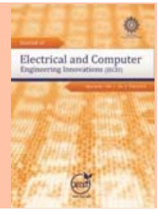
http://jecei.srttu.edu

SRTTU

Journal of Electrical and Computer Engineering Innovations

JECEI, Vol. 1, No. 2, 2013

Regular Paper



Automatic Sleep Stages Detection Based on EEG Signals Using Combination of Classifiers

R. Kianzad[1,*] and H. Montazery Kordy[1]

[1]Faculaty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

[*]Corresponding Author: kianzad@nit.ac.ir

ARTICLE INFO

ARTICLE HISTORY:

Received 15 September 2012
Revised 15 August2013
Accepted 22 August 2013

KEYWORDS:

Sleep stages classification
EEG signals
Wavelet packets
Classifier combination
Majority voting

ABSTRACT

Sleep stages classification is one of the most important methods for diagnosis in psychiatry and neurology. In this paper, a combination of three kinds of classifiers are proposed which classify the EEG signal into five sleep stages including Awake, N-REM (non-rapid eye movement) stage 1, N-REM stage 2, N-REM stage 3 and 4 (also called Slow Wave Sleep), and REM. Twenty-five all night recordings from Physionet database are used in this study. EEG signals were decomposed into the frequency sub-bands using wavelet packet tree (WPT) and a set of statistical features was extracted from the sub-bands to represent the distribution of wavelet coefficients. Then, these statistical features are used as the input to three different classifiers: (1) Logistic Linear classifier, (2) Gaussian classifier and (3) Radial Basis Function classifier. As the results show, each classifier has its own characteristics. It detects particular stages with high accuracy but, on the other hand, it has not enough success to detect the others. To overcome this problem, we tried the *majority vote* combination method to combine the outputs of these base classifiers to have a rather good success in detecting all sleep stages. The highest classification accuracy is obtained for Slow Wave Sleep as 81.68% in addition to the lowest classification accuracy of 43.68% for N-REM stage 1. The overall accuracy is 70%.

1. Introduction

Sleep is a physical and mental resting state of brain in which most of external stimulus stopped from the sense. In normal humans, about thirty percent of their life is spent for sleep. The disorders going on during sleep phase such as insomnia, narcolepsy, sleep walking and nocturnal breathing disorders naturally give inconvenience to the subject, and need to be treated [1]. Sleep is not an eventless process. On the contrary, many events occur in the body during this state: heartbeat slows down, blood pressure falls, muscles relax and the body's metabolic rate decreases.

Polysomnography (PSG) is based on study of sleep and wakefulness from the simultaneous recording of multiple bioelectric signals including the electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG). In practice, it is used to study/record in detail all the biophysiological changes that occur in the human body when the person is asleep. A system of standardized rules established in the conventional Rechtschaffen and Kales (R&K) human sleep/wake stage scoring manual [2] enables the visual recognition by sleep specialists of up to six different stages: Awake, non-rapid eye movement (N-REM) sleep stages 1,2,3 and 4, and REM. N-REM stages 3 and 4 represent the Slow Wave Sleep (*SWS*). The successive visual interpretation, by 30-second epochs, of 8-24h PSG recordings leads to the representation of the temporal distribution of sleep/wake stages called a hypnogram. PSG is thus a powerful tool in the diagnosis of sleep disorders, which are common with about 5% of the general population affected [3].

Generally, Awake stage is seen in the beginning of the sleep and it can be defined as a transition stage from the full alertness to the half-sleepy situation. Alpha rhythms that are the signals with frequencies between 8 and 13 Hz are seen in EEG signal in addition to the eye movements and high muscle tone during this stage. N-REM stage 1 (*S1*) is characterized with mixed frequency activity but 4-7 Hz frequencies are dominant. Also, slow eye movements in EOG and vertex sharp waves in EEG can be other indicators for that stage. In N-REM stage 2 (*S2*), on the other hand, sleep spindles and K-complexes are searched in the EEG activity. Among the sleep stages, Slow Wave Sleep can be detected with the observation of frontal region EEG signals with frequencies 0.5-2 Hz and amplitude values higher than 75 μV. Lastly, REM stage is characterized mainly with rapid eye movements [4]. Fig. 1 shows some typical 30-second epochs of the EEG signal in different sleep stages.

