Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 23 (2013) 84 – 95

Procedia Computer Science

4th International Conference on Computational Systems-Biology and Bioinformatics, CSBio2013

Epileptic seizure identification from electroencephalography signal using DE-RBFNs ensemble

Satchidanada Dehuri[a]*, Alok Kumar Jagadev[b], and Sung-Bae Cho[c]

[a]Department of Systems Engineering, Ajou University, San 5, Woncheon-dong, Suwon 443-749, South Korea
[b]Department of Computer Science and Engineering, SOA University, Bhubaneswar, India
[c]Soft computing Laboratory, Department of Computer Science, Yonsei University, Seoul 120-749, South Korea.

Abstract

In this paper, an ensemble of radial basis function neural networks (RBFNs) optimized by differential evolution (DE) (DE-RBFNs) is presented for identification of epileptic seizure by analyzing the electroencephalography (EEG) signal. The ensemble is based on the bagging approach and the base learner is DE-RBFNs. The EEGs are decomposed with wavelet transform into different sub-bands and some statistical information is extracted from the wavelet coefficients to supply as the input to ensemble of DE-RBFNs. A benchmark publicly available dataset is used to evaluate the proposed method. The classification results confirm that the proposed ensemble of DE-RBFNs has greater potentiality to identify the epileptic disorders.

Selection and peer-review under responsibility of the Program Committee of CSBio2013

Keywords: EEG; classification; radial basis function neural networks; differential evolution; bagging

1. Introduction

The EEG signal is usually used for the purpose of recording the electrical activities of the brain signal that typically arises in the human brain[1]. An EEG signal is a measurement of currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. During the activation of brain cells, the synaptic currents are produced within the dendrites. As a result of which these currents generates a magnetic

* Corresponding author. Tel.: +82-31-219-3574.
E-mail address: satchi@ajou.ac.kr.

1877-0509
Selection and peer-review under responsibility of the Program Committee of CSBio2013
doi:10.1016/j.procs.2013.10.012

field measurable by electromyogram (EMG) machines and a secondary electrical field over the scalp measurable by EEG systems[2].

The recording of the electrical activity is basically done by placing electrodes on the scalp, which measures the voltage fluctuations in the brain[3]. The neurons in the brain are the source of electric charge, so they exchange ions with the extracellular milieu. Ions of same charge repel each other and in this way they are pushed out of the neurons when a number of ions are pushed out at the same time they push each other and form a way this is called as volume conduction[3]. When this wave reaches the electrode they push or pull the ions on the surface of the electrode which create potential difference and this voltage difference recorded over time gives EEG signals[4]. There are certain unwanted signals that are generated during the EEG signal recorded over the time that are called artifacts. These artifacts are later pre-processed as they create noise. Different types of feature reduction techniques are used to remove these artifacts[24,40].

Epilepsy is a critical neurological disease stemming from temporary abnormal discharges of the brain electrical activity, leading to uncontrollable movements and trembling. About 1% of the world population suffers from epilepsy[17]. Therefore, the diagnosis of epilepsy allows the choice of medicine or surgical treatment[5]. Since the EEG records show the brain electrical activities, they can provide valuable insight into disorders of the brain activity. In this context, the EEG recordings measured in seizure-free intervals from the epilepsy patients are considered as important components for the diagnosis or prediction process[6,7,14]. Although the occurrence of epileptic seizures seems unpredictable[9], more efforts are focused on the development of computational models for automatic detection of epileptic discharges, which then can be used to predict the onset of seizure[7].

Over decades neural networks (NNs) have been widely used in many biomedical signal analysis because they model the signal very efficiently and make a decision to classify the signal[6,10,13,15]. Therefore, they provide an important support for the medical diagnostic decision process. In a classification system with NN, first step is related to the feature extraction from the raw data with minimal loss of potential information by using different methods such as frequency domain features, time-frequency features, wavelet transform (WT), leading to the extracted feature vectors[7,8]. In the second step, some statistics over the vectors are used to reduce the dimensionality of these vectors. Final step is to apply the feature vectors as inputs to NNs[12]. Both the architecture of the NN and the training algorithm play key roles to obtain satisfactory results from NNs[14]. Over the years NN models with different architectures have been used such as multilayer perceptron neural network (MLPNN)[17,39], adaptive neuro-fuzzy inference system (ANFIS)[18], radial basis function neural network[15], and recurrent neural network (RNN)[16].

