Design,” I, Springer-Verlag, New York, **(1991)**.

[22] D. H. Wolpert, “The Lack of a Prior Distinctions between Learning Algorithms.” Neural Comput, vol. 8, no. 7, **(1996)**, pp. 1341–90.