With growing innovative developments in machine learning and signal processing, several attempts to develop automated sleep stagers have been realized with increasing number of studies for about 20 years. In [5], Becq et al. by comparing five classifiers—linear and quadratic classifiers, k nearest neighbours, parzen kernels and neural networks—gave a sight in deciding which classifier is more efficient in sleep scoring. They saw that best classifier was observed as neural network with a true scoring ratio of 72 %. When the previous studies are examined, it will be seen that the performance of the automatic sleep stage classification applications that use all of the 5 sleep stages as sleep classes is generally in the 70–85 % interval.

2. Proposed Structure

In this paper, we propose a new sleep stage classification system that performs data acquisition, feature extraction, classification, and combination of classifiers in an automatic manner using a single EEG channel. First, the sleep EEG recordings are obtained from Sleep-EDF database. The nature of EEG signal is non-stationary and the discrete wavelet transform is useful for analysis of non-stationary signals. Also, wavelet transform has been previously used for sleep staging [6] and alertness level detection [7]. Here, we apply the wavelet packet to extract appropriate features from EEG signals. Then, three base classifiers are utilized to detect the sleep stages and benefits of each classifier are monitored. Finally, for increasing the total accuracy rates of sleep stages classification, we combine the outputs of these classifiers using the *majority vote* combination method.

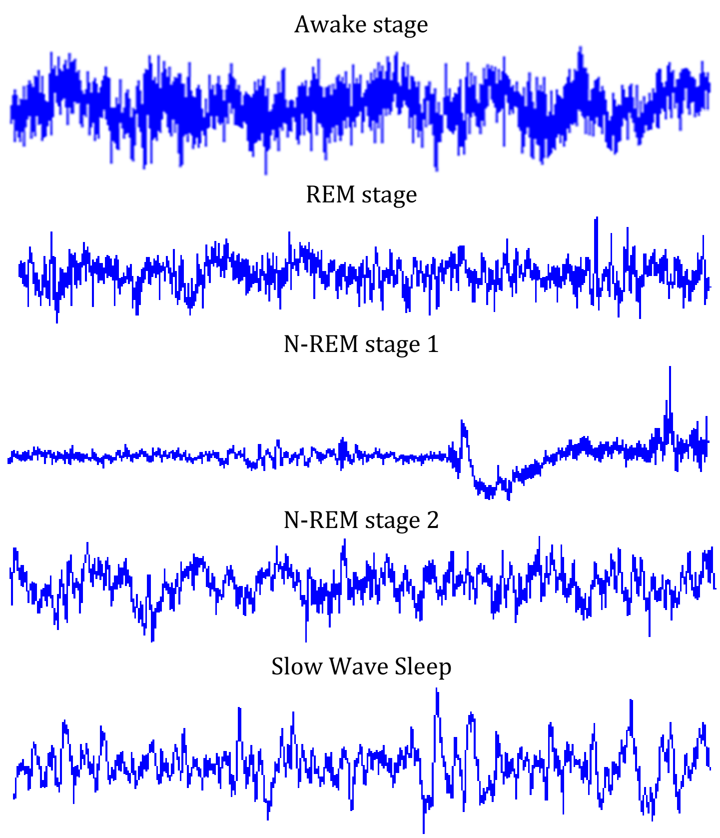


Figure 1: 30-second epochs of EEG signal in different sleep stages

3. MATERIALS AND METHODS

A. Sleep EEG Data

Data were available on the Physionet website for downloading on EDF (European Data Format). Subjects were randomly selected over a 6-month period from patients referred to the Sleep Disorders Clinic at St Vincent's University Hospital, Dublin, for possible diagnosis of obstructive sleep apnea, central sleep apnea or primary snoring [8]. In each epoch, a score has been attributed which is assigned from a set constituted of K=5 classes representing the five different stages in regards with the conventional criteria of R&K: 0-Awake, 1-REM stage, 2-NREM stage 1, 3-NREM stage2, 4and5-Slow Wave Sleep.

The total number of epochs included was 20,774. As it can be seen in Table 1, the number of epochs classified in each sleep/wake stage is different. *S2* lasts a long time, whereas *SWS* is rather short. To avoid classification errors related to differences in the sample size of each class, the database was randomly reduced to a smaller one where each class is composed of about the same number of epochs. The numbers of epochs classified in each sleep stage for the reduced database are presented in the second row of Table 1.

TABLE 1
NUMBER OF EPOCHS IN THE SLEEP STAGES

Stages / Data	Awake	REM	S1	S2	SWS
Full database	4707	3016	3403	6985	2663
Reduced Database	1318	1328	1362	1397	1332

B. Wavelet Based Feature Extraction

Discrete Wavelet Transform (DWT) analyzes the signal at different frequency bands, with different resolutions by decomposing the signal into a coarse approximation and detail information. The DWT employs two sets of functions called scaling functions and wavelet functions, which are associated with low-pass and high-pass filters, respectively. The decomposition of the signal into the different frequency bands is simply obtained by successive high-pass and low-pass filtering of the time domain signal.