The trend of identifying a ‘single best’ model was popular long back ago, whether the model is based on machine learning[28,30] or statistics that is more accurate for a given medical decision application[22,27]. However, this reliance on a single model may be ill-advised. The studies of ensemble of predictors have realized that the accuracy of single best model can be improved by reducing the generalization error of a single model between 5~70%[9,20,21]. There are three major strategies can be adopted for forming ensembles of learners. The easiest one is the cross-validation neural networks ensemble, where all members of ensemble are learnt with the same training sample[35]. The second and third approaches rely on perturbed training set to ensemble learners. In other words, the base learners are learning from different variants of the original training data. In bagging, a unique training set from the original data is created by sampling with replacement over a uniform probability distribution to train each base learner[20,23]. The approach boosting is based on re-sampling, but unlike bagging, sampling over a probability distribution dependents on the misclassification rate for each sample[23]. Boosting is an iterative algorithm. As it progresses, the composition of the training instances becomes largely dominated by hard-to-classify samples[23]. On the other hand, stacking is combining base learners built by different learning algorithms. It has limited usage as because it is difficult to analyze theoretically. By considering all, one of the primaries concerned of this paper is to design a simple, unbiased, and efficient classifier for epileptic seizure detection to aid medical practitioner. Our method is based on bagging. The individual learner is evolved radial basis function neural networks using differential evolution (DE-RBFNs)[32]. Radial basis function networks (RBFNs)[26] have been studied in many disciplines including EEG signal

analysis[15]. Radial basis function networks have attracted the attention of many researchers because of its: (i) universal approximation[29], (ii) compact topology[31], and (iii) faster learning speed[28].

The rest of the paper is set out as follows. The discrete wavelet transforms and the method of extracting features is discussed in Section 2. Our approach of ensemble of classifiers is discussed in Section 3. Experimental design and results are analyzed in Sections 4. Conclusions are drawn at Section 5.

2. Discrete wavelet transform and extracted features

Discrete wavelet transform (DWT) signal is very much similar to sub-band coding and pyramidal coding or multi-resolution analysis[25]. The DWT uses multi-resolution filter banks and special wavelet filters for the analysis and reconstruction of signals[19]. It includes iteration of filters with rescaling. The transform involves successive low pass and high pass filtering of the discrete time-domain signal. The resolution of the signal (that is a measure of the amount of detail information in the signal) is determined by the filtering, and the scale is determined by up and down sampling[25]. This was called the Mallat algorithm or Mallat-tree decomposition. This algorithm includes decomposition of a signal into the approximations and details. It converts an input signal into one high-pass wavelet coefficient series and one low-pass wavelet coefficient series (of length reduced to half). The procedure of decomposition of a signal by DWT, where filters H[n] and L[n] correspond to high-pass and low-pass filters, respectively[10] can be illustrated through a dendrogram like tree structure. In level-1, the signal is fed simultaneously through H[n] and L[n] filters with the cut-off frequency being the ¼th of the sampling frequency. The outputs of L[n] and H[n] filters corresponds to detail (D1) and approximation (A1) coefficients of the level-1, respectively. The second level coefficients can be obtained by applying this multi-resolution procedure to first level. At each level of decomposition, through filtering and down sampling, the frequency resolution is doubled and the time resolution is halved. The frequency content of the original signal can be represented by the coefficients A1, D1, A2, and D2 within the band $0 - f_S/4$, $f_S/4 - f_S/2$, $0 - f_S/8$, and $f_S/8 - f_S/4$, respectively, where f_S is the sampling frequency of the original signal. The process of multi-resolution decomposition helps to capture and localize transient events in both time and frequency domain.

It is to be noted that the selection of the appropriate wavelet function and the number of decomposition levels are very important in the process of multi-resolution decomposition. One of the solutions of deciding the number of decomposition levels is the dominant frequency components of the signal. In this work, the number of decomposition levels is chosen 4, which is recommended by others' work[9,39] too. Generally, tests are performed with different types of wavelet function and the one which gives maximum efficiency is selected for the particular application. Further, Daubechies wavelet of order 4 (db4) is selected, because its smoothing feature can suitable for detecting changes of the EEGs, as proven in other work[9].

Table 1. Frequency bands with four-level DWT decomposition of EEG signals

Sub-Signal/Decomposition Level	Frequency Band (Hz)
D1/1	43.4–86.8
D2/2	21.7–43.4
D3/3	10.8–21.7
D4/4	5.4–10.8
A4/4	0–5.4

The frequency bands corresponding to 4-level DWT decomposition with selected coefficients of the EEG signal are illustrated in Table 1. An EEG signal is decomposed into four detail sub-signals D1–D4 and one final approximation sub-signal A4. The sub-signals D1, D2, D3, D4, and A4 corresponds to levels 1, 2, 3, 4, and 4, respectively. Further, each sub-signal corresponds to different frequency bands.

