International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 3 Issue 8, August 2014

Automatic Sleep Stage Detection of an EEG Signal Using an Ensemble Method

Bharti Kamlesh Kanoje
MTech CSE-IT
Vishwakarma Institute of Technology
666 Upper Indira nagar
Bibvewadi pune

Ashwini S Shingare
Assistant Professor
Vishwakarma Institute of Technology
666 Upper Indira nagar
Bibvewadi Pune

Abstract— **Now a days, sleep related problems increases day by day, which is a problem of public health, where lot many people suffering from sleep disorder that affects their normal day to day life. For diagnosis of sleep disorder many methods was available for detection of sleep disorder where sleep stages are used as evidence for further detection. There are two standards available for scoring R&K and AASM and according to Rechtschaffen and Kales (R&K) rule; Electroencephalogram (EEG) signal is analyzed by dividing each signal into periods of 30 seconds small parts called segments. Previous method of sleep detection method uses manual visualization where specialist divides signal into 30 second epoch and assign sleep stages according to their wave frequencies and characteristics and this method is a time consuming method with errors. For this purpose, automatic sleep stage classification is used which reduce the manual tasks and improve stage detection and classification accuracy. Sleep stage detection comprises detection and classification of the data into classes such as awake NREM and REM stages. Sleep identification involves feature extraction from each segments, classification of each selected features and ensemble of result of each classifier. For feature extraction of an EEG signals, Discrete wavelet packet transform (DWPT) method is used, base classifiers like Artificial Neural Networks, K-Nearest Neighbor, M-SVM applied for classification and then ensemble method was deployed in our approach for getting better accuracy by combining the results of all classifier using majority voting method.**

Keywords **—Electroencephalogram (EEG) signal, Sleep staging, Feature extraction, Ensemble learning, Artificial Neural Network (ANN), k-Nearest Neighbor (KNN), Discrete Wavelet Packet Transform (DPWT), Standard Deviation (SD), Support Vector machine (SVM).**

I INTRODUCTION

Sleep plays a vital role in quality of life, good health, protects mental and physical health, of human being throughout their life. Now a day there are many people who suffers from sleep disorders like physiological and neurological disorders such as apnea, insomnia, and narcolepsy etc., which has great socioeconomic concern. Sleep analysis is used for diagnosis and treatment of such type of disorders. The analysis of sleep can be done by studying the electroencephalogram (EEG) signals, which is usually based on periods of 30 seconds length segments called epoch and this method of sleep staging is also called polysomnography (PSG), which is a common procedure for the detection and diagnosis of sleep disorders.

Sleep study is normally done based on one of two available standards: the Rechtschaffen and Kales (R&K) [1] and the new standard developed by the American academy of sleep medicine (AASM) [2]. Human sleep is characterized by 5 stages such as Awake (W), Non-REM1, Non-REM2, Non-REM3 and Rapid Eye Movement (REM) sleep. Characteristics of sleep are described as below.

1.1 Wake Stage

The Wake stage is characterized by low amplitude and mixed frequency of EEG signal. Generally, Wake stage is seen in the beginning of the sleep and it can be defined as a transition stage from the full alertness state to the half-sleepy state. This stage consists of alpha and beta activities which are not similar to the slow activities seen in SWS, along with the eye movements and high muscle tone.

ISSN: 2278 – 1323

1.2 NREM Stage 1

In stage1 of NREM, the EEG signal has the highest amplitude nearly 50-70µV with frequency ranges of 2–7 Hz. This stage acts as a transition stage between wakefulness and sleep, which usually lasts between 1 to 5 minutes. This stage consists of a low-voltage EEG tracing with well-defined alpha and theta activity, occasional vertex spikes, and slow rolling eye movements (SEMs). Stage N1 constitutes 2% to 5% of total sleep time and exhibits the absence of sleep spindles, K-complexes, and REM sleep waves [6].

1.3 NREM Stage 2

N1 stage is followed by the stage 2 which is also called as deeper sleep stage, which lasts for approximately 10 to 20 minutes. Stage N2 is act as a "baseline" of sleep and it is characterized by the occurrence of sleep spindles and K-complexes. K-complex is a transient waveform with a negative deflection called upward deflection followed by positive sharp wave called downward deflection. This relatively large waveform has duration of between 0.5 and 1.5s duration and generally larger than 75 µV amplitude and a relatively low-voltage, mixed frequency EEG background. Alternatively, high voltage delta waves may comprise up to 20% of stage 2 epochs. Complete sleep duration may consist of 45-55% of Stage 2 [4].