Selection of suitable wavelet and the number of levels of decomposition is important in analysis of signals using the DWT. The typical way is to visually inspect the data first, and if the data are kind of discontinuous, Haar (db1) or other sharp wavelet functions are applied; otherwise a smoother wavelet can be employed. The number of levels is chosen based on the dominant frequency components of the signal. The levels are also selected such that those parts of the signal that correlate well with the frequencies required for classification of the signal are retained in the wavelet coefficients [9].

Since the EEG signal used in the research has the sampling rate of 128 Hz, a DWT of 8 levels was designed for this purpose. Daubechies order 2 (db2) wavelet transform was applied to 30-second epochs of EEG signal [10]. Using the flexibility of the *wavelet packet tree* (*WPT*), we chose a set of the frequency sub-bands to characterize the EEG events. The frequency ranges of the EEG signal were broken down into Delta (below 3.5 Hz), Theta (4-7 Hz), Alpha (8-13 Hz), and Beta (14-30 Hz) bands. In the sleep EEG, because of presence of sleep spindles (12-14 Hz), there is another frequency band, that is, spindle frequency band. The wavelet coefficients (nodes index) corresponding to each frequency band are shown in Table 2.

In order to further reduce the dimensionality of the extracted feature vector, statistics over the set of the wavelet coefficients was used. The following statistical features were used to represent the time-frequency distribution of the EEG signals [9]:

- Standard deviation of the coefficients in each frequency band.
- Ratio of the absolute mean values of adjacent frequency bands.
- Mean quadratic value or Energy of the wavelet coefficients in each frequency band.
- Total Energy of the 5 frequency bands.
- Mean of the absolute values of the coefficients in each frequency band.

TABLE 2
WAVELET NODES INDEX CORRESPONDING TO FREQ. BANDS

Freq. Bands	Wavelet Coefficients
Delta	128,129,130,256
Theta	66,67,131,132,137
Alpha	69,138
Spindle	70,71,72
Beta	9,36,37,38

C. Classifiers

After extraction of appropriate features from sleep EEG signals, three different base classifiers were trained for detecting each stage. We used half of the reduced database, as shown in Table 1, for training the classifiers and the rest for testing them. Each classifier calculates the frontiers of each class in a different way:

I) Logistic Linear Classifier

This classifier computes the linear classification for the input features by *maximizing the likelihood criterion* ($\mathcal{L}$) using the logistic (*sigmoid*) function. A likelihood function (often simply the likelihood) is a function of the parameters of a statistical model. The likelihood of a set of parameter values, θ, given outcomes *x,* is equal to the probability of those observed outcomes given those parameter values, that is

$$\mathcal{L}(\theta|x) = P(x|\theta) \quad (1)$$

Maximum likelihood estimation, finds the parameters that maximize this likelihood function. To use the method of maximum likelihood, we first specify the joint density function for *n* random samples. By assuming *statistical independence* between the different samples, we have [11]:

$$P(x_1, x_2, \dots, x_n|\theta) = P(x_1|\theta) * P(x_2|\theta) * \dots * P(x_n|\theta) \quad (2)$$

by considering Eq. (2)we conclude

$$\mathcal{L}(\theta|x_1, \dots, x_n) = P(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} P(x_i|\theta) \quad (3)$$

The maximum likelihood (*ML*) method estimates θ so that the likelihood function takes its maximum value, i.e.

$$\widehat{\theta_{ML}} = \arg\max_{\theta} \prod_{i=1}^{n} P(x_i|\theta) \quad (4)$$

A necessary condition that $\widehat{\theta_{ML}}$must satisfy in order to be a maximum is the gradient of the likelihood function with respect to θ to be zero, i.e. [11]

$$\frac{\partial \prod_{i=1}^{n} P(x_i|\theta)}{\partial \theta} = 0 \quad (5)$$

Because of the monotonicity of the logarithmic function, it is more convenient to work with the logarithm of the likelihood function, called the **log-**likelihood:

$$L(\theta) \equiv \ln \prod_{i=1}^{n} P(x_i|\theta) \quad (6)$$

Thus, for the first classification, we applied Logistic (sigmoid) linear classifier and, in Table 3, the results have been shown as a confusion matrix, i.e. we determined exactly how many epochs have been correctly detected at each sleep stage and the others that have been misclassified:

TABLE 3
CLASSIFICATION RESULTS BY LOGISTIC LINEAR CLASSIFIER

TRUE LABELS	ESTIMATED LABELS					SUCCESS RATE (%)
	AWAKE	REM	S1	S2	SWS	
AWAKE	1013	63	193	39	10	76.86
REM	52	1080	59	122	15	81.32
S1	273	273	525	275	16	38.55
S2	52	132	105	957	151	68.50
SWS	5	7	11	112	1197	89.86

II) Gaussian Classifier

As mentioned before, we need the class likelihood to make a decision in classification. Now, if we assume that the input data (i.e. feature vector), is Gaussian distributed ($\mathcal{N}$), the probability of input data *x* in Eq. (3), becomes:

$$P(x|\omega_i) = \mathcal{N}(x|\mu_i, \sigma_i) \quad (7)$$

$$\mu = \frac{1}{M} \sum x_i \quad (8)$$

$$\sigma = \frac{1}{M-1} (x - \mu)^T . (x - \mu) \quad (9)$$

where$M$ is the number of samples of input data; ω_i the ith class, μ the mean and σ the variance of that class. In other words, we fit a Gaussian model to each class.

After performing parameter estimation for mean and variance of each class and forming the likelihood function, we constituted the Gaussian classifier. Table 4, shows the obtained classification results.

III) Radial Basis Function Classifier

A Radial Basis Function (RBF) network is a feed-forward artificial neural network classifier with one hidden layer that uses *radial basis function* as activation functions. In pattern classification applications, the Gaussian function is preferred [12]. Mixtures of Gaussians have been considered in various scientific fields. The Gaussian activation function for RBF networks is given by:

$$\phi_j(X) = exp\left[-\left(X - \mu_j\right)^T \Sigma_j^{-1}\left(X - \mu_j\right)\right] \quad (10)$$

for $j=1,\dots, L$, where X is the input feature vector, L is the number of RBF units in hidden layer, μ_j and Σ_j are the mean and the covariance matrix of the jth Gaussian function, respectively.

In this work, our RBF classifier has three layers: an input layer (the input vector X is used as input to all radial basis functions), a hidden layer with five RBFs (i.e. L=5) and a linear output layer. After forming the classifier, we trained the network with Mathworks' Neural Network toolbox and the results of classification, shown in Table-5, have been derived.

TABLE 4
CLASSIFICATION RESULTS BY GAUSSIAN CLASSIFIER

TRUE LABELS	ESTIMATED LABELS					SUCCESS RATE (%)
	AWAKE	REM	S1	S2	SWS	
AWAKE	1013	63	193	39	10	76.86
REM	52	1080	59	122	15	81.32
S1	273	273	525	275	16	38.55
S2	52	132	105	957	151	68.50
SWS	5	7	11	112	1197	89.86

TABLE 5
CLASSIFICATION RESULTS BY RBF CLASSIFIER

TRUE LABELS	ESTIMATED LABELS					SUCCESS RATE (%)
	AWAKE	REM	S1	S2	SWS	
AWAKE	24	1	0	1293	0	1.82
REM	0	252	0	1076	0	18.97
S1	0	29	0	1333	0	0
S2	0	10	0	1387	0	99.28
SWS	0	0	0	1332	0	0

D. *Combination of Classifiers*

As we can see in the last three Tables, each classifier has its own positive and negative points, i.e. none of them is able to detect all sleep stages with a quite good success rate. For example, as shown in Table 3, Logistic classifier has not strength to detect *S1* with more success rate of 38.55%. On the other hand, Gaussian classifier could detect *S1* with a relatively better diagnosis rate (61.89%), but *S2* has not been well detected by this classifier; since the confusion matrix shows the 35.22% correct detection (Table 4). Remarkably in Table 5, the results show that RBF classifier has detected *S2* with a great discrimination rate of 99.28%, but other stages have not been well detected at all. To have a relatively good detection of all sleep stages, we decided to combine these base classifiers.

A combination or an ensemble of classifiers is a set of classifiers whose individual outputs are combined to classify new samples. A combination of classifiers is often much more accurate than the individual classifiers that make them up. One reason for this could be that the training data may not provide sufficient information for choosing a single best classifier and a combination is the best compromise. Another reason could be that the learning algorithms used may not be able to solve the difficult search problem posed [13].