From these coefficients the following important features can be extracted: i) maximum of wavelet coefficient in each sub band, ii) minimum of wavelet coefficient in each sub band, iii) mean of wavelet coefficient in each sub band, and iv) standard deviation of wavelet coefficient in each sub band.

Figure 1 shows the discrete wavelet transform of a sample EEG signal using DB2 and decomposed up to level 4.

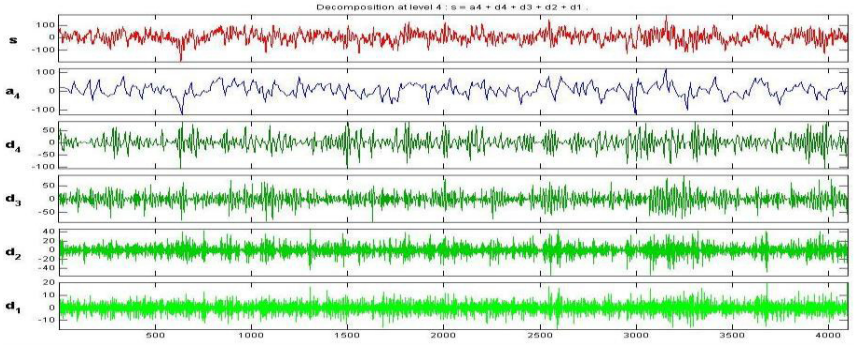

Fig. 1. Discrete wavelet transforms of a sample EEG signal using DB2 and decomposed up to level 4

3. Our method

The proposed method of identifying epileptic seizure is developed by combining the components like: i) EEG signal decomposition using DWT (c.f., Subsection 2), ii) Feature extraction (c.f., Subsection 2), iii) Bagging: Base learner is DE evolved RBFN.

As first two components have already been introduced in Subsection 2, therefore, the central focus of this section is on discussion of third component. However, it is also important to discuss, where the newness of this paper can fit in the general framework of the EEG signal analysis. Figure 2 discuss the general framework for EEG signal analysis, especially to identify the epileptic seizure.

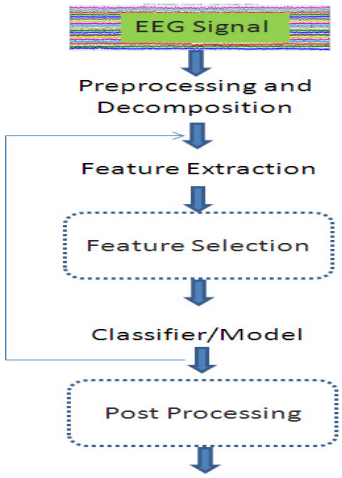

Fig. 2. Process of EEG Signal Analysis

In addition, the framework also introduces the feedback path for classification to feature extraction, the reason is that if the extracted features are not significantly contributing towards the accuracy of the model then either it needs addition or removal of some features. Let us discuss DE-RBFNs, the core part of the ensembler.

In the context of universal approximation, it has been proved that "a radial basis function networks can approximate arbitrarily well any multivariate continuous function on a compact domain if a sufficient number of radial basis function units are given". Note, however that the number of kernels (k) chosen, need not equal to the number of training patterns (n). In general it is better to have k much less than n i.e., $k << n$. Besides the gain in computational complexity, the reduction in the number of kernels is beneficial for the generalization capability of the resulting model

In RBFNs, other extensions are possible e.g., adapting centers, weighted norms, devising learning rules, and network with novel and different types of basis functions and multiple scales. A variety of learning procedure for RBFNs has been developed[32]. It is normally divided into two phases: 1) the adjustment of the connection weight vector; and 2) the modification of parameter of RBF units such as center and spreads[38]. In this paper, we have adopted our prior work[33], which was focus on two aspects such as devising a new kernel with a suitable modification of existing one and determination of hidden centers, along with spreads by differential evolution. The motivation of using differential evolution (DE)[33,34] over other evolutionary algorithms (EAs) such as genetic algorithms (GAs)[35] is that in DE string encoding are typically represented as real valued vectors, and the perturbation of solution vectors is based on the scaled difference of two randomly selected individuals of the current population. Unlike GA, the resulting step size and orientation during the perturbation process automatically adopt to the fitness function landscape.