1.4 NREM Stage 3

Stage 3 is referred to a period during which at least 20% and not more than 50% of the sleep consists of EEG waves with frequencies of 2 Hz or smaller and amplitudes of more than 75 mV (delta waves). This stage normally appears only in the first one-third of the sleep episode, and it usually compromises 4-6% of whole sleep time. Stage 4 is similar to Stage 3, except that delta waves cover 50% or more of the record. Sleep stage 4 usually represents 12-15% of total sleep time. Stages 3 and 4 together are also known as "deep sleep" or, slow wave sleep (SWS) and this is the most restorative part of sleep. These two stages merely represent the same behavior hence according to the new standard developed by the American academy of sleep medicine (AASM), both are combined to represent single stage as NREM 3. In our experimental evolution, AASM standard was used for scoring purpose [2].

1.5 REM Stage

REM is the sleep stage in which dreaming occurs and it constitutes up to 20-25% of total sleep. It is well characterized by the rapid eye movements under the closed eyelids and has low voltage EEG patterns. In this sleep stage, the brain activity of human beings is reversed from stage 4 to a stage1. The REM stage contains low voltage, mixed frequency EEG signal with saw-tooth wave-like pattern, EMG with low amplitude, and high amplitude EOG signal from both eyes [5]. Table 1 shows the characteristic of the brain waves of an EEG signal with their features, frequency, amplitude, and different sleep stages [8].

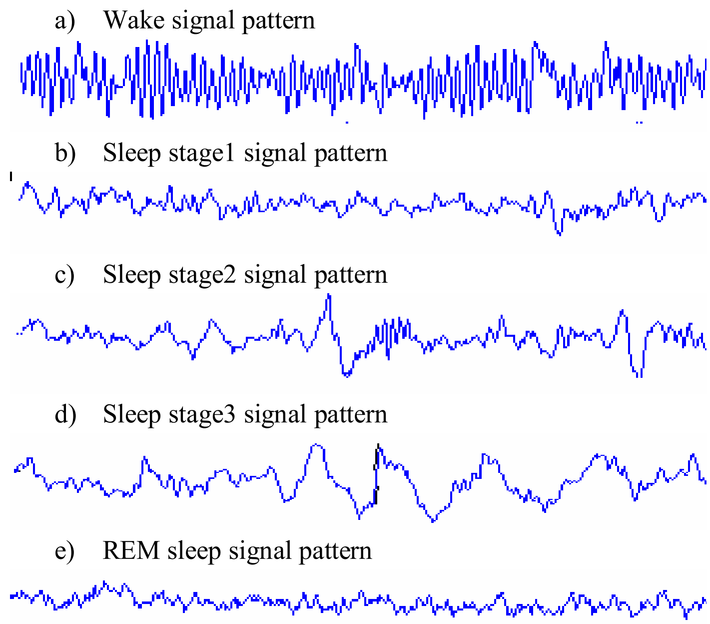

Figure 1: Different patterns of an EEG signal from different sleep stages

Figure 1 shows the pattern of EEG signal from different sleep stages [3]. For this experiment, first the input signal is given to the system and then signal is segmented into 30 seconds epoch, on each segments Discrete Wavelet packet transform was applied and wavelet coefficients are extracted from each segments. For classification of the extracted features three base classifiers are used, such as ANN, KNN and SVM. Again on the classified data ensemble classifier method was applied aiming to improve sleep stage classification accuracy.

Table I: Characteristics wave of EEG signals and their respective sleep stages

Feature	Frequency	Amplitude	Stages
Alpha	8-13 HZ	20-60 µV	Present in awake, stage I and REM
Beta	13+ HZ	2-20 µV	Dominant in Activity awake
Theta	4-8 HZ	50-75 µV	Present in stage I,II,III and IV
Delta	0-4 HZ	75+ µV	Present in stage III and IV
Sleep Spindles	12-14 HZ		Present in stage II
K-Complex	0.5-1 HZ		Present in stage II

ISSN: 2278 – 1323

II MATERIALS AND METHODOLOGY

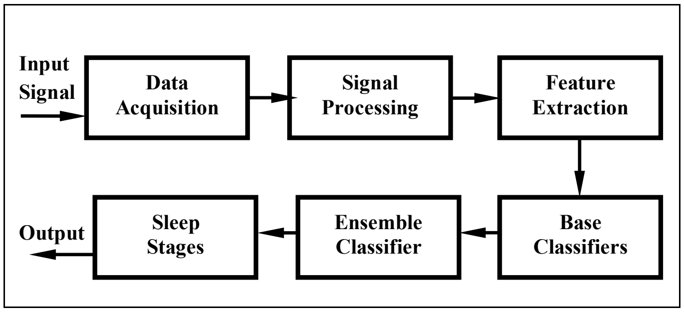

Figure 2: Block diagram of sleep stage classification

As shown in the figure 2, the data taken from subjects whose stage information is tabulated in Table 2 were used. The sleep epochs were prepared by divided signal into 30 seconds segments. On each segments wavelet transform technique is applied to extract the features. In this proposed system, discrete wavelet packet transform method was applied on each segment to extract the features. After feature extraction, feature selection for further classification is done where some statistical features are selected to reduce the formation of large data set. The selected feature vector is passed to the three base classifiers, which are designed to discriminate between different sleep stages. Training of these classifiers was done in following manner, first acquired signal is divided into each 30 seconds in signal processing step and then DWPT is applied to extract the features and those features are passed to the base classifiers. The output from the base classifier is given as input to the ensemble classifier to improve classification accuracy for sleep stages. The proposed system can be elaborated in various sectional parts as illustrated below.