Combining classifiers is an established research area shared between statistical pattern recognition and machine learning. It is variously known as committees of learners, mixtures of experts, multiple classifier systems, consensus theory, etc. There are many different combination methods available, such as: *Majority vote (MAJ), Maximum (MAX), Minimum (MIN), Average (AVR), Product (PRO)*. Once the classifiers in the ensemble are trained, these combination methods do not require any further training [14]. Among all the combination methods, majority vote is by far the simplest for implementation. It does not assume prior knowledge of the behavior of the base classifiers and it does not require training on large quantities of representative classification results from the base classifiers [15].

E. *Majority Voting*

In majority voting, each classifier in the ensemble is asked to predict the class label of the sleep stage considered. Once all the classifiers have been queried, the class that receives the greatest number of votes is returned as the final decision of the ensemble.

The PRTools [16] for Matlab has been used for the implementation of the combination approach. According to this toolbox, If $V = [V1, V2, V3]$ is a stacked set of the base classifiers (i.e. all classifiers have the same feature vector as input) that are trained for the same classes, then $W = V * votec$, is the voting combiner, i.e. it selects the class with the highest vote of the base classifiers. The direct classifier outputs *D*, are posterior probability estimates:

$$D(i,j) = (v+1)/(n+c) \tag{11}$$

Where *v* is the number of votes which epoch *i* receives for the *j*th class, *n* is the total number of classifiers (i.e. *n*=3), and *c* the total number of classes (i.e. *c*=5).

4. RESULTS

The features were extracted from 30-second epochs of sleep EEG signal. The number of epochs is shown in Table 1. These features were then classified by three base classifiers, individually. The results are shown in Table 3, 4 and 5 respectively. Since, N-REM stage 1 was not well detected by Logistic classifier; we combined the outputs of this classifier with Gaussian classifier which detected this stage with much accuracy. The results of combination are shown in Table 6.

TABLE 6
COMBINATION OF LOGISTIC AND GAUSSIAN CLASSIFIERS

TRUE LABELS	ESTIMATED LABELS					SUCCESS RATE (%)
	AWAKE	REM	S1	S2	SWS	
AWAKE	1062	62	181	10	3	80.58
REM	71	1146	62	38	11	86.29
S1	296	375	602	84	5	**44.20**
S2	70	313	385	527	102	37.72
SWS	31	27	68	86	1120	84.10

As we were expected, the success rate of detecting *S1* obtained by combination of Logistic and Gaussian classifier, increased from 38.55% to 44.20%, but the accuracy of detecting *S2* significantly decreased, because of the weakness of Gaussian classifier in scoring this stage.

To overcome this problem, we could combine the outputs of Logistic and Gaussian classifiers with RBF classifier which is the best classifier in detecting *S2*. Table 7, shows the results of combining these three classifiers. This Table indicates that combination of RBF classifier outputs with the two other classifiers, improved the detection rate of *S2* by about 35% (from 37.72% to 72.73%). Although detecting accuracy of *Awake* and *REM* stages partly decreased (-9.56% and -6.47% respectively), but they are still acceptable.

TABLE 7
COMBINATION OF LOGISTIC, GAUSSIAN AND RBF CLASSIFIERS

TRUE LABELS	ESTIMATED LABELS					SUCCESS RATE (%)
	AWAKE	REM	S1	S2	SWS	
AWAKE	936	56	270	51	5	71.02
REM	44	1060	77	136	11	79.82
S1	223	263	595	275	6	43.68
S2	38	145	122	1016	76	**72.73**
SWS	9	6	11	218	1088	81.68

5. CONCLUSIONS AND FUTURE WORK

In this study, we have compared the performance of three base classifiers and combined the outputs of them to automatically score 30-second (3840-sample) epochs of sleep EEG signal from 25 individuals. These signals were divided into the five R&K sleep-wake stages: Awake, N-REM (non-rapid eye movement) stage 1 (*S1*), N-REM stage 2 (*S2*), Slow Wave Sleep (*SWS*), and REM. We utilized the wavelet packet to extract wavelet based features by breaking down the signal into the five most relevant EEG frequency bands. Then, some statistics was used in order to reduce the dimensionality of the extracted feature vector. During the phase of classification, wefirst applied Logistic Linear classifier. The weakness of this classifier, is detecting *S1* with rather low success rate. Then, we deployed Gaussian classifier which, contrary to the Logistic classifier, could detect *S1* with high accuracy but was unable to diagnose *S2* well. Therefore, the RBF classifier was applied to the dataset and it could detect *S2* perfectly, but it had a very poor performance in detecting the other stages. Thus, by considering all these facts, we decided to combine the outputs of these base classifiers to have a rather good success rate in detecting all sleep stages. We used the *majority voting* as the combination approach. After combination, the results showed that we have almost reached to our desirable accuracies. The highest classification accuracy was obtained for Slow Wave Sleep as 81.68% in addition to the lowest classification accuracy of 43.68% for N-REM stage 1. The overall accuracy was obtained 70% for all classes. Future work will be focus on refining the wavelet based data vector and adding new set of features, if needed, to have improvement in discrimination of N-REM stage 1.