3.1 Modified kernels

A RBF network with k hidden units and single node in output layer can be expressed as in equation (1):

$$\Phi(\vec{x}) = w_0 + \sum_{i=1}^{k} w_i . \exp\left(-\frac{\|\vec{x} - c_i\|^2}{\sigma_i^2} \right), \tag{1}$$

i.e., $\Phi(\vec{x}) : R^d \to R$, where w_i , w_0 , and c_i are the connection weights, bias connection weight, centers, respectively and x is the input vector to RBF network respectively. The Gaussian kernel is local in the sense that $\lim_{\|x\| \to \infty} f(\|x - c_i\|) = 0$, i.e., changing parameters of one neuron has a small effect for input values that are far away from the centers of that neuron (i.e., such type of case may be arise in the case of noise sample). The algorithm based on centers (i.e., mean vector) is very sensitive to outliers as because an object with an extremely large value may substantially distort the distribution of data. Therefore, to distinguish such sensitivity, we consider a data object which is nearest neighbor (nn) to mean vector (i.e., mean vector may not be a data object). Hence, the new modified kernel is:

$$\Phi(\vec{x}) = w_0 + \sum_{i=1}^{k} w_i . \exp\left(-\frac{\|\vec{x} - m_i\|^2}{\sigma_i^2} \right), \quad \text{where } m_i \approx nn(c_i) . \tag{2}$$

3.2 Learning procedure

As mentioned there are two phases within the learning procedure [40]. In phase one differential evolution is employed to reveal the centers and spread of the RBFNS. Although centers, spread and weights can be evolved using DE, but here we restrict ourselves with evolving only centers and spreads. This ensures efficient representation of an individual of DE. If we encode all the parameters such as centers, spread and weights into a individual chromosome, the chromosome length is too long and the search space becomes too large, which results in a very slow convergence rate. Since the performance of the RBFNs mainly depends on medoid and spread of the kernel, we just encode the medoids and spread into an individual chromosome for stochastic search.

Suppose the maximum numbers of kernels are set to $K_{\max}$, then the structure of the individual is represented as follows:

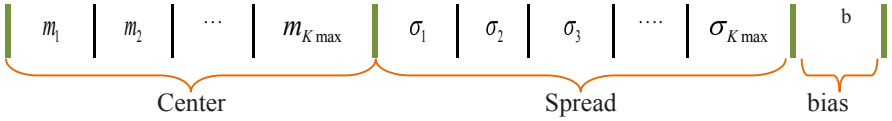

Fig. 3. Structure of the individual

In other words each individual has three constituent parts such as center, spread, and bias. The length of an individual is 2 $K_{\max}$ +1. The fitness function which is used to guide the search process is defined in equation (3).

$$f(x) = \frac{1}{N}\sum_{j=1}^{N}(t_j - \phi(\vec{x}_j)), \tag{3}$$

where N is the total number of training sample, t_j is the actual output of j^{th} sample and $\phi(\vec{x}_j)$ is the estimated output of DE-RBFNs. Once the centers and spreads are fixed, the task of determining weights in second phase of learning reduces to solving a simple linear system. In this work the pseudo inverse method is adopted to find out a set of optimal weights. Figure 4 illustrate the two phase learning procedure adopted in this work.

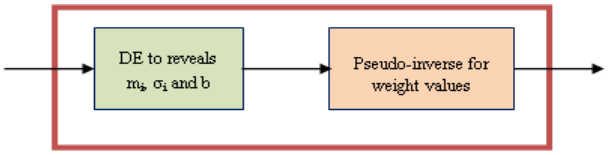

Fig. 4. Two phase learning scheme

The algorithmic framework of DE-RBFN is described as follows:

Initially, a set of n_p individuals (i.e., n_p is the size of the population) pertaining to networks medoids, spreads, and bias is initialized randomly, the individuals have the form:

$$x_i^{(t)} = \left\langle x_{i1}^{(t)}, x_{i2}^{(t)}, \ldots, x_{id}^{(t)} \right\rangle, \quad i = 1, 2, \ldots, n_p,$$

where $d = 2K_{\max} + 1$ and t is the iteration number.