1 DATA COLLECTION

The data to be used in the sleep experimental procedure was downloaded from the site www.physionet.org/physiobank/database/sleep-edf, where all sleep database available. EEG signals recordings were taken from Caucasian males and females ages up to 21-35 years old without any medication for 24 hours and these signals are sampled at 100 Hz. The Physio-Bank ATM stores the sleep data in the European Data Format (EDF) [7]. The scoring of signal also called hypnograms were scored using Rechtschaffen and Kales rules which is based on 30-second segments from Fpz-Cz / Pz-Oz EEG channel. In this experiment Pz-Oz channel is used for evaluations. Whole 24 hours recordings were divided to 30 seconds long epochs and each epoch was scored as Awake, Non-REM1, Non-REM2, Non-REM3 and REM by a sleep expert. For this experimental evaluation record name Sc4012eO were used which is a 24 hour data taken from single channel Pz/Oz EEG channel. Table 2 shows number of epoch in each stage, of the subject used for experiment.

Table II: Subjects epoch Information

Data set/state	**Sc4012eO**
Awake	1824
Stage 1 + REM	268
Stage 2	660
Stage3 + 4 (SWS)	96
Total segments	2848

2 SIGNAL PROCESSING

The purpose of pre-processing step was to enhance the EEG signal by normalizing. EEG signal was segmented into epoch of 30 seconds each with each epoch corresponds to different sleep stages. Usually, this signal is composed of mixture of brain signals like alpha wave, beta, theta wave, delta, spindle, saw-tooth, and K complexes waves and all these waves represents different sleep stages with different characteristics.

3. FEATURE EXTRACTION

3.1 Time-frequency analysis

EEG signals are mostly dynamic, transient which contain spikes with bursts of activities, also called non-stationary signal corrupted with noise. For this experimental analysis of EEG signal, which has main focus on event related potential so that it is not only necessary to know the frequency components of the signal but also the times at which this event occurs. As time analysis, does not provide frequency details and spectral analysis does not provide the time at which frequency changes occur, and for addressing these type of issues time-frequency analysis is needed for non-stationary signal. Wavelet transform is such a tool which provides both time and frequency of the signal and the ability to manipulate and compute large feature data set into compressed features. All these features of the signals characterize the behavior of the EEG signal and used for recognition and diagnostic purposes [6].

3.2 Wavelet transform

The Wavelet Transform technique is used to address the problem of non-stationary signals which represents a time function in the form of simple, fixed building blocks, called wavelets and these wavelets are actually a family of functions which are derived from the mother wavelet by using two operations translation and dilation. Dilation also called as

ISSN: 2278 – 1323

scaling of the signal which compresses or stretches the mother wavelet and another one translation which shifts signal along the time axis. The Wavelet transform can further categorized into continuous and discrete transform where continuous wavelet transforms (CWT) is represents as given below:

$$CWT(a,b)=\int_{-\infty}^{+\infty} x(t)\,\psi_{a,b}^{*}(t)dt \qquad (1)$$

here *x(t)* represents the signal to be analyzed, *a* represents scaling factor and *b* represent the translation factor also called shifting coefficient respectively, and the ψ^{*} superscript asterisk used to denotes the complex conjugation. $\psi_{a,b}(t)$ is obtained by scaling the wavelet at time *b* and scale *a* where ψ(t) represents the wavelet [8].

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi((t-b)/a) \qquad (2)$$

3.3 Discrete Wavelet Packet Transform

Discrete wavelet packet transform (DWPT) is extended from the wavelet transform (WT). In discrete wavelet transform signals repeatedly split, filter from the low pass filter bands which result the wavelet system in a logarithmic frequency resolution, where low frequencies represents narrow bandwidths and the high frequencies have wide bandwidths for that sub bands. The discrete wavelet packet transform, not only split the signal from the low pass filter where output represents the detail coefficients but also high pass filter bands which gives approximation coefficients. The computational complexity of the DWPT is O(N(log(N))where discrete wavelet transform gives O(N) [3].