6. REFERENCES

[1] M. E. Tagluk, N. Sezgin, and M. Akin, “Estimation of sleep stages by an artificial neural network employing EEG, EMG and EOG,” J Med Syst, vol. 34, pp. 717–725, 2010.

[2] A. Rechtschaffen and A. Kales, “A manual of standardized terminology, techniques and scoring systems for sleep Stages of human subjects,” US Government Printing Office, Washington, 1969.

[3] M. Reite, D. Buysse, C. Reynolds, and W. Mendelson, “The use of polysomnography in the evaluation of insomnia,” *Sleep*, vol. 18, pp. 58–70, 1995.

[4] “American Academy of Sleep Medicine Task Force. Sleep related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research,” *Sleep*, vol. 22(5), pp. 667–689, 1999.

[5] G. Becq, S. Charbonnier, F. Chapotot, A. Buguet, L. Bourdon, and P. Baconnier., “Comparison between five classifiers for automatic scoring of human sleep recordings,” *Studies in Computational Intelligence (SCI)*, vol. 4, pp. 113–127, 2005.

[6] E. Oropesa, H. L. Cycon, and M. Jobert, “Sleep stage classification using wavelet transform and neural network,” *International Computer Science Institute (ICSI)*, 1999.

[7] M. Kiymik, M. Akin, and A. Subasi, “Automatic recognition of alertness level by using wavelet transform and artificial neural network,” *J. Neuroscience Methods*, vol. 139, pp. 231–240, 2004.

[8] http://www.physionet.org/pn3/ucddb/

[9] A. Subasi, “Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients,” *Expert Systems with Applications*, vol. 28, pp. 701–711, 2005.

[10] C. Burros, R. Goliath, and H. Guo, “Introduction to wavelets and wavelet transforms,” Prentice Hall Pub, 1998.

[11] S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4rd ed., Elsevier, 2009, pp. 34–36.

[12] T. Poggio and F. Girosi, “Networks for approximation and learning,” *Proc. IEEE*, vol. 78, no. 9, pp. 1481–1497, 1990.

[13] M. N. Murty and V. S. Devi, Pattern Recognition, vol. 0, Springer: London, 2011, pp.188–206.

[14] C. A. Shipp and L. I. Kuncheva, “Relationships between combination methods and measures of diversity in combining classifiers,” *Information Fusion*, vol. 3, pp. 135–148, 2002.

[15] L. Lam and C. Y. Suen, “Application of majority voting to pattern recognition: an analysis of its behavior and performance,” *IEEE Transaction on systems, man, and cybernetics-Part A: systems and humans*, vol. 27, no. 5, September 1997.

[16] P. Robert, W. Duin, and D. Tax, PRTools: *A Matlab toolbox for pattern recognition*, 2012, software available at http://prtools.org/software/.

BIOGRAPHIES

Reza Kianzad was born in Tehran, Iran, in 1988. He received his B.S. degree in 2010 from Science and Research branch of Islamic Azad University, Iran, Tehran and M.Sc. degree in Biomedical Engineering from Babol Noshirvani University of Technology, Iran, Babol in 2013. His interests are bio-signal processing and classification methods.
Email:kianzad@nit.ac.ir

Hussain Montazery Kordy received the B.S. degree in electronic engineering from Mazandaran University, Babol, in 2000 and M.S. degree in biomedical engineering from Sharif University of Technology, in 2003 and the Ph.D. degree from Tarbiat Modarres University, Tehran, Iran, in 2009. Since 2010, he is a member of Electrical and Computer Engineering with Babol Noshirvani University of Technology (NIT), Babol, Iran, where he is currently an Assistant Professor of Biomedical Engineering. His teaching interests involve the medical instrumentation, biological system modeling, pattern recognition, and time-frequency signal processing. Also, his research focuses on computer aided diagnosis, feature selection and extraction, biomedical signal processing, and wavelet based signal analysis.