In each iteration, e.g., iteration t, for individual $x_i^{(t)}$ undergoes mutation, crossover, and selection as follows:

Mutation: For vector $x_i^{(t)}$ a perturbed vector $V_i^{(t+1)}$ called donor vector is generated according to equation (4).

$$V_i^{(t+1)} = x_{r_1}^{(t)} + m_f(x_{r_2}^{(t)} - x_{r_3}^{(t)}), \tag{4}$$

where m_f is the mutation factor, lies in the interval (0,2], the indices r_1, r_2, and r_3 are selected randomly from $\{1,2,3,\ldots,n_p\}$, such that $r_1 \neq r_2 \neq r_3 \neq i$.

Crossover: The trial vector $u_i^{(t+1)} = \left\langle u_{i1}^{(t+1)}, u_{i2}^{(t+1)}, \ldots, u_{id}^{(t+1)} \right\rangle$ is generated as given in the equation (5).

$$u_{ij}^{(t+1)} = \begin{cases} v_{ij}^{(t+1)} \ if \ rand \le c_r \ or \ i = rand(1,2,\ldots,d) \\ x_{ij}^{(t)} \ if \ rand > c_r \ and \ i \neq rand(1,2,\ldots,d) \end{cases} \tag{5}$$

where j=1,..d, rand is random number generated in the range (0,1), c_r is the user specified crossover constant from the range (0,1) and rand (1,2,…,d) is a random integer from [1,2,…,d].

The random index is used to ensure that the trial vector differs by, at least one element from $x_i^{(t)}$. The resultant trial (child) vector replaces its parent if it has higher accuracy (a form of selection), otherwise the parent survives unchanged into the next iteration of the algorithm. Finally, we use selection operation and obtain the target vector $x_i^{(t+1)}$ as given in equation (6).

$$x_i^{(t+1)} = \begin{cases} u_i^{(t+1)} & if \ f\left(x_i^{(t+1)}\right) \le f\left(x_i^{(t)}\right) \\ x_i^{(t)}, & otherwise \end{cases} \quad (6)$$

j=1,2,…,d.

After getting medoids, spread, and bias from above pseudo-code now the weights of the network are computed by pseudo-inverse method as described in below equation.

Y=WΦ,

$$W = \left(\Phi^T \Phi\right)^{-1} \Phi^T Y$$

Pseudo-code of De-RBFNs-Bagging

Input:

Dataset D={(a_1, c_1), (a_2,c_2),…., (a_n, c_n}}, where a_i i=1, 2,..,n is a vector of attributes.

Base Classifier and Learning: DE-RBFNs.

Number of Learning Rounds: M

Output: H(a)=arg $Max_c \sum_{i=1}^{M} 1(c = h_i(a))$, the value of 1(z)=1 if z is true, otherwise 0.

Computational Steps:

For i=1 to M

D_i= *Bootstrap(D);*

h_i= *DE-RBFN(D_i);*

End

4. Experimental study

This section describes the dataset, parameter, evaluation criteria for experimental study in addition to results and analysis.

4.1. Description of dataset

The dataset described by Andrzejak et al. in[37], which is publicly available in[38] was used to validate the proposed method. The dataset consists of five sets (denoted as A, B, C, D, and E), each containing 100 single-channel EEG segments of 23.6sec. duration, with sampling rate of 173.6 Hz. After visual inspection for artifacts (such as eye

movement or muscle activity) these segments were selected and cut-out from continuous multi-channel EEG recordings. Segments belongs to sets A and B are taken from surface EEG recording that were carried out on five healthy volunteers by a standardized electrode placement scheme. Volunteers were in states like eyes open (A) and close (B), respectively. Similarly, from an EEG archive of pre-surgical diagnosis, sets C, D, and E are originated. However, in set D, segments were recorded from the epileptogenic zone, and those in set C from the hippocampal formation of the opposite hemisphere of the brain. While sets C and D contained only activity measured form seizure free intervals, set E only contained seizure activity. Based on an average common reference, all EEG signals were recorded with the same 128-channel amplifier system. The data were digitized at 173.61 samples/sec. using 12 bit resolution and they have the spectral bandwidth of the acquisition system that varies from 0.5 Hz to 85 Hz. Figure 1 shows typical EEG segments (one from each of the five described sets).

In this case three different classification problems are created from the above dataset in order to compare the performance of our method with other approaches.

Experiment 1: The aim of this experiment is the diagnosis of epileptic seizure. Here all EEGs from the dataset were used and they are classified into two different classes. Sets like A, B, C, and D is included in non-seizure class and set E is included in the seizure class. This classification has a strong resemblance with the clinical application.