The wavelet decomposition splits the original signal into two parts high and low frequency subspaces, V and W, with V being the subspace that includes the low frequency information and W includes the high frequency information of the signal. Fig. 2 shows the decomposition of the low frequency subspace where V was repeated. Wavelet transform only partitions the low frequency contents finely, and DWPT is a generalized version of WT, which also decomposes the high frequency bands and leads to a complete wavelet packet tree, which is shown in Figure 3, where $U_{n,j}$ is the nth subspace where n is the frequency factor, n = 0, 1, 2, . . .,2j -1 of wavelet packet at the jth scale, and $U_{j,k}^{n}(t)$ its corresponding orthonormal basis [24],

$U_{j,k}^{n}(t)=2^{-j/2}u^{n}(2^{-j}$t-k)

where k is the shift factor , it satisfies with (3) and (4)

$$u_{j,0}^{n}(t)=h_0(k)u_{j-1,k}^{i} \qquad \text{n is even} \qquad (3)$$

$$u_{j,0}^{n}(t)=\sum_k h1(k)u_{j-1,k}^{i} \qquad \text{n is odd} \qquad (4)$$

Where j, k ∈ Z, n = 0,1,2,. . .,2^j - 1, $h_0(k)$, $h_1(k)$ is a couple of quadruple mirror filters (QMF) which is scales and satisfies with equation (3)

$$h_0(k)=(-1)^{1-k}h_0(1-k) \qquad (5)$$

When the scale is sampling sequence of *f*(t), (*f*(k ∧ t)) is directly used as the coefficient of $U_0^0 d_0^0(k)$ in approximation.

Equations (4) and (5) shows the coefficient of DWPT at j^{th} and k^{th} level by the quadruple wavelet packet transformation

$$d_j^n(k)=\sum_m h_0(m-2k)d_{j-1}^{\frac{n}{2}}(m)\text{n is even} \qquad (6)$$

$$d_j^n(k)=\sum_m h_1(m-2k)d_{j-1}^{(\frac{i-1}{2})}(m) \qquad \text{n is odd} \qquad (7)$$

This decomposition coefficient of j^{th} level can be obtained by the $(j-1)^{th}$ level. After it is decomposed by j levels, the frequency ranges of all subspaces at the j^{th} level U_j^n are

$$\left\{\left[0,\frac{fs}{2^{j+1}}\right];\left[\frac{fs}{2^{j+1}},\frac{2fs}{2^{j+1}}\right];\left[\frac{fs}{2^{j+1}}\frac{3fs}{2^{j+1}}\right];\ldots\left[\frac{(2^j-1)fs}{2^{j+1}}\frac{fs}{2}\right]\right\}$$

Where, *f*s is the sampling frequency [10].

4. FORMATION AND SELECTION OF FEATURES

For this experiment a discrete wavelet packet transform (DWPT) with of depth 7 levels was used. On each 30 seconds segments of EEG signal daubechies of order2 (db2) wavelet transform was applied and from each segments the following features frequencies ranges are selected from each sub-bands and co-efficient of sub bands shown in fig 4 [11].

<table>
<tr><td colspan="8">$U_0^0(V_0)$</td></tr>
<tr><td colspan="4">$U_1^0(V_1)$</td><td colspan="4">$U_1^1(W_1)$</td></tr>
<tr><td colspan="2">$U_2^0(V_1)$</td><td colspan="2">$U_2^1(W_2)$</td><td colspan="2">U_2^2</td><td colspan="2">U_2^3</td></tr>
<tr><td>$U_3^0(V_3)$</td><td>$U_3^1(W_3)$</td><td>U_3^2</td><td>U_2^3</td><td>U_3^4</td><td>U_3^5</td><td>U_3^6</td><td>U_3^7</td></tr>
<tr><td colspan="8">..........</td></tr>
</table>

Figure 3: Tree Structure of Discrete Wavelet Packet Transform

1. {0.39 - 3.13 Hz}, Delta,
Wavelet coefficient = [C38, C30, C31, C32] = S1
2. {3.13 - 8.46 Hz}, Theta,
Wavelet coefficient = [C33, C34, C22, C23, C35] = S2
3. {8.46 - 10.93 Hz}, Alpha,
Wavelet coefficient = [C36, C25] = S3
4. {10.93 - 15.63 Hz}, Spindle,
Wavelet coefficient = [C26, C27, C28] = S4
5. {15.63 - 21.88 Hz}, Beta1,
Wavelet coefficient = [C16, C17] = S5
6. {21.88 - 37.50 Hz}, Beta2,
Wavelet coefficient = [C18, C5] = S6

Features which are used for classification listed below which represents the time–frequency analysis of the EEG signals [3, 12].

1. Mean Energy (E1, E2… E6) of wavelet packet coefficients for each of the 6 bands,
2. Total Energy (*E7*), which is the sum of energy in the above mentioned 6 frequency bands
3. Ratio of different Energy values (E8, E9, E10).