Experiment 2: Like experiment 1, the aim of this is to identify epileptic seizure. In this case two sets are examined for binary classification. Set A is treated as a normal class and set E is treated as a seizure class.

Experiment 3: The aim of this experiment is to classify samples of seizure and non-seizure excluding healthy with eyes closed. In this case four sets from the dataset were used, sets A, C, and D is belongs to non-seizure excluding healthy with eyes closed class and set E belongs to seizure class.

Table 2. Describes dataset of all experiments concisely

Experiment	Classes	Segments	Total Channels
#1 (ABCD-E)	Non-Seizure (ABCD)	400	500
	Seizure (E)	100	
#2 (A-E)	Normal (A)	100	200
	Seizure (E)	100	
#3 (ACD-E)	Non-Seizure Excluding Healthy with Eyes Closed (ACD)	300	400
	Seizure (E)	100	

4.2. Evaluation criteria

Standard performance evaluation criteria in the fields of medical expert systems include accuracy, area under the ROC curve, and type-I and type-II errors. For a two-class problem, most of these metrics can be easily derived from a 2x2 confusion matrix as that given in Table 4, where each entry (i, j) contains the number of seizure or non-seizure samples.

Table 3. Confusion matrix for epileptic seizure identification problem

		Predicted	
		Seizure	Non-Seizure
Actual	Seizure	a	b
	Non-Seizure	c	d

Table 4. Parameters used in the experimental study

Parameter	Experiment #1	Experiment #2	Experiment #3
Maximum Iteration	100	100	100
Population	50	30	40
Mutation	0.2	0.2	0.2
Crossover	0.5	0.5	0.5

In many EEG signal analysis applications often employ the accuracy as the criterion for performance evaluation. It represents the proportion of the correctly predicted cases (Seizure and non-seizure) on a particular dataset. However, empirical and theoretical evidences show that this measure is strongly biased with respect to data imbalance and proportions of correct and incorrect predictions. Therefore, the area under the ROC curve (AUC) has been suggested as an appropriate performance evaluator without regard to class distribution or misclassification costs. The AUC criterion for a binary problem can be defined as the arithmetic average of the mean predictions for each class:

$$AUC = \tfrac{1}{2}(sensitivity + specificity), \tag{7}$$

where sensitivity $= \left(a \big/ (a+b) \right)$ measures the percentage of seizure applicants that have been predicted correctly, whereas specificity $= \left(d \big/ (c+d) \right)$ corresponds to the percentage of non-seizure applicants predicted as non-seizure.

On the other hand, the accuracy ignores the cost of different error types (non-seizure being predicted as seizure, or vice versa). This is the reason why it also becomes especially interesting to measure the error on each individual class by using the type-I and type-II errors:

Type-I Error $\quad \dfrac{c}{c+d}$ $\qquad$ Type-II Error $\quad \dfrac{b}{a+b}$

Type-I error (or miss) is the rate of non-seizure cases being categorized as seizure. When this happens, the misclassified non-seizure cases will undergoes for necessary treatment and creates a heavy financial and health burden of a patient. Type-II error (or false-alarm) defines the rate of seizure applicants being predicted as non-seizure. When this happens, the misclassified seizure cases are refused for further necessary treatment and therefore, the further consequence is severe for a patient. Therefore, if the medical practitioner of a medical institution is too generous, this will be exposed to high risk.

4.3. Parameters

In our experiment, every dataset is divided into two mutually exclusive parts: 2/3 is for training set and 1/3 for test set. The parameters' value used for validating our proposed method is listed in Table 5.

In addition to the tabulated parameters pertaining to DE, we have fixed 5 neurons in the hidden layer and one neuron in the output layer for all experiments. In the case of experiments 1, 2, and 3, our empirical evaluation recommends to use 5, 3, and 4 base learners respectively.

5. Results and analysis

Tables 5-7 shows the confusion matrix obtained during testing of the proposed method. In the case of experiment 1 the number of testing samples is coming round to 167 i.e.,1/3rd of total number of samples whereas it is 67 and 133 for the case of experiment 2 and 3 respectively.

Table 5. Confusion Matrix of Bagging (DE-RBFN) (Experiment 1)

		Predicted	
		ABCD	E
Actual	ABCD	128	2
	E	2	35

Table 6. Confusion Matrix of Bagging (DE-RBFN) (Experiment 2)

		Predicted	
		A	E
Actual	A	33	0
	E	0	33

Table 7. Confusion Matrix of Bagging (DE-RBFN) (Experiment 3)

		Predicted	
		ACD	E
Actual	ACD	97	1
	E	0	35

Similarly, the Type-I and Type-II errors are enumerated in Table 8. In experiment 1, 0.0541 rates of non-seizure cases being categorized as seizure and 0.0154 rates of seizure applicants being predicted as non-seizure. In experiment 2 the error rate of Type-I and Type-II is zero. However, in experiment 3 no seizure cases are misclassified as non-seizure but only one non-seizure case is classified as seizure case. Therefore, the medical practitioner of a medical institution should not be too generous; this will be exposed to high risk.

Table 8. Type I and II errors

Experiment	Type	Error
#1	I	0.0541
	II	0.0154
#2	I	0.0
	II	0.0
#3	I	0.0
	II	0.0102

The comparative performance with other algorithms based on the criterions like specificity and sensitivity are presented in Table 9.

Table 9. Comparative Performance based on Specificity and Sensitivity in Terms of Percentage

Algorithm	Experiment #1		Experiment #2		Experiment #3	
	Sen.	Spec.	Sen.	Spec.	Sen.	Spec.
Proposed Method	98.46	94.59	100	100	98.98	100
Guo, et al. [49]	98.61	94.60	99.40	100	98.55	95.61

(Note: Sensitivity- Sen.; Specificity: Spec.)

Further, it is to be noted from Table 9 that the overall accuracy of the proposed method is improved in the case of experiment 2 and 3 compared to Guo et al.[39] approach, however, it is competitive in experiment 1.

6. Conclusions

In this work, we have developed an ensemble of differential evolution evolved RBFNs for classification algorithm to classify EEG signals. EEG signals were decomposed into sub-bands through the DWT. The basic statistics over the wavelet coefficients are extracted to use as input to the classifier. We have performed three different experiments to obtain the performance of the proposed ensemble in the classifications of normal and seizure segments, non-seizure excluding healthy with eyes closed and seizure, and non-seizure and seizure segments. Based on the various evaluation criteria, we have reported a promising performance. Further, the proposed classification algorithm showed that ensemble approach is better than individual classifier. Hence considering notable experimental results this paper suggests that ensemble can be used as a diagnostic decision support mechanism in the epilepsy treatment.

Acknowledgement

Authors are gratefully acknowledge the support of the Original Technology Research Program for Brain Science through the National Research Foundation (NRF) of Korea (NRF: 2010-0018948) funded by the Ministry of Education, Science, and Technology.

References

1. Garrett D, Peterson D, Anderson C, Thaut M. Comparison of linear and non-linear methods for EEG signal classification. *IEEE Transactions on Neural Systems and Rehabilitative Engineering*, 2003;**11**(2), 141-144.
2. Dingle A, Johns R, Carroll G, Fright W. A multi-stage system to detect epileptiform activity in the EEG. *IEEE Transactions on Biomedical Engineering,* 1993;**40**(12):1260-1268.
3. Niedermeyer E, Lopes da Silva F. *Electroencephalography: Basic principles, clinical applications, and related fields*. Lippincot Williams & Wilkins; 2004.
4. Towle VL, Bolaños J, Suarez D, Tan K, Grzeszczuk R, Levin DN, Cakmur R, Frank SA, Spire JP. The spatial location of EEG electrodes: locating the best fitting sphere relative to cortical anatomy. *Electroencephalogr Clin Neurophysiol*, 2003; **86** (1):1–6.
5. Aurlien H, Gjerde IO, Aarseth JH, Karlsen B, Skeidsvoll H, Gilhus NE. EEG background activity described by a large computerized database. *Clinical Neurophysiology*, 2004; **115** (3): 665–673.
6. Ogulata SN, Sahin C, Erol R. Neural network-based computer aided diagnosis in classification of primary generalized epilepsy by EEG signals. *Journal of Medical Systems*, 2009;**33**:107–112.
7. Adeli H, Zhou Z, Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of Neuroscience Methods*, 2003;**123**:69–87..
8. Subasi A. Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Systems with Applications*, 2005; **28**:701–711.