ISSN: 2278 – 1323

Where E8 is the ratio of energy in the Alpha band to the combined power in delta and theta bands, E9 is the ratio of energy in the delta band to the combined power in alpha and theta bands and E10 is the ratio of energy in theta band to the combined power in delta and alpha bands.

4. Mean of the absolute values of the coefficients in each 6 sub-band, and

5. Standard deviation of the coefficients in each 6 sub-band.

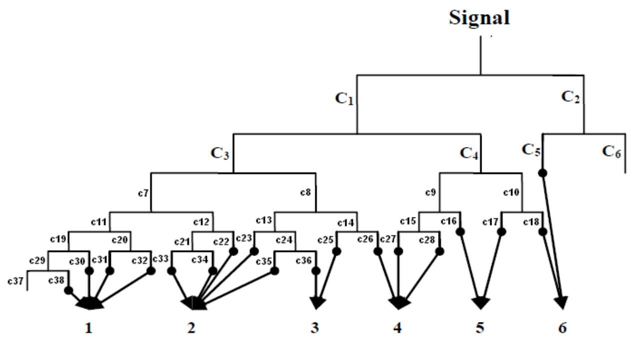

Figure 4: Wavelet packet Transform tree and selected coefficient of each sub band

5. CLASSIFICATION USING BASE CLASSIFIER

5.1 Artificial Neural Networks

There are various machine learning approaches available but neural networks have been recognized as one of the most efficient classifiers for classification of human sleep stage problems because of their learning capability and robustness. An ANN based on network of artificial neurons with an input layer, some number of hidden layers connected with each other via weights and an output layer where the network tries to find best output values for any input given to this structure. There exist connections between these layers and it is the values of these connection weights, which is learned with the training process [12]. The number of neurons in the input layer is defined by the number of input features which are extracted from the segments, here in our experiment 12 inputs with hyperbolic tangent transfer function for the neurons, one hidden layer with 8 neurons with logarithmic sigmoid transfer function and one output layer with 4 neurons for discrimination between sleep stages Awake, Stage1+REM, Stage2 and SWS was used. The neural network was trained using feed-forward back-propagation gradient algorithm with momentum and adaptive learning rate (traingdx) [3]. We also focus on two parameters on which training process depends; these are learning rate (lr) and momentum constant (mc), where learning rate (lr) used to determine the change in the weight along with mc but this mc values changes according to gradient of errors occurred in the weights. Learning rate affects execution of the networks, if this rate is low then system results in a slow training process, so the learning rate was kept high to reduce the training time and to keep learning state stable [13]. The momentum constant helps our network not to get trapped in local minima and not to reach global minima while training phase. The network was trained 6 times and every time 6 different random initialization feature data sets used and the network which give best results for this set was kept. In the classification stage, the output will give “1” if the input match with the target output vector and the otherwise it gives ‘‘0’’ as the corresponding input patterns epoch does not belongs to classes given [14].

5.2 k-Nearest Neighbor Classifier

The KNN method is an instance based nonparametric pattern classification approach which has relatively robust performance. It is also called a lazy learning algorithm where the training data points are stored only and all the training data are used during the testing phase. The KNN algorithm divided into two phases: Training and Classification phase. In training phase, the training data which consists of a set of vectors along with their class label are stored in multidimensional feature vector space. In our training phase, our training data are the segments whose features are extracted by using DWPT method and these features vectors are stored in our current folder which are not used to do any generalization and these data are used only during the testing phase. In the classification phase, KNN compares the input feature vector with stored vectors by calculating the Euclidean distance metric (EU) which is used for computing a distance between two points and then the test feature vector get assigned by a class label which is the nearest class of stored feature vector hence the name “k-nearest neighbor” and K is a user-defined constant. Training phase of k-NN algorithm is very fast but testing phase consume some time. The main idea of KNN method is to assign new test feature vector, a class to which the majority of its K nearest neighbors belongs [15, 16].

5.3 Multi-Class SVM Classifier

Support Vector Machines (SVM) is binary classification algorithms that analyzes data and recognizes patterns which are used in classification. SVM takes a set of training samples where each sample are marked as belonging to one of two class, and uses a training algorithm to create a model that assigns new samples into one of the two group. When the data are linearly separable, hyper plane is computed which maximizes the margin between the training examples and the class boundary and kernel function is used for mapping process when the data are not linearly separable. In this experiment for classifying four classes the network is train with Multi-class SVM classifier (M-SVM) [18]. For multi-class scenario the problem is decomposed into an M-class problem into a series of two-class problems, and it uses one-against-all method, where it consider the underline class as

ISSN: 2278 – 1323

one group and the remaining class as another group by constructing M binary classifiers to distinguish each class from the other class. While classifying new test data, the test data gets assigned to the class whose decision function has largest value [17, 19].

6 CLASSIFICATION USING ENSEMBLE CLASSIFIER

Ensemble learning is a machine learning paradigm where multiple learners are trained to solve the same problem [20]. It contains a number of learners which are called base learners generated from training data by a base learning algorithm, here base learner are Artificial neural network, k-nearest neighbor, m-SVM to produce heterogeneous learners. An ensemble of classifier is a set of multiple classifiers, so that individual decisions are combined to classify new test samples. The main idea is to combine a set of classifier each of which solve the same tasks in order to obtain robustness, accurate and reliable decisions that cannot be obtain by single classifier. In this experiment averaging method is used where the principle used is to average their predictions for classification and this combined model gives better result than any of the single model because; its variance is reduced [21]. An ensemble is constructed by first forming a number of base learners, which can be generated in a parallel style where all the individual classifiers are worked separately on the same training data set and then their results are combined with an aggregation rule such as majority voting for classification. Here bagging ensemble method is used which is a well-known method, that processes samples concurrently and it is based on the concept of bootstrapping, where each classifier is trained on a sample of instances taken with replacement from the training set and aggregating method. To classify a new instance, each classifier returns the class prediction for the unknown instance and bagged classifier returns the class that has been predicted most often by voting method [22].

7 PERFORMANCE EVALUATIONS

The accuracy and other performance measures of classification of the different classes differ slightly but we can say they are almost of the same except for stage REM + stage1. This happens due to the similarity between these two stages. Further, the parameters which are used to calculate the performance of the proposed method are the Sensitivity (SE), Specificity (SP), and Accuracy (AC) of the classification is given below as:

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})}$$

$$\text{Accuracy} = \frac{(\text{TN} + \text{TP})}{(\text{TN} + \text{TP} + \text{FP} + \text{FN})}$$

Where (TP) is the number of True Positive where the algorithm detects the correct sleep stage, (TN) is the true negative, (FN) is number of false negative, where algorithm fails to identify sleep stage, and false positive (FP) if the algorithm detects non-sleep stage as sleep stage [23].

III RESULTS AND DISCUSSION

The main aim of automatic sleep stage classification is to discriminate between Awake, Stage1+REM, Stage2 and SWS stages by using a single channel EEG signals. The feature extraction is done by wavelet packet transform which applied on each 30-second segments of Pz-Oz channel EEG signals and features are calculated from wavelet packet coefficients. These feature vectors act as input for the classifiers. For ensemble of the classifier bagging method is used which uses bootstrap aggregating method. Here we are using majority voting method for predicting the final result.

For testing purpose Sc4012eO subject record was used which is about 24 hour and this is data taken from single channel Pz/Oz EEG channel. Description about subject information is shown in table 2 which contains 30 seconds segmented data, where each segment presents a stage and this analysis was performed manually by the sleep specialist. Our classifiers operated on this signal data given as input to the system. The result of each classifier is given below as:

From the table 3 and fig 5, it observed that ANN can discriminate between Awake with 96.9%, Stage2 with 87.6% and SWS with 95.8%, but not able to discriminate the REM+Stage1. The overall classification average of ANN is 85.6%. From table 4 and fig 6 it is observed that the classification of stages with KNN is 94.4% for Awake, 35.8% for REM+Stage1, 78.0% for stage2 and 92.7%for SWS stage with average classification accuracy of 85.0%. KNN can classify some segments of REM+Stage1 accurately. From table 5 and fig 7 it is inferred that M-SVM gives 93.2% sensitivity for Awake stage, 75.4% sensitivity for REM+Stage1 which is better than both ANN and KNN. 86.7% for stage 3 and 15.6% classification accuracy for SWS which is less percent classification accuracy as compared to classification accuracy given by ANN and KNN classifier for the stage REM+stage1.The ensemble method gives 94.8% for Awake, 75.4% for REM+Stage1, 75.9% for stage2 and 95.8% for stage SWS. The average accuracy of classification is 88.7% from table 6 and fig 8, which is better than all these three base classifier which is used for this experiment hence we can say that by using ensemble we get better accuracy for classification of each sleep stage.

ISSN: 2278 – 1323

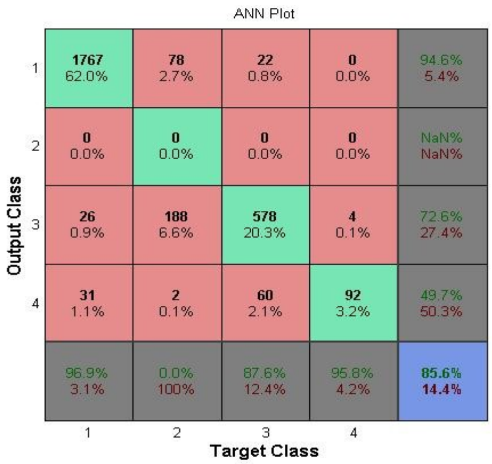

Figure 5: Confusion matrix for ANN Classifier

Table III: Artificial Neural Network classification result

Stages	Awake	N1+REM	N2	SWS	Avg. accuracy
Sensitivity	96.9%	0.0%	87.6%	95.8%	
Specificity	94.6%	0.0%	72.6%	49.7%	85.6%
Accuracy	62.0%	0.0%	20.3%	3.2%	