9. Subasi A. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 2007;**32**:1084–1093.
10. Subasi A, Ercelebi E. Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine*, 2005;**78**:87–99.
11. Hazarika N, Chen JZ, Tsoi AC, Sergejew A. Classification of EEG signals using the wavelet transform. *Signal Processing*, 1997;**59**(1): 61–72.
12. Ubeyli, ED. Combined neural network model employing wavelet coefficients for EEG signals classification. *Digital Signal Processing*, 2009; **19**: 297–308.
13. Ubeyli ED. Decision support systems for time-varying biomedical signals: EEG signals classification. *Expert Systems with Applications*, 2009;**36**: 2275–2284.
14. Alkan A, Koklukaya E, Subasi A. Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods*, 2005;**148**:167–176.
15. Aslan K, Bozdemir H, Sahin S, Ogulata SN, Erol R. A radial basis function neural network model for classification of epilepsy using EEG signals. *Journal of Medical Systems*, 2008;**32**:403–408.
16. Petrosian A, Prokhorov D, Homan R, Dashei R, Wunsch D. Recurrent neural network based prediction of epileptic seizures in intra and extracranial EEG. *Neurocomputing*, 2000; **30**:201–218.
17. Pradhan N, Sadasivan PK, Arunodaya GR. Detection of seizure activity in EEG by an artificial neural network: A preliminary study. *Computers and Biomedical Research*, 1996;**29**:303–313.
18. Srinivasan V, Eswaran C, Sriraam N. Artificial neural network based epileptic detection using time-domain and frequency-domain features. *Journal of Medical Systems*, 2005;**29**(6):647–660.
19. Subasi A. Epileptic seizure detection using dynamic wavelet network. *Expert Systems with Applications*, 2005;29:343–355.
20. Breiman L. Bagging Predictors. *Machine Learning*, 1996;**24**:123–140.
21. Salchenberger LM, Cianr EM, Lash NA. Neural networks: A new tool for thrift failures. *Decision Sciences*, 1992;**23**:899-916.
22. Lee TS, Chiu CC, Chou YC, Lu CJ. Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics and Data Analysis*, 2006;**50**:1113–1130.
23. Dietterich, TG. Experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. *Machine Learning*, 2000;**40**:139-157.
24. Durand, S., Froment, J. 2001, "Artifacts free signal denoising with wavelets," *IEEE, Published in the 2001 International Conference on Acoustics, Speech and Signal Processing*, 6, 3685-3688.
25. Ocak H. Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm. *Signal Processing*, 2008; **88**: 1858–1867.
26. Powell, MJD, 1985. "Radial Basis Functions for Multi-variable Interpolation: A Review," *IMA Conference on Algorithms for the Approximations of Functions and Data*, RMOS Shrivenham, UK.
27. Broomhead, DS, Lowe D. Multivariable functional interpolation and adaptive networks. *Complex systems*, 1988;**2**:321-355.
28. Moody J, Darken CJ. Fast learning networks of locally-tuned processing units. *Neural Computation*, 1989;**6**(4):281-294.
29. Park J, Sandberg JW. Universal approximation using radial basis function networks. *Neural Computation*, 1991;**3**:246-257.
30. Lee S, Kil RM. A Gaussian potential function network with hierarchically self organizing learning. *Neural Networks*, 1991;**4**:207-224.
31. Gomm JB, Yu DL. Selecting radial basis function network centers with recursive orthogonal least squares training. *IEEE Transactions of Neural Networks*, 2000;**11**(3):306–314.
32. Dash CSK, Behera A, Dehuri S, Cho SB. Differential evolution based optimization of kernel parameters in radial basis function networks for classification. *International Journal of Applied Evolutionary Computation*, 2012;**4**(1):56-80.
33. Stron R. System design by constraint adaption and differential evolution. *IEEE Transactions on Evolutionary Computation*, 1999;**3**: 22-34.
34. Das S, Suganthan PN. Differential evolution: A survey of the state-of- the-art. *IEEE Transactions on Evolutionary Computation*, 2011;**15**(1):4-31.
35. Goldberg, DE. *Genetic algorithm in search optimization and machine learning*. Addison-Wesley: Reading: MA; 1989.
36. Kuncheva, LI. Combining classifiers: soft computing solutions. In S. K. Pal and A. Ghosh (Eds.), *Pattern Recognition from Classical to Modern Approaches*, World Scientific, 2003,p. 427-449.
37. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 2001; **64**(6):1-8.
38. EEG time series are available under (http://www.meb.uni-bonn.de/ epileptologie/science/physik/eegdata.html).
39. Guo L, Rivero D, Dorado J, Rabunal JR, Pazos A. Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. *Journal of Neuroscience Methods*, 2010;**191**:101-109.
40 Subha DP, Joseph PK, Acharya R, Lim CM. EEG signal analysis: a survey. *Journal of Medical Systems*, 2010;**34**:195-212.