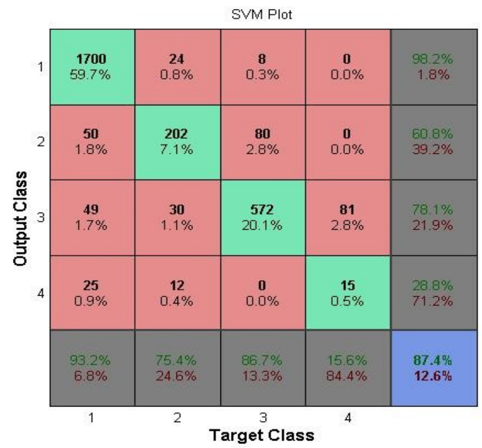

Figure 7: Confusion matrix for M-SVM classifier

Table IV: M-Support Vector Machine classification result

Stages	Awake	N1+REM	N2	SWS	Avg. accuracy
Sensitivity	93.2%	75.4%	86.7%	15.6%	
Specificity	98.2%	60.8%	78.1%	28.8%	87.4%
Accuracy	59.7%	7.2%	20.1%	0.5%	

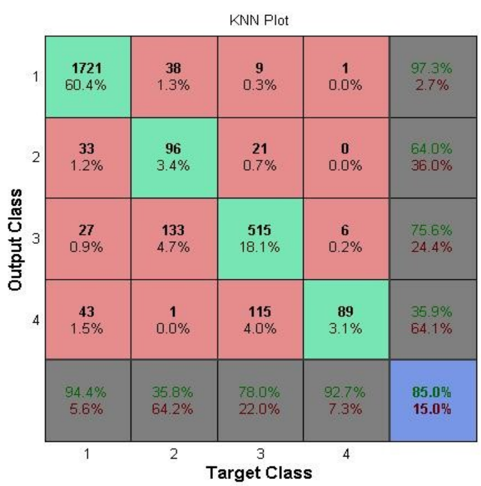

Figure 6: Confusion Matrix for KNN classifier

Table 4: Results given by k-Nearest Neighbor classifier

Stages	Awake	N1+REM	N2	SWS	Avg. accuracy
Sensitivity	94.4%	35.8%	78.0%	92.7%	
Specificity	97.3%	64.0%	75.6%	35.9%	85.0%
Accuracy	60.4%	3.4%	18.1%	3.1%	

Figure 8: Confusion matrix for Ensemble method

Table 6: Ensemble classification Result

Stages	Awake	N1+REM	N2	SWS	Avg. accuracy
Sensitivity	94.8%	75.4%	75.9%	95.8%	
Specificity	96.9%	60.8%	91.9%	49.7%	88.7%
Accuracy	60.7%	7.1%	17.6%	3.2%	

ISSN: 2278 – 1323

IV CONCLUSIONS AND FUTURE SCOPE

EEG signals are non-variant signal and for sleep stage classification we need to know not only time analysis but also frequency analysis of the signal, so for feature extraction purpose we are using Discrete Wavelet packet Transform method with 7 levels of decomposition and statistical feature are used to reduce the feature vector space. On these selected features base classifier such as ANN, KNN and SVM algorithms are applied. According to the result given in classification table 6, it could be inferred that Ensemble of classification generates highest classification result as compared to the base classifiers. The result of the base classifier are combined using ensemble method which uses majority voting scheme where the classes are predicted based on voting for final classification result. From table 6 we can say that our ensemble method gives 88.7% average accuracy which is better than all these three base classifier for the same data set. In this proposed method, we are combining the advantages of all the used base classifiers to overcome the disadvantages of individual classifiers.

In our method we used only single EEG channel Pz/Oz for experimentation, in future we can apply this method for multichannel signal, and again to improve classification accuracy it would be possible to use more than three base classifiers with ensemble method.

REFERENCES

[1] A. Rechtschaffen and A. Kales, "A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects", No. 204 in National Institutes of Health Publications, and Los Angeles: Brain Information service/Brain Research Institute U.S. Government Printing Office, Washington DC 1968.

[2] C. Iber, Ancoli-Israel S, Chesson A, and Quan SF, “The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Technical Specifications,” 1st ed.: Westchester, Illinois: *American Academy of Sleep Medicine*, 2007.

[3] Farideh Ebrahimi, Mohammad Mikaeili, Edson Estrada, Homer Nazeran, “Automatic Sleep Stage Classification Based on EEG Signals by Using Neural Networks and Wavelet Packet Coefficients” 30th Annual International IEEE EMBS Conference Vancouver, British Columbia, Canada, August 20-24, 2008.

[4] Aykut Erdamar, Fazıl Duman, Sinan Yetkin, “A wavelet and teager energy operator based method for automatic detection of K-Complex in sleep EEG”, Expert Systems with Applications 39 (2012) 1284–1290 .

[5] F. Ebrahimi, M. Mikaili, E. Estrada, H. Nazeran “Assessment of Itakura Distance as a Valuable Feature for Computer-aided Classification of Sleep Stages” Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite International, Lyon, France August 23-26, 2007.

[6] S. Sanei and J. Chambers, “EEG Signal Processing”, John Wiley & Sons Ltd, England, ISBN-13 978-0-470-02581-9, 2007.

[7] B. Kemp and J. Olivan, "European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data," Clin Neurophysiology*,* vol. 114, pp. 1755-61, Sep 2003.

[8] Dean Cvetkovic, Elif Derya Übeyli, Irena Cosic, “Wavelet transform feature extraction from human PPG, ECG, and EEG signal responses to ELF PEMF exposures: A pilot study”, Digital Signal Processing 18 (2008) 861–874.

[9] Maryam Vatankhah, Mohammad-R. Akbarzadeh-T, and Ali Moghimi”An Intelligent System for Diagnosing Sleep Stages Using Wavelet Coefficients” EEE (February 7, 2010) 978-1-4244-8126-2.

[10] Wu Ting, Yan Guo-zheng, Yang Bang-hua, Sun Hong, “Measurement EEG feature extraction based on wavelet packet decomposition for brain computer interface”, 41 (2008) 618–625.

[11] Edgar Oropesa, Hans L. Cycon , Marc Jobert, ”Sleep Stage Classification using Wavelet Transform and Neural Network” International Computer Science Institute, TR-99-008 March 30, 1999.

[12] Lukas Zoubek, Sylvie Charbonnier, Suzanne Lesecq, Alain Buguet, Florian Chapotot, “Feature selection for sleep/wake stages classification using data driven methods”, science direct Biomedical Signal Processing and Control 2 (2007) 171–179.

[13] Seral Ozsen, “Classification of sleep stages using class-dependent sequential feature selection and artificial neural network”, Springer Neural Computing & Application (2013) 23:1239–1250.

[14] Yu-Liang Hsu a, Ya-Ting Yang b, Jeen-Shing Wanga, n, Chung-YaoHsu c “Automatic sleep stage recurrent neural classifier using energy features of EEG signals”, Science Direct, Neuro computing 104 (2013) 105–114.

[15] Lukas Zoubek, Sylvie Charbonnier, Suzanne Lesecq, Alain Buguet, Florian Chapotot, “Feature selection for sleep/wake stages classification using data driven methods”, science direct Biomedical Signal Processing and Control 2 (2007) 171–179.

[16] Bülent Yılmaz, Musa H Asyalı, Eren Arıkan, Sinan Yetkin, Fuat Özgen Yılmaz etal, “Sleep stage and obstructive apneaic epoch classification using single-lead ECG”, Bio Medical Engineering 2010.

[17] Suicheng Gu and Yuhong Guo,”Learning SVM Classifiers with Indefinite Kernels”, Department of Computer and Information Sciences, USA Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, PA 19122.

[18] Fabien Lauer, Yann Guermeur, “MSV Mpack: A Multi-Class Support Vector Machine Package”, Journal of Machine Learning Research 12 (2011) 2293-2296.

[19] Yi Liu and Yuan F. Zheng, “One-Against-All Multi-Class SVM Classification Using Reliability Measures”, Department of Electrical and Computer Engineering Columbus, 43210.

[20] “Ensemble Learning”, by Zhi-Hua Zhou.

[21] Lior Rokach, “Ensemble Methods for Classifiers“, Department of Industrial Engineering Tel-Aviv University.

[22] Kristína Machová, František Barčák, Peter Bednár, “A Bagging Method using Decision Trees in the Role of Base Classifiers”, Acta Polytechnica Hungarica Vol. 3, No. 2, 2006.

[23] Luay Fraiwana, Khaldon Lweesy, Natheer Khasawnehc, Heinrich Wenzd, Hartmut Dickhause, “Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier”, computer method and programs in biomedicines I08 (2013) 10– 17 November 2011.

[24] J.Z. Xue, H. Zhang, C.X. Zheng, Wavelet packet transform for feature extraction of EEG during mental tasks, Presented at Proceedings of the Second International Conference on Machine Learning and Cybernetics, Xi’an, 2003

ISSN: 2278 – 